

Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2019»

Москва, 29 мая — 1 июня 2019 г.

ЛЕКСИЧЕСКИЙ СОСТАВ ТЕКСТОВ УЧЕБНИКОВ РУССКОГО ЯЗЫКА ДЛЯ МЛАДШЕЙ ШКОЛЫ: КОРПУСНОЕ ИССЛЕДОВАНИЕ¹

Лапошина А. Н. (antonina.laposhina@gmail.com),
Веселовская Т. С. (TSVeselovskaya@pushkin.institute),
Лебедева М. Ю. (m.u.lebedeva@gmail.com),
Купрещенко О. Ф. (ofkupr@gmail.com)

Государственный институт русского языка
им. А. С. Пушкина (Москва, Россия)

В статье представлены первые результаты сравнительного корпусного исследования современных учебников русского языка для младшего школьного возраста. Приводятся статистические данные об объеме и разнообразии лексики, используемой в учебниках, результаты анализа лексики на вхождение в частотные и тематические группы.

Ключевые слова: русский язык, учебник русского языка, корпус учебников, учебные тексты, сложность текста, русский язык в начальной школе, доказательная педагогика, сравнение корпусов

¹ Работа выполнена при финансовой поддержке РФФИ, проект 17-29-09156

LEXICAL ANALYSIS OF THE RUSSIAN LANGUAGE TEXTBOOKS FOR PRIMARY SCHOOL: CORPUS STUDY

Laposhina A. N. (antonina.laposhina@gmail.com),
Veselovskaya T. S. (TSVeselovskaya@pushkin.institute),
Lebedeva M. U. (m.u.lebedeva@gmail.com),
Kupreshchenko O. F. (ofkupr@gmail.com)

Pushkin State Russian Language Institute (Moscow, Russia)

Annotation: This paper presents the first results of a comparative corpus-based research of the modern Russian language textbooks for primary school children. Volume and diversity statistics of textbooks' vocabulary, the results of the vocabulary's analysis included in frequency and thematic groups are given.

Keywords: Russian language, textbook of the Russian language, corpus of textbooks, educational texts, text complexity, Russian language in elementary school, evidence-based pedagogy, corpora comparison

Постановка проблемы

Роль учебника русского языка для начальной школы трудно переоценить: он является не только источником первичной информации о системе языка, но и развивает навыки, необходимые для успешного освоения других предметов школьной программы: умение читать, понимать и анализировать тексты, грамотно писать, выражать свои мысли.

Качество учебных текстов — их доступность, разнообразность, актуальность и занимательность для данного возраста — является основным критерием оценки качества учебника, причем как с позиций теории учебника [Беспалько, с. 97–100], [Михеева с. 176], так и с пользовательских позиций². Попытки объективно оценить качество учебных текстов предпринимаются мировым научным сообществом уже не один десяток лет [Collins-Thompson, 2014], [DuBay, 2007]. Среди исследований последних лет, направленных на тексты российских учебников, стоит отметить исследование И. А. Оборневой, посвященное адаптации формул читабельности для русского языка [Оборнева, 2006], работу на материале учебных текстов по химии [Шпаковский, 2012], исследование сложности текста на материале учебников истории [Иванов et al., 2018]. Однако все вышеупомянутые исследования посвящены анализу читабельности текста учебника как формальному критерию понятности/доступности текста.

² т. н. «родительские» форумы, отзывы об учебниках в интернет-магазинах.

Таким образом, за рамками проведенных исследований остаются другие параметры, которые могут влиять на доступность и увлекательность учебника — тематическая принадлежность текстов, стилистические особенности, объем агнонимичной, абстрактной лексики.

Помочь в получении объективных данных о текстовом наполнении учебников может частотный анализ коллекции учебников: в настоящее время корпусные методы широко используются в исследовании и создании учебных материалов [Boulton, 2017]; [McEnery et al, 2010]; [Tribble, 2015]. Тем не менее, примеров создания и разметки корпуса учебной литературы даже в мировой практике не так много: стоит упомянуть исследование на материале учебников английского языка [Islam, 2014], корпус учебников японского языка [Sato, 2008], сравнение методических школ Южной и Северной Кореи на материале учебников английского для детей этих стран [Kim, 2017]. Однако внимание исследователей ещё не было направлено на учебные материалы по русскому языку.

