




DOI: 10.22363/2313-2299-2025-16-3-821-844

EDN: STUQVT

УДК 811.161.1'42-057.87-054.68:372.881.161.1

Научная статья / Research article

Механизм сложности при адаптации текста

М.И. Андреева¹  , М.И. Солнышкина¹ , Сарра Саадна² ¹Казанский (Приволжский) федеральный университет, Казань, Российская Федерация²Российский университет дружбы народов, Москва, Российская Федерация mariia99andreeva@yandex.ru

Аннотация. В работе предложен и апробирован алгоритм расчета и контрастирования лингвистической и когнитивной сложности художественного (первичного) и учебного (вторичного, адаптированного для изучения русского как иностранного) текстов. В качестве материала использована коллекция сопоставляемых текстов, общий объем составил 22232 словоупотребления. Расчет лингвистической сложности текста при помощи Rulingva (rulingva.kpfu.ru) выявил 14 статистически значимых параметров, дифференцирующих контрастируемые пары текстов на трех уровнях: морфологическом (количество двусложных слов, глаголов, существительных, числительных, среднее количество существительных, глаголов и наречий в предложении, «дескриптивность», «нарративность»), лексическом (неповторяющиеся слова, абстрактность) и синтаксическом (читабельность, длина предложения, длина слова в слогах или в символах). Расчет когнитивной сложности анализируемых пар текстов на основе пропозиционального анализа (по В. Кинчу) и последующая нормализация продемонстрировали рост информационной плотности в 7 из 9 пар сравниваемых текстов. Исследование подтвердило гипотезу о том, что адаптированный текст, обладая более низкой степенью лингвистической сложности, сохраняет высокую степень нормализованной информационной плотности, свойственной первичному тексту.

Ключевые слова: русский язык как иностранный, вторичный текст, художественный текст, корпус, лингвистическая сложность, жанр, когнитивная сложность

Вклад авторов: Андреева М.И. — сбор материала, концепция и дизайн исследования, осуществление исследования, написание статьи; Солнышкина М.И. — руководство исследованием, анализ полученных данных, написание статьи, Сарра Саадна — проведение исследования, обработка данных.

Финансирование: исследование выполнено за счет гранта Российского научного фонда № 24-78-10129, <http://rscf.ru/project/24-78-10129/>.

Заявление о конфликте интересов: авторы заявляют об отсутствии конфликта интересов.

История статьи: дата поступления: 28.03.2024; дата приема в печать: 15.04.2025.

Для цитирования: Андреева М.И., Солнышкина М.И., Саадна С. Механизм сложности при адаптации текста // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2025. Т. 16. № 3. С. 821–844. <https://doi.org/10.22363/2313-2299-2025-16-3-821-844>

© Андреева М.И., Солнышкина М.И., Саадна С., 2025



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

Complexity Mechanism in Leveled Texts

Mariia I. Andreeva¹ , Marina I. Solnyshkina¹ , Sarra Saadna² 

¹Kazan (Volga Region) Federal University, Kazan, Russian Federation

²RUDN University, Moscow, Russian Federation

✉ mariia99andreeva@yandex.ru

Abstract. The paper presents the research conducted to assess and contrast linguistic and cognitive complexity of primary (fiction) and secondary (leveled to teach Russian as a foreign language) texts. The research dataset comprises dataset of texts with the total amount of 22,232 tokens. Measurement of linguistic complexity performed with Rulingva (rulingva.kpfu.ru) revealed 14 statistically significant parameters discriminating the primary and secondary (leveled) texts. The parameters-discriminators fall into three groups: morphological (the number of two-syllabic words, verbs, nouns, numerals, the average number of nouns, verbs and adverbs per sentence, ‘descriptivity’, ‘narrativity’), lexical (hapax legomena, Abstractness) and syntactic (readability, sentence length, word length in syllables or in characters). Assessment of cognitive complexity of the contrasted pairs of texts included the two-step algorithm: propositional analysis and information density measurement. The propositional analysis (by W. Kintsch) was aimed at identifying and calculating the number propositions, while the information density measurement implied normalization of the number of propositions by the length of the text (In words) followed by the contrastive analysis. The study revealed increase in the normalized information density in 7 out of 9 leveled texts and confirmed the hypothesis that though leveled texts have a lower degree of linguistic complexity they retain a higher degree of normalized information density inherent in the primary text.

Keywords: Russian as a foreign language, secondary text, fiction text, corpus, linguistic complexity, genre, cognitive complexity

Authors’ contribution: Andreeva M.I. — data collection, research idea and design, research implementation, writing the article; Solnyshkina M.I.— research supervision, obtained data analysis, writing the article, Sarra Saadna — research implementation, processing of the data.

Funding: the research was supported by the Russian Science Foundation, grant 24-78-10129, <http://rscf.ru/project/24-78-10129/>.

Conflicts of interest: The authors declare no conflict of interest.

Article history: received: 28.03.2024; accepted: 15.04.2025.

For citation: Andreeva, M.I., Solnyshkina, M.I., & Saadna, S. (2025). Complexity Mechanism in Leveled Texts. *RUDN Journal of Language Studies, Semiotics and Semantics*, 16(3), 821–844. (In Russ.). <https://doi.org/10.22363/2313-2299-2025-16-3-821-844>

Введение

Всплеск интереса к вопросам понимания сопровождается в настоящее время разработкой многочисленных подходов к оценке сложности или меры трудности текста [1], осуществляемой как автоматически, т.е. на основе валидированных лингвистических параметров [2], так и на основе экспертных оценок [3–5]. Целью обеих процедур является обеспечение языковой доступности текста, поскольку в противном случае утрачивается цель коммуникации, а воздействие текста может становиться деструктивным [6].

Среди объектов такого рода исследований учебный текст занимает особое место в силу его полифункциональности, т.е. совмещения информативной, культурологической и развивающей функций [7]. Важное значение в данном аспекте приобретает использование адаптированных художественных текстов для учебных целей в практике преподавания русского языка как иностранного (далее РКИ), поскольку именно художественные тексты, обеспечивая изучение системы языка, представляют материал для познания норм коммуникации в различных ситуациях [8]. При этом ученые признают, что алгоритмы оценки сложности вторичного текста, имея высокую значимость для прикладной лингвистики и лингводидактики, продолжают оставаться в стадии разработки [9].

Представленное исследование осуществлено с целью выявления механизмов лингвистической адаптации (преобразования) текста, создающих предпосылки для адекватного восприятия и понимания текста определенной целевой аудиторией. На примере художественных и адаптированных экспертами текстов, используемых в практике преподавания РКИ, тестируется гипотеза о том, что учебный текст, преобразованный на основе художественного, обладая более низкой степенью лингвистической сложности, сохраняет высокую степень когнитивной сложности, свойственной первичному тексту.

Обзор литературы

Текст: художественный vs учебный текст

Обладая информационной избыточностью и занимая «промежуточное положение между деловым стилем и устно-разговорной речью: нижняя граница равна 76,3 %, а верхняя — 82,6 %» [10], художественный текст выступает в качестве пропозиционально-логического инварианта [11], на основе которого создаются многочисленные варианты — вторичные тексты, обладающие разной пропозициональной структурой и иллокутивной силой [12].

Адаптация учебного текста, трактуемая как «упрощение, приспособление, облегчение <...> текста в соответствии с уровнем языковой компетенции обучающихся» [13], манифестируется в трех основных механизмах: свертывании (сужении, сокращениях, компрессии, редукции), развертывании (расширении) и замене единиц текста [14]. Таким образом, преобразование художественного текста для учебных целей осуществляется за счет количественных и качественных изменений морфологических, лексических, синтаксических и дискурсивных параметров.

Традиционно считается, что в отличие от переводческих трансформаций, нацеленных на сохранение уровня лингвистической и смысловой сложности оригинала [15; 16], трансформации, осуществляемые при адаптации текста для учебных целей, нацелены на снижение уровня лингвистической сложности [17–21].

Сложность текста. Научная традиция современной парадигмы дискурсивной комплексологии предполагает оценку двух типов сложности: лингвистической, манифестируемой в языковых единицах всех уровней текста, и когнитивной (информационной), трактуемой как совокупность воздействующих на читателя элементов содержания текста. Лингвистическая и когнитивная «сложности» печатного текста суть два аспекта одной категории, поскольку когнитивная сложность манифестируется и видоизменяется исключительно через языковые маркеры.

Лингвистическая сложность текста. Способы и методы оценки лингвистической сложности текста хорошо описаны в отечественной и зарубежной лингвистике в рамках параметрического подхода или при помощи нейронных сетей [22]. Автоматическая оценка лингвистической сложности текста позволяет не только избежать субъективности экспертизы, осуществляемой специалистом, но и призвана обеспечить (1) список языковых единиц, отсутствующих в словарном запасе целевой аудитории, а также (2) варианты их синонимических замен [23]. Имея данные о референсных диапазонах метрик синтаксических и морфологических параметров текста, современные анализаторы оценивают, насколько текст соответствует установленным нормам языковой сложности того или иного уровня [24; 25]. Особый вес и высокую предиктивную силу при оценке морфологической сложности текста, имеют доли частей речи: чем выше доля глаголов, тем выше его нарративность и тем легче воспринимается текст [25], в то время как высокая доля имен существительных, наоборот, свидетельствует о высокой мере трудности восприятия текста [26]. Установлено, что лексическое разнообразие учебных текстов колеблется в диапазоне 0.3–0.6, а в случаях, когда значение превышает данный уровень, текст не обеспечивает достаточной когезии и лексических повторов, а понимание усложнено [26; 27].

Когнитивная сложность текста. Оценка когнитивной сложности текста как объема передаваемой текстом информации [28; 29] предполагает анализ смыслового пространства текста, его информационного единства, информационной плотности или компрессивности [30–32], а также способности текста к «уплотнению информации» [33]. Существующие подходы к когнитивной сложности трактуют данную категорию как конструкт [34], оценка которого осуществляется основе двух основных подходов: личностно-ориентированного и текстоцентричного.

