

УДК 811.111
ББК 81.2 АНГЛ.

ЛЕКСИЧЕСКИЕ БЛОКИ В ОТРАСЛЕВОМ ИНТЕРНЕТ-ДИСКУРСЕ (на материале блогов в сфере геоинформационных технологий)

Т.В. Толстова

Аннотация. В статье анализируются основные типы лексических блоков — автоматически выделенных повторяющихся сочетаний из нескольких слов — на материале профессиональных интернет-блогов в области геоинформационных технологий. Несмотря на то, что данный жанр реализуется в письменной форме, он служит для межличностной и групповой коммуникации, участие в которой требует от коммуникантов владением медиаграмотностью, выражающейся в осознании коммуникантами принадлежности к профессиональному сообществу, в наличии вопросов и проблем, вызывающих общий интерес, а также в существовании общего языка в виде специфического набора языковых средств, употребляемых и понимаемых всеми членами данного сообщества. К таким средствам относятся не только отдельные лексические единицы, связанные с профессиональной терминологией, но и «готовые» блоки слов, в число которых, помимо устойчивых фразеологических сочетаний, зафиксированных в словарях, входят и более протяженные последовательности слов. Материалом исследования послужил электронный корпус, состоящий из более ста блогов, опубликованных за период с 2013 по 2017 годы. Инструменты корпусной лингвистики, в частности, N-граммы, позволили выявить блоки различной длины — от семи до двух единиц. С помощью когнитивного анализа были определены приоритетные виды сочетаний — лексические (преимущественно профессиональные и терминологические), лексико-грамматические и грамматические (главным образом модальные), а также их коммуникативные функции — информативная, когезивная (реализуемая в виде дискурсивных маркеров) и персуазивная.

Ключевые слова: лексические блоки, чанки, корпусная лингвистика, медиаграмотность, профессиональные блоги на английском языке.

399

LEXICAL CHUNKS IN INDUSTRY-BASED INTERNET DISCOURSE
(A Case of Blogs in Geographic Information Systems)**T.V. Tolstova**

Abstract. *The article aims at analyzing various types of lexical chunks — automatically generated repetitive strings of words — from professional blogs in the field of Geographic Information Systems. Although this genre occurs in written form, it is used for interpersonal and group communication maintaining which requires from its participants to have some degree of media literacy. Media literacy consists in communicants awareness of their belonging to a certain professional community, in questions and issues of common interest, as well as in a specific set of linguistic features utilized and understood by all members of this community. These features include not only single lexical units related to professional terminology but also pre-fabricated chunks that involve, along with set phraseological units provided in dictionaries, longer strings of words. The material of the study is an electronic corpus consisting of over a hundred texts of blogs published from 2013 to 2017. The tools of Corpus Linguistics, such as N-grams, have retrieved sequences from seven to two words. The cognitive analysis has identified several major types of chunks — lexical (specifically, professional and terminological) bundles, lexico-grammatical, and grammatical (mainly modal) bundles, as well as their communicative functions — informing, cohesion (via discourse markers) and persuasion.*

Keywords: *lexical strings, chunks, corpus linguistics, media literacy, professional blogs in English.*

400

Важнейшим видом человеческого общения¹ является профессиональное общение, в котором участники обмениваются информацией, обогащают, уточняют или расширяют свои знания в той или иной области. Разновидностью профессионального общения выступают интернет-блоги, где авторы делятся информацией о новинках той или иной отрасли, о проблемах, с которыми они сталкиваются, об эффективных спо-

собах решения этих проблем. У читателей блогов есть не только возможность быть в курсе основных событий, но и задавать вопросы, делать комментарии, участвовать в обсуждениях интересующих их тем.

Слово «блог» (от англ. *blog*, сокращение от *web log* или *weblog*) означает онлайн-дневник, где автор (отдельный человек, группа или организация) регулярно делает записи о своих действиях, мыслях, убеждениях [2].

¹ Общение — сложный, многоплановый процесс установления и развития контактов между людьми, порождаемый потребностями совместной деятельности и включающий в себя обмен информацией, выработку единой стратегии взаимодействия, восприятие и понимание другого человека [1].

