

Учебный центр русского языка (Россия, г. Москва)

**Казахский национальный университет имени Аль-Фараби,
Центр «Русский язык и культура» (Казахстан, г. Алма-Ата)**

**Узбекский государственный университет мировых языков
(Узбекистан, г. Ташкент)**

НОВЫЕ И ТРАДИЦИОННЫЕ АСПЕКТЫ В ПОЛИКУЛЬТУРНОМ ПРОСТРАНСТВЕ ПРЕПОДАВАНИЯ РКИ: АКТУАЛЬНЫЕ ВОПРОСЫ ТЕОРИИ И ПРАКТИКИ

**VII Московский международный
культурно-образовательный
форум по РКИ**

Москва, 21-22 ноября 2024 года

Сборник материалов

**Москва
2024**

Главный редактор:
кандидат филол. наук **И.Е. Карпенко**

Ответственный секретарь:
ст. преподаватель УЦРЯ МГУ **Е.Г. Ковалева**

Новые и традиционные аспекты в поликультурном пространстве преподавания РКИ: актуальные вопросы теории и практики: Сборник материалов VII Московского международного культурно-образовательного форума по РКИ. 21-22 ноября 2024 г. / Под ред.: И.Е. Карпенко – Москва: УЦРЯ МГУ, 2024. – 208 с.

В сборник вошли материалы VII Международного форума по РКИ, традиционно объединившего профессионалов – русистов из России и из-за рубежа в поисках ответов на важные вопросы теории и практики преподавания русского языка в актуальной ситуации динамично меняющегося мира. А также о новых вызовах, связанных с искусственным интеллектом, цифровизацией образования, появлением новых подходов в обучении (в частности, нейропедагогики) и, несомненно, с темой продвижения русского языка. В представленных статьях рассматриваются проблемы общей методики преподавания русского языка и отдельные аспекты преподавания РКИ, вопросы лингводидактики и особенности обучения РЯ в разных странах; предлагаются новые подходы, педагогические технологии и оригинальные авторские методики, а также анализируются особенности работы с детьми – инофонами, билингвами и мультилингвами в поликультурном пространстве современного мира. Кроме того, авторы учебников и учебных пособий рассказывают коллегам о новых изданиях по теме РКИ.

© 2024 Издательство «Учебный центр русского языка Московского государственного университета»

ISBN 978_5_6040687_8_6

IV. ЛИНГВИСТИЧЕСКИЕ И КОГНИТИВНЫЕ АСПЕКТЫ ИЗУЧЕНИЯ РУССКОГО ЯЗЫКА ЧЕРЕЗ ПРИЗМУ РКИ

М.И. Андреева

maria.andreeva@kazangmu.ru

кандидат филологических наук,

доцент кафедры иностранных языков Казанского ГМУ,

старший научный сотрудник НИЛ «Мультидисциплинарные исследования текста»,

Казанский федеральный университет

Казань, Россия

М.И. Солнышкина

mesoln@yandex.ru

доктор филологических наук, профессор,

профессор кафедры теории и практики преподавания иностранных языков,

руководитель НИЛ «Мультидисциплинарные исследования текста»,

Казанский федеральный университет

Казань, Россия

Н.И. Муфазалова

mufazalovanailya@gmail.com

ассистент кафедры теории и практики преподавания иностранных языков,

Казанский федеральный университет

Казань, Россия

ПРЕДИКТОРЫ КОГНИТИВНОЙ И ЛИНГВИСТИЧЕСКОЙ СЛОЖНОСТИ ТЕКСТОВ РКИ

Аннотация. В статье представлены результаты продолжающегося исследования, нацеленного на выявление предикторов информационной (когнитивной) и лингвистической сложности учебных текстов по русскому языку как иностранному. Исследование проводится на материале корпуса 54-х учебных текстов РКИ (37 344 слов), а также оригинальных текстов, на основе которых была осуществлена адаптация учебных текстов РКИ. Общий объем 54-х оригинальных текстов составил 69 231 слово. Алгоритм исследования включает: 1) отбор и предобработку оригинальных и адаптированных текстов; 2) оценку лингвистической сложности на основе метрик лингвистических параметров, рассчитываемых текстовым профайлером *RuLingva* (rulinga.kpfu.ru/); 3) оценку когнитивной сложности текстов при помощи пропозиционального анализа; 4) идентификацию предикторов сложности на основе сравнительного анализа. Результаты исследования будут использованы для изучения трудности восприятия текстов различной сложности респондентами – носителями русского языка как иностранного.