Методы, разработанные в рамках *личностно-ориентированного подхода*, предполагают оценку готовности и/или способности читателя (категорий читателей) к восприятию текста и косвенно связаны с информативностью текста для читателя, т.е. «степенью / мерой ожидаемости / неожиданности, известности / неизвестности предъявляемых адресату смысловых образований» [35. С. 7]. Методы в рамках данного подхода, в частности, включают (а) индивидуальную оценку читателями нагрузки, возникающей при

чтении [36], (б) расчет темпа выполнения различных видов интеллектуальной деятельности [37], (в) ожидаемый от субъекта уровень «глубины знаний» по Уэббсу [38]; (г) таксономию когнитивных процессов по Б. Блуму [39; 40].

Текстоцентрические методы оценки когнитивной сложности ориентированы на оценку содержащейся в тексте информации, его информационной насыщенности, «семантического веса» [41] или «напряженности» как абсолютного показателя текста [42]. Текстоцентрические методы предполагают либо применение стилостатистических приемов, включающих оценку языковых маркеров когнитивной нагрузки, валидированных в более ранних исследованиях [43], либо с использованием информационно-интеллектуальных технологий, в том числе «машинное/глубокое обучение», применение нейронных сетей, трансформеров и проч. [34; 44].

Оценка изменения / сохранения когнитивной сложности учебного текста и его оригинала может быть осуществлена на основе пропозиционального анализа обоих вариантов текста. Методология такого рода анализа была разработана В. Кинчем и Т.А. ван Дейком [45], а позднее успешно валидирована в серии психолингвистических исследований [46; 47]. Пропозициональный анализ в описательных и контрастивных исследованиях репрезентации текста активно используется в современной когнитивной психологии и доказал свою валидность как метод [48. С. 82]. При оценке понимания текста респондентом пропозициональный анализ применяется для оценки объема понимания текста и имеет в качестве обязательного этапа выявление отношения количества и качества (состава) воспроизведенных пропозиций к исходным пропозициям в тексте. Согласно модели понимания В. Кинча, пропозиция трактуется как основная содержательная единица текста и наименьшая единица знания, выделяемая в высказывании [49], вербализация пропозиции осуществляется предикатом и как минимум одним аргументом [50; 51]. Пропозициональный анализ текста оригинала и его адаптированной для учебных целей версии позволяет оценить информационную плотность каждого из них для последующего контрастирования (сравнения).

Материал и методы

Материал исследования. Исследование осуществлено на материале двух коллекций текстов: (1) первичных текстов — фрагментов художественных произведений русских авторов (n=12) и (2) вторичных — адаптированных для преподавания русского как иностранного (РКИ) версий художественных произведений (n=12) (см. Таблицу 1). Общий объем корпуса составил 22232 словоформ. Первичные тексты — фрагменты художественных известных российских писателей: А.П. Чехова (Тексты О_1¹ — О_5,

¹ В скобках указан код текста в авторском корпусе исследования. Подробнее см. Таблицу 1.

О_8, О_10), В.П. Астафьева (О_6), И.А. Бунина (О_7), И.С. Тургенева (О_9), Ф.М. Достоевского (О_11), В.В. Набокова (О_12).

Источником всех изучаемых текстов РКИ послужили учебные пособия, используемые в практике преподавания РКИ. В качестве критериев отбора использовались следующие: (1) вторичность, т.е. наличие первичного варианта, художественного произведения; (2) уровень сложности, заявленный авторами учебника, в диапазоне А1 — В2; (3) симметрия заданий к тексту, обеспечивающая «изучающий» вид чтения и установку читателя на полное понимание, запоминание и дальнейшее использование текстовой информации [52; 53]. Исследовательский корпус представлен в Таблице 1.

Таблица 1 / Table 1

Объем корпуса исследования / Volume of the Research Case

Коллекция первичных текстов / Corpus of primary texts	Коллекция вторичных текстов (РКИ) / Corpus of secondary RFL texts
О_1 ² (668) ³	A2_1 ⁴ (421)
О_2 ⁵ (1096)	A2_2 ⁶ (583)
О_3 ⁷ (903)	A2_3 ⁸ (652)
О_4 ⁹ (991)	A2_4 ¹⁰ (760)
О_5 ¹¹ (1363)	A2_5 ¹² (594)

² Чехов А.П. Злой мальчик. Режим доступа: https://royallib.com/book/chehov_anton/zloy_malchik.html (дата обращения 04.06.2025).

³ В скобках указан объем текста в словоформах.

⁴ Типовые тесты по русскому языку как иностранному. Базовый уровень. Общее владение / В.Е. Антонова и др. М. : ЦМО МГУ; СПб. : Златоуст, 2007. Чехов А.П. Злой мальчик.

⁵ Чехов А.П. Месть. Режим доступа: https://royallib.com/book/chehov_anton/mest.html (дата обращения 04.06.2025).

⁶ Типовые тесты по русскому языку как иностранному. Базовый уровень. Общее владение / В.Е. Антонова и др. М. : ЦМО МГУ; СПб. : Златоуст, 2007. Чехов А.П. Месть.

⁷ Чехов А.П. Глупый француз. Режим доступа: https://royallib.com/book/chehov_anton/glupiy_frantsuz.html (дата обращения 04.06.2025).

⁸ Финагина Ю.В. Русский язык как иностранный. Пособие по чтению / под ред. Н.А. Дмитренко. СПб. : НИУ ИТМО; ИХиБТ, 2014. Чехов А.П. Глупый француз.

⁹ Чехов А.П. История одного торгового предприятия. Режим доступа: https://royallib.com/book/chehov_anton/istoriya_odnogo_torgovogo_predpriyatiya.html (дата обращения 04.06.2025).

¹⁰ Финагина Ю.В. Русский язык как иностранный. Пособие по чтению / под ред. Н.А. Дмитренко. СПб. : НИУ ИТМО; ИХиБТ, 2014.

¹¹ Чехов А.П. На даче. Режим доступа: <https://ilibrary.ru/text/4211/p.1/index.html> (дата обращения 04.06.2025).

¹² Финагина Ю.В. Русский язык как иностранный. Пособие по чтению / под ред. Н.А. Дмитренко. СПб. : НИУ ИТМО; ИХиБТ, 2014.

Окончание таблицы 1 / Ending of the Table 1

Коллекция первичных текстов / Corpus of primary texts	Коллекция вторичных текстов (РКИ) / Corpus of secondary RFL texts
O_6 ¹³ (892)	B2_6 ¹⁴ (487)
O_7 ¹⁵ (552)	B2_7 ¹⁶ (493)
O_8 ¹⁷ (495)	B2_8 ¹⁸ (464)
O_9 ¹⁹ (599)	B2_9 ²⁰ (441)
O_10 ²¹ (550)	B2_10 ²² (494)
O_11 ²³ (3279)	B2_11 ²⁴ (872)
O_12 ²⁵ (3094)	B2_12 ²⁶ (1489)
TOTAL: 14482	TOTAL: 7750

Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна. / Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

¹³ Астафьев В.П. Ода русскому огороду. Режим доступа: https://librebook.me/oda_russkomu_ogorodu/vol1/1 (дата обращения 04.06.2025).

¹⁴ Учебно-тренировочные тесты по русскому языку как иностранному. Вып. 2. Чтение / под общ. ред. М.Э. Парецкой. СПб. : Златоуст; Ростов-на-Дону : Южный федеральный университет, 2009.

¹⁵ Бунин И.А. Поздний час. Режим доступа: <https://ilibrary.ru/text/1816/p.1/index.html> (дата обращения 04.06.2025).

¹⁶ Учебно-тренировочные тесты по русскому языку как иностранному. Вып. 2. Чтение / под общ. ред. М.Э. Парецкой. СПб. : Златоуст; Ростов-на-Дону : Южный федеральный университет, 2009.

¹⁷ Чехов А.П. Дачники. Режим доступа: https://royallib.com/book/chehov_anton/dachniki.html (дата обращения 04.06.2025).

¹⁸ Учебно-тренировочные тесты по русскому языку как иностранному. Вып. 2. Чтение / под общ. ред. М.Э. Парецкой. СПб. : Златоуст; Ростов-на-Дону : Южный федеральный университет, 2009.

¹⁹ Тургенев И.С. Дворянское гнездо. Режим доступа: https://royallib.com/book/turgenev_ivan/dvoryanskoe_gnezdo.html (дата обращения 04.06.2025).

²⁰ Учебно-тренировочные тесты по русскому языку как иностранному. Вып. 2. Чтение / под общ. ред. М.Э. Парецкой. СПб. : Златоуст; Ростов-на-Дону : Южный федеральный университет, 2009.

²¹ Чехов А.П. Ионыч. Режим доступа: https://royallib.com/book/chehov_anton/ionich.html (дата обращения 04.06.2025).

²² Чехов А.П., Николаева Ю.В. Ионыч. Библиотека Златоуста. Режим доступа: http://www.litres.ru/pages/biblio_book/?art=9188620&lfrom=204420127. А.П. Чехов. Ионыч. Санкт Петербург : Златоуст, 2015.

²³ Достоевский Ф.М. Идиот. Режим доступа: https://royallib.com/book/dostoevskiy_fedor/idiot.html (дата обращения 04.06.2025).

²⁴ Достоевский Ф.М. Идиот (адаптированный текст) Библиотека Златоуста http://www.litres.ru/pages/biblio_book/?art=8871474&lfrom=204420127 Фёдор Достоевский. Идиот. Роман в 4 частях. Санкт Петербург : Златоуст, 2015. Достоевский Ф.М. Идиот.