Таким образом, данная работа преследует две локальные цели. Во-первых, это постановка проблемы, обнаружение лакуны в объективных данных о содержании учебников русского языка. Во-вторых, иллюстрация варианта решения этой проблемы на примере исследования двух параллелей учебников русского языка для младшей школы методами корпусной лингвистики.

Более широкой, глобальной целью исследования является описание современного состояния методики преподавания русского языка посредством количественного анализа содержания одного из главных компонентов обучающего процесса — учебника.

1. Материал исследования

Данное сравнительное исследование лексического состава учебников русского языка было проведено на материале созданного пилотного корпуса учебников русского языка для детей младшего школьного возраста (далее — корпус учебников). Отбор учебников проходил на основании Федерального перечня учебников³, отзывов сообществ учителей и родителей.

Простейшим элементом корпуса является законченный, визуально отделяемый блок текста. Разметка представляет собой метатекстовую информацию о каждом блоке текста, необходимую для анализа содержания фрагментов и сравнения методических приёмов и содержит следующую информацию:

- тип аппарата учебника (текст, наполнение упражнения, формулировка задания и др.);
- тип текста (поэзия или проза);
- авторство текста (аутентичный, адаптированный, сконструированный);

Пример разметки корпуса представлен в **Таблице 1**.

³ Приказ Министерства просвещения РФ №345 от 28.12.2018

Таблица 1. Пример метатекстовой разметки корпуса учебников

Текстовый блок	Аппарат учебника	Тип текста	Авторство	Имя автора
Летит над речкой зимородок. Очень красивая птица: брюшко оранжевое, спинка ярко-зелёная, а нос длинный и прямой, как палочка.	Наполнение упражнения. Текст	Проза	Аутентичный	Г. Скребицкий
Определите тему и главную мысль текста. Придумайте к нему заголовок.	Формулировка задания	—	—	—

2. Объём и воспроизводимость лексики в текстах учебников русского языка

Объём размеченного корпуса, на материале которого проводится данное исследование, составляет две полные параллели учебников 1–4 класса, около 300 тысяч токенов, 42 300 предложений, 15 000 текстовых блоков. На **Рис. 1** представлено распределение уникальной лексики в учебниках 1–4 классов двух анализируемых параллелей.

Для удобства визуализации и описания данных здесь и далее мы обозначим параллель учебников под ред. В. П. Канакиной и В. Г. Горецкого [Канакина et al, 2013–2014] как Канакина и учебники под ред. М. Л. Каленчук и Н. А. Чураковой [Чуракова et al, 2013–2017] как Чуракова.

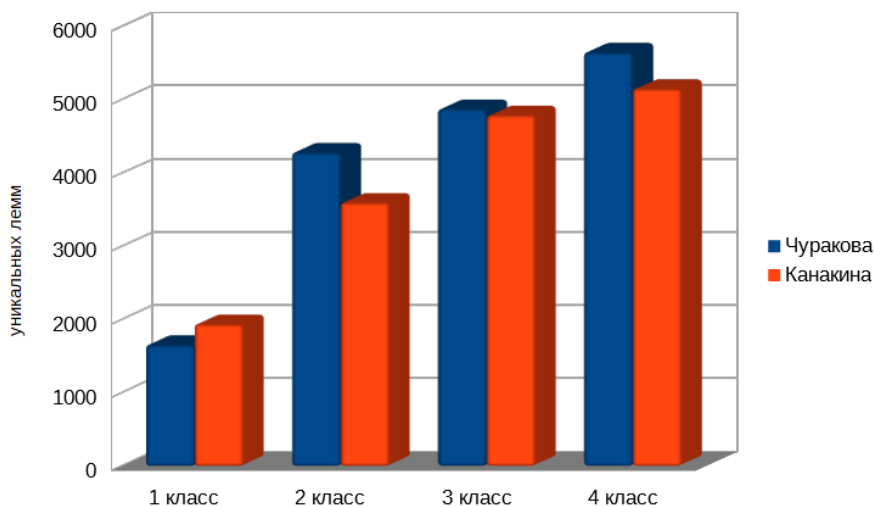


Рис. 1. Кол-во уникальных лемм в учебниках в зависимости от класса

Для параллели Канакиной общее число уникальных лемм составляет 8363, для Чураковой — 9720. При переходе от класса к классу совершенствуется навык чтения учащихся, растет их словарный запас, вместе с тем увеличивается и объем входящей в учебники лексики. На графике (Рис. 1) наглядно представлен процесс прироста лексики, при котором максимальный скачок в обоих учебниках приходится на переход из первого класса во второй. Так, например, в параллели Чураковой количество уникальных лемм в первом и втором классах отличается в 2,5 раза. Далее скорость роста объема лексики замедляется.