Одни блоги могут выступать просто в роли новостных фильтров, когда авторы выбирают интересующие их события из сети Интернет, кратко комментируют их и снабжают соответствующими ссылками. Другие — содержат собственный оригинальный материал блогеров. Как правило, блоги включают в себя не только письменный текст, но и иллюстрации, а также ссылки на другие сайты. Дана Бойд [3], рассматривая блоги с точки зрения культуры и практики, отмечает, что они «размыли границу между устной и письменной речью, между телесностью и пространственностью, между публичным и частным», и видит блоггинг как «набор разнообразных приемов, использование которых приводит к созданию разнообразного контента помимо среды, которую мы называем блогами».

Умение читать блоги, писать собственные блоги и общаться через блоги с единомышленниками или коллегами, подразумевает наличие так называемой «медиаграмотности», которая, по мнению Т.Г. Добросклонской [4, с. 262], охватывает «все стороны сложного процесса взаимодействия со средствами массовой информации: от необходимых технических навыков и умения находить нужную информацию до способности выявлять различные идеологические интерпретации и понимания лингво-медийных технологий воздействия».

Применительно к профессиональным блогам медиаграмотность предполагает осознание коммуникантами принадлежности к определенному сообществу, наличие вопросов и проблем, вызывающих общий интерес, а также использование общего языка в

виде специфического набора лингвистических средств, используемых и понимаемых всеми членами данного сообщества. Примером таких средств могут служить лексические блоки, состоящие из нескольких (как правило, от двух до пяти) слов [5]. В лингвистической литературе для обозначения таких блоков используются различные термины: *lexical phrases, pre-fabricated patterns, routine formulae, formulaic sequences, lexicalized stems, chunks* [6].

В отечественном языкознании термин «словосочетание» означает «соединение двух или более знаменательных слов (вместе с относящимися к ним служебными словами или без них), служащее для выражения единого, но расчлененного понятия или представления» [7–9]. Этот термин применяется к самым различным сочетаниям слов и, как правило, предполагает всесторонний анализ сложных и многообразных переходов между тем, что называется «свободное словосочетание», с одной стороны, и «фразеологическая единица» — с другой. Естественная речь — это также диалектическое единство коллокации (лексико-фразеологических связей между словами) и коллигации (формальных или грамматических правил, управляющих введением в речь слов как частей речи), которое обеспечивает говорящего большим набором предельных синтагматических единиц для их функционирования в процессе коммуникации [8; 10]. Изучение и описание сочетаний слов в тексте осуществляются в рамках синтагматики — области языкознания, которая занимается изучением «линейного соположения слов и морфем в

речи, а также исследует различные типы словосочетаний: от производных слов до сложных многоэлементных образований» [11; 12].

С целью выявления отношений между следующими друг за другом лексическими единицами существуют специальные процедуры сегментации, позволяющие отличить и отделить одну единицу от другой на основании свойства ее повторяемости и контраста с соседними единицами [13]. Процедура синтагматического анализа состоит в членении связного текста на синтагмы с последующим определением их состава и способа соединения их компонентов. Данный дедуктивный подход к членению языковых последовательностей состоит в определении границ рассматриваемых единиц на основе знаний и языковой интуиции исследователя. В данной работе используется противоположный — индуктивный — подход, заключающийся в анализе единиц, уже сегментированных компьютерной программой на основе частотности их встречаемости в тексте, с целью выявления приоритетных типов данных единиц — лексических, грамматических, лексикограмматических, а также определения их коммуникативных функций.

В наши дни с развитием корпусной лингвистики², позволяющей производить компьютеризированную обработку больших массивов текстов в электронном виде [15–18], линейные последовательности слов вычлняются автоматически для разного количества элементов. Инструменты корпусной лингвистики

[19–21] позволяют легко извлекать такие блоки из больших массивов текстов. Недостатком этого процесса является то, что компьютерная программа автоматически выделяет и регулярно повторяющиеся фрагментарные сочетания, образовавшиеся в результате того, что несколько слов «случайно» оказались рядом несколько раз. Исключив (вручную, путем анализа каждого блока) такие фрагментарные сочетания, можно получить две группы: синтаксически неполные, но значащие сочетания (*allow you to, as opposed to*) и полнозначные (с семантической и прагматической точки зрения) словосочетания — *to take advantage, broad range of, geographic information systems*.

Преыдущие корпусные исследования лексических блоков проводилось преимущественно на материале разговорной речи и академических текстов [22–24]. Объектом данного исследования являются частотные сочетания слов (лексические блоки) на материале корпуса текстов профессиональных блогов в области геоинформационных технологий.