²⁵ Набоков В.Д. Звонок. Режим доступа: https://royallib.com/book/nabokov_vladimir/zvonok.html (дата обращения 04.06.2025).

²⁶ Казакова О.А., Фрик Т.Б. Практикум по культуре речевого общения на русском языке: учебное пособие. Томск: Изд-во Томского политехнического университета, 2009.

Алгоритм исследования включает пять этапов: (1) сбор корпуса исследования, в состав которого входят две коллекции текстов исследования: (а) оригинальные (первичные) тексты фрагментов повестей и романов классиков русской литературы (n=12) и (б) адаптированные (вторичные) тексты из учебников по РКИ уровней А2 и В2 (n=12); (2) расчет значений лингвистических параметров текстов каждой из коллекций при помощи автоматического анализатора RuLingva²⁷; (3) статистический анализ метрик вышеуказанных оппозиций для выявления статистически значимых различий между контрастируемыми текстами (по Вилконсону) при помощи программы STATISTIKA; (4) оценка когнитивной сложности оригинальных (первичных) и адаптированных (вторичных) текстов при помощи пропозиционального анализа и расчета абсолютной и нормализованной информационной плотности текстов; (5) контрастирование индексов когнитивной сложности и выявление (а) симметрии когнитивной сложности изучаемых оригинальных и вторичных (адаптированных) текстов.

128 лингвистических параметра, рассчитываемых при помощи RuLingva, включают параметры четырех кластеров: дескриптивные, морфологические, лексические и дискурсивные с грануляцией на (1) документ; (2) предложение; (3) абзац и (4) 1000 словоформ²⁸. Список параметров включает, но не ограничивается следующими: среднее количество слов в предложении, среднее количество слогов в слове, среднее количество знаков в слове, существительные, среднее количество существительных в предложении, глаголы, среднее количество глаголов в предложении, среднее количество прилагательных в предложении, местоимения, числительные, лексическое разнообразие, родительный падеж существительных, творительный падеж существительных, прошедшее время глаголов, соотношение глаголов и существительных, количество односложных слов, количество двусложных слов, количество трехсложных слов, количество четырехсложных слов, среднее количество наречий в предложении гапакс легомена²⁹. Отдельно оцениваются два индекса читабельности текста по двум формулам: (а) по формуле F(SIS) — читабельность нехудожественных текстов:

$$(FKG (SIS) = 0,063 \times \text{средняя длина предложения} + 2,0 \times \\ \times \text{средняя длина слов в слогах} — 2,3),$$

а также по формуле И.В. Оборновой — для художественных текстов:

²⁷ RuLingva. Режим доступа: <https://rulingva.kpfu.ru/> (дата обращения 04.06.2025).

²⁸ При сегментации объемных текстов длина отрывков может колебаться в пределах от 985 до 1010, поскольку RuLingva сохраняет неразрывность предложения.

²⁹ Напax legomena суть слова, которые появляются в тексте (корпусе текстов) только один раз: Fayzullayev Sh., Kovalevskii A. Напax legomena via stochastic processes // Glottometrics. 2024. Vol. 56. P. 22–39. https://doi.org/10.53482/2024_56_415

$$\begin{aligned} FK(Obor) &= 206,836 — (1,52 \times \text{средняя длина предложения}) — \\ &— (65,14 \times \text{среднее число слогов}) [54]. \end{aligned}$$

Для сравнения параметров первичных и вторичных текстов использовался Т-критерий Вилкоксона, валидированный для сравнения показателей, измеряемых в различных вариантах одного и того же текста, и позволяющий установить эффективность воздействия адаптации на оригинальный текст [55]. В представленном исследовании применение Т-критерия Вилкоксона значимо для оценки спектра изменений в каждой из пар О (оригинальный, первичный) — РКИ (учебный, вторичный, используемый в РКИ) текстов.

Когнитивная сложность текстов оценивается при помощи пропозиционного анализа. Парадигма структурного подхода в рамках семантического синтаксиса выделяет главные пропозиции и субпропозиции, классифицируемые на актанты и сирконстанты [56]. Оценка главных пропозиций осуществляется на основе предикативных структур, вербализуемых глаголами, причастиями, деепричастиями, отглагольными существительными и прилагательными. Таксономия актантов включает субъект, объект, адресат и инструмент, а сирконстанты указывают местоположение, время, причину действий [56]. Пропозициональная сложность относится к числу идей, закодированных говорящими для передачи данного сообщения [57; 58]. Таким образом, чем больше идейных единиц закодировано при повествовании истории, тем выше ее пропозициональная сложность.

Анализ

На *Этапе I* первичные (художественные, оригинальные, далее — О) и вторичные (учебные, используемые в практике РКИ, далее — РКИ) тексты после сканирования и извлечения из источников прошли предобработку: из них были удалены символы, контаминирующие однородность дискурса, например, сноски, иноязычные «острова», объем которых варьируется от морфемы до фразы, пробелы, были исправлены опечатки. Метаразметка документов в корпусе предусматривает обязательность включения следующих блоков информации: автор, название произведения, номер текста в корпусе, тип текста (О vs РКИ). Корпус хранится на внутреннем репозитории Казанского федерального университета, доступ к нему предоставлен ограниченному числу научных сотрудников для выполнения исследовательских работ.

На *Этапе II* каждый из текстов корпуса был подвергнут лингвистическому профилированию при помощи сервера RuLingva. Для сравнительного анализа текстов разной длины использовались нормализованные значения на 1000 словоформ (ipf), рассчитываемые по формуле $F_{норм} = \text{метрика} / \text{количество словоформ в отрывке} * 1000$.

На *Этапе III* с помощью программы STATISTIKA был осуществлен статистический анализ метрик параметров сравниваемых текстов (см. Таблица 2).

Таблица 2 / Table 2

Статистически значимые параметры-дискриминанты лингвистической сложности текстов О и РКИ (по Вилконсону) / Statistically significant discriminant parameters of linguistic complexity of O and RKI texts (according to Wilconson)

№	Параметры / Parameters	p-value
1	Среднее количество слов в предложении / Average number of words per sentence	0,003*
2	Среднее количество слогов в слове / Average number of syllables per word	0,017*
3	Среднее количество символов в слове / Average number of characters per word	0,041*
4	Существительные / Nouns	0,583
5	Среднее количество существительных в предложении / Average number of nouns per sentence	0,003*
6	Глаголы / Verbs	0,022*
7	Среднее количество глаголов в предложении / Average number of verbs per sentence	0,028*
8	Среднее количество прилагательных в предложении / Average number of adjectives per sentence	0,004*
9	Числительные / Numericals	0,023*
10	Индекс FKG (Obor) / Flesh-Kincaid score (Obor)	0,003*
11	Индекс абстрактности / Abstractness score	0,002*
12	Отношение глаголов к существительным / Verb / Noun ratio	0,037*
13	Двусложные слова / Two-syllable words	0,023*
14	Среднее количество наречий в одном предложении / Average number of adverbs per sentence	0,008*
15.	Среднее количество гапакс легомена в предложении / Hapax legomena per sentence	0,002*

Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна. / Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna

Параметры, обладающие статистической значимостью ($<0,050$), маркированы*

Таблица 2 содержит 15 статистически значимых параметров, дифференцирующих исследуемые группы текстов.

К ним относятся следующие кластеры параметров:

Морфологические параметры-дискриминанты: существительные, среднее количество существительных в предложении, глаголы, среднее количество глаголов в предложении, среднее количество прилагательных в предложении, числительные, отношение глаголов к существительным, среднее количество наречий в одном предложении. Проиллюстрируем динамику изменений в рассматриваемых текстах на примере количества глаголов (см. Рис. 1).

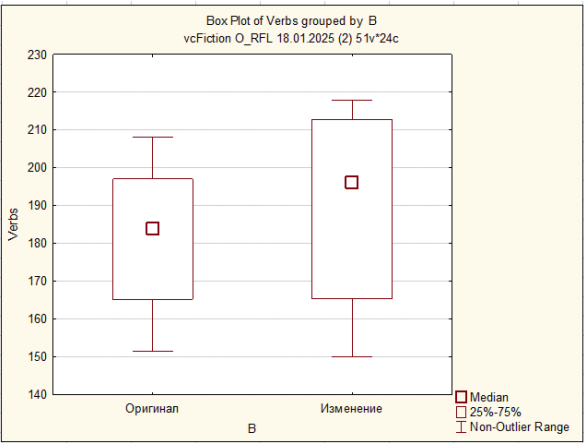


Рис. 1. Количество глаголов в текстах О — РКИ (уровней A2 и B2)
Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна.
Fig. 1. The number of verbs in original and RFL (levels A2, B2) texts
Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

Как видим, количество глаголов снизилось за счет внесенных в текст лексических изменений, как правило, связанных с заменами на менее частотные глаголы (*назначен* (O_10) → *приехал работать* (B2_10), а также замены наречий или адвербиальных комплексов на глагольные комплексы: *впрочем* (O_5) → *может быть* (A2_5). Например, *поселился* (O_10) → *стал жить* (B2_10), *сходились* (O_10) → *начали приходить* (B2_10), *молчать нет сил* (O_5) → *не могу молчать* (A2_5).

Аналогичным образом изменяется дескриптивность, т.е. соотношение прилагательных к существительным, в текстах РКИ, созданных на основе оригинальных художественных произведений (см. рис. 2).

