DOI: 10.26907/2074-0239-2020-61-3-72-80
УДК 811.161.1'075.8

ЛЕКСИЧЕСКИЕ ПАРАМЕТРЫ УЧЕБНОГО ТЕКСТА (НА МАТЕРИАЛЕ ТЕКСТОВ УЧЕБНОГО КОРПУСА РУССКОГО ЯЗЫКА)

© Екатерина Мартынова, Марина Солнышкина, Амина Мерзлякова, Диана Гизатулина

LEXICAL PARAMETERS OF THE ACADEMIC TEXT (BASED ON THE TEXTS OF THE ACADEMIC CORPUS OF THE RUSSIAN LANGUAGE)

Ekaterina Martynova, Marina Solnyshkina, Amina Merzlyakova, Diana Gizatulina

The article presents the results of the analysis of the influence exerted by three lexical parameters (abstractness, frequency, and lexical diversity) on the complexity of educational texts. The representative corpus of academic texts, compiled for the study, comprises textbooks on social sciences for grades 5–11 (139971 tokens) and Biology (81712 tokens). The increasing complexity of the texts from grades 5–11 is confirmed by the steadily increasing Flash-Kincaid readability indices in both Social Science and Biology texts. The discrepancies between textbooks readability indices and the grade level are insignificant and do not exceed 1.0. Vocabulary Frequency as a text parameter does not correlate with complexity: the increasing complexity of texts across the grades is not accompanied by a stable increase in the corresponding frequency indices. The degree of abstractness of the texts increases in both lines of textbooks, however, texts on Social Studies have a higher degree of abstractness than Biology texts. The lexical diversity of texts in Biology textbooks is lower than the corresponding indicator in texts on social sciences, although the correlation of this parameter with text complexity requires extra studies. The results of the research and its algorithm can be used both in developing the theory of text complexity and in ranking academic texts for particular categories of readers.

Keywords: lexical diversity, frequency, readability, abstractness, text complexity.

В статье представлены результаты анализа влияния трех лексических параметров (абстрактность, частотность, лексическое многообразие) на сложность учебного текста. Исследование выполнено на репрезентативном корпусе текстов учебников для 5-11 классов по обществознанию (139971 словоупотреблений) и по биологии (81712 словоупотреблений). Возрастающая от класса к классу сложность текстов подтверждена стабильно повышающимися индексами читабельности Флеша-Кинкейда как в текстах по обществознанию, так и в текстах по биологии. Несоответствия читабельности учебника номеру класса незначительны и не превышают 1.0. Частотность как параметр словарного состава изученных текстов не коррелирует со сложностью: усложнение текстов не сопровождается стабильным повышением индекса частотности текстов. Степень абстрактности изучаемых текстов растет в обеих линейках учебников, при этом тексты по обществознанию имеют более высокую степень абстрактности, чем тексты по биологии. Лексическое многообразие текстов учебников по биологии ниже, чем соответствующий показатель в текстах по обществознанию, однако взаимосвязь данного параметра со степенью сложности требует дополнительных исследований. Результаты и алгоритм исследования могут быть использованы как при изучении общей проблематики сложности текста, так и при оценке соответствия текста лингвистическим способностям конкретной категории читателей.

Ключевые слова: лексическое разнообразие, частотность, читабельность, абстрактность, сложность текста.

Вместе с изменениями в отечественном и мировом образовательных пространствах возникла реальная необходимость осуществления полного комплексного анализа учебных текстов, используемых на различных ступенях образования. По-

следнее, в свою очередь, продиктовано значимостью обеспечения качества образовательного процесса, во многом обусловленное соблюдением принципа систематичности и последовательности «от простого к сложному» [Солнышкина,

Калинкина, Зиганшина с. 141]. Для авторов учебных материалов данный принцип реализуется в необходимости представления информации с нарастающим, но не флуктуирующим уровнем сложности. В самом общем представлении сложность текста определяется морфологическими (длина и количество морфем), лексическими (частотность, абстрактность, уникальность, этноспецифичность, неоднозначность, референциальная специфичность и др.), синтаксическими (комплексность модели, порядок слов, длина предложения и др.) характеристиками текста, а также его информационной (когнитивной) сложностью [Солнышкина, Казачкова, Харькова, с. 75].

Очевидно, что сложность учебного текста представляет собой особую лингвистическую категорию, отличную, например, от сложности рекламного текста [Уша с. 157]. Последнее связано со значимостью полного понимания текста обучающимися в условиях лингводидактического дискурса, а с другой стороны, с необходимостью обеспечения постоянного и равномерного увеличения сложности текстов в учебном процессе. Взаимодействие коммуникантов в рамках лингводидактического дискурса обеспечивается созданием единого когнитивного пространства учителя и обучающегося, разделяемого участниками диалога информационной среды [Гюрджян, с. 96]. Успешная коммуникация предполагает, что все участники обладают семью типами знания, которые категоризируются следующим образом: специфицирующее, пропозициональное, референциальное, предикационное, экзистенциальное, а также знание кода и знание ситуации общения, то есть дискурсивной ситуации [Йокояма, с. 131]. Во избежание когнитивного конфликта предполагается, что коммуникант должен хранить в памяти и уметь использовать (а) экзистенциальное знание, то есть знание о «бытии», существовании «терма Х», то есть объекта Х; (б) предикационное знание, то есть знание о существовании «терма Х» при некоторых обстоятельствах Y (параметры, действия, изменения, трансформации и т. д.); (в) пропозициональное знание или знание пропозиций, то есть глубинных синтаксических структур «Х ((не) является) Y»; (г) знание о способности пропозиции быть заполненной специфическими термами (специфицирующее знание); (д) знание о том, что каждому объекту (терму) соответствует определенный перечень параметров, характеристик (референциальное знание); (е) знание кода, то есть языка коммуникации; (ж) знание о дискурсивной ситуации (субъект, местоположение, время, коагенс(ы), интенции и т. д.) [Гюрджян, с. 97]. При

передаче / получении информации для коагенсов лингводидактического дискурса наибольшую значимость имеет готовность и умение каждого из них воспринимать и понимать пропозиции другого [Грайс, с. 222–224], что в условиях передачи специализированного знания (обучение) делает обязательным не только знание кода, но и референциальные знания, то есть знание параметров референта и способность правильно использовать их номинации в различных коммуникативных ситуациях. Сама ситуация процесса восприятия высказывания (текста) описана А. А. Леонтьевым как способность перевести смысл одного текста в текст другой формы [Леонтьев, с. 141]. Таким образом утверждается, что полное понимание достигается только при «переводе», переформулировке первичного (оригинального) текста во вторичный. При этом вторичный неизбежно будет отличаться от первого как в отношении формы (расширение, сокращение, замена), так и в отношении значения (добавочные денотативные смыслы, повторы, добавочные коннотативные смыслы).