Материалы и методы. Электронный корпус состоит из 111 текстов блогов профессионального сайта *GIS Café* (<https://www.giscafe.com/>), адресованного специалистам в области геоинформационных технологий и просто лицам, интересующимся данной областью, за 2013–2017 годы. Более подробная статистическая информация о корпусе представлена в табл. 1.

Списки блоков разной длины были автоматически составлены с по-

² Корпусная лингвистика – «научная деятельность, связанная с составлением и анализом корпусов» [14, с. 50]. Корпусы – «большие собрания машиночитаемых текстов, состоящие из тысяч или миллионов слов» [там же, с. 48].

Таблица 1

Статистическая информация о корпусе *GIS Café*

Количество текстов	111
Общее количество слов (tokens)	104.631
Количество индивидуальных слов (types)	8.132
Среднее количество слов в тексте	942
Количество предложений	4.754
Среднее количество предложений в тексте	43
Средняя длина предложения (кол-во слов)	22
Общее количество лексических блоков	14.264
Среднее количество блоков на 1 текст	128
Среднее количество блоков на 1 предложение	3

мощью инструмента *kfNgram* программы KWICFINDER, версия 0.98x (12/06) [25]. *kfNgram* — это приложение, работающее в операционной системе Windows для создания *N*-грамм — групп из *N* последовательных слов, где *N* может быть любым положительным целым числом. Полученные списки, состоящие из 2–7 слов, по аналогии с количеством исполнителей в музыкальных ансамблях обозначены как *септет* (7 слов), *секстет* (6 слов), *квинтет* (5 слов), *квартет* (4 слова), *трио* (3 слова) и *дуэт* (2 слова). Затем из корпуса были исключены фрагментарные (неполные / недостаточные / неоднозначные) блоки, имеющие случайный ха-

рактер и не обладающие целостным семантическим или грамматическим значением. Оставшиеся блоки были проанализированы и разделены на группы в соответствии с типами связи между элементами и с их коммуникативными функциями.

Результаты исследования и их обсуждение. Автоматический подсчет лексических блоков в соответствии с количеством слов (с учетом фрагментарных блоков), показал подавляющее преобладание дуэтов (12.273 единицы), как показано на рис. 1.

Септеты представлены всего двумя фрагментарными блоками, связанными с частым употреблени-

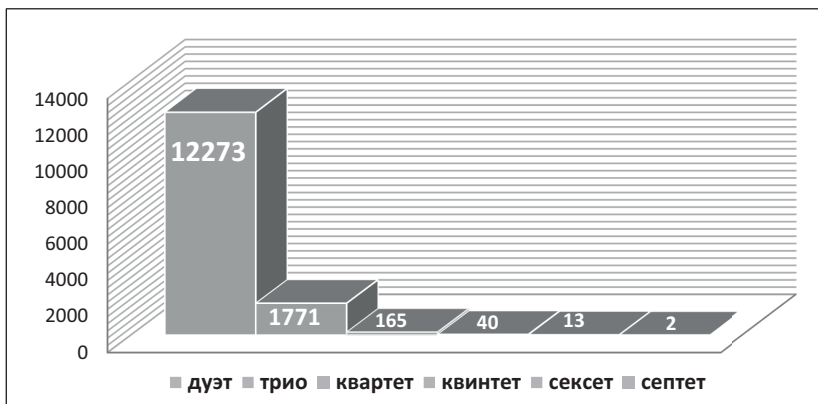


Рис. 1. Общее распределение лексических блоков по количеству слов

Таблица 2.

Лексические блоки-секстеты

	Лексический блок	Количество
1	# reasons you should attend hxgn	6
2	reasons you should attend hxgn live	6
3	should attend hxgn live #### here	6
4	top # reasons you should attend	6
5	you should attend hxgn live ####	6
6	be sure to check out my	6
7	for more information be sure to	6
8	information be sure to check out	6
9	more information be sure to check	6

ем (10 случаев) названия компании и имени ее президента. Более интересными представляются фрагментарные блоки-секстеты:

Из табл. 2 видно, что фрагментарные блоки №№ 1-4 выстраиваются в более длинную последовательность: *top # reasons you should attend hxgn live ####*. Данная фраза согласуется с рекламной функцией блога — убедить читателя посетить какое-либо профессиональное мероприятие. Аналогичным образом соединяются фрагментарные блоки №№ 5-6: *for more information be sure to check out my*. Здесь рекламная персуазивная функция сочетается еще с одной функцией блога — служить ссылкой на другой источник информации. Похожие результаты наблюдаются для **квинтетов** — фрагментарные блоки также выстраиваются в последовательные цепочки из секстетов: *be sure to check out | for more information be sure | information be sure to check | more information be sure to*.