Логично предположить, что некоторый процент этой входящей лексики будет пересекаться, встречаться сразу в нескольких классах. Эти цифры могут служить показателем согласованности учебников внутри одной параллели между собой: такой общей базой является терминологический (*предложение, звук*) и операционный (*прочитать, списать*) аппарат учебника, отрабатываемые словарные слова, опционально может присутствовать общая сюжетная линия. Так, например, для параллели Чураковой число слов, общих для всех четырех классов даже с учетом общей сюжетной линии и системы персонажей составляет 8% (872 леммы из 9720), цифра же для параллели Канакиной, не объединенной общей сюжетной канвой, несколько выше — 14,8% (1240 лемм из 8360). Это может свидетельствовать как о более высоком уровне преемственности лексики в параллели Канакиной, её воспроизводимости из учебника в учебник, так и о тематическом однообразии текстов параллели, и требует дополнительного исследования.

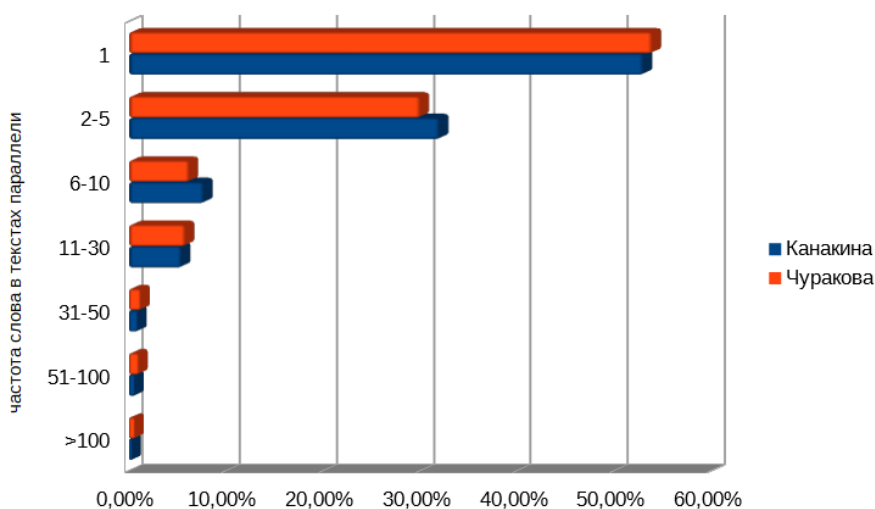


Рис. 2. Распределение слов по абсолютной частоте встречаемости во всех текстах параллели

Рис. 2 демонстрирует распределение всей лексики текстов по абсолютной частоте встречаемости. Общая схема распределения частот для двух параллелей весьма схожа. Так, две группы с наивысшей частотой (более 100)

представляют собой предлоги и союзы (т. н. стоп-слова, частотные для всех русских текстов), в параллели Чураковой также к этой группе относятся имена основных персонажей учебников. Показательно, что единственное неслужебное слово, попавшее в эту частотную группу у обеих авторов — *лес*. Эта информация нам пригодится в дальнейшем. В группе с частотой от 51 появляется лексика, оказавшаяся центральной в учебниках русского языка (*земля, человек, идти, русский, берёза, птица*). Самую большую группу (более 50% всех уникальных лемм) на **рис. 2** составляют слова, которые встречаются во всей параллели 1 раз.

Вполне ожидаемо самыми частотными словами в учебниках русского языка оказываются лингвистические термины (*звук, глагол*) и требуемые учебные операции (*спиши, прочитай*). Набор этой лексики весьма ограничен: несмотря на то, что формулировки заданий и справочная информация занимают более 60% объема учебника, лексика этих блоков составляет 11% от всех уникальных лемм учебника.

Напротив, текстовые фрагменты, включенные в учебник в качестве иллюстрации лингвистических явлений, отработки орфографии и т. д., составляют около 30% всего объема учебника, однако являются самой насыщенной и лексически разнообразной категорией: в обеих параллелях 89% всех уникальных лемм встречается именно в текстах. Если набор терминов хотя бы в какой-то мере регулируется учебными программами, то выбор текстовых фрагментов полностью зависит от авторов учебника, что и представляет большой интерес для исследования. Поэтому в дальнейших разделах речь будет идти о подсчётах, выполненных на материале собственно текстовых фрагментов учебников.