Языковые особенности учебного текста не раз становились предметом исследования лингвистов и методистов [Захарова, с. 10–12], [Абдуллаева, с. 719], [Гюрджян, с. 96]. В современной отечественной парадигме учебным текстом именуется «особый вид вторичного речевого произведения, обладающий рядом признаков, существенно отличающих его от других видов вторичных текстов. Вторичность учебного текста носит особый характер, поскольку он не заменяет полностью первоисточник, а служит его продолжением и развитием в соответствии с коммуникативной задачей» [Пешкова, Яхиббаева, с. 57]. Расширяя данное положение, О. Ю. Редькина указывает на три основных характеристики учебного текста, функционирующего в рамках лингводидактического дискурса: 1) учебный текст произведен или вторичен; 2) во вторичном тексте имеет место утрата, замена или размывание образа автора первичного текста, то есть исходного научного текста, на базе которого (-х) создавался вторичный (учебный) текст; 3) степень потенциальной вероятности генерирования нескольких учебных (вторичных) текстов на основе исходного научного текста довольно высока, если адаптация (генерирование учебных текстов) осуществляется разными субъектами [Редькина, с. 126].

Наиболее общие требования к учебному тексту сформулированы в Федеральном госстандарте, который предлагает следующую характеристику учебных изданий: это «издания, содержащие систематизированные сведения научного

или прикладного характера, изложенные в форме, удобной для изучения и преподавания, и рассчитанные на учащихся разного возраста и степени обучения» [ГОСТ 7.60 – 2003. Издания. Основные виды. Термины и определения]. Типологическая характеристика языка учебных изданий также включает указание на необходимость редакторской оценки терминосистемы обучающего издания и его соответствие «уровню развития определенной области науки и техники, действующим стандартам и нормам языка» [Антонова, с. 264–270]. При этом оценка сложности современных отечественных учебников и учебных пособий осуществляется несколькими способами: 1) на основе экспертной оценки [Петрова, Окладникова, с. 91]; 2) на основе формул читабельности [Оборнева, с. 15], [Экспресс-анализ методов оценки сложности текстов учебно-методических материалов]; 3) при помощи комплексных методик, включающих расчет ряда лексических и синтаксических параметров [Наумов, Выхованец, с. 121], [Савельев, с. 66].

Отечественные и зарубежные исследования подтверждают необходимость особого обращения к лексическим параметрам текста при оценке его сложности [Solnyshkina, Harkova, Kazachkova, с. 107]. При этом ученые используют различные лексические параметры для оценки лексической сложности текста: (1) доля слов в тексте, входящих в лексические минимумы для читателей определенных категорий; (2) частотность словарного состава текста [Лапошина]; (3) отношение количества слов с этнокультурными, терминологическими и коннотативными элементами в семантике к количеству полных слов в тексте [Самсонов, Чмыхова, Давыдов, с. 6]; (4) лексическая близость слов текста (Lexical Tightness) [Flor, Klebanov, Sheehan, с. 32]; (4) частотность определенных коллокаций [Solnyshkina, Ivanov, Solovyev]. Однако наиболее часто используемый параметр – это лексическое многообразие (Type-Token Ratio, TTR) текста, рассчитываемый как отношение неповторяющихся в тексте слов ко всем словам текста. Основная идея, лежащая в основе данной теории сложности состоит в том, что в более сложных текстах авторы используют более разнообразный словарный запас, а сложность самих текстов растет по мере увеличения доли уникальных (неповторяющихся) слов в тексте [Ляшевская, с. 412].

Рассмотрим, каким образом изменяются лексические параметры русскоязычных учебных текстов по мере их усложнения в лингводидактическом дискурсе.

В качестве языкового корпуса для осуществления исследования были использованы две вы-

борки учебных текстов для 5-11 классов: тексты учебников по обществознанию [Учебники по Обществознанию] и тексты по биологии [Биология. 5 кл.], [Биология: учеб. пособие для 7-го кл.], [Биология: Человек и его здоровье], [Общая биология: Учебник для 10-11 кл.]. Общий объем учебного корпуса составил 795760 словоупотреблений (см. Таблицу №1).

Таблица 1.

Учебный корпус русского языка

Класс	Обществознание	Биология
5	37232	31269
6	33353	
7	46955	40622
8	90579	58958
9	87203	33474
10	114432	81712
11	139971	

Оценка лексических параметров производилась при помощи онлайн сервиса RusAC (<http://tykau.pythonanywhere.com/>), осуществляющего морфологический и лексический анализ текстов на основе следующих 27 параметров: две средние величины (длина предложения и длина слова), абсолютные метрики: общее количество слов, количество предложений, количество частей речи (существительных, глаголов, прилагательных, местоимений, наречий), морфологическая дистрибуция категории падежа для существительных и времени для глагола и, наконец, индексы, рассчитываемые при помощи специальных формул: индекс читабельности, индекс частотности, уровень (степень) абстрактности, коэффициент лексического многообразия [Солнышкина, Мартынова, Андреева, с. 50]. Для обеспечения репрезентативной выборки при расчете указанных параметров последовательно применялись следующие выборки: 1) полные тексты учебников; 2) 10 равномерно удаленных друг от друга текстов, объем каждого из которых составляет не менее 5% текста. Во втором случае анализу подвергалось 50 % исходного текста, содержащего, по оценкам экспертов [Biber, т. 23, с. 130–138], 70% всего лексического состава текста.