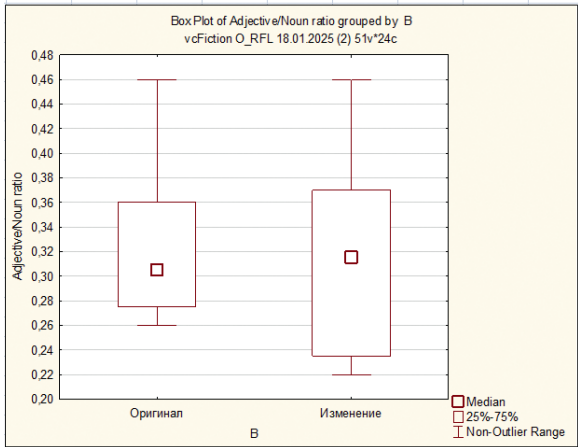


Рис. 2. Дескриптивность, т.е. соотношение прилагательных к существительным в текстах О и РКИ
Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна .
Fig. 2. Descriptivity, adjectives/noun ratio in original and RFL (levels A2, B2) texts
Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

Например, *досуг* (O_10) → *свободное время* (B2_10), *душенька* (O_10) → *дорогая* (B2_10), *смеющиеся глаза* (O_10) → *весёлые глаза* (B2_10), *приятеля* (O_2) → *хорошего знакомого* (A2-2), *индюк* (O_2) → *старый дурак муж* (A2-2), *душа моя* (O_2) → *моя милая* (A2-2), *твой пузан* (O_2) → *твой старый толстый муж* (A2-2), *вдруг он остановился и хлопнул себя по лбу* (O_2) → *наконец у него появилась отличная идея* (A2-2), *улыбался* (O_2) → *прекрасное настроение* (A2-2), *индюка да Собакевича* (O_2) → *старый глупый дурак* (A2-2), *Дегтярев* (O_2) → *молодой человек* (A2-2).

В некоторой степени меняется и нарративность, рассчитываемая на основе соотношения глаголов к существительным в текстах О и РКИ (Рис. 3). Как было указано (Рис. 1), количество глаголов в текстах РКИ увеличивается, снижение нарративности как соотношение глаголов к существительным (см. Рис. 3), очевидно, происходит за снижения количества существительных.

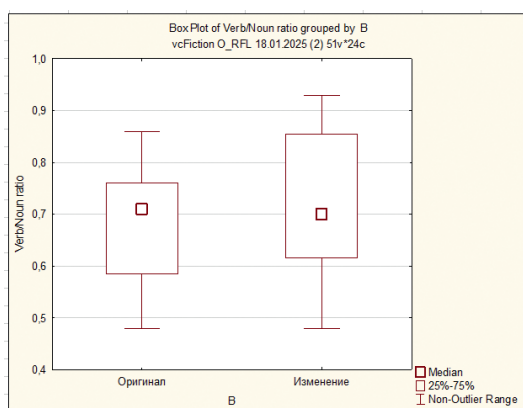


Рис. 3. Нарративность, соотношение глаголов к существительным в текстах О и РКИ
Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна.

Fig. 3. Narrativity, verb/noun ratio in original and RFL texts
Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

Снижение количества существительных объясняется в том числе их элиминацией. Например, (1) Дай еще икры! крикнул сосед, утирая **салфеткой** масленые губы. (O_3) → Дай ещё икры! — крикнул сосед, вытирая масленые губы (A2_3); (2) Может быть, это мне снится? изумился клоун, откидываясь на **спинку** стула. (O_3) → «Может быть, это мне снится? — удивился клоун (A2_3).

Лексическая адаптация текста. Свертывание (сужения, сокращения, компрессия, редукции)

Анализ текстов оригиналов и вторичных текстов РКИ выявил многочисленные лексические замены единиц с низкой частотностью и высоким уровнем сложности на более частотные синонимы.

Отдельного внимания заслуживает замена числительных и использование цифр в текстах РКИ. Например, *в шесть часов* → *в 6 часов*, *двести рублей* → *200 рублей*. Очевидно, это связано со стремлением снизить когнитивную трудность восприятия текста.

Тексты РКИ практически не содержат фразеологических единиц, которые потенциально могут быть особо сложными для понимания. В качестве замен авторы-составители предлагают слова и словосочетания более высокой частотности. Например, *Он шел пешком, не спеша (своих лошадей у него еще не было) и все время напевал: Когда еще я не пил слез из чаши бытия...* (O_10) → *Он шёл пешком (своих лошадей у него ещё не было), не спешил и всё время пел* (B2_10); *скалит зубы* (O_2) → *улыбается* (A2_2); *ваша галантерейная лавка взлетит на воздух* (O_2) → *ваш магазин разрушат* (A2_2).

На *Этапе IV* при помощи пропозиционального анализа была осуществлена оценка когнитивной сложности оригинальных (первичных) и адаптированных (вторичных) текстов, а также нормализация данных (% в Таблице 3). Нормализация количества пропозиций выполнялась для исключения влияния длины текста на результат расчетов (см. Таблица 3, выделено в серой заливке).

Таблица 3 / Table 3

Абсолютное и нормализованное (%) количество главных пропозиций в текстах О и РКИ / Absolute and normalized (%) number of main propositions in O and RKI texts

Тексты / Text	Глагол / Verb	Модал. Глагол / Modal verb	Прич. / Past Partic.	Деепр. / Pres. Partic.	Кач. прилаг. / Adj.	Отглагол. сущ. / Verbal noun	Адвер. Обор. / Adv.	Сущ. / Noun	Имен. констр. / Nom. phrase
O_1↔A2_1	119↔87	7↔2	2	7	7↔2	15↔1	7↔2	2	
%	17,8↔20,7	1↔0,5			1↔0,5	2,2↔0,2	1↔0,5		
O_2↔A2_2	192↔111	10↔6	4	20↔1	11↔6	12↔1	15↔6		
%	17,5↔19	0,9↔1		1,8↔0,2	1↔1	1,1↔0,2	1,3↔1		
O_3↔A2_3	154↔125	17↔14	4↔3	15↔6	21↔12	1↔1	6↔4	18↔14	15↔11
%	17↔19,1	1,9↔2,1	0,4↔0,5	1,7↔1	2,3↔1,8	0,1↔0,2	0,7↔0,6	2↔2,1	1,7↔1,7
O_4↔A2_4	143↔120	9↔9	8↔8	8↔4	1↔1	8↔7	5↔5	7↔7	17↔9
%	14,4↔15,8	0,9↔1,1	0,8↔1	0,8↔0,5	0,1↔0,1	0,8↔0,9	0,5↔0,7	0,7↔0,9	1,7↔1,1
O_5↔A2_5	241↔107	14↔7	2↔2	23↔13	25↔8	19↔8	5↔4		
%	17,7↔18	1↔2,3	0,1↔0,3	1,7↔2,1	1,8↔1,3	1,4↔1,3	0,4↔0,7		
O_10↔B2_10	91↔86		1↔1	5↔2	8↔8	7↔5		5↔3	21↔19
%	16,5↔17,4		0,2↔0,2	0,9↔0,4	1,5↔1,6	1,3↔1		0,9↔0,6	3,8↔3,8
O_11↔B2_11	157↔126	18↔11	13↔9	7↔3	30↔25	3↔3	20↔12	34↔24	57↔43
%	12,3↔14,4	1,4↔1,3	1↔1	0,5↔0,3	2,3↔2,9	0,2↔0,3	1,6↔1,4	2,6↔2,8	4,5↔4,9
O_12↔B2_12	363↔296	19↔19	5	29↔13	20↔9	46↔26	16↔15		25↔13
%	16,1↔19,8	0,8↔1,2		1,3↔0,8	0,9↔0,6	2↔1,7	0,7↔1		1,1↔0,9

Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна. / Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

Наибольшей редукции подверглись сирконстанты оригинального текста. Например, инструмент — *посылать письма по почте* (O_2) → *посылать письма* (A2_2), экспериенсер — *Лев Саввич чувствовал себя* (O_2), время — *Вечно, дуралей, проигрывает, Время, когда он <...> давно уже прошло* (O_2), образ действия — *на цыпочках, отечески похлопал по плечу* (O_2), подняла вверх руку (O_1) → *подняла руку* (A2_1), *как-то нечаянно схватил руку* (O_1) → *случайно взял руку* (A2_1). Например, Возвратясь домой, **Лев Саввич чувствовал себя злым и неудовлетворенным** (O2) → Когда Турманов вернулся домой, он стал думать, как отомстить Дегтярёву (A2_2); **Поднявшись, он на цыпочках, солидно покачиваясь, пробрался между столов, прошел через гостиную, где танцевала молодежь <...> Но тут за стеной слышались голоса.** (O_2) → Он подошёл к двери в другую комнату и вдруг услышал разговор (A2_2).

При адаптации также были сокращены или элиминированы модификаторы актантов. Например, *прочими рыболовными принадлежностями, очень многое* (O_1), *какой-нибудь фокус, какой фортель я придумал* (O_2), *у меня у самого* (O_3) → *у меня* (A2_3), *пять бараньих котлет* (O_3) → *пять котлет* (A2_3). Например, (1) У меня у **самого** в Бретани был дядя Франсуа, который на пари съедает две тарелки супу и пять **бараньих** котлет... (O3) → У меня в Бретани был дядя Франсуа, который на пари съедает две тарелки супу и пять котлет...; (2) Я должен сказать вам многое, Анна Семеновна... **Очень** многое... (O_1) → Я должен сказать вам многое, Анна Семёновна... (A2_2).

Синтаксические параметры. Для сравнения синтаксических параметров оригинальных и адаптированных текстов были рассмотрены следующие параметры: длина предложения, читабельность, длина слов (см. Рис. 4 и Рис. 5).

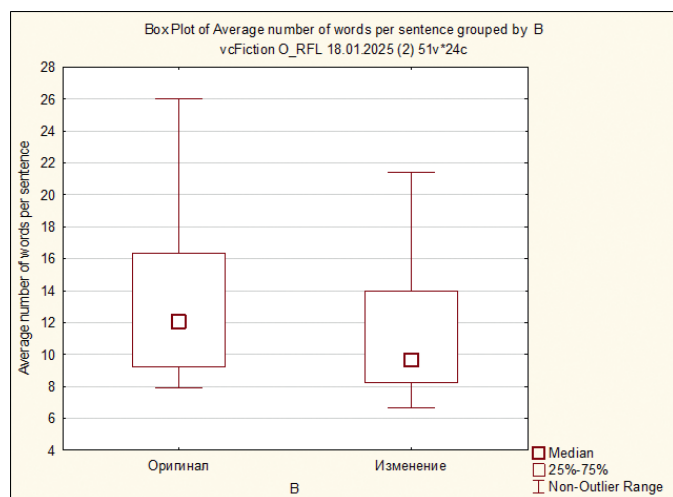


Рис. 4. Средняя длина предложения (в словах) текстов О и РКИ
Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна.