В **квартетах** можно выделить несколько групп блоков в зависимости от смысловых и грамматических связей между словами. В первую очередь были исключены фрагментарные

блоки, не имеющие смысловой целостности сами по себе и не содержащие в себе связанных элементов, напр., *geospatial track for more* (в отличие от, скажем, *be sure to check*, где компонент *be sure to* является словосочетанием, обычно используемым в повелительном наклонении для выражения настоятельной рекомендации, напоминания и т.п.).

Наиболее частотными стали блоки, указывающие на авторов блога: *#### by Susan Smith* (50 употреблений), *##th #### by Susan* (35), *#### by Mladen Stojic* (10), *#### by Ron Bisio* (10) и т.д. Нарушение коллигации (в данном случае использование предлога *by* перед доминирующим/опорным словом, означающим автора) привело бы к нарушению языковой нормы. Вслед за этими блоками идут сочетания слов, связанные с датой опубликования блога — *April ##th* (8) и т.п. (всего 59 употреблений), использующие американский стиль — название месяца + число в форме порядкового числительного). Всего коллигационные сочетания представлены 52 квартетами (344 употребления). В блоке *according to company materials* (12) сочетание *according to* является дискурсивным

маркером, указывающим на источник информации.

Коллокационные сочетания представлены двумя группами — идиоматическими выражениями общего языка (*to take advantage of* (10), *a broad range of* (7), *a wide variety of* (6) и т.д.) и словосочетаниями, связанными с профессиональной тематикой, в том числе терминологическими сочетаниями — *the UK BIM alliance* (7), *wind and storm surge* (7), *the Singapore land authority* (6), *Geographic Information Systems GIS* (5) и т.д.

Грамматические блоки (под которыми здесь подразумеваются аналитические грамматические формы, а также модальные сочетания) характеризуются большим количеством модальных сочетаний (14 единиц), выражающих в основном возможность: *will be able to* (12 употреблений), *to be able to* (11), в том числе с пассивным инфинитивом: *can be used to* (6), *can be used by* (4), *can be used for* (4), *can be used in* (4), которые могут также рассматриваться как коллигационные сочетания. С функциональной точки зрения последние сочетания служат для описания использования того или иного оборудования или программного обеспечения в контексте его назначения (*can be used to*, *can be used for*), пользователя (*can be used by*), области применения (*can be used in*).

Количество 3-элементных блоков (**трио**) значительно превышает предыдущие — 1771 блок (7727 употреблений). В связи с этим были рассмотрены лишь наиболее частотные из них — те, которые встречаются в корпусе минимум 5 раз и составляют 44,6% (3450 из 7727) от общего числа трио и 24,8% от количества индивидуальных трио — 440 из 1771). Здесь

обращает на себя внимание большее (по сравнению с 6-4-элементными блоками) число коллокаций общего (непрофессионального) английского языка, представленных двумя главными группами: а) дискурсивными маркерами цели (*in order to*), результата (*as a result*, *a result of*, *at the end*), добавления информации (*in addition to*, *as well as*) и б) выражениями количества и разнообразия (*a lot of*, *a number of*, *on top of*, *and much more*, *a variety of*, *a broad range*, *broad range of*, *a wide variety*, *all sorts of*). При заимствовании информации из других источников для выражения благодарности характерно использование сочетания “*courtesy of*” (*image courtesy of*, *courtesy of Bentley*).

Среди профессиональных сочетаний, помимо специальных терминов (*business mapping software*, *the geospatial industry*, *geographic information systems*, *field data collection*, *big data module*), выделяются сочетания, построенные по модели «существительное» + *and* + «существительное»: *positions and attributes*, *time and costs*, *oil and gas*, *BIM and GIS*, *wind and storm*, *smartphones and tablets* и другие.