Ключевые слова: РКИ, сложность текста, художественный текст, учебный текст, пропозиции

Понимание во многом зависит от способности коммуникантов выбирать такие средства, которые не только соответствуют интенциям продуцента информации, но также когнитивным и лингвистическим способностям реципиента. При институциональном общении (врач – пациент, адвокат – клиент, консультант/учитель – обучающийся) могут возникать ситуации, требующие снижения сложности текста, т.е. его адаптацию. Неспособность представлять информацию в соответствующем для реципиента диапазоне сложности неизбежно приводит к коммуникативным сбоям. Очевидно, что сама адаптация текста «под читателя» есть не только признание способности текста видоизменять форму с сохранением содержания, но и способности продуцента текста управлять данным процессом. В спектре текстовых изменений при его адаптации ярко отражаются **предикторы сложности**, т.е. те параметры, изменение которых меняет объем воспринимаемой информации.

В представленном исследовании вторичным именуем текст, «основной категорией» которого «общей для всех их видов, является категория сходства/различия» с первичным [Вербицкая 2000]. Сама природа вторичного текста, его суть определяется композиционно-образным и лингвостилистическим сходством с «протословом», т.е. оригинальным первичным текстом, однако благодаря произведенным в ходе адаптации изменениям вторичный текст иначе «воспринимается читателем, оказывает иное эстетическое воздействие» [Вербицкая 2000]. Важным в ракурсе представленного исследования является и указание на то, что характерным при генерировании вторичных текстов является экспликация имплицитного. Способы экспликации могут быть разнообразны, однако общим для всех них является перевод с уровня идейно-художественного содержания на семантический [Вербицкая 2000]. В свете вышеизложенного актуальным является выявление спектра параметров, дифференцирующих оригинальные и вторичные тексты с целью типологизации способов адаптации художественных текстов для учебных целей русского как иностранного (РКИ). Представленное исследование тестирует три основные гипотезы:

(1) экспликация имплицитных смыслов вторичного учебного текста осуществляется на локальном (одно или смежные предложения) и на глобальном (текст) уровнях;

(2) сокращение количества пропозиций при адаптации учебного текста осуществляется преимущественно за счет элиминации субпропозиций;

(3) элимируемые из оригинального текста предикторы сложности могут быть типологизированы на материале достаточно репрезентативной выборки учебных текстов.

Исследование выполнено на **материале** двух коллекций текстов:

(1) первичных текстов – фрагментов художественных произведений русских авторов (n=54) и

(2) адаптированных версий художественных произведений, т.е. вторичных текстов (n=54), используемых авторами учебников РКИ в учебном процессе. Авторский исследовательский корпус содержит 106 575 словоформ (см. *Таблицу 1*).

Вторичные тексты (РКИ)		Первичные (оригинальные) тексты		Всего	
тексты	словоформы	тексты	словоформы	тексты	словоформы
54	37344	54	69231	108	106575

Таблица 1. Объем корпуса исследования

Художественные тексты принадлежат перу известных российских писателей: А.С. Пушкину (1799-1837), И.А. Гончарову (1812-1891), И.С. Тургеневу (1818-1883), Ф.М. Достоевскому (1821-1881), Л.Н. Толстому (1828-1910), А.П. Чехову (1860-1904), В.В. Вересаеву (1867-1945), А.И. Куприну (1870-1938), М.А. Булгакову (1891-1940), К.Г. Паустовскому (1892-1968), В.В. Набокову (1899-1977) и написаны преимущественно в жанрах романа и повести.

Учебные тексты представляют собой сокращенные фрагменты, подвергшиеся авторами учебников адаптации, нацеленной на снижение уровня трудности для обучающихся. Заявленная авторами учебников сложность текстов колеблется в диапазоне от А2-С2.