Fig. 4. Average sentence length (In words) of original and RFL texts
Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

Рис. 4. наглядно демонстрирует более высокий диапазон метрик длины предложения, т.е. предложения оригинального художественного текста длинее, чем адаптированных текстов РКИ (ср. 9–16 слов vs 8–14 слов)

При адаптации текста изменяется и его читабельность: в целом для адаптированного текста характерны менее низкий уровень длины предложения (см. Рис. 5).

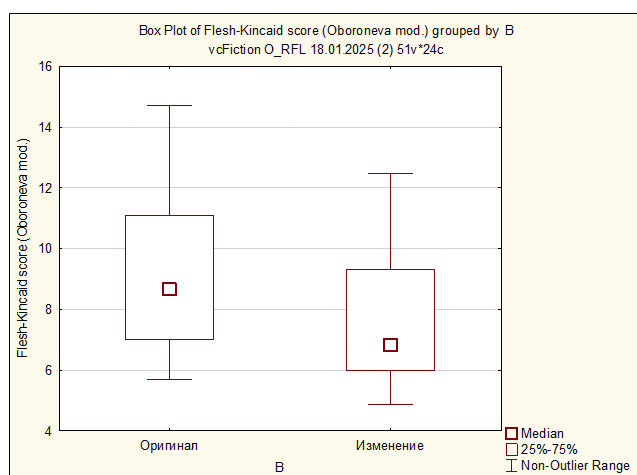


Рис. 5. Индекс читабельности текстов О и РКИ
 Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна.

Fig. 5. Readability index of original and RFL texts
 Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

Преимущественное большинство текстов в коллекции имеют читабельность 6–11. Видим, что адаптированные тексты РКИ имеют более низкий индекс читабельности ($FKGL(Ob) = 6$). Для оригинальных художественных текстов как нижняя ($FKGL(=Ob) = 7$), так и верхняя границы ($FKGL(Ob) = 11$) немного выше. (см. Рис. 5).

Оценка когнитивной сложности и информационной плотности

На *Этапе V* осуществлялось контрастирование данных когнитивной сложности изучаемых оригинальных и вторичных (адаптированных) текстов. Для данной оценки были избраны два параметра: (а) количество пропозиций; (б) нормализованная информационная плотность, рассчитываемая как количество пропозиций к количеству слов.

Как видим, нормализованная информационная плотность в преимущественном большинстве текстов РКИ увеличилась (см. Рис. 6). Исключение составляют тексты 1 и 2, в которых количество пропозиций снизилось: 1: 24, 9 (О) — 22, 3 (РКИ); 2: 24,1 (О) — 22, 5 (РКИ).

Что касается количества субпропозиций (Рис. 7), то, как видим, нормализованная плотность увеличилась во всех изучаемых текстах. Последнее означает, что доля пропозиций в расчете на слово увеличилась.

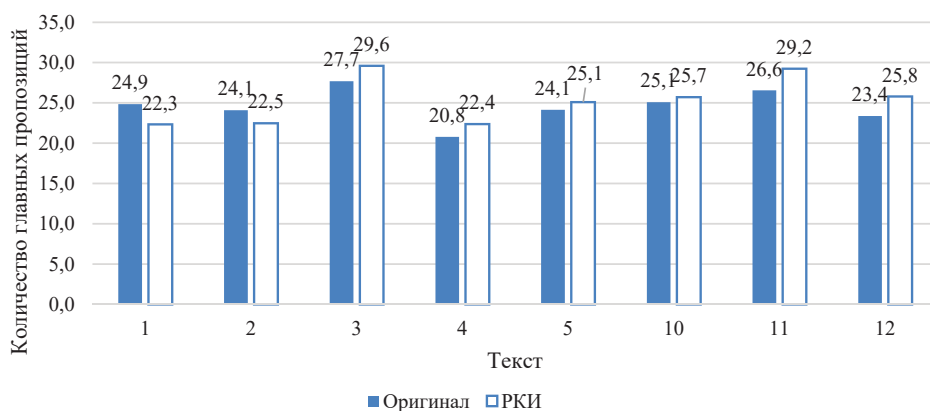


Рис. 6. Нормализованное количество главных пропозиций в текстах О — РКИ(A2,B2)

Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна.

Fig. 6. Normalized number of main propositions the original and adapted for RFL texts (A2, B2)

Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

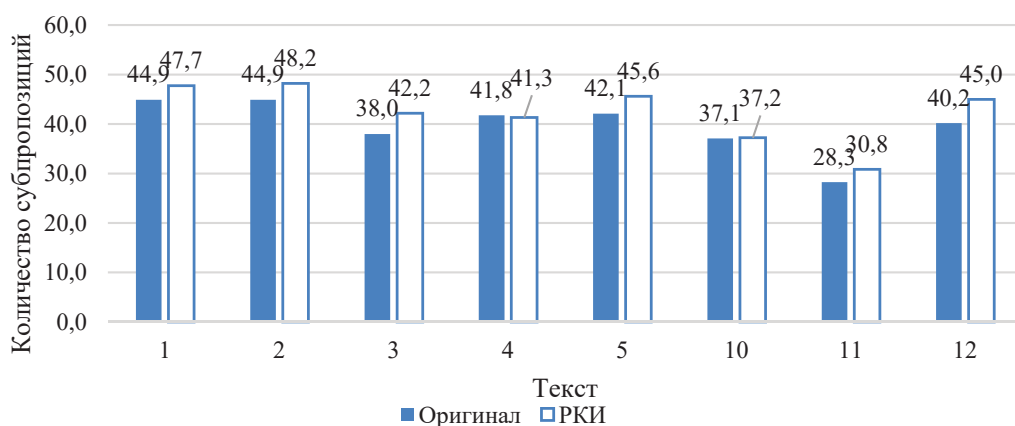


Рис. 7. Нормализованное количество субпропозиций в текстах О — РКИ(A2,B2)

Источник: составлено М.И. Андреевой, М.И. Солнышкиной и С. Саадна.

Fig. 7. Normalized number of subpropositions the original and adapted for RFL texts (A2, B2)

Source: Compiled by Mariia I. Andreeva, Marina I. Solnyshkina, Sarra Saadna.

Как видим, и это важно отметить, нормализованная когнитивная сложность в адаптированных версиях для изучения РКИ несколько выше, чем в оригинале.

Дискуссия

Результаты представленного исследования расширяют данные предыдущих исследований [59; 60] и уточняют список лингвистических параметров-дискриминантов, дифференцирующих первичные и вторичные тексты. Параметрами-дискриминантами художественных и учебных текстов РКИ являются следующие: длина предложения, количество глаголов, индекс читабельности, количество двусложных слов и гапакс легомена. Адаптированный

или вторичный текст трактуется как текст, «основной категорией» которого «общей для всех их видов, является категория сходства/различия» [58]. Признавая, что сама природа вторичного текста, его суть определяется композиционно-образным и лингвостилистическим сходством с «протословом», т.е. оригинальным первичным текстом, исследователь признает, что вторичный текст иначе «воспринимается читателем, оказывает иное эстетическое воздействие» [58].

Заключение

Оценка сложности текста — актуальная проблема современной мультидисциплинарной научной парадигмы, в арсенале которой спектр лингвистических, квантитативных и психолингвистических методов и подходов. Представленное исследование объективирует спектр типичных для учебного текста трансформаций как пропозициональных, так и поверхностных структур текста и определяет предикторы сложности текста. В представленных исследованиях лингвистическая сложность текста как аддитивная категория оценивается на основе 128 валидированных лингвистических параметров, рассчитываемых при помощи автоматического анализатора RuLingva. Различия в лингвистической сложности оригинального и адаптированного для обучения РКИ текстов манифестируются в морфологических (количестве двусложных слов, количестве глаголов, соотношении прилагательного / существительного, соотношении глагола / существительного), лексических (гапакс легомена), а также и синтаксических (индексе читабельности, длине предложения) параметрах).

Выявление специфики лексико-семантических трансформаций текста при адаптации текста и информационной плотности осуществляется с использованием пропозиционального анализа и последующей нормализации на длину текста, необходимой для исключения влияния длины текста на его информационную плотность.

Анализ показал, что индекс нормализованной информационной плотности адаптированного художественного текста, рассчитываемый на основе отношения количества пропозиций на длину текста (в словах), несколько выше аналогичного параметра первичных (художественных) текстов. Таким образом, преобразованный на основе художественного, учебный текст, обладая более низкой степенью лингвистической сложности, сохраняет высокую степень информационной плотности, свойственной первичному тексту.

Результаты исследования будут востребованы для автоматизации пропозиционального анализа текста, а сопоставительный анализ вариантов реализаций глубинных структур текста позволит приблизиться к разработке методики выявления пропозиций.