В отношении лексико-грамматических блоков, как и для квартетов, для трио наблюдается преобладающее число модальных сочетаний возможности (*to be able to*, *being able to*, *can be used*, *that can be*, *allows you to*) и необходимости (*need to be*, *needs to be*).

Подавляющее большинство лексических блоков составляют **дуэты** (12.273 отдельных блока и 50.659 единиц в целом), поэтому оправданным представляется ограничение числа исследуемых единиц только теми, которые встречаются в корпусе не менее 25 раз, что составляет 148

Таблица 3

Значения десяти наиболее частотных дуэтов

№	Лексический блок	Количество	Семантика
1	of the	370	(общие) отношения между объектами
2	in the	346	пространственные отношения
3	can be	179	возможность
4	to the	174	пространственные отношения
5	on the	162	пространственный отношения
6	for the	139	целевые отношения
7	to be	133	целевые отношения
8	such as	121	иллюстративно-пояснительные отношения
9	with the	121	отношения сопровождения
10	is a	116	определяющие отношения

отдельных сочетаний и 7841 употреблен в целом.

Помимо ожидаемого большого количество фрагментарных блоков (*and the, at the, of a, has been* и т.п.), сочетаний «артикль» + «существительное/прилагательное» (*the data, the new*), глаголов с инфинитивной частицей (*to do, to create*), удалось обнаружить значительное число терминологических сочетаний (*data collection, reality modelling, geospatial data*) и фразеологических сочетаний в роли дискурсивных маркеров (*according to, in addition, for example*), а также коллигационных сочетаний (*part of, based on, data from, access to*) и грамматических сочетаний (*able to, you can, need to, be able, they can*).

Если посмотреть (см. табл. 3) на десятку «лидеров» среди дуэтов (которые встречаются в корпусе 1861 раз — 3,7% от общего числа дуэтов), то, несмотря на кажущуюся фрагментарность, они демонстрируют достаточно четкие, хотя и более абстрактные значения:

Очевидно, что это лишь самое общее выражение значений данных синтаксически и прагматически неполных лексических блоков, однако

ими также не стоит пренебрегать, поскольку они очерчивают некое семантическое пространство, в котором существуют блогеры как вид профессионального дискурса.

Рассматривая лексические блоки разной длины (от двух до семи), можно наблюдать достаточно полное присутствие разных типов отношений (лексических, лексико-грамматических, грамматических) в более коротких единицах — от двух до четырех. Более того, количество компонентов в лексических блоках влияет на их тип: среди квартетов преобладают лексические сочетания, тогда как в трио и дуэтах значительно чаще встречаются коллигационные (лексико-грамматические) сочетания, при этом, если в трио грамматических блоков значительно меньше, чем лексических, то в дуэтах их поровну (см. рис. 2).

Лексические сочетания представлены главным образом профессионализмами и терминами (60% от общего количества), через которые реализуется важнейшая коммуникативная функция блогов — информирование читателя о наиболее важных событиях и достижениях в мире геоинформационных технологий.

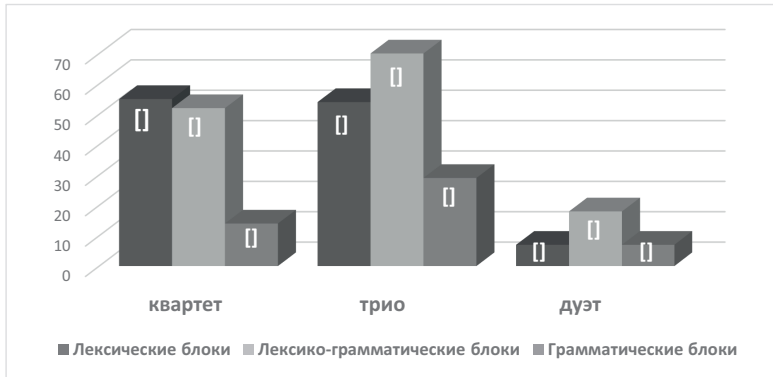


Рис. 2. Зависимость типов лексических блоков от количества слов

Грамматические блоки представлены в основном модальностью возможности и модальностью необходимости. Последняя является средством убеждения и побуждения читателя блога совершить действие, выбрать тот или иной вариант и т.п. Модальность возможности, преобладающая в корпусе, служит главным образом для описания характеристик и достоинств каких-либо устройств, программ, технологий и косвенным образом также выполняет персуазивную функцию — убедить читателя воспользоваться этими технологиями, что связано с реализацией рекламных целей профессионального блога.