Алгоритм исследования.

На **ЭТАПЕ I** все вторичные и первичные (оригинальные) тексты после сканирования и извлечения их из источников прошли предобработку: из них были удалены лишние символы, пробелы, исправлены опечатки. Метаразметка включенных в корпус файлов включала указание на автора, название произведения, номер текста в корпусе, тип текста (первичный, О – вторичный, РКИ) и объем (в словоформах). Все тексты РКИ дополнительно получили маркировку уровня сложности, заявленную авторами учебника, например, Dost Idiot_1_RFL_B(2)_1288 маркирует фрагмент из романа Ф.М. Достоевского «Идиот» из учебника русского РКИ уровня В2 и содержащий 1288 словоформ.

На **ЭТАПЕ II** каждый из текстов корпуса был подвергнут лингвистическому профилированию при помощи сервера *RuLingva* (rulingva.kpfu.ru). *RuLingva* является продуктом НИЛ «Мультидисциплинарные исследования текста» (clck.ru/3Dt2b2, <https://rulex.kpfu.ru/rsci>) и в настоящее время рассчитывает значения 80-ти лингвистических параметров текста с грануляцией на (1) документ; (2) предложение; (3) абзац и (4) около 1000 словоформ¹¹.

Классификация параметров включает описательные (количество словоформ, лемм, слогов, предложений, среднее количество слов в предложении, слогов в слове, букв в слове, односложные, двусложные, трехсложные, четырехсложные слова); параметры читабельности (Индекс Флеша-Кинкейда – SIS для академических текстов, Индекс Флеша-Кинкейда – Ob для художественных текстов); морфологические параметры (количество знаменательных слов, служебных слов, лексическое разнообразие, лексическая плотность, водность); имена существительные (количество, падежи), глаголы (количество, «нарративность», доли временных форм), прилагательные (количество, «описательность»); наречия (количество); местоимения (количество, лицо, число); числительные (количество); лексические параметры (средний ранг частотности, средняя частотность, абстрактность); количество гапакс легоменов (уникальных слов); термины филологии, математики, информатики, естественнонаучные термины, физики,

¹¹ Точное количество слов может варьироваться, поскольку при нормализации на 1000 сохраняется неразрывность предложения.

изо, музыки; доли слов уровней A1-C2; доля слов, отсутствующих в лексических минимумах РКИ; дискурсивные параметры (локальный и глобальный повторы существительных, аргументов, глаголов, местоимений, наречий, прилагательных, числительных).

Для последующего сравнительного анализа нормализация значений «сырых» параметров осуществляется по формуле: $F^{норм} = \text{метрика} / \text{количество словоформ в отрывке} * 1000$.

На *ЭТАПЕ III* с помощью программы *STATISTIKA* на основе Т-критерия Вилкоксона была осуществлена оценка статистически значимых различий сравниваемых текстов (см. *Таблица 2*).

Т-критерий Вилкоксона является одним из наиболее используемых инструментов непараметрической статистики для сравнения показателей, измеряемых в одной и той же выборке при различных условиях. Данный критерий позволяет установить эффективность воздействия на измеряемый объект [Бондарева, Мазепа 2021] и широко используется при интерпретации результатов педагогических экспериментов контрольной и экспериментальной групп [Бондарева, Мазепа 2021], эффективности действия медицинских препаратов [Шитова 2017], выявления отклонений технологического процесса и критических событий [Коченгин 2020], обнаружения космических объектов [Рябова, Замега, 20220]. В представленном исследовании применение Т-критерия Вилкоксона значимо для оценки спектра изменений в каждой из пар сравниваемых О – РКИ текстов.