3. Частотный анализ лексики учебников

Учебник русского языка готовит ученика к грамотному использованию родного языка в разных ситуациях и контекстах общения, формирует целостный, социально ориентированный взгляд на мир⁴. В таком случае логично предположить, что учебник должен содержать в себе усредненную модель русского языка, в равной степени отражать различные темы, постепенно наращивая объем этой лексики от простой к более сложной. Однако знакомство с учебными материалами позволяет выдвинуть предварительную гипотезу о тематической несбалансированности текстов в современных учебниках русского языка и господстве темы природы. Проверить эту гипотезу на практике возможно с помощью частотного анализа лексики учебников и сравнения полученных данных с условным «стандартным» русским языком. В качестве такой отправной точки для сравнения мы будем использовать данные Нового частотного словаря русской лексики О. Н. Ляшевской и С. А. Шарова [Ляшевская, Шаров, 2009], далее — частотный словарь.

⁴ Приказ Минобрнауки России № 373 от 06.10.2009.

3.1. Частотная лексика в учебниках русского языка

Обратимся вначале к высокочастотной лексике текстовых фрагментов учебников. **Таблица 1** представляет собой список самых популярных самостоятельных частей речи по частотному словарю, в параллели Канакиной и параллели Чураковой.

Таблица 1. Самые частотные слова самостоятельных частей речи

Частотный словарь	ipm	Параллель Канакиной	ipm	Параллель Чураковой	ipm
быть	12 160	быть	7 540	быть	6 821
год	3 727	слово	6 263	лес	3 427
мочь	2 912	лес	3 873	солнце	2 537
человек	2 723	язык	3 667	сказать	2 504
сказать	2 396	земля	3 172	вода	2 372
ещё	2 323	русский	2 884	день	2 372
уже	2 179	вода	2 801	маленький	2 239
время	2 015	птица	2 554	река	2 207
говорить	1 755	солнце	2 554	жить	2 207
знать	1 713	весна	2 472	идти	2 142
стать	1 621	день	2 389	земля	2 076
дело	1 412	белый	2 389	собака	2 010
жизнь	1 389	снег	2 348	снег	1 845

Первое, что обращает на себя внимание — разница в распределении слов по частям речи. В списках по учебникам значительно меньше глаголов и больше существительных: если в частотном словаре в таблице представлено 3 существительных и 6 глаголов, то в параллели Канакиной соотношение равно 10 и 1, Чураковой — 8 и 4. Это может сигнализировать о преобладании в учебниках описательных текстов над повествовательными, а следовательно, малом количестве действий. Так, примеры 1 и 2 демонстрируют характерные для учебников тексты описания природы без участия глаголов. По этому параметру параллель Чураковой оказывается ближе к корпусу, чем параллель Канакиной.

- (1) *Зимний тёплый день. На земле мягкий снег.*
(Канакина et al, 2013–2014, 2 кл., ч. 2, стр. 139)
- (2) *А сегодня — что за день! Солнце, птицы! Блеск и счастье! Луг росист...*
(Канакина et al, 2013–2014, 4 кл., ч. 2, стр. 138)

Данные таблицы 1 обнаруживают также явные тематические доминанты текстов учебников. В абсолютные лидеры по частотности в обеих параллелях выходит *лес* (211 ipm по частотному словарю), а также именованя стихий и природных явлений: *земля, вода, солнце, снег* (494, 484, 165, 125 ipm по частотному словарю соответственно). На этом этапе становятся видны и личные предпочтения авторов касательно животных: параллель Канакиной обнаруживает интерес к птицам (79 упоминаний вариантов *птица/птичка/пичужка* + 560 упоминаний

конкретных видов птиц, см. подробнее [раздел 3.2](#)), параллель Чураковой на первый взгляд отдаёт предпочтение собакам, однако при более детальном рассмотрении выявляется небольшое превосходство кошек (95 упоминаний на варианты *собака/пёс/щенок/собачка* и 104 на *кот/кошка/котёнок/котик*).

Таким образом, данные [Таблицы 1](#), во-первых, подтверждают гипотезу о значительном уклоне русских учебных текстов в сторону описания природы и погодных явлений, а во-вторых, демонстрируют явный частеречный дисбаланс, который требует отдельного изучения.