Подводя итоги, можно сделать выводы о том, что наиболее результативными с точки зрения анализа видами лексических блоков являются квартеты, трио и дуэты, что связано не только с их более высокой частотностью, но и более четкими и разнообразными типами связей и функциями. Лексические связи (большинство которых представлены в виде профессиональных и терминологических сочетаний) преобладают в квартетах. Блоки, сочетающие в себе лексические и грамматические характери-

стики, составляют большую часть трио и дуэтов. Последние также отличаются тем, что лексические и грамматические блоки (преимущественно модальные сочетания возможности и долженствования) в них представлены одинаково, а также большим количеством неполных сочетаний, выражающих более абстрактные отношения между объектами (общие, пространственные, целевые).

Среди основных коммуникативных функций, выполняемых лексическими блоками, прежде всего можно выделить информативную (сообщение новых сведений через профессиональную лексику и терминологию). Когезивная функция обеспечения связности текста (выполняемая дискурсивными маркерами источника информации, цели, результата и добавления информации) облегчают восприятие текста читателем и позволяют ему сосредоточиться на содержательной части. Еще одной важной функцией является функция убеждения / побуждения, служащая рекламным целям и выражающаяся не только императивом, но модальными средствами — прежде всего, модальностью необходимости.

**СПИСОК ИСТОЧНИКОВ
И ЛИТЕРАТУРЫ**

1. Краткий психологический словарь [Текст] / Сост. Л.А. Карпенко; под общ. ред. А.В. Петровского, М.Г. Ярошевского. — М.: Политиздат, 1985. — 431 с.
2. Encyclopedia Britannica [Электронный ресурс]. — URL: <https://www.britannica.com/topic/blog#accordion-article-history> (дата обращения: 22.01.2019).
3. *Boyd, D.* A Bloggers blog: Exploring the definition of a medium [Электронный ресурс] / D. Boyd // Reconstruction. — Vol. 6(4), 2006. — URL: <http://reconstruction.cserver.org/064/boyd.shtml> (дата обращения: 22.01.2019).
4. *Добросклонская, Т.Г.* Вопросы изучения медиатекстов (опыт исследования современной английской медиаречи) [Текст] / Т.Г. Добросклонская. — М.: Едиториал УРСС, 2005. — 288 с.
5. *McCarthy, M.* Spoken fluency revisited [Текст] / M. McCarthy // English profile journal. — Vol. 1(1), 2010. — P. 1-15.
6. *O'Keefe, A., McCarthy, M., Carter R.* From corpus to classroom [Текст] / A. O'Keefe, M. McCarthy, R. Carter. — Cambridge: Cambridge University Press, 2007. — 315 p.
7. *Ахманова, О.С.* Словарь лингвистических терминов [Текст] / О.С. Ахманова. — М.: Едиториал УРСС, 2004. — 576 с.
8. *Александрова, О.В.* Проблемы экспрессивного синтаксиса (на материале английского языка) [Текст] / О.В. Александрова. — М.: Высшая школа, 1984. — 212 с.
9. *Тер-Минасова, С.Г.* Синтагматика речи: онтология и эвристика: Общая и английская синтагматика составных номинативных групп [Текст] / С.Г. Тер-Минасова. — М.: URSS, 2009. — 200 с.
10. *Кочеткова, Н.А.* Статистические языковые методы. Коллокации и коллигации [Электронный ресурс] / Н.А. Кочеткова // Новые информационные технологии в автоматизированных системах. — 2013. — №16. — С. 301-305. URL: <https://cyberleninka.ru/article/n/statisticheskie-yazykovye-metody-kollokatsii-i-kolligatsii> (дата обращения: 22.01.2019).
11. *Добросклонская, Т.Г.* Синтагматика глобально-клишированных текстов [Текст] / Т.Г. Добросклонская // Терминоведение и профессиональная лингводидактика. — М.: Московский лицей, 1993. — С. 11-15.
12. *Влавацкая, М.В.* Система базовых понятий комбинаторной лингвистики [Текст] / М.В. Влавацкая // Мир науки, культуры, образования. — 2011. — № 1 (26). — С. 32-38.
13. *Кубрякова, Е.С.* Синтагматика [Текст] / Е.С. Кубрякова // Большая российская энциклопедия. Т. 30. — М.: Большая российская энциклопедия, 2015. — С. 240-241.
14. *Baker, P., Hardie, A., McEnery, T.* A Glossary of Corpus Linguistics. [Текст] / P. Baker, A. Hardie, T. McEnery. — Edinburgh: Edinburgh University Press, 2006. — 187 p.
15. *Hunston, S.* Corpora in applied linguistics [Текст] / S. Hunston. — Cambridge: Cambridge University Press, 2002. — 241 p.
16. *Sinclair, J.* Corpus, concordance, collocation [Текст] / J. Sinclair. — Oxford: Oxford University Press, 1991. — 200 p.
17. *Stubbs, M.* Text and corpus analysis: Computer-assisted studies of language and culture [Текст] / M. Stubbs. — Oxford: Blackwell Publishers, 1996. — 267 p.
18. *Tognini-Bonelli, E.* Corpus Linguistics at Work [Текст] / E. Tognini-Bonelli. — Amsterdam: John Benjamins Publishing, 2001. — 223 p.
19. *Назарова, Т.Б.* Когнитивный подход и корпусные исследования: соперничество, союз или слияние? [Текст] / Т.Б. Назарова // Вопросы прикладной лингвистики. — 2012. — Вып. 7. — М.: РУДН, 2012. — С. 48-54.
20. *Толстова, Т.В.* Жанр и корпус: современные подходы к изучению и преподаванию языка [Текст] / Т.В. Толстова. — Самара: Изд-во Самарского университета, 2018. — 196 с.
21. *Толстова, Т.В.* Грамматическая семиотика отраслевого дискурса на примере геоинформационных технологий [Текст] / Т.В. Толстова // Эволюция и трансформация дискурсов. Сборник научных статей. — Самара: Изд-во Самарского университета, 2017. — С. 96-103.
22. *Biber, D., Conrad, S.* Lexical bundles in conversation and academic prose [Текст] /