		Переменная	Ср. стат-ое значение для оригинальных текстов	Ср. стат-ое значение для текстов РКИ	Т-критерий Вилкоксона р-значение*
Морф-ие параметры	1	Сущ.	240,8 ± 39,5	234,8 ± 38,2	0,00
	2	Ср. кол-во сущ. в предлож.	3,0 ± 1,0	2,7 ± 0,8	0,00
	3	Глаголы	199,3 ± 23,0	205,3 ± 23,2	0,00
	4	Ср. кол-во глаг. в предлож.	2,4 ± 0,6	2,3 ± 0,5	0,00
	5	Ср. кол-во прилаг. в предлож.	0,9 ± 0,4	0,8 ± 0,3	0,00
	6	Ср. кол-во нареч. в предлож.	0,9 ± 0,4	0,8 ± 0,3	0,00
	7	Местоимения	167,2 ± 40,1	175,8 ± 38,6	0,00
	8	Родительный падеж (сущ.)	43,3 ± 12,7	40,7 ± 13,7	0,01
	9	Творительный падеж (сущ.)	26,9 ± 8,1	24,1 ± 9,4	0,00
	10	Двусложные слова	321,0 ± 27,9	333,9 ± 34,4	0,00
	11	Четырехсложные слова	90,7 ± 19,3	86,5 ± 20,2	0,01
	12	Гапакс легомены	0,5 ± 0,5	0,8 ± 0,6	0,00
	13	Прош. вр. (глагол.)	121,3 ± 24,8	127,3 ± 26,8	0,01
Читабельность	14	Читабельность FKGL (Об.)	8,2 ± 2,0	7,4 ± 1,7	0,00

* $p < 0,01$ — статистически значимое различие

Таблица 2. Лингвистические параметры текстов двух подкорпусов

Как видим, статистически значимые различия выявлены для 14 параметров (см. Табл. 2). Адаптация 54-х оригинальных текстов для учебных целей сопровождалась

снижением значений преимущественного большинства параметров, включая длину слова (строки 10, 11, 14) и читабельность и, в сущности, подтверждает выводы комплексологов о значимости морфологических параметров при оценке сложности учебного текста [Gatiyatullina et al 2024].

Увеличение значений наблюдается для глаголов, местоимений и гапакс легоменов. Результатом роста количества глаголов в изучаемых текстах (строка 3) является рост «нарративности», которая способствует более легкому восприятию текста [Вахрушева и др. 2021].

Увеличение количества местоимений (строка 7) рассматривается в современной дискурсивной комплексологии не только как маркер диалогичности текста, но и роста референциальной связности текста, предполагающей «повтор слов или концептов», формирующих эксплицитные связи внутри текста и снижающих уровень трудности восприятия текста читателем [McNamara 2013]¹². Например, O_Chehov_Ionych : *«Как-то зимой на улице его представили Ивану Петровичу; поговорили о погоде, о театре, о холере, последовало приглашение»*. O_Chehov_Ionych_B2: *«Однажды зимой на улице его представили Ивану Петровичу, они поговорили о погоде, о театре, и его пригласили в гости»*.

Несколько неожиданным оказалось увеличение количества гапакс легоменов как единиц, содержащиеся только в одном документе [Козырев 2009]. Очевидно, что рост значения данного параметра связан с вынужденной необходимостью разработчиков учебных материалов вводить контаминированные, сокращенные варианты именованных сущностей или при сокращении текста утрачивается повторяемость слова, она приобретает «уникальность» для данного текста. Например, в оригинальном тексте O_Chehov_Ionych единица *«Иван Петрович Туркин»* использована 4 раза, в Chehov Ionych B2 данное имя использовано только один раз. *«Князь Мышкин»* в O_Dostoevsky_Idiot_1 используется пять раз, и только один раз в O_Dostoevsky_Idiot_B2. В тексте контаминации K_Tolstoy_Karenina_1_C2 именованные сущности не являются гапакс легоменами, поскольку единицы *«Анна Каренина»*, *«Вронский»*, *«Левин»* и *«Кити»* используются от 3-х до 5-ти раз. Аналогичным образом в K_Turgenev_Dvor gnezdo_1_C2 фамилии *Лаврецкий*, *Калитины* и имя *Лиза* используются не более одного раза.