Остановимся подробнее на тематической принадлежности наиболее часто встречающихся слов в учебниках. Примерные тематические группы расположены ниже по убыванию количества представителей и их общей частоте, в скобках указаны абсолютные частоты в параллелях авторов Чураковой (число слева) и Канакиной (число справа):

- Природа и животные: *лес* (136/118), *береза* (26/117), *заяц* (34/37), *река* (56/45);
- Погода и сезоны: *солнце* (80/86), *снег* (72/79), *зима* (54/44);
- Деревенская лексика: *село* (30/15), *сад* (32/22), *деревня* (13/25), *корова* (22/11), *сено* (4/9), *огород* (20/9), *грядка* (7/6), *изба* (6/6);
- Школьная лексика: *класс* (17/18), *читать* (16/23), *учительница* (15/4), *карандаш* (10/17);
- Бытовая лексика: *хлеб* (15/23), *шапка* (11/10), *мыло* (12/4), *подушка* (13/5), *магазин* (8/7);
- «Детская» лексика: *мама* (36/26), *мяч* (26/5), *игрушка* (8/3), *кататься* (3/3), *качели* (2/0), *велосипед* (6/2), *коньки* (2/6), *лыжи* (2/4);
- Городская лексика: *город* (28/42), *парк* (3/2), *автобус* (4/0), *музей* (4/1), *квартира* (4/3), *театр* (0/5); *зоопарк* (7/4), *цирк* (3/1);

На первый план выходят два блока лексики, тематически связанные с природой и погодой: они значительно опережают остальные группы по частотности. Также интересно отметить преобладание деревенской лексики над городской. Этот факт может быть связан как с использованием материала сказок, место действия которых обычно связано с деревней, так и с тем, что жизнь в деревне сильнее связана с природой, которая в свою очередь является самой частотной темой. Интересно отметить, что «детская» лексика, связанная с играми, развлечениями и спортом и городская лексика представлены значительно скуднее.

Также стоит отметить частотные по корпусу слова, ни разу не появившиеся в учебниках (например, *документ, информация, система, проблема, центр, цена*). Безусловно, они часто встречаются во «взрослых» публицистических текстах и деловых документах и трудно говорить о методической необходимости их добавления в учебник. Однако эта информация лишний раз подчеркивает стилистическую ориентацию детских учебников на художественные и описательные тексты.

3.2. Много в учебниках, мало в корпусе: осознанный выбор авторов

Редкая по частотному словарю на основе корпуса русского языка, но часто появляющаяся в учебниках лексика особенно интересна для анализа, так как она ярче всего отражает осознанный (или неосознанный) выбор авторов, касающийся тематики текстового наполнения учебника, воспроизводимости лексики, включения слов различных исторических эпох и т. д. **Таблица 2** представляет собой топ-10 слов, часто появляющихся в учебниках, но не вошедших в список 10 тысяч частотных слов русского языка.

Учебники параллели Чураковой связаны между собой центральными персонажами — Машей и Мишей, которые осваивают правила русского языка под руководством волшебницы Анишит Йокоповны и её помощников Асырк и Торк⁵. Эта сюжетная линия хорошо просматривается в таблице — персонажи заняли 4 из 10 популярных низкочастотных слов параллели. В общей сложности имена персонажей насчитывают 1500 упоминаний за параллель, эту информацию стоит учитывать при расчете объема лексики.

Таблица 2. Топ-10 слов из параллели, не входящих в 10 тыс. частотных слов русского языка

Параллель Канакиной		Параллель Чураковой	
лемма	абс. частота	лемма	абс. частота
прилетать	29	Анишит Йокоповна	239
метель	25	волшебница	104
скворец	24	норка	41
дождик	23	Асырк	30
иней	22	соловей	23
клюв	22	ива	20
вьюга	20	Торк	18
душистый	19	барашек	17
ласточка	19	дождик	17
соловей	18	радуга	16

В обеих параллелях снова подтверждается предположение о главенствующих тематиках текстов — природа, погодные явления и животный мир, однако и тут можно заметить ряд любопытных особенностей. Так, параллель Канакиной обнаруживает явное тяготение к описанию зимнего времени года (*метель, иней, вьюга*), осадкам (*дождик, снег, иней*) и птицам (*прилетать, клюв, скворец* и др.), в то время как параллель Чураковой — к лесу (*пенёк, опушка, норка*). Хотелось бы отдельно отметить масштаб интереса к текстам о птицах в учебниках Канакиной: всего в параллели учебников было найдено упоминание 46 видов птиц с общей частотой встречаемости более 560 раз за параллель. Помимо привычных представителей крылатой фауны, таких как *ворона, воробей, сорока,*

⁵ Выбор имен персонажей становится немного понятнее, если прочитать их наоборот.