Рассмотрим динамику трех лексических параметров (частотность, абстрактность, лексическое разнообразие (TTR) на фоне изменения (роста) читабельности в текстах учебников для 5-11 классов (см. Таблицу 2) и охарактеризуем данные параметры.

Читабельность текстов (ФК) рассчитывается при помощи индекса Флеша-Кинкейда, модифицированного для русского языка. В качестве метрик при расчетах использованы два стандартных параметра – длина предложения и длина слова:

FKG mod ($SIS = 0.36 \times ASL + 5.76 \times ASW - 11.97$), где ASL – среднее количество слов в предложении; ASW – среднее количество слогов в слове [Solovyev, Ivanov, Solnyshkina, с. 3050]. Индекс Флеша-Кинкейда удобен тем, что ранжирует тексты по количеству лет формального образования потенциального читателя изучаемого текста. Величина данного индекса находится в диапазоне от 1.0 (минимальная читабельность) до 20.0 и выше (кандидат наук).

Частотность или *объективная частотность лексики* (Ч) как параметр сложности признан многими учеными [Solnyshkina, Narkova, Kazachkova, с. 105] и рассчитывается на основе списков частотных слов [Лапошина], создаваемых при помощи выявления частоты использования конкретного слова в репрезентативных и сбалансированных корпусах текстов. Метрика данного параметра определяется отношением числа словоупотреблений конкретного слова (вхождений в корпус) к общему количеству всех слов (словоупотреблений) в данном корпусе. Для сравнения частот данные нормализуются на миллион вхождений: ipm – item per million. Считается, что наиболее частотные единицы менее сложны для обработки головным мозгом и восприятия [Rudell, с. 460].

Общепризнанным параметром, детерминирующим сложность текста, является *абстрактность* или *доля абстрактных слов* (А) в тексте. Расчеты данного параметра представляют особую трудность в силу относительности самой категории абстрактности: одно и то же слово может иметь различную степень абстрактности в зависимости от контекста [Солнышкина, Казачкова, Исмаева, с. 75]. Например, степень абстрактности слова *сфера* будет разной в зависимости от того, в сочетании с каким словом оно использовано: степень абстрактности в сочетании *земная сфера* ниже, чем в сочетании *сфера здравоохранения*. Как в современной отечественной, так и зарубежной практике степень абстрактности / конкретности языкового знака определяется

двумя способами: 1) при помощи метода семантического дифференциала, когда участников исследования, носителей языка, просят оценить степень абстрактности слова в определенном диапазоне (например, от 1 до 5); 2) при помощи анализа контекста слов в корпусе: наиболее высокая вероятность использования абстрактных слов – в контексте с другими абстрактными [Computing Concreteness Ratings of Russian and English Most Frequent Words: Contrastive Approach, с. 405]. Например, словарь абстрактности частотных слов русского языка фиксирует наиболее низкую степень абстрактности для слова *стол* (1.2) и наиболее высокую – для слов *дух* и *душа* (4.3) [Результаты опроса «Конкретность-абстрактность»]. Контексты данных слов в изучаемом корпусе отличаются степенью абстрактности: для слова *стол* наиболее частотными являются следующие коллокации: *накрывать на стол, за столом (пили фруктовую воду и сок), (во время обеда или ужина) за стол садилось*, а для слова *душа* – в «коллективной душе» *толпы, на душу населения, душевное волнение*.

Лексическое многообразие (ЛМ) признается многими учеными как параметр сложности текста, однако считается, что тексты большего объема демонстрируют более низкие показатели лексического многообразия [Борунов с. 84]. В качестве причины ученые указывают на то, что при введении «уникальных» слов автору текста приходится вновь и вновь повторять функциональные слова, снижающие уровень лексического многообразия. В качестве способа решения данной проблемы рекомендуется сравнивать тексты одинаковой длины, например, отрывки по 1000 слов [Захарова, Савина, с. 25], [Савельев с. 65].

В представленном исследовании метрики частотности (Ч), абстрактности (А) и лексического многообразия (ЛМ) были рассчитаны для каждого из двух блоков учебных текстов (по обществознанию и биологии) и сопоставлены с уровнем читабельности (ФК) (см. Табл.2).

Таблица 2

Лексические параметры учебных текстов на фоне роста индекса читабельности

класс	Обществознание					Биология				
	ФК	Ч	А	ЛМ	ЛМ (1000)	ФК	Ч	А	ЛМ	ЛМ (1000)
5	5,40	260,9	2,89	0,73	0, 61	6,86	76,14	1,17	0,55	0, 58
6	6,51	257,8	3,56	0,72	0, 62					
7	7,48	218,5	3,66	0,69	0, 62					
8	8,63	232,8	3,84	0,61	0, 63					
9	9,44	209,5	3,74	0,52	0, 62					
10	10,28	217,4	4,17	0,55	0, 61					
11	10,73	206,4	3,84	0,55	0, 62	9,67	74,81	1,45	0,49	0,59

Как видим, отрицательная динамика читабельности наблюдается в текстах обеих линеек учебников: как в текстах по обществознанию, так и в текстах по биологии (см. Табл. 2). Незначительные различия (менее 1.0) по классам и несоответствия читабельности учебника и номера класса (менее 1.0) не влияют на общую динамику: в целом параметр читабельности текстов от 5 до 11 класса растет вместе с номером класса и соответствует лингвистическому и когнитивному возрасту класса. Исключение составляет учебник по биологии для 10-11 классов, ФК которого ниже плановой на 1, 33.

Частотность слов в линейках учебников обеих дисциплин имеет тенденцию к снижению от 260,9 (обществознание) и 76,14 (биология) в 5 классах до 206,4 (обществознание) и 74,81 (биология) в учебных текстах 11 классов. Однако данный параметр следует признать некоррелирующим с читабельностью изучаемых текстов: по мере усложнения текстов флуктуации индекса частотности не являются стабильными. Обращает на себя внимание и различия частотности словарного состава учебников обществознания и биологии: очевидно, что авторы учебников по биологии используют базовую лексику, в то время как тексты учебников по обществознанию содержат большее количество редких слов.