Список литературы

1. *Dascalu M. et al.* Discourse Complexity: Driving Forces of the New Paradigm // Научный результат. Вопросы теоретической и прикладной лингвистики. 2023. Т. 9. № 1. С. 4–10. <https://doi.org/10.18413/2313-8912-2023-9-1-0-1> EDN: FGGAMR
2. *Solovyev V. et al.* Complexity of Russian Academic Texts as the Function of Syntactic Parameters // International Conference on Computational Linguistics and Intelligent Text Processing. Cham : Springer Nature Switzerland, 2018. С. 168–179. https://doi.org/10.1007/978-3-031-23793-5_15
3. *Борисова Е.Г.* Грамматика слушающего: прогнозирование понимания связного текста и способы его коррекции // Русистика и компаративистика. М. : Книгодел, 2021. С. 262–277. <https://doi.org/10.25688/2619-0656.2021.15.15> EDN: CDZMCP
4. *Карамнов А.С.* Метод экспертных оценок в лингводидактике // Вестник науки Сибири. 2013. № 4(10). С. 224–232. EDN: RSWBMT
5. *Змеева Н.Б.* Ранжирование учебных текстов методом экспертных оценок // Проблемы современной науки и образования. 2016. № 4(46). С. 133–138. EDN: VOGYPJ
6. *Кулибина Н.В.* Адаптировать нельзя понять. Принципы адаптации художественных текстов в соответствии с «Общевропейскими компетенциями владения иностранным языком» // Русский язык за рубежом. 2013. № 5(240). С. 22–30. EDN: RTYEGT
7. *Яхиббаева Л.М.* Учебный текст как особый вид вторичного текста и составляющая учебного дискурса // Вестник Башкирского университета. 2008. Т. 13. № 4. С. 1029–1031. EDN: JKPEAJ
8. *Дзюба Е.В., Ткаченко Е.В.* Художественный текст в учебнике РКИ: объект изучения или средство обучения? // Филологический класс. 2021. Т. 26. № 3. С. 243–254. <https://doi.org/10.51762/1FK-2021-26-03-21> EDN: BFMAAE
9. *Симоненко М.А., Багринцева О.Б.* Пропозициональный подход при анализе денотативной структуры вторичных текстов (на материале пересказов текста И.А. Бунина «Роман Горбуна») // Филологические науки. Вопросы теории и практики. 2019. Т. 12. № 3. С. 310–313. <https://doi.org/10.30853/filnauki.2019.3.66> EDN: VVMNJM
10. *Пиотровская А.А., Пиотровский Р.Г., Разживин К.А.* Энтропия русского языка // Вопросы языкознания. 1962. № 6. С. 115–130.
11. *Филиппов В.С.* Инварианты текста // Филология и культура. Материалы IV международной конференции. Тамбов, 2003. С. 69–72.
12. *Чернявская В.Е.* Дискурс // Стилистический энциклопедический словарь русского языка. М., 2003. С. 54–55.
13. *Азимов Э.Г., Щукин А.Н.* Новый словарь методических терминов и понятий (теория и практика обучения языкам). М. : Икар, 2009. EDN: XQRFTT
14. *Дьякова А.А.* Вторичная репрезентация текстового содержания при интердискурсивной адаптации // *Lingua mobilis*. 2009. № 3(17). С. 79–88. EDN: MSXYUR
15. *Бархударов Л.С.* Язык и перевод (Вопросы общей и частной теории перевода). М. : Междунар. отношения, 1975.
16. *Карасик В. и др.* Типы вторичных текстов // Языковая личность: Проблемы обозначения и понимания. М., 1997. С. 69–70.
17. *Валгина Н.С.* Теория текста. М. : Логос, 2003.
18. *Гак В.Г.* Языковые преобразования. М. : Языки русской культуры, 1998.
19. *Кох В.А.* Предварительный набросок дискурсивного анализа семантического типа // Новое в зарубежной лингвистике. М. : Прогресс, 1978. № 8. С. 149–171.
20. *Мурзин Л.Н., Штерн А.С.* Текст и его восприятие. Свердловск : изд-во УГУ, 1991.
21. *Фрумкина Р.М.* Психолингвистика. М. : Академия, 2007.
22. *Morozov D.A., Glazkova A.V., Iomdin B.L.* Text complexity and linguistic features: Their correlation in English and Russian // *Russian Journal of Linguistics*. 2022. Vol. 26. № 2. P. 426–448. <https://doi.org/10.22363/2687-0088-30132> EDN: IWWUTZ

23. *Khosmood F., Levinson R.* Automatic synonym and phrase replacement show promise for style transformation // 2010 Ninth International Conference on Machine Learning and Applications. IEEE, 2010. P. 958–961. <https://doi.org/10.1109/ICMLA.2010.153>
24. *Dascalu M. et al.* ReaderBench: A multi-lingual framework for analyzing text complexity // Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017, Proceedings 12. Springer International Publishing, 2017. P. 495–499. https://doi.org/10.1007/978-3-319-66610-5_48
25. *McNamara D.S. et al.* Automated evaluation of text and discourse with Coh-Metrix. Cambridge : Cambridge University Press, 2014. <https://doi.org/10.1017/CBO9780511894664>
26. *Solnyshkina M.I., Kupriyanov R.V., Shoeva G.N.* Linguistic profiling of text genres: adventure stories vs. textbooks // Научный результат. Вопросы теоретической и прикладной лингвистики. 2024. Т. 10. № 1. С. 115–132. <https://doi.org/10.18413/2313-8912-2024-10-1-0-7EDN: YABVVQ>
27. *Чурунина А.А., Солнышкина М.И., Ярмачев И.Э.* Лексическое разнообразие как предиктор сложности учебников по русскому языку // Русистика. 2023. Т. 21. № 2. С. 212–227. <http://doi.org/10.22363/2618-8163-2023-21-2-212-227> EDN: ZYHNDWM
28. *Майер Р.В.* Проблема определения дидактической сложности учебных текстов // Ученые записки Крымского федерального университета имени В.И. Вернадского. Социология. Педагогика. Психология. 2020. Т. 6(72). № 1. С. 37–47. EDN: UURISI
29. *Кудж С.А., Цветков В.Я.* Факторы когнитивной сложности // ИТНОУ: Информационные технологии в науке, образовании и управлении. 2018. № 6(10). С. 34–41. EDN: YOYXWP
30. *Бредихин С.Н., Серебрякова С.В., Лиховид А.А.* Способы компрессии когнитивной информации в научно-популярном тексте // Актуальные проблемы филологии и педагогической лингвистики. 2019. № 3. С. 139–145. <https://doi.org/10.29025/2079-6021-2019-3-139-145> EDN: MAZDIA
31. *Матвеева О.В.* «Информационная насыщенность» как характерная особенность специальных текстов // Актуальные проблемы гуманитарных и естественных наук. 2014. № 4–1. С. 363–368. EDN: SDWKPL
32. *Алексеева И.С.* Введение в переводоведение. М. : Академия, 2004. EDN: QQZJET
33. *Тихомирова Л.С.* Плотность текста в когнитивном сознании // Гуманитарные исследования в Восточной Сибири и на Дальнем Востоке. 2011. № 2(14). С. 99–104. EDN: OBUEUZ
34. *Chen Y.Y. et al.* Understanding the Cognitive Complexity in Language Elicited by Product Images // arXiv preprint arXiv:2409.16521. 2024.
35. *Чернявская В.Е.* Текст в когнитивно-дискурсивной парадигме: к вопросу о градуальном характере текстуальности // Acta Linguistica Petropolitana. Труды института лингвистических исследований. 2007. Т. 3. № 1. С. 303–314. EDN: YUFBBR
36. *Bolton M.L., Biltekoff E., Humphrey L.* The mathematical meaninglessness of the NASA task load index: A level of measurement analysis // IEEE Transactions on Human-Machine Systems. 2023. Vol. 53. № 3. P. 590–599. <https://doi.org/10.1109/thms.2023.3263482> EDN: YUQIOI
37. *Азрус А.Р., Ахутина Т.В., Корнеев А.А.* Варианты дефицита функций I блока мозга у детей с трудностями обучения // Вестник Московского университета. Серия 14. Психология. 2014. № 3. С. 34–46. EDN: SKSFRB
38. *Webb N.L.* Issues related to judging the alignment of curriculum standards and assessments // Applied measurement in education. 2007. Т. 20. № 1. С. 7–25. <https://doi.org/10.1080/08957340709336728>
39. *Anderson L.W., Krathwohl D.R.* A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition. Addison Wesley Longman, Inc., 2001.
40. *Bloom B.S. (ed.).* Taxonomy of educational objectives: Handbook II. David McKay, 1956.
41. *Samsonov N., Chmykhova E., Davydov D.* Разработка и апробация лингвистической методики оценки когнитивной сложности научно-учебного текста // Психологические исследования. 2015. Т. 8. № 41. С. 6. <https://doi.org/10.54359/ps.v8i41.548> EDN: VHFQHX