синица, на страницах учебников обнаруживаются и более незаурядные примеры, такие как *рябчик, сойка, киви, оляпка, дряба*.

Таким образом, предположение о тематической сбалансированности текстов учебника также не находит подтверждения в данных **Таблицы 2**.

3.3. «Одноразовая» лексика

Данные **Рисунка 2** показали, что более 50% всей лексики анализируемых учебников встречается 1 раз за всю параллель. Предположение, что это повседневная лексика, встречающаяся в текстах, но не требующая повторения и отработки, не подтверждается: 48% этой лексики не входит в частотный список 10000. Рассмотрим некоторые случаи употребления такой единичной лексики подробнее:

1. Задания, направленные на изучение или уточнение лексического значения этих слов. Чаще всего это задания на определение «старых» и «новых» слов, где встречаются историзмы (*бурлак*), устаревшая лексика (*ветрило, длань, ямщик, скань*). Интересен и тот факт, что некоторые слова, задуманные авторами как примеры «новых» слов, уже стали историзмами (*дискета, радиотелефон*), но всё ещё присутствуют в современных переизданиях учебника.

2. Иллюстрация специфических языковых явлений. Так, например, слово *гусятница* приводится для отработки суффикса *ниц*, а пример 1, вызвавший резонанс на родительских форумах — парность согласного звука *З* по твердости-мягкости:

- (3) *Зябнет*
Зубр,
Зайчонок
Зябнет,
Зуй⁶ зазяб,
Зазяб и
Зяблик.

(Чуракова et al, 2013–2017, 1 кл., стр.45).

3. Выбор авторами фрагмента аутентичного текста или самостоятельное составление текстов с редкой неповторяющейся лексикой для упражнений на анализ текста (найти главную мысль, озаглавить и т. д.)

- (4) *Мальчики посадили луковицы гладиолусов и клубни георгинов.* (текст авторов учебника, Канакина et al, 2013–2014, 2 кл., ч. 1, стр. 21).
- (5) *Каждый год, принимаясь за огород, тётя Шура кладёт на меже тряпцу и носит в неё всё, что блеснёт под лопатой: бусину, черепок, косянкую пуговицу, наконечник стрелы...* (фрагмент аутентичного текста В. Пескова, Чуракова et al, 2013–2017, 4 кл., ч.2, стр. 63).

⁶ Толковый словарь Д. Н. Ушакова: Зуй, зуй, мн. зуй, муж. (обл.). Местное название некоторых птиц из рода куликов. Пример: болотный зуй.

Описанные примеры показывают, что появление редкой лексики один раз за параллель может быть обусловлено специфическим заданием (группа примеров 1), или быть личным авторским решением. Зачастую наличие устаревшей лексики и историзмов связано с продолжением традиции литературоцентричности в обучении русскому языку и использованием в качестве учебного материала текстов русской классики. Между тем, иллюстрация новых лингвистических явлений (группа примеров 2) или аналитическая работа с текстом (группа примеров 3) на незнакомых школьнику словах и ситуациях, далеких от его жизни, представляется спорным методическим решением. Кроме того, такая лексика с большой вероятностью вызовет трудности при чтении. Поэтому вопрос о целесообразности и объеме включения такой лексики в учебники должен исследоваться отдельно и подтверждаться экспериментально, например, методами айтрекинга или экспериментами на скорость чтения с заданиями на понимание.

4. Выводы

Данное исследование является первым шагом в направлении исследования лексики учебников и в большей степени намечает проблемные точки, чем предлагает готовые ответы. Однако подведем промежуточные итоги исследования лексического состава учебников русского языка для младшей школы на материале двух параллелей учебников.

Текстовые фрагменты являются самым лексически разнообразным блоком учебника, в них содержится 89% всей уникальной лексики. Более 50% этой лексики появляется один раз за всю параллель. Явными тематическими доминантами текстов оказываются природа, погодные явления и животные. Этот вывод вряд ли можно назвать революционным, однако он впервые подкреплен реальными цифрами. Деревенская лексика представлена в учебниках шире, чем городская и детская. Скучно представлена лексика, свойственная для публицистического стиля. Разница в частотах глаголов и существительных с частотным словарём может сигнализировать о перевесе текстов описательного характера.