Анализ выявил появление гапакс легоменов в так называемых контаминированных текстовых блоках в составе текстов РКИ, которые разработчики учебных материалов вынуждены создавать для обеспечения контекста изучаемого текста. Например, для обеспечения понимания студентами текста «Ионыч» (А.П. Чехов) авторы вводят хронотоп повествования в сокращенном виде: включают краткую информацию об авторе, годе и обстоятельствах написания рассказа, главных героях, изложение основного содержания рассказа, комментариев предлагаемого читателю адаптированного отрывка. Например, в K_Turgenev_Dvor gnezdo_1_C2: *«Роман “Дворянское гнездо” был написан в 1858 году. Это одно из самых поэтических произведений Тургенева. Герой романа, Фёдор Иванович Лаврецкий, – человек редких душевных качеств: благородный, великодушный, обладающий незаурядным умом и чутким сердцем»*.

¹² Перевод на русский язык осуществлен авторами

На *ЭТАПЕ IV* был произведен (а) пропозициональный анализ пар текстов О – РКИ, а также (б) сравнение состава и количества пропозиций в оппозициях О – РКИ. Пропозициональный анализ осуществлялся по методике, разработанной в [Петрова 2021] на основе интегрированных теорий дискурсивного подхода [Дейк 1989], метода семантического анализа, основными понятиями которого являются «пропозиция» и «база текста», денотативный анализ [Бабенко, Казарин 2005; Дейк 1989; Новиков 1983]. В составе каждого предложения выделялись главные пропозиции и субпропозиции. Первые репрезентированы глаголами, глагольными конструкциями, причастиями, деепричастиями, качественными прилагательными, отглагольными существительными, адвербиальными оборотами. Последние включают следующие роли: актант (субъект, объект), модификатор актанта, сирконстант (времени, места, образа действия, цели), модификатор сирконстанта [Петрова 2021].

Проиллюстрируем пропозициональный анализ отрывком из «Ионыча» А.П. Чехова (О_Chehov Ionych) и адаптированного на его основе текста (Chehov Ionych B2) (см. *Таблицу 3*).

	O_Chehov Ionych	Chehov Ionych B2
Кол-во слов	548	491
Читабельность FKGL (Ob)	13,16	10,03

Таблица 3. Объем и читабельность фрагмента текста «Ионыч» (А.П. Чехов)

Как видим, текст стал короче, его читабельность снизилась на три уровня. Предложение из оригинального текста (O_Chehov Ionych) «*Эта семья жила на главной улице, возле губернатора, в собственном доме*» содержит одну главную (жила) и семь субпропозиций (актант (агенс) – *семья*; модификатор актанта – *эта*; сирконстанты места – *на улице, возле губернатора, в доме*; модификаторы сирконстанта – *главной, собственном*). Аналогичное предложение из текста РКИ (Chehov Ionych B2) «*Эта семья жила на главной улице, у них был свой дом*» содержит модифицированные семантические роли и включает две главных (жила, был) и семь субпропозиций (актанты (агенсы) – *семья, у них*; актант (объект) – *дом*; модификатор актанта – *эта, свой*; сирконстант места – *на улице*; модификатор сирконстанта – *главной*).

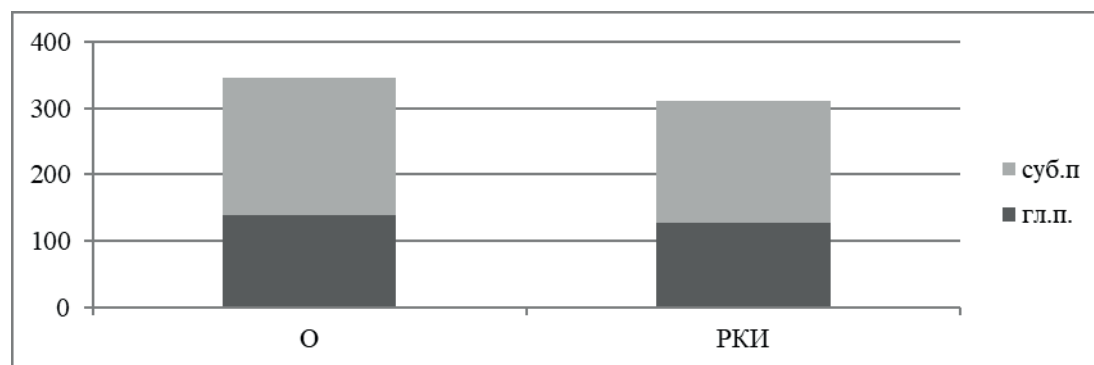


Рисунок 1. Пропозициональный анализ фрагмента оригинального и адаптированного учебного текста РКИ «Ионыч» (А.П. Чехова)

Рисунок 1 показывает, что в тексте РКИ наблюдается снижение количества пропозиций.