- D. Biber, S. Conrad // Hasselard, H., Oksefjell, S. (eds) *Out of corpora: Studies in honor of Stig Johansson*. — Amsterdam: Rodopi, 1999. — P. 181-189.
23. Biber, D., Conrad, S., Cortés, V. If you look at...: Lexical bundles in university teaching and textbooks [Текст] / D. Biber, S. Conrad, V. Cortés // *Applied linguistics*. — Vol. 25(3), 2004. — P. 371-405.
 24. Ricardo, A., Nausa, T. Lexical bundles in academic presentations given by PhD level EAP students [Текст] / A. Ricardo, T. Nausa // *Gist education and learning research journal*. — Vol. 7, 2013. — P. 259-270.
 25. Fletcher, W. 2007. KfNgram [Электронный ресурс]. — URL: <http://kwicfinder.com/kfNgram/kfNgramHelp.html> (дата обращения: 22.01.2019).
- REFERENCES**
1. Akhmanova O.S., *Slovar lingvisticheskikh terminov*, Moscow, Editorial URSS, 2004, 576 p. (in Russian)
 2. Aleksandrova O.V., *Problemy ekspressivno-gosintaksisa (na materiale angliiskogo yazyka)*, Moscow, Vysshaya shkola, 1984, 212 p. (in Russian)
 3. Baker P., Hardie A., McEnery T., *A Glossary of Corpus Linguistics*, Edinburgh, Edinburgh University Press, 2006, 187 p.
 4. Biber D., Conrad S., Lexical bundles in conversation in conversation and academic prose, in: Hasselard H., Oksefjell S. (Eds.) *Out of corpora: Studies in honor of Stig Johansson*, Amsterdam, Rodopi, 1999, pp. 181-189.
 5. Biber D., Conrad S., Cortés V., If you look at...: Lexical bundles in university teaching and textbooks, *Applied linguistics*, 2004. Vol. 25(3), pp. 371-405.
 6. Boyd D., A Bloggers blog: Exploring the definition of a medium, *Reconstruction*, 2006, Vol. 6(4), available at: <http://reconstruction.eserver.org/064/boyd.shtml> (accessed: 22.01.2019).
 7. Dobrosklonskaya T.G., *Sintagmatika globalno-klisirovannykh tekstov, Terminovedenie i professionalnaya lingvodidaktika*, Moscow, Moskovskii litsei, 1993, pp. 11-15. (in Russian)
 8. Dobrosklonskaya T.G., *Voprosy izucheniya mediatekstov (opyt issledovaniya sovremennoi angliiskoi mediarechi)*, Moscow, Editorial URSS, 2005, 288 p. (in Russian)
 9. *Encyclopedia Britannica*, available at: <https://www.britannica.com/topic/blog#accordion-article-history> (accessed: 22.01.2019).
 10. Fletcher W., 2007. KfNgram, available at: <http://kwicfinder.com/kfNgram/kfNgramHelp.html> (accessed: 22.01.2019).
 11. Hunston S., *Corpora in applied linguistics*, Cambridge, Cambridge University Press, 2002, 241 p.
 12. Kochetkova N.A., *Statisticheskie yazykovye metody. Kollokatsii i kolligatsii, Novye informatsionnye tekhnologii v avtomatizirovannykh sistemakh*, 2013, No. 16, pp. 301-305, available at: <https://cyberleninka.ru/article/n/statisticheskie-yazykovye-metody-kollokatsii-i-kolligatsii> (accessed: 22.01.2019). (in Russian)
 13. *Kratkii psikhologicheskii slovar*, sost. L.A. Karpenko, ed. A.V. Petrovskii, M.G. Yaroshevskii, Moscow, Politizdat, 1985, 431 p. (in Russian)
 14. Kubryakova E.S., *Sintagmatika, Bolshaya rossiiskaya entsiklopediya, T. 30*, Moscow, Bolshaya rossiiskaya entsiklopediya, 2015, pp. 240-241. (in Russian).
 15. McCarthy M., Spoken fluency revisited, *English profile journal*, 2010, Vol. 1(1), pp. 1-15.
 16. Nazarova T.B., Kognitivnyi podkhod i korpusnye issledovaniya: sopernichestvo, soyuz ili sliyanie? *Voprosy prikladnoi lingvistiki*, 2012, Vyp. 7, Moscow, RUDN, pp. 48-54. (in Russian)
 17. O'Keeffe A., McCarthy M., Carter R., *From corpus to classroom*, Cambridge, Cambridge University Press, 2007, 315 p.
 18. Ricardo A., Nausa T., Lexical bundles in academic presentations given by PhD level EAP students, *Gist education and learning research journal*, 2013, Vol. 7, pp. 259-270.
 19. Sinclair J., *Corpus, concordance, collocation*, Oxford: Oxford University Press, 1991, 200 p.
 20. Stubbs M., *Text and corpus analysis: Computer-assisted studies of language and culture*, Oxford: Blackwell publishers, 1996, 267 p.
 21. Ter-Minasova S.G., *Sintagmatika rechi: ontologiya i evristika: Obshchaya i angliiskaya*