В качестве перспективных направлений работы представляется коллокативный анализ учебных текстов, более подробное количественное исследование принадлежности лексики учебников к определенной тематической группе, исторической эпохе, типу текста и функциональному стилю русского языка.

5. Благодарности

Авторский коллектив сердечно благодарит студентов 1 курса магистратуры филологического факультета и С. И. Ельникову, заведующую кафедрой методики преподавания РКИ Государственного института русского языка им. А. С. Пушкина, за помощь в разметке корпуса, а также анонимных рецензентов за ценные комментарии.

Литература

1. *Boulton A.* (2017), Corpora in language teaching and learning.//Language Teaching. 50.4, pp. 483–506.
2. *Collins-Thompson K.* (2014), Computational assessment of text readability: a survey of current and future research. In: François, Thomas and Delphine Bernhard (eds.), Recent Advances in Automatic Readability Assessment and Text Simplification. //Special issue of International Journal of Applied Linguistics, pp. 97–135.
3. *Dax Thomas* (2005), Type-token Ratios in One Teacher’s Classroom Talk: An Investigation of Lexical Complexity.
4. *DuBay W. H.* (2007), Smart Language: Readers, Readability, and the Grading of Text. ERIC.
5. *Islam, Md. Zahurul* (2014), Multilingual Text Classification using Information-Theoretic Features PhD Thesis; Goethe University Frankfurt.
6. *Ivanov V. V., Solnyshkina M. I., Solovyev V. D.* (2018), Efficiency of Text Readability Features in Russian Academic Texts//Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2018”, Moscow, pp. 284–293.
7. *Jeong-ryeol Kim* (2017), A Comparative Look at South and North Korean English Education Based on the Analysis of Textbook Corpus//Corpus Linguistics Research. Vol. 3 (2017.09). pp. 1–22.
8. *McEnery, Tony; Xiao, Richard.* (2010), What corpora can offer in language teaching and learning. / Handbook of Research in Second Language Teaching and Learning. ed. / E. Hinkel. Vol. 2 London & New York: Routledge, pp. 364–380.
9. *Norberg, Cathrine & Nordlund, Marie* (2018), A Corpus-based Study of Lexis in L2 English Textbooks. Journal of Language Teaching and Research. Vol. 9, No. 3, pp. 463–473.
10. *Sato, Satoshi & Matsuyoshi, Suguru & Kondoh, Yohsuke* (2008), Automatic Assessment of Japanese Text Readability Based on a Textbook Corpus.Proceedings of the International Conference on Language Resources and Evaluation, LREC 2008, pp. 654–660.
11. *Tribble, Christopher* (2015), “Teaching and language corpora.” Multiple affordances of language corpora for data-driven learning, 69, pp. 37–62.
12. *Bespalko V. P.* (1988), Theory of the textbook: Didactic aspect [Teoriya uchebnika: Didakticheskij aspekt]. — M.: Pedagogy [M.: Pedagogika].
13. *Bim I. L.* (1975), To the development of a foreign language textbook theory [K razrabotke teorii uchebnika inostrannogo yazyka], Russian language abroad [Russkij yazyk za rubezhom], Moscow, № 4.
14. *Glinkina L. A.* (2008), Illustrated Dictionary of Forgotten and Difficult Words of the Russian Language [Illyustrirovannyj tolkovyy slovar’ zabytykh i trudnykh slov russkogo yazyka]. World of Encyclopedia Avanta + [Mir entsiklopediy Avanta+].
15. *Lerner I. Ya.* (1974), Criteria for the complexity of some elements of the textbook [Kriterii slozhnosti nekotorykh elementov uchebnika]. Problems of the school textbook [Problemy shkol’nogo uchebnika], vol. I, pp. 23–25.