Степень абстрактности изучаемых текстов растет в обеих линейках учебников, и по биологии, и по обществознанию, однако их индексы в значительной степени отличаются: в диапазоне от 2,89 до 3,84 в текстах по обществознанию и в значительно меньшем диапазоне – от 1,17 до 1,45 – в текстах по биологии. Данный факт свидетельствует о том, что лексика текстов по обществознанию имеет более высокий уровень абстрактности по сравнению с лексикой текстов по биологии, что при прочих равных параметрах может представлять дополнительную сложность для читателей.

Как и прогнозировалось, уровень лексического многообразия в данном корпусе по мере увеличения длины текста демонстрирует тенденцию к снижению: от 0,73 до 0,55 в учебниках по обществознанию и несколько ниже – от 0,55 до 0,42 в учебниках по биологии. При расчете на 1000 слов данный показатель остается относительно стабильным и находится в пределах 0,61 – 0,63 для учебников по обществознанию и в пределах 0,58 – 0,61 для учебников по биологии. В целом лексическое многообразие текстов учебников по биологии ниже, чем соответствующий показатель текстов по обществознанию.

Таким образом, сравнительное исследование лексических параметров (частотность, абстрактность и лексическое многообразие) учебных текстов гуманитарного и общенаучного циклов продемонстрировало их взаимозависимость с читабельностью текста и ряд внутриязыковых различий: при почти равной читабельности тексты по обществознанию характеризуются более высокой степенью абстрактности, частотность используемой в них лексики ниже, что может вызывать дополнительные трудности при чтении. Перспектива исследования определяется необходимостью выявления корреляции синтаксических и семантических параметров текста со сложностью и видится в выявлении спектра данных параметров, а также диапазона их метрик для учебных текстов.

Список литературы

- Абдуллаева С. Х. Лексическая стилистика учебного текста по светотехнике // Молодой ученый. 2015. № 7 (87). С. 719–721. URL: <https://moluch.ru/archive/87/17064/> (дата обращения: 04.10.2020).
- Антонова С. Г. и др. Редакторская подготовка изданий: Учебник / С. Г. Антонова, В. И. Васильев, И. А. Жарков, О. В. Коланькова, Б. В. Ленский, Н. З. Рябина, В. И. Соловьев; Под общ. ред. Антоновой С. Г., д.ф.н. М.: Издательство МГУП, 2002. 468 с.
- Биология. 5 кл. : учеб. для организаций, осуществляющих образовательную деятельность // С.Н. Ловягин, А.А. Вахрушев, А.С. Раутиан. М.: Баласс, 2015. 176 с.
- Биология: учеб. пособие для 7-го кл. общеобразоват. учреждений с рус. яз. обучения // В. Н. Тихомиров [и др.]; под ред. В. Н. Тихомирова. Минск: Нар. асвета, 2010. 199 с.
- Биология: Человек и его здоровье. Учеб. для 9 кл. общеобразоват. учреждений // А. М. Цузмер, О. Л. Петришина; Под ред. В. Н. Загорской и др. 26. изд. М.: Просвещение, 2001. 240 с.
- Общая биология: Учебник для 10-11 кл. общеобразовательных учреждений // С. Ю. Вертьянов, под ред. Ю. П. Алтухова. Москва: Свято-Троицкая Сергиева Лавра, 2012. 352 с.
- Борунов А. Б. Разнообразие речи и методы его измерения в тексте (лингвостатистический подход) // Litera. 2017. № 4. С. 81–86.
- Глушань В. М., Бузняков А. Ю., Бузнякова Е. С., Жданов С. А., Михайлов Н. В. Экспресс-анализ методов оценки сложности текстов учебно-методических материалов // О Некоторых Вопросах и Проблемах Психологии и Педагогики: сборник научных трудов по итогам международной научно-практической конференции, 2017. Издательство: Инновационный центр развития образования и науки. URL: <https://izron.ru/articles/o-nekotorykh-voprosakh-i-problemakh-psikhologii-i-pedagogiki-sbornik-nauchnykh-trudov-po-itogam-mezhsektiya-9-innovatsionnye->

protsessy-i-informatsionnye-tehnologii-v-obrazovanii/ekspress-analiz-metodov-otsenki-slozhnosti-tekstov-uchebno-metodicheskikh-materialov/ (дата обращения: 04.10.2020).

ГОСТ 7.60–2003. Издания. Основные виды. Термины и определения. URL: www.ifap.ru/library/gost/7602003.pdf (дата обращения: 04.09.2020).

Грайс Г. П. Логика и речевое общение // Новое в зарубежной лингвистике. М.: Прогресс, 1985. Вып. 16. Лингвистическая прагматика. С. 217–237.

Гурджян Н. С. Речевые и языковые манифестации когнитивного конфликта в английском и русском языках (контрастивный аспект) // Научная мысль Кавказа. 2008. №1 (53). С. 95–99. URL: <https://cyberleninka.ru/article/n/rechevye-i-yazykovye-manifestatsii-kognitivnogo-konflikta-v-angliyskom-i-russkom-yazykah-kontrastivnyy-aspekt> (дата обращения: 04.10.2020).

Захарова Г. Ф. Лексико-стилистические особенности учебных текстов для обучения чтению научно-технической литературы на английском языке // Гуманитарный вестник. 2015. №1 (27). С. 3–17. URL: <https://cyberleninka.ru/article/n/leksiko-stilisticheskie-osobennosti-uchebnyh-tekstov-dlya-obucheniya-chteniyu-nauchno-tehnicheskoy-literatury-na-angliyskom-yazyke> (дата обращения: 04.10.2020).

Захарова Е. Ю., Савина О. Ю. Лексическое разнообразие текста и способы его измерения // Вестник Тюменского государственного университета. Гуманитарные исследования. Humanitates. 2020. Том 6. № 1 (21). С. 20–34.