42. Дридзе Т.М. Язык и социальная психология. М. : Высш. школа, 1980.
43. Куприянов Р.В., Букач О.В., Александрова О.И. Оценка когнитивной сложности учебных текстов: эмпирическая валидация лингвистических параметров // *Russian Journal of Linguistics*. 2023. Т. 27. № 3. С. 641–662. <https://doi.org/10.22363/2687-0088-35817> EDN: NAZBGA
44. Das S., Das Mandal S.K., Basu A. Cognitive complexity analysis of learning-related texts: A case study on school textbooks // *Methodologies and Intelligent Systems for Technology Enhanced Learning*, 10th International Conference. Springer International Publishing, 2020. P. 74–84. https://doi.org/10.1007/978-3-030-52538-5_9 EDN: HKOTPP
45. Kintsch W., Van Dijk T.A. Toward a model of text comprehension and production // *Psychological review*. 1978. Vol. 85. № 5. P. 363. <https://doi.org/10.1037/0033-295X.85.5.363>
46. Graesser A.C., McNamara D.S. Computational analyses of multilevel discourse comprehension // *Topics in cognitive science*. 2011. Т. 3. № 2. С. 371–398. <https://doi.org/10.1111/j.1756-8765.2010.01081.x>
47. Солнышкина М.И., Харькова Е.В., Эбзеева Ю.Н. Влияние содержания на понимание текста: пропозициональный анализ дискурса // *Russian Journal of Linguistics*. 2023. Т. 27. № 4. С. 938–956. <https://doi.org/10.22363/2687-0088-35915> EDN: ZAJKUE
48. Kintsch W. *Comprehension: A paradigm for cognition*. Cambridge : Cambridge University Press, 1998.
49. Kintsch W. *The Representation of Meaning in Memory*, Poiomac : Erlbaum, 1974.
50. Thorndyke P.W. Cognitive structures in comprehension and memory of narrative discourse // *Cognitive Psychology*. 1977. Vol. 9. P. 77–110. [https://doi.org/10.1016/0010-0285\(77\)90005-6](https://doi.org/10.1016/0010-0285(77)90005-6)
51. Anderson J.A. *Cognitive psychology and its implications*. N.Y. : W.H. Freeman, 1990.
52. Бирюкова Ю.Н. Виды чтения при обучении переводу научно-технических текстов // *Вестник Российского университета дружбы народов. Серия: Вопросы образования: языки и специальность*. 2014. № 2. С. 103–109. EDN: SFGYNN
53. Фоломкина С.К. *Обучение чтению на иностранном языке в неязыковом вузе*. М., 1987.
54. Оборнева И.В. Автоматизированная оценка сложности учебных текстов на основе статистических параметров: дис. ... канд. пед. наук. М., 2006. EDN: NNRVKP
55. Бондарева Е.В., Мазепа Е.А. Об использовании критерия Вилкоксона в практике педагогических исследований // *Математическая физика и компьютерное моделирование*. 2021. Т. 24. № 3. С. 26–32. <https://doi.org/10.15688/mpcm.jvolsu.2021.3.3> EDN: EXNRKA
56. Стародумова Е.А. *Синтаксис современного русского языка*. Владивосток : Дальневосточный федеральный университет, 2005. EDN: GWUTOY
57. Zaki H., Ellis R. *Learning vocabulary through interacting with a written text // Learning a second language through interaction*. John Benjamins Publishing Company, 2008. P. 151–170. <https://doi.org/10.1075/sibil.17.10zak>
58. Rod E., Barkhuizen G. *Analysing Learner Language*. Oxford : Oxford University Press, 2005.
59. Андреева М.И., Солнышкина М.И., Муфазалова Н.И. Предикторы когнитивной и лингвистической сложности текстов РКИ. Новые и традиционные аспекты в поликультурном пространстве преподавания РКИ: актуальные вопросы теории и практики: Сб. материалов VII Московского международного культурно-образовательного форума по РКИ. 21–22 ноября 2024 г. / Под ред.: И.Е. Карпенко М. : УЦРЯ МГУ, 2024. С. 117 —125.
60. Вербицкая М.В. *Теория вторичных текстов: на материале современного английского языка: автореф. дисс. ... д-ра филол. наук*. М., 2000. EDN: ZKUIDX

References

1. Dascalu, M. (2023). Discourse Complexity: Driving Forces of the New Paradigm. *A scientific result. Questions of Theoretical and Applied Linguistics*, 9(1), 4–10. <https://doi.org/10.18413/2313-8912-2023-9-1-0-1> EDN: FGGAMR

2. Solovyev, V., Solnyshkina, M., Ivanov, V., & Timoshenko, S. (2018). Complexity of Russian Academic Texts as the Function of Syntactic Parameters. In: *International Conference on Computational Linguistics and Intelligent Text Processing* (pp. 168–179). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-23793-5_15
3. Borisova, E.G. (2021). Listener's Grammar: Predicting the Understanding of a Coherent Text and Ways to Correct It. In: *Russian Studies and Comparative Studies* (pp. 262–277). Moscow: Knogodel. <https://doi.org/10.25688/2619-0656.2021.15.15> (In Russ.). EDN: CDZMCP
4. Karamnov, A.S. (2013). The Method of Expert Assessments in Linguodidactics. *Vectors of well-being: economy and society*, 4(10), 224–232. (In Russ.). EDN: RSWBMT
5. Zmeeva, N.B. (2016). Ranking of Educational Texts by the Method of Expert Assessments. *Problems of modern science and education*, 4(46), 133–138. (In Russ.). EDN: VOGYPJ
6. Kulibina, N.V. (2013). It is Impossible to Adapt to Understand. Principles of Adaptation of Literary Texts in Accordance with the “Pan-European Competencies of Foreign Language Proficiency”. *Russian Language Abroad*, 5(240), 22–30. (In Russ.). EDN: RTYEGT
7. Yahibbayeva, L.M. (2008). Educational Text as a Special Type of Secondary Text and a Component of Educational Discourse. *Bulletin of the Bashkir University*, 13(4), 1029–1031. (In Russ.). EDN: JKPEAJ
8. Dzyuba, E.V., & Tkachenko, Yu.G. (2021). An Artistic Text in an RCT Textbook: an Object of Study or a Means of Teaching? *Philological Class*, 26(3), 243–254. <https://doi.org/10.5176/2/1FK-2021-26-03-21> (In Russ.). EDN: BFMAAE
9. Simonenko, M.A., & Bagrintseva, O.B. (2019). A Propositional Approach to the Analysis of the Denotative Structure of Secondary Texts (Based on the Material of Retellings of IABunin's Text “The Novel of the Hunchback”). *Philological Sciences. Questions of Theory and Practice*, 12(3), 310–313. <https://doi.org/10.30853/filnauki.2019.3.66> (In Russ.). EDN: VVMNJM
10. Piotrovskaya, A.A., Piotrovsky, R.G., & Razzhivin, K.A. (1962). The Entropy of the Russian Language. *Voprosy Jazykoznanija*, 6, 115–130. (In Russ.).
11. Filippov, V.S. (2003). Text invariants. In: *Philology and culture. Proceedings of the IV International Conference* (pp. 69–72). Tambov. (In Russ.).
12. Chernyavskaya, V.E. (2003) Discourse. In: *Stylistic Encyclopedic Dictionary of the Russian language* (pp. 54–55). Moscow. (In Russ.).
13. Azimov, E.G., & Shchukin, A.N. (2009). *A New Dictionary of Methodological Terms and Concepts* (Theory and Practice of Language Teaching). Moscow: Ikar. (In Russ.). EDN: XQRFTT
14. Dyakova, A.A. (2009). Secondary Representation of Textual Content During Interdiscursive Adaptation. *Lingua Mobilis*, 3(17), 79–88. (In Russ.). EDN: MSXYUR
15. Barkhudarov, L.S. (1975). *Language and Translation (Issues of General and Particular Theory of Translation)*. Moscow: International Relations. (In Russ.).
16. Karasik, V. (1997). Types of Secondary Texts. In: *Linguistic Personality: Problems of Designation And Understanding* (pp. 69–70). Moscow. (In Russ.).
17. Valgina, N.S. (2003). *Text theory*. Moscow: Logos. (In Russ.).
18. Gak, V.G. (1998). *Linguistic transformations*. Moscow: Languages of Russian Culture. (In Russ.).
19. Koch, V.A. (1978). A Preliminary Sketch of the Semantic Type of Discursive Analysis. In: *New in Foreign Linguistics* (8. pp. 149–171). Moscow: Progress. (In Russ.).
20. Murzin, L.N., & Stern, A.S. (1991). *The Text and Its Perception*. Sverdlovsk: UGU publ. (In Russ.).
21. Frumkina, R.M. (2007). *Psycholinguistics*. Moscow: Akademiya. (In Russ.).
22. Morozov, D.A., Glazkova, A.V., & Iomdin, B.L. (2022). Text Complexity and Linguistic Features: Their Correlation in English and Russian. *Russian Journal of Linguistics*, 26(2), 426–448. <https://doi.org/10.22363/2687-0088-30132> EDN: IWWUTZ