16. *Lyashevskaya O. N., Sharov S. A.* (2009), Frequency Dictionary of the Modern Russian Language (on the materials of the National Corpus of the Russian Language) [Chastotnyy slovar' sovremennogo russkogo yazyka (na materialakh Natsional'nogo korpusa russkogo yazyka)]. M.: Azbukovnik [Azbukovnik].
17. *Matskovsky M. S.* (1976), The Problem of Readability of Printed Material [Problema chitabel'nosti pechatnogo materiala]. The Meaning Perception of a Voice Message in Mass Communication [Smyslovoye vospriyatiye rechevogo soobshcheniya v usloviyakh massovoy kommunikatsii]. M.: Science, pp. 126–141.
18. *Mikk, J. A.* (1981), Optimizing the complexity of an educational text: to help authors and editors [Optimizatsiya slozhnosti uchebnogo teksta: v pomoshch' avtoram i redaktoram]. M.: Education [Prosveshcheniye].
19. *Mikheeva S. A.* (2015), System of formalized criteria for evaluating a school textbook [Sistema formalizovannykh kriteriyev otsenki shkol'nogo uchebnika], Educational Studies [Voprosy obrazovaniya], issue 4, pp. 147–183.
20. *Obornneva I. V.* (2006), Automated assessment of the complexity of educational texts based on statistical parameters [Avtomatizirovannaya otsenka slozhnosti uchebnykh tekstov na osnove statisticheskikh parametrov]. Thesis for Ph.D. [Dissertatsiya na soiskaniye step. K.p. n.], Moscow.
21. *Pedagogical and ergonomic requirements for means of education* [Pedagogiko-ergonomicheskkiye trebovaniya k sredstvam obucheniya]. — SPb, M: Christmas +, IOSO RAO, 2000. — 64 p.
22. *Plungyan V. A.* (2005), Why do we make the National Corpus of the Russian language? [Zachem my delayem Natsional'nyy korpus russkogo yazyka?] Domestic Notes [Otechestvennyye zapiski]. № 2, pp. 296–308.
23. *Order of the Ministry of Education and Science of Russia dated October 06, 2009 № 373 (as amended on 12/31/2015) “On approval and implementation of the federal state educational standard of primary general education”* (Registered in the Ministry of Justice of Russia December 22, 2009 N 15785) [Priказ Minobrnauki Rossii ot 06.10.2009 № 373 (red. ot 31.12.2015) «Ob utverzhdenii i vvedenii v deystviye federal'nogo gosudarstvennogo obrazovatel'nogo standarta nachal'nogo obshchego obrazovaniya» (Zaregistrirvano v Minyuste Rossii 22.12.2009 N 15785)].
24. *Order of the Ministry of Education of the Russian Federation of December 28, 2018 № 345 “On the federal list of textbooks recommended for use in the implementation of state-accredited educational programs of primary general, basic general and secondary general education”* [Priказ Ministerstva prosveshcheniya RF ot 28.12.2018 №345 «O federal'nom perechne uchebnikov, rekomendovannykh k ispol'zovaniyu pri realizatsii imeyushchikh gosudarstvennyuyu akkreditatsiyu obrazovatel'nykh programm nachal'nogo obshchego, osnovnogo obshchego, srednego obshchego obrazovaniya»].
25. *Frumin I. D., Dobryakova M. S., Barannikov K. A., Remorenko I. M.* (2018), Universal Competences and New Literacy: What to Teach Today for Success tomorrow. Preliminary findings of the international report on trends in the transformation of school education [Universal'nyye kompetentnosti i novaya gramotnost': chemu uchit' segodnya dlya uspekha zavtra. Predvaritel'nyye

vyvody mezhdunarodnogo doklada o tendentsiyakh transformatsii shkol'nogo obrazovaniya]. National Research University Higher School of Economics, Institute of Education. Modern education analytics [Natsional'nyy issledovatel'skiy universitet «Vysshaya shkola ekonomiki», Institut obrazovaniya. Sovremennaya analitika obrazovaniya]. № 2 (19). М.: HSE.

26. *Shpakovsky Yu.* (2012), Estimation of the difficulty of perception and optimization of the complexity of an educational text. On the material of texts in chemistry [Otsenka trudnosti vospriyatiya i optimizatsiya slozhnosti uchebnogo teksta. Na materiale tekstov po khimii]. LAP Lambert Academic Publishing.

Источники анализируемого материала

27. *Канакина В. П.* Русский язык. Коллекция учебников 1–4 класс. Учеб, для общеобразоват. организаций с прил. на электрон, носителе / В. П. Канакина, В. Г. Горецкий. 5-е изд. — М.: Просвещение, 2013–2014 гг.
28. *Чуракова Н. А., Каленчук М. Л.* Русский язык. Коллекция учебников 1–4 класс: учебник / Н. А. Чуракова, М. Л. Каленчук, Т. А. Байкова, О. В. Малаховская — М.: Академкнига/Учебник, 2013–2017 гг.
29. *Канакина В. П.* Русский язык. Рабочие программы. Предметная линия учебников системы «Школа России». 1–4 классы: пособие для учителей общеобразоват. организаций / [В. П. Канакина, В. Г. Горецкий, М. В. Бойкина и др.]. — М.: Просвещение, 2014. — 340 с.