Йокояма О. Б. Когнитивная модель дискурса и русский порядок слов. М.: Языки славянской культуры, 2005. 424 с.

Лапошина А. Н. Анализ релевантных признаков для автоматического определения сложности русского текста как иностранного. 2017. URL: <http://www.dialog-21.ru/media/3993/laposhina.pdf> (дата обращения: 04.10.2020).

Леонтьев А. А. Основы психолингвистики. М.: Смысл, 1997. 287 с.

Ляшевская О. Н. К определению сложности русских текстов // XVII Апрельская международная научная конференция по проблемам развития экономики и общества: в 4 кн. Под редакцией: Е. Г. Ясин Кн. 4. М.: Издательский дом НИУ ВШЭ, 2017. С. 408–419.

Наумов И. С., Выхованец В. С. Оценка трудности и сложности учебных задач на основе синтаксического анализа текстов // УБС. 2014. № 48. С. 97–131. URL: <https://cyberleninka.ru/article/n/otsenka-trudnosti-i-slozhnosti-uchebnyh-zadach-na-osnove-sintaksicheskogo-analiza-tekstov> (дата обращения: 04.10.2020).

Оборнева И. В. Автоматизированная оценка сложности учебных текстов на основе статистических параметров: автореф. дис. ... канд. пед. наук: Москва, 2006. 20 с.

Петрова И. Ю., Окладникова С. В. Методика расчета базовых показателей читабельности тестовых материалов на основе экспертных оценок // Прикасп.

журн.: упр. и высокие технологии. 2009. № 3 (7). С. 85–95.

Пешкова Н. П., Яхиббаева Л. М. Первичность и вторичность как онтологические категории учебного текста // Вестн. ЮжноУрал. гос. ун-та. 2009. Лингвистика. № 2 (135). С. 56–60.

Редькина О. Ю. Производность лингводидактического дискурса и ее следствия // Вестник ЧелГУ. 2019. №10 (432). С. 124–129. URL: <https://cyberleninka.ru/article/n/proizvodnost-lingvodidakticheskogo-diskursa-i-ee-sledstviya> (дата обращения: 05.09.2020).

Результаты опроса «Конкретность-абстрактность» // НИЛ «Интеллектуальные технологии управления текстами» Казанского федерального университета. URL: https://kpfu.ru/portal/docs/F_178053613/MNSK.zip (дата обращения: 04.10.2020).

Савельев Д. А. Исследование сложности предложений, составляющих тексты правовых актов органов власти Российской Федерации // Право. Журнал Высшей школы экономики. 2020. № 1. С. 50–74.

Самсонов Н. Б., Чмыхова Е. В., Давыдов Д. Г. Разработка и апробация лингвистической методики оценки когнитивной сложности научно-учебного текста // Психологические исследования. 2015. Т. 8, № 41. С. 6. URL: <http://psystudy.ru> (дата обращения: 04.10.2020).

Солнышкина М. И., Казачкова М. Б., Исмаева Ф. Х. Абстрактность / конкретность как лингвистическая проблема // Ученые записки Национального общества прикладной лингвистики. М.: НОПриЛ. 2019. № 2. С. 74–83.

Солнышкина М. И., Казачкова М. Б., Харьковская Е. В. Инструменты измерения сложности текстов на английском языке // Иностр. языки в школе. 2020. № 3. С. 15–21. URL: <https://www.elibrary.ru/item.asp?id=42609743> (дата обращения: 04.10.2020).

Солнышкина М. И., Калинкина Т. Е., Зиганшина Ч. Р. Конвенции профессиональной коммуникации // Вестник ТГГПУ. 2015. №3 (41). С. 138–145. URL: <https://cyberleninka.ru/article/n/konventsii-professionalnoy-kommunikatsii> (дата обращения: 04.10.2020).

Солнышкина М. И., Мартынова Е. В., Андреева М. И. Пропозициональное моделирование для оценки информативности текста // Ученые записки Национального общества прикладной лингвистики, № 3 (31), 2020. С. 47–57.

Учебники по Обществознанию // НИЛ «Интеллектуальные технологии управления текстами» Казанского федерального университета. URL: <https://kpfu.ru/philology-culture/struktura-instituta/otdelenie-russkoj-i-zarubezhnoj-filologii-imeni/kafedra-inostrannyh-yazikov/nil-39intellektualnyetehtologii-upravleniya/issledovaniya> (дата обращения: 04.10.2020).

Уша Т. Ю. Язык школьного учебника: проблема понимания учащимся-инофоном учебного текста, терминологической лексики, формулировок заданий // Теория и практика общественного развития, 2015. № 15. С. 156–159. URL: <https://cyberleninka.ru/article/n/yazyk-shkolnogo-uchebnika-problema-ponimaniya>

uchaschimsya-inofonom-uchebnogo-teksta-terminologicheskoy-leksiki-formulirovok-zadaniy (дата обращения: 04.10.2020).

Biber, D. University language. A corpus-based study of spoken and written registers. // *Studies in Corpus Linguistics*, ISSN 1388-0373 ; V. 23, 2006. 262 p.

Michael Flor, Beata Beigman Klebanov, Kathleen M. Sheehan Lexical Tightness and Text Complexity // *Proceedings of the 2th Workshop of Natural Language Processing for Improving Textual Accessibility (NLP4ITA)*. 2013. pp. 29–38.

Rudell, A. P. Frequency of word usage and perceived word difficulty: Ratings of Kucera and Francis words // *Behaviour Research Methods, Instruments, & Computers*. № 25 (4), 1993. pp. 455–463.

Solnyshkina, M. I., Harkova, E. V., Kazachkova, M. B. The Structure of Cross-Linguistic Differences: Meaning and Context of 'Readability' and its Russian Equivalent 'Chitabelnost' // *Journal of Language & Education*, 6 (1), 2020. pp. 103–119. URL: <https://jle.hse.ru/article/view/7176/12052> (дата обращения: 04.10.2020).