23. Khosmood, F., & Levinson, R. (2010). Automatic Synonym and Phrase Replacement Show Promise for Style Transformation. In: *2010 Ninth International Conference on Machine Learning and Applications* (pp. 958–961). IEEE. <https://doi.org/10.1109/ICMLA.2010.153>
24. Dascalu, M., Gutu, G., Ruseti, S., Paraschiv, I.C., Dessus, P., McNamara, D.S., ... & Trausan-Matu, S. (2017). ReaderBench: A multi-lingual framework for analyzing text complexity. In: *Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning, EC-TEL 2017*, Tallinn, Estonia, September 12–15, 2017, Proceedings 12 (pp. 495–499). Springer International Publ.
25. McNamara, D.S., Graesser, A.C., McCarthy, P.M., & Cai, Z. (2014). *Automated Evaluation of Text And Discourse with Coh-Matrix*. Cambridge: Cambridge University Press.
26. Solnyshkina, M.I., Kupriyanov, R.V., & Shoeva, G.N. (2024). Linguistic Profiling of Text Genres: Adventure Stories vs. Textbooks. *A Research Result. Issues of Theoretical and Applied Linguistics*, 10(1), 115–132. <https://doi.org/10.18413/2313-8912-2024-10-1-0-7EDN: YABVVQ>
27. Churunina, A.A., Solnyshkina, M.I., & Yarmakeev, I.E. (2023). Lexical Diversity as a Predictor of the Complexity of Textbooks in the Russian Language. *Russian Language Studies*, 21(2), 212–227. <http://doi.org/10.22363/2618-8163-2023-21-2-212-227> (In Russ.). EDN: ZYHDM
28. Mayer, R.V. (2020). The Problem of Determining the Didactic Complexity of Educational Texts. *Scientific Notes of the Vernadsky Crimean Federal University. Sociology. Pedagogy. Psychology*, 6(72)(1), 37–47. (In Russ.). EDN: UURISI
29. Kudzh, S.A., & Tsvetkov, V.Ya. (2018). Factors of Cognitive Complexity. *ITNOW: Information Technologies in Science, Education and Management*, 6(10), 34–41. (In Russ.). EDN: YOYXWP
30. Bredikhin, S.N., Serebryakova, S.V., & Likhovid, A.A. (2019). Methods of Compression of Cognitive Information in a Popular Science Text. *Actual Problems of Philology and Pedagogical Linguistics*, 3, 139–145. <https://doi.org/10.29025/2079-6021-2019-3-139-145> (In Russ.). EDN: MAZDIA
31. Matveeva, O.V. (2014). “Information Saturation” as a Characteristic Feature of Special Texts. *Actual Problems of the Humanities and Natural Sciences*, 4–1, 363–368. EDN: SDWKPL
32. Alekseeva, I.S. (2004). *Introduction to translation studies*. Moscow: Academy. (In Russ.). EDN: QQZJET
33. Tikhomirova, L.S. (2011). The Density of Text in Cognitive Consciousness. *Humanitarian Studies in Eastern Siberia and the Far East*, 2–14, 99–104. (In Russ.).
34. Chen, Y.Y., Hakimi, S., Van, M., Chen, F., Hong, M., Klenk, M., & Wu, C. (2024). Understanding the Cognitive Complexity in Language Elicited by Product Images. In: *arXiv preprint arXiv:2409.16521*.
35. Chernyavskaya, V.E. (2007). Text in the Cognitive-Discursive Paradigm: on the Issue of the Graduated Nature of Textuality. *Acta Linguistica Petropolitana. Proceedings of the Institute of Linguistic Research*, 3(1), 303–314. (In Russ.). EDN: YUFBRR
36. Bolton, M.L., Biletkoff, E. & Humphrey, L. (2023). The Mathematical Meaninglessness of the NASA Task Load Index: A Level of Measurement Analysis. *IEEE Transactions on Human-Machine Systems*, 53(3), 590–599. <https://doi.org/10.1109/thms.2023.3263482> EDN: YUQIOI
37. Agris, A.R., Akhutina, T.V., & Korneev, A.A. (2014). Variants of Brain Block i Function Deficiency in Children with Learning Difficulties. *Bulletin of the Moscow University. Episode 14. Psychology*, 3, 34–46. (In Russ.). EDN: SKSFRB
38. Webb, N.L. (2007). Issues Related to Judging the Alignment of Curriculum Standards and Assessments. *Applied Measurement in Education*, 20(1), 7–25. <https://doi.org/10.1080/08957340709336728>
39. Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Rath, J., & Wittrock, M. (eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Abridged Edition).
40. Bloom, B.S. (Ed.). (1956). *Taxonomy of educational objectives: Handbook II*. David McKay.

41. Samsonov, N., Chmykhova, E., & Davydov, D. (2015). Development and Approbation of the Linguistic Technique of Cognitive Complexity Assessment for the Educational Text. *Psychological research*, 8(41), 6. <https://doi.org/10.54359/ps.v8i41.548> (In Russ.). EDN: VHFQHX
42. Dridze, T.M. (1980). *Language and social psychology*. Moscow: HS. (In Russ.).
43. Kupriyanov, R.V., Bukach, O.V., & Aleksandrova, O.I. (2023). Cognitive Complexity Measures for Educational Texts: Empirical Validation of Linguistic Parameters. *Russian Journal of Linguistics*, 27(3), 641–662. (In Russ.). <https://doi.org/10.22363/2687-0088-35817> EDN: NAZBGA
44. Das, S., Das Mandal, S.K., & Basu, A. (2020). Cognitive Complexity Analysis of Learning-Related Texts: a Case Study on School Textbooks. In *Methodologies and Intelligent Systems for Technology Enhanced Learning, 10th International Conference* (pp. 74–84). Springer International Publ. https://doi.org/10.1007/978-3-030-52538-5_9 EDN: HKOTPP
45. Kintsch, W., & van Dijk, T.A. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review*, 85(5), 363–394. <https://doi.org/10.1037/0033-295X.85.5.363>
46. Graesser, A. & McNamara, D. (2011). Computational Analyses of Multilevel Discourse Comprehension. *Topics in Cognitive Science*, 3, 371398. <https://doi.org/10.1111/j.1756-8765.2010.01081>
47. Solnyshkina, M.I., Harkov, E.V., & Ebzeeva, Yu.N. (2023). The Influence of Content on the Understanding of a Text: a Propositional Analysis of Discourse. *Russian Journal of Linguistics*, 27(4), 938–956. <https://doi.org/10.22363/2687-0088-35915> (In Russ.). EDN: ZAJKUE
48. Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
49. Kintsch, W. (1974). *The Representation of Meaning in Memory*. Poimac: Erlbaum.
50. Thorndyke, P.W. (1977). Cognitive Structures in Comprehension and Memory of Narrative Discourse. *Cognitive Psychology*, 9(1), 77–110.
51. Anderson, J.A. (1990). *Cognitive Psychology and Its Implications*. N.Y.: W.H. Freeman. [https://doi.org/10.1016/0010-0285\(77\)90005-6](https://doi.org/10.1016/0010-0285(77)90005-6)
52. Biryukova, Yu.N. (2014). Types of Reading in Teaching Translation of Scientific and Technical Texts. *Multilingualism and Transcultural Practices*, 2, 103–109. (In Russ.). EDN: SFGYNN
53. Folomkina, S.K. (1987). *Teaching Reading in a Foreign Language in a Non-Linguistic University*. Moscow. (In Russ.).
54. Osborneva, I.V. (2006). *Automated Assessment of the Complexity of Educational Texts Based on Statistical Parameters* [PhD thesis]. Moscow. (In Russ.). EDN: NNRVKP
55. Bondareva, E.V., & Mazepa, E.A. (2021). On the Use of the Wilcoxon Criterion in the Practice of Pedagogical Research. *Mathematical Physics and Computer Modeling*, 24(3), 26–32. <https://doi.org/10.15688/mpcm.jvolsu.2021.3.3> (In Russ.). EDN: EXNRKA
56. Starodumova, E.A. (2005). *Syntax of the Modern Russian Language*. Vladivostok: Far Eastern Federal University. (In Russ.). EDN: GWUTOY
57. Zaki, H., & Ellis, R. (1999). Learning Vocabulary through Interacting with a Written Text. In: Ellis, R. (ed), *Learning a Second Language through Interaction* (pp. 153–169). Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.17.10zak>
58. Rod, E. & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford: Oxford University Press.
59. Andreeva, M.I., Solnyshkina, M.I., Mufazalova, N.I. (2024). Predictors of Cognitive and Linguistic Complexity of RCT Texts. In: *New and traditional aspects in the multicultural space of teaching RCT: current issues of theory and practice: Collection of materials of the VII Moscow International Cultural and Educational Forum on RFL*. November 21–22, 2024 (pp. 117–1125). Moscow: UCR MSU. (In Russ.).
60. Verbitskaya, M.V. (2000). *Theory of Secondary Texts: Based on the Material of Modern English* [Doctor of Philology thesis]. Moscow. (In Russ.). EDN: ZKUIDX

Сведения об авторах:

Андреева Мария Игоревна, кандидат филологических наук, доцент, старший научный сотрудник НИЛ «Мультидисциплинарные исследования текста», Казанский (Приволжский) федеральный университет (420008, Российская Федерация, г. Казань, ул. Кремлевская, д. 18, корп. 1); *сфера научных интересов*: корпусная лингвистика, сложность текстов, семантика, лексикология; *e-mail*: mariia99andreeva@yandex.ru
ORCID: 0000-0002-5760-0934. SPIN-код: 9243-6995. Researcher ID: ABF-7003-2020; Scopus ID: 57195974758.

Солнышкина Марина Ивановна, доктор филологических наук, профессор, профессор кафедры теории и практики преподавания иностранных языков, руководитель НИЛ «Мультидисциплинарные исследования текста», Казанский (Приволжский) федеральный университет (420008, Российская Федерация, г. Казань, ул. Кремлевская, д. 18, корп. 1); *сфера научных интересов*: сложность текста, дискурсивная комплексология; *e-mail*: mesoln@yandex.ru
ORCID: 0000-0003-1885-3039; E-library SPIN-код: 6480-1830; Researcher ID: E-3863-2015; Scopus ID: 56429529500.

Сарра Саадна, аспирант кафедры иностранных языков филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: дискурсивные исследования, корпусная лингвистика; *e-mail*: 1042218192@rudn.ru
ORCID: 0009-0004-3832-6660.

Information about the authors:

Mariia I. Andreeva, Associate Professor, PhD in Philology, Senior Researcher of the 'Multidisciplinary text research' Research Lab, Institute of Philology and Intercultural Communication, Kazan Federal University (18 / 1, Kremlevskaya st., Kazan, Russian Federation, 420008); *Research interests*: corpus linguistics, text complexity, semantics, lexicology; *e-mail*: mariia99andreeva@yandex.ru
ORCID: 0000-0002-5760-0934. SPIN-code: 9243-6995. Researcher ID: ABF-7003-2020; Scopus ID: 57195974758.

Marina I. Solnyshkina, Dr.Sc. (Philology), Professor of the Department of Theory and Practice of Teaching Foreign Languages, Head of 'Multidisciplinary text research' Research Lab, Institute of Philology and Intercultural Communication, Kazan Federal University (18 / 1, Kremlevskaya str., Kazan, Russian Federation, 420008); *Research interests*: linguistic complexology, corpus linguistics, and lexicography; *e-mail*: mesoln@yandex.ru
ORCID: 0000-0003-1885-3039; SPIN-code: 6480-1830; Researcher ID: E-3863-2015; Scopus ID: 56429529500.

Sarra Saadna, PhD student at the Foreign languages department, Faculty of Philology, RUDN University (6, Miklukho-Maklaya str., Moscow, Russian Federation, 117198); *Research interests*: discourse studies; corpus linguistics; *e-mail*: 1042218192@rudn.ru
ORCID: 0009-0004-3832-6660.