	Актант (агенс)	Актант (объект)	Модиф. актантов	Сир-т (время)	Сир-т. (место)	Сир-т (обр. дейст)	Сир-т (цель)	Модиф. сир-т.
РКИ	42	56	29	15	22	17	1	5
Оригинал	44	63	28	17	25	23	1	4

Таблица 4. Количество субпропозиций в текстах «Ионыч» (А.П. Чехова)

Сокращение количества главных пропозиций объясняется в том числе уменьшением отглагольных существительных в текстах РКИ. Например, *завести* «<...>он должен проводить своё **свободное время** в обществе». В оригинальном предложении «он был **назначен** земским врачом» (О) глагол **назначать** объясняется как «он **приехал работать** врачом» (В2).

В данном отрывке также находим многочисленные синонимы, как правило, более высокой частотности. Ср.: «губернский город» (О) – «провинциальный город» (В2), «это предвещало» (О) – «это значило» (В2), «поселился» (О) – «стал жить» (В2), «благверная» (О) – «жена» (В2), «пробормотал» (О) – «сказал» (В2), «выражение» (О) – «лицо» (В2), «с наступлением вечера» (О) – «вечером» (В2).

Глобальный уровень экспликации достигается размещением перед отрывком учебного текста РКИ, расширением информации об авторе, героях, кратком содержании произведения и пояснении предлагаемого читателю отрывка. Например, K_Turgenev_Dvor gnezdo_1_C2: «В отрывке из романа описывается первое свидание Лаврецкого и Лизы, их короткое счастье».

Выявлено, что количество пропозиций снижается в учебном тексте за счет удаления субпропозиций. Например, сирконстантов места «дом возле губернаторского» (О), «возле себя» (О); цели «с благотворительной целью» (О); времени «и это всякий раз»; образа действия «само собой», «мало-помалу».

Сравнительное исследование предикторов информационной (когнитивной) и лингвистической сложности учебных текстов РКИ и их источников – оригинальных (первичных) текстов художественных произведений продемонстрировало, что экспликация имплицитных смыслов в тексте РКИ осуществляется на локальном и глобальном уровнях. Изменение количества пропозиций во вторичном тексте осуществляется, с одной стороны, за счет элиминации субпропозиций места, времени и обстоятельств, а с другой стороны, за счет введения контаминированных фрагментов в начале адаптированного текста, обеспечивающих прагмалингвистический контекст учебного текста. Снижение уровня сложности осуществляется преимущественно за счет изменения морфологических параметров текста, оказывающих кумулятивное влияние на его читабельность и доступность для восприятия.

Исследование выполнено при финансовой поддержке РНФ; грант 24-78-10129: «Когнитивная и лингвистическая сложность текста».

Литература

1. Бабенко Л. Г., Казарин Ю. В. Лингвистический анализ художественного текста. Теория и практика: учебник, практикум. – М.: Флинта; Наука, 2005.