Solnyshkina, M., Ivanov V., Solovyev V. Characterizing Text Complexity with Core Vocabulary Distributional Patterns: Corpus-based Approach // *AIST*, 2018. URL: <http://ceur-ws.org/Vol-2268/paper12.pdf>. (дата обращения: 04.10.2020).

Solovyev, V., Ivanov, V., Solnyshkina, M. Assessment of reading difficulty levels in Russian academic texts: Approaches and metrics // *Journal of intelligent & fuzzy systems*. V. 34. № 5, 2018. pp. 3049–3058.

Solovyev, V., Andreeva, M., Solnyshkina, M., Zamaletdinov, R., Danilov, A. and Gaynutdinova, D., Computing Concreteness Ratings of Russian and English Most Frequent Words: Contrastive Approach // *12th International Conference on Developments in eSystems Engineering (DeSE)*, Kazan, Russia, 2019, pp. 403–408.

References

Abdullaeva, S. Kh. (2015). *Leksicheskaya stilistika uchebnogo teksta po svetotekhnike* [Lexical Style of the Educational Text on Lighting Engineering]. Molodoi uchenyi. No. 7 (87), pp. 719–721. URL: <https://moluch.ru/archive/87/17064/> (accessed: 04.10.2020). (In Russian)

Antonova, S. G. i dr. (2002). *Redaktorskaya podgotovka izdaniy: Uchebnik* [Editorial Preparation of Publications: A Textbook]. S. G. Antonova, V. I. Vasil'ev, I. A. Zharkov, O. V. Kolan'kova, B. V. Lenskii, N. Z. Riabinina, V. I. Solov'ev; Pod obshch. red. Antonovoi S. G., d.f.n. 468 p. Moscow, izdatel'stvo MGUP. (In Russian)

Biber, D. (2006). *University Language. A Corpus-Based Study of Spoken and Written Registers*. *Studies in Corpus Linguistics*, ISSN 1388-0373; V. 23, 262 p. (In English)

Biologiya. (2015). 5 kl. ucheb. dlia organizatsii, osushchestvliaiushchih obrazovatel'nuii deiatel'nost' [Biology. 5th grade. A Textbook for Educational Organizations]. S. N. Loviagin, A. A. Vahrushev, A. S. Rautian. 176 p. Moscow, Balass. (In Russian)

Biologiya. (2010). Ucheb.posobie dlia 7-go kl. obshcheobrazovat. uchrezhdenii s B63 rus. iaz. obucheniia [Biology. A Textbook for the 7th Grade of Public Schools with the Russian Language as a Teaching Language]. V. N. Tihomirov [et al.] ; edited by V. N. Tihomirov. 199 p. Minsk, Nar.Asveta. (In Russian)

Biologiya: Chelovek i ego zdorov'e. (2001). Ucheb.dlja 9 kl. obshheobrazovat. Uchrezhdenij [Human and its Health]. A. M. Cuzmer, O. L. Petrishina; Pod red. V. N. Zagorskoj i dr. - 26.izd. 240 p. Moscow, Prosveshchenie. (In Russian)

Obshchaia biologiya. (2012). Uchebnik dlia 10–11 kl. obshcheobrazovatel'nyh uchrezhdenii [General Biology: A Textbook for 10–11th Grades in Public Schools]. S. Yu. Vert'ianov, pod red. Yu. P. Altuhova. 352 p. Moscow, Sviato-Troickaia Sergieva Lavra. (In Russian)

Borunov, A. B. (2017). *Raznoobrazie rechi i metody ego izmereniia v tekste (lingvostatisticheskii podkhod)* [Speech Diversity and Methods of Its Measurement in the Text (Linguostatistical Approach)]. *Litera*. No. 4, pp. 81–86. (In Russian)

Giurdzhian, N. S. (2008). *Rechevye i iazykovye manifestatsii kognitivnogo konflikta v angliiskom i russkom iazykakh (kontrastivnyi aspekt)* [Speech and Language Manifestations of Cognitive Conflict in English and Russian (A Contrastive Aspect)]. *Nauchnaia mysl' Kavkaza*. No. 1 (53), pp. 95–99. URL: <https://cyberleninka.ru/article/n/rechevye-i-iazikovye-manifestatsii-kognitivnogo-konflikta-v-angliiskom-i-russkom-iazыkah-kontrastivnyi-aspekt> (accessed: 04.10.2020). (In Russian)

Glushan, V. M., Buzniakov, A. Iu., Buzniakova, E. S., Zhdanov, S. A., Mikhailov, N. V. (2017). *Ekspress-analiz metodov otsenki slozhnosti tekstov uchebno-metodicheskikh materialov* [Express Analysis of Methods for Assessing the Complexity of Texts of Educational Material]. O Nekotorykh Voprosakh i Problemakh Psikhologii i Pedagogiki: sbornik nauchnykh trudov po itogam mezhdunarodnoi nauchno-prakticheskoi konferentsii, Izdatel'stvo: Innovatsionnyi tsentr razvitiia obrazovaniia i nauki. URL: <https://izron.ru/articles/onekotorykh-voprosakh-i-problemakh-psikhologii-i-pedagogiki-sbornik-nauchnykh-trudov-po-itogam-mezh/sektsiya-9-innovatsionnye-protsessy-i-informatsionnye-tekhnologii-v-obrazovanii/ekspress-analiz-metodov-otsenki-slozhnosti-tekstov-uchebno-metodicheskikh-materialov/> (accessed: 04.10.2020). (In Russian)

GOST 7.60 (2003) Izdaniia. Osnovnye vidy. Terminy i opredeleniia [Editions. Basic Types. Terms and Definitions]. 41 p. URL: www.ifap.ru/library/gost/7602003.pdf (accessed: 04.09.2020). (In Russian)

Grais, G. P. (1985). *Logika i rechevye obshchenie* [Logic and Speech Communication]. *Novoe v zarubezhnoi lingvistike*. Vyp. 16. Lingvisticheskaya pragmatika. Pp. 217–237. Moscow, Progress. (In Russian)