- sintagmatika sostavnykh nominativnykh grupp*, Moscow, URSS, 2009, 200 p. (in Russian)
22. Tognini-Bonelli E., *Corpus Linguistics at Work*, Amsterdam: John Benjamins Publishing, 2001, 223 p.
23. Tolstova T.V., Grammaticheskaya semiotika otraslevogo diskursa na primere geoinformatsionnykh tekhnologii, *Evolyutsiya i transformatsiya diskursov*. Sbornik nauchnykh statei, Samara, Izdatelstvo Samarskogo universiteta, 2017, pp. 96-103. (in Russian)
24. Tolstova T.V., *Zhanr i korpus: sovremennye podkhody k izucheniyu i prepodavaniyu yazyka*, Samara, Izdatelstvo Samarskogo universiteta, 2018, 196 p. (in Russian)
25. Vlavatskaya M.V., Sistema bazovykh ponyatii kombinatornoi lingvistiki, *Mir nauki, kultury, obrazovaniya*, 2011, No. 1 (26), pp. 32-38. (in Russian)

Толстова Татьяна Витальевна, кандидат филологических наук, доцент, кафедра иностранных языков и русского как иностранного, Самарский национальный исследовательский университет им. академика С.П. Королева, tvtolstova@yandex.ru

Tolstova T.V., PhD (Philology), Associate Professor, Department of Foreign Languages and Russian as a Foreign Language, S.P. Korolev Samara National Research University, tvtolstova@yandex.ru