2. 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
3. Ван Дейк Т. А. Язык. Познание. Коммуникация / пер. с англ. М.: Прогресс, 1989.
4. Вахрушева А. Я., Солнышкина М. И., Куприянов Р. В., Гафиятова Э. В., Климагина И. О. Лингвистическая сложность учебных текстов // Вопросы журналистики, педагогики, языкознания. 2021. Т. 40. № 1. – С. 88-99. DOI 10.18413/2712-7451-2021-40-1-89-99.
5. Вербицкая М.В., Теория вторичных текстов: На материале современного английского языка: Автореф. дисс. ... д-ра филол. наук. М., 2000.
6. Козырев В. А. «Слово о полку Игореве» и его изучение в герценовской школе русской диалектологии // Universum: Вестник Герценовского университета. – 2009. – №. 5. – С. 40-46.
7. Коченгин А. Е., Шихин В. А. Непараметрический критерий Вилкоксона как элемент процедуры метода распознавания критических событий // Прикладная информатика. 2020. Т. 15. № 2(86). DOI 10.37791/2687-0649-2020-15-2-74-88.
8. Матвеева Н.В. Механизмы формирования смысла и содержания текста в процессе его понимания: На материале анализа вторичных письменных текстов школьников 12 -16 лет: Дис. канд. филол. наук: Уфа, 2004.
9. Новиков А. И. Семантика текста и ее формализация. М.: Наука, 1983.
10. Орлов А. И. Двухвыборочный критерий Вилкоксона – анализ двух мифов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 104. – С. 91-111.
11. Петрова А. А., Солнышкина М. И. Неподготовленный устный пересказ как вторичный текст: референциально прагматические и пропозициональные характеристики // Russian Journal of Linguistics. 2021. Т. 25. № 1. С. 221–249. DOI 10.22363/2687-0088-2021-25-1-221-249.
12. Рябова Т. П., Замега Э. Н. Применение критерия Вилкоксона для обнаружения неопределенных космических объектов // Индустриальная Россия: Вчера, сегодня, завтра: сборник научных статей по материалам IV Международной научно-практической конференции. 2020.
13. Терехова М.Ю. Лингвостилистический статус учебных видов вторичного текста. Автореф. дис. ... канд. филол. наук. М., 1988.
14. Шитова К. Г. Использование критерия Вилкоксона для оценки выраженности изменения показателей // Информационные технологии в процессе подготовки современного специалиста: межвузовский сборник научных трудов. Вып. 21. Липецк: Липецкий государственный педагогический университет имени П. П. Семенова-Тян-Шанского, 2017.
15. Gatiyatullina G. M., Solnyshkina M.I., Kupriyanov, C. R. Lexical density as a complexity predictor: the case of Science and Social Studies textbooks // Research Result. Theoretical and Applied Linguistics. 2023. Is. 9 (1). P. 11–26. DOI 10.18413/2313-8912-2023-9-1-0-2.
16. McNamara D. S. The epistemic stance between the author and reader: A driving force in the cohesion of text and writing // Discourse Studies. – 2013. – Т. 15. – №. 5. – С. 579-595.

M.I. Andreeva

Candidate of Philology, associate professor of the department
of foreign languages, Kazan State Medical University,
senior researcher, research laboratory 'Multidisciplinary text research' KFU
Kazan, Russia
email: maria.andreeva@kazangmu.ru

M.I. Solnyshkina

Doctor of Philology, professor,
Professor of the Department of theory and practice of teaching foreign languages, KFU
Head of the research laboratory 'Multidisciplinary text research'
Kazan, Russia
email: mesoln@yandex.ru

N.I. Mufazalova

assistant of the Department of theory and practice
of teaching foreign languages, KFU
Kazan, Russia
email: mufazalovanailya@gmail.com

PREDICTORS OF COGNITIVE AND LINGUISTIC COMPLEXITY OF RFL TEXTS

Abstract. The article presents the results of an ongoing study aimed at identifying predictors of informational (cognitive) and linguistic complexity of educational texts on Russian as a foreign language. The research is based on the corpus of RFL educational texts comprising 37344 words, and original texts, adapted for RFL educational purposes, comprising 69231 words. The research algorithm includes (1) selection and preprocessing of original and adapted texts; (2) assessment of linguistic complexity based on metrics of linguistic parameters calculated by the RuLingva text profiler (rulinga.kpfu.ru/); (3) assessment of the cognitive complexity of texts using propositional analysis; (4) identification of predictors of complexity based on comparative analysis. The results of the study will be used to study the difficulty of text perception by respondents speakers of Russian as a foreign language.

Keywords: RFL, text complexity, fiction text, educational text, propositions.