Iokoiama, O. B. (2005). *Kognitivnaia model' diskursa i russkii poriadok slov* [Cognitive Model of Discourse and Russian Word Order]. 424 p. Moscow, Iazyki slavianskoi kul'tury. (In Russian)

Laposhina, A. N. (2017). *Analiz relevantnykh priznakov dlia avtomaticheskogo opredeleniia slozhnosti*

rusского teksta kak inostrannogo [Analysis of Relevant Features for Automatic Determination of the Complexity of a Russian Text as a Foreign One]. URL: <http://www.dialog-21.ru/media/3993/laposhina.pdf> (accessed: 04.10.2020). (In Russian)

Leont'ev, A. A. (1997). *Osnovy psikholingvistiki* [Basics of Psycholinguistics]. 287 p. Moscow. Smysl. (In Russian)

Liashevskaya, O. N. (2017). *K opredeleniiu slozhnosti russkikh tekstov* [On Determining the Complexity of Russian Texts]. XVII Aprel'skaya mezhdunarodnaya nauchnaya konferentsiya po problemam razvitiya ekonomiki i obshchestva: v 4 kn. Pod redaktsiei: E. G. Iasin Kn. 4, pp. 408–419. Moscow, Izdatel'skii dom NIU VShE. (In Russian)

Michael Flor, Beata Beigman Klebanov, Kathleen M. Sheehan. (2013). *Lexical Tightness and Text Complexity*. Proceedings of the 2th Workshop of Natural Language Processing for Improving Textual Accessibility (NLP4ITA). Pp. 29–38. (In English)

Naumov, I. S., Vykhovanets, V. S. (2014). *Otsenka trudnosti i slozhnosti uchebnykh zadach na osnove sintaksicheskogo analiza tekstov* [Evaluating the Difficulty and Complexity of Learning Tasks Based on Text Parsing]. UBS. No. 48, pp. 97–131. URL: <https://cyberleninka.ru/article/n/otsenka-trudnosti-i-slozhnosti-uchebnykh-zadach-na-osnove-sintaksicheskogo-analiza-tekstov> (accessed: 04.10.2020). (In Russian)

Oborneva, I. V. (2006). *Avtomatizirovannaya otsenka slozhnosti uchebnykh tekstov na osnove statisticheskikh parametrov: avtoref. dis. ... kand.ped. nauk* [Automated Assessment of the Complexity of Educational Texts Based on Statistical Parameters: Ph.D. Thesis Abstract]. Moscow, 20 p. (In Russian)

Peshkova, N. P., Iakhimbaeva, L. M. (2009). *Pervichnost' i vtorichnost' kak ontologicheskie kategorii uchebnogo teksta* [Primary and Secondary as Ontological Categories of the Educational Text]. Vestn. IuzhnoUral.gos. un-ta. Lingvistika. No. 2 (135), pp. 56–60. (In Russian)

Petrova, I. Iu., Okladnikova, C. V. (2009). *Metodika rascheta bazovykh pokazatelei chitabel'nosti testovykh materialov na osnove ekspertnykh otsenok* [Methodology for Calculating the Basic Readability Indicators of Test Materials Based on Expert Assessments]. Priasp. zhurn.: upr. i vysokie tekhnologii. No. 3 (7), pp. 85–95. (In Russian)

Red'kina, O. Iu. (2019). *Proizvodnost' lingvodidakticheskogo diskursa i ee sledstviia* [Derivation of Linguodidactic Discourse and Its Consequences]. Vestnik ChelGU. No. 10 (432), pp. 124–129. URL: <https://cyberleninka.ru/article/n/proizvodnost-lingvodidakticheskogo-diskursa-i-ee-sledstviya> (accessed: 05.09.2020). (In Russian)

Rezultaty oprosa "Konkretnost'-abstraktnost'" [Results of the "Concreteness-Abstractness" Survey]. NIL "Intellektual'nye tekhnologii upravleniia tekstami" Kazanskogo federal'nogo universiteta. URL: https://kpfu.ru/portal/docs/F_178053613/MNSK.zip (accessed: 04.10.2020). (In Russian)

Rudell, A. P. (1993). *Frequency of Word Usage and Perceived Word Difficulty: Ratings of Kucera and Fran-*

cis Words. Behaviour Research Methods, Instruments, & Computers. No. 25 (4), pp. 455–463. (In English)

Samsonov, N. B., Chmykhova, E. V., Davydov, D. G. (2015). *Razrabotka i aprobatsiia lingvisticheskoi metodiki otsenki kognitivnoi slozhnosti nauchno-uchebnogo teksta* [Development and Testing of a Linguistic Methodology for Assessing the Cognitive Complexity of a Scientific and Educational Text]. Psikhologicheskie issledovaniia. T. 8, No. 41, pp. 6. URL: <http://psystudy.ru> (accessed: 04.10.2020). (In Russian)

Savel'ev, D. A. (2020). *Issledovanie slozhnosti predlozhenii, sostavliaiushchikh teksty pravovykh aktov organov vlasti Rossiiskoi Federatsii* [Study of the Complexity of Proposals that Make Up the Texts of Legal Acts of the Authorities of the Russian Federation]. Pravo. Zhurnal Vysshei shkoly ekonomiki. No. 1, pp. 50–74. (In Russian)

Solnyshkina, M. I., Harkova, E. V., Kazachkova, M. B. (2020). *The Structure of Cross-Linguistic Differences: Meaning and Context of 'Readability' and its Russian Equivalent 'Chitabelnost'*. Journal of Language & Education, 6 (1), pp. 103–119. URL: <https://jle.hse.ru/article/view/7176/12052> (accessed: 04.10.2020). (In English)

Solnyshkina, M. I., Kalinkina, T. E., Ziganshina, Ch. R. (2015). *Konventsii professional'noi kommunikatsii* [Conventions of Professional Communication]. Vestnik TGGPU. No. 3 (41), pp. 138–145. URL: <https://cyberleninka.ru/article/n/konventsii-professionalnoy-kommunikatsii> (accessed: 04.10.2020). (In Russian)

Solnyshkina, M. I., Kazachkova, M. B., Ismaeva, F. Kh. (2019). *Abstraktnost'/konkretnost' kak lingvisticheskaya problema* [Abstractness/Concreteness as a Linguistic Problem]. Uchenye zapiski Natsional'nogo obshchestva prikladnoi lingvistiki. No. 2, pp. 74–83. Moscow, NOPriL. (In Russian)

Solnyshkina, M. I., Kazachkova, M. B., Khar'kova, E. V. (2020). *Instrumenty izmereniia slozhnosti tekstov na angliiskom iazyke* [Tools for Measuring the Complexity of English Texts]. Inostr. iazyki v shkole. No. 3, pp. 15–21. URL: <https://www.elibrary.ru/item.asp?id=42609743> (accessed: 04.10.2020). (In Russian)

Solnyshkina, M. I., Martynova, E. V., Andreeva, M. I. (2020). *Propozitsional'noe modelirovanie dlia otsenki informativnosti teksta* [Propositional Modeling to Assess the Information Content of the Text]. Uchenye zapiski Natsional'nogo obshchestva prikladnoi lingvistiki. No. 3 (31), pp. 47–57. (In Russian)

Solnyshkina, M., Ivanov V., Soloviev V. (2018). *Characterizing Text Complexity with Core Vocabulary Distributional Patterns: Corpus-based Approach*. AIST. URL: <http://ceur-ws.org/Vol-2268/paper12.pdf> (accessed: 04.10.2020). (In English)

Soloviev, V., Ivanov, V., Solnyshkina, M. (2018). *Assessment of reading difficulty levels in Russian academic texts: Approaches and metrics*. Journal of intelligent & fuzzy systems. V. 34. No. 5, pp. 3049–3058. (In English)

Soloviev, V., Andreeva, M., Solnyshkina, M., Zamaletdinov, R., Danilov A., and Gainutdinova, D. (2019). *Computing Concreteness Ratings of Russian and English Most Frequent Words: Contrastive Approach*. 12th International Conference on Developments in

eSystems Engineering (DeSE), Kazan, Russia, pp. 403–408. (In English)

Uchebniki po obshchestvoznaniyu [Social Sciences textbooks]. NIL “Intellektual’nye tekhnologii upravleniya tekstami” Kazanskogo federal’nogo universiteta. URL: <https://kpfu.ru/philology-culture/struktura-instituta/otdelenie-russkoj-i-zarubezhnoj-filologii-imeni/kafedra-inostrannykh-yazykov/nil-39intellektualnye-tehnologii-upravleniya/issledovaniya> (accessed: 04.10.2020). (In Russian)

Usha, T. Yu. (2015). *Iazyk shkol’nogo uchebnika: problema ponimaniia uchashchimsia-inofonom uchebnogo teksta, terminologicheskoi leksiki, formulirovok zadaniy* [The Language of the School Textbook: The Problem of Understanding the Educational Text, Terminological Vocabulary, and Task Formulations for Students Who Are Foreign Speakers]. *Teoriia i*

praktika obshchestvennogo razvitiia. No. 15, pp. 156–159. (In Russian)

Zakharova, E. Yu., Savina, O. Yu. (2020). *Leksicheskoe raznoobrazie teksta i sposoby ego izmereniia* [Lexical Diversity of the Text and Ways to Measure It]. *Vestnik Tiumenskogo gosudarstvennogo universiteta. Gumanitarnye issledovaniia. Humanitates*. Tom 6. No. 1 (21), pp. 20–34. (In Russian)

Zakharova, G. F. (2015). *Leksiko-stilisticheskie osobennosti uchebnykh tekstov dlia obucheniia chteniiu nauchno-tehnicheskoi literatury na angliiskom iazyke* [Lexical and Stylistic Features of Educational Texts for Teaching Reading of Scientific and Technical Literature in English]. *Gumanitarnyi vestnik*. No. 1 (27), pp. 3–17. URL: <https://cyberleninka.ru/article/n/leksiko-stilisticheskie-osobennosti-uchebnykh-tekstov-dlya-obucheniya-chteniyu-nauchno-tehnicheskoy-literatury-na-angliiskom-yazyke> (accessed: 04.10.2020). (In Russian)

The article was submitted on 12.10.2020

Поступила в редакцию 12.10.2020

Мартынова Екатерина Владимировна,
младший научный сотрудник
НИЛ «Интеллектуальные технологии
управления текстами»,
старший преподаватель,
Казанский федеральный университет,
420008, Россия, Казань,
Кремлевская, 18.
Katerinamarty@yandex.ru

Солнышкина Марина Ивановна,
доктор филологических наук,
профессор,
Казанский федеральный университет,
420008, Россия, Казань,
Кремлевская, 18.
mesoln@yandex.ru

Мерзлякова Амина Фариковна,
аспирант,
ассистент,
Казанский федеральный университет,
420008, Россия, Казань,
Кремлевская, 18.
amirus1985@mail.ru

Гизатулина Диана Юрьевна,
аспирант,
педагог IT-лицея,
Казанский федеральный университет,
420008, Россия, Казань,
Кремлевская, 18.
giz-diana@mail.ru

Martynova Ekaterina Vladimirovna,
Associate Scientist of “Intelligent Text
Management Technologies” laboratory,

Assistant Professor
Kazan Federal University,
18 Kremlyovskaya Str.,
Kazan, 420008, Russian Federation.
Katerinamarty@yandex.ru

Solnyshkina Marina Ivanovna,
Doctor of Philology,
Professor,
Kazan Federal University,
18 Kremlyovskaya Str.,
Kazan, 420008, Russian Federation.
mesoln@yandex.ru

Merzlyakova Amina Farikovna,
graduate student,
Assistant Professor,
Kazan Federal University,
18 Kremlyovskaya Str.,
Kazan, 420008, Russian Federation.
amirus1985@mail.ru

Gizatulina Diana Yuryevna,
graduate student,
Teacher of IT lyceum,
Kazan Federal University,
18 Kremlyovskaya Str.,
Kazan, 420008, Russian Federation.
giz-diana@mail.ru