

**ОСОБЕННОСТИ ПОДБОРА КЛЮЧЕВЫХ СЛОВ В АНГЛОЯЗЫЧНЫХ СТАТЬЯХ (НА ПРИМЕРЕ ИСТОРИЧЕСКИХ ЖУРНАЛОВ, ВХОДЯЩИХ В SCOPUS)<sup>1</sup>**

**Аннотация.** Статья посвящена особенностям подбора ключевых слов в англоязычных статьях на тему мировой истории, опубликованных в научно-исторических журналах, индексируемых в международной библиографической и реферативной базе данных Scopus. Материалом исследования послужили ключевые слова из 12 статей, опубликованных в четырех англоязычных научно-исторических журналах: «The Journal of Medieval History», «The Journal of Pacific History», «Cultural and Social History», «The International Journal of History of Sport». Ключевые слова изучались с точки зрения их числа, структуры и частоты упоминания в соответствующем тексте статьи. При этом автором разграничивается понятие «ключевое слово» и «тег». Тег – это такое слово или словосочетание, которое входит в перечень ключевых слов к статье, но при этом ни разу не встречается ни в заглавии, ни в аннотации, ни в тексте статьи. Исследование показало, что среднее количество ключевых слов в статье из англоязычного научно-исторического журнала составляет от 5 до 6, что соответствует и оперативной памяти человека и требованиям журналов. Ключевые слова, как правило, имеют очень высокий показатель частоты встречаемости в тексте статьи. Сравнение полученных результатов с данными ранее проведенного автором исследования в той же методологии на русскоязычном материале позволило прийти к заключению, что в англоязычных статьях в журналах гуманитарной направленности теги используются реже, чем в соответствующих русскоязычных статьях.

**Ключевые слова:** ключевое слово, тег, частота встречаемости, научный текст, научная статья, научно-исторический журнал, англоязычный журнал.

**Введение**

Актуальность веденного исследования обуславливается двумя факторами. Первым является доминирующая роль английского языка в сфере высшего образования и науки и в мировых коммуникационных технологиях, а именно его использование для написания или перевода статей, публикуемых в научных рецензируемых журналах [7, с. 47]. Вторым является обязательное указание ключевых слов (далее – КС) в публикуемых статьях с целью формирования определенного представления о теме научного текста и содержанием статьи, а также облегчения поиска статьи в библиографических и реферативных базах данных [13, с. 82]. Кроме того, набор ключевых слов, начиная с Л. В. Сахарного и А. С. Штерн [14], изучается как текст-примитив, как разновидность вторичного текста (см., например, [16]).

Цель данного исследования – определение особенностей подбора ключевых слов для раскрытия содержания научной статьи, публикуемой в англоязычном научно-историческом журнале. В рамках исследования были проанализированы тексты 12 научных статей из 4 англоязычных научно-исторических журналов (по 3 статьи из каждого журнала): «Journal of Medieval History», «The Journal of Pacific History», «Cultural and Social History» и «The International Journal of History of Sport», которые индексируются в международной библиографической и реферативной базе Scopus, содержащей в своих списках (по состоянию на 2018 г.) более 23 тысяч научных журналов по состоянию на 2018 год [7, с. 47].

Выбор журналов по проблемам истории обусловлен тем фактом, что, по сведениям других исследователей (например, [3]), в журналах гуманитарного профиля довольно часто наблюдается ситуация, когда в числе ключевых слов к статьям используются единицы, которые ни разу не встречаются в тексте самой статьи, – так называемые теги [1]. Кроме того, ранее нами было проведено исследование, связанное с изучением специфики подбора ключевых слов в научных статьях исторической тематики в русскоязычных журналах. Необходимо отметить, что англоязычный журнал в принципе не привязан к какой-либо конкретной стране, государственным языком которой является английский язык, если нет прямых указаний на это в заголовке журнала или темах его статей. В текстах проанализированных статей нами были найдены слова и словосочетания, которые номинально не относятся к предваряющим статью ключевым словам, но при этом состоят с ними в определенных лексических отношениях (являются синонимами или антонимами, гиперонимами или гипонимами и т. д.) и помо-

<sup>1</sup>Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 20-012-00290 «Устный и письменный нарратив как вторичный текст: особенности порождения разными категориями носителей русского языка»).

гают в определенной степени раскрыть содержание статьи. В рамках настоящего исследования слова, которые состоят в указанных выше отношениях либо с тегами, либо с ключевыми словами, употребляемыми не более 3 раз в тексте, т. е. мало упоминаемыми ключевыми словами (далее – МКС), обозначаются как «близкие по теме слова» (далее – БТ).

Исследование проходило в три этапа.

1. В заголовке, аннотации и тексте каждой статьи был произведен поиск предваряющих каждую статью ключевых слов (КС), а также поиск состоящих в лексических отношениях с тегами или мало употребляемыми ключевыми словами (МКС) близких по теме слов (БТ).

2. В каждой статье произведен подсчет числа упоминаний каждого КС и выявленного БТ, определено среднее количество упоминаний КС (с учетом БТ и без их учета), на основе чего сделаны выводы о возможной зависимости частоты упоминания КС от отнесения к таковым БТ. Вычисления и проверка корреляции проводились с применением программы LibreOffice Calc.

3. Проанализированы особенности подбора каждого из КС (в том числе тегов) с учетом их частотности и возможного наличия БТ в тексте каждой статьи.

### 1. Основные понятия, используемые в исследовании

В статье используются следующие термины и понятия.

**Научный стиль речи**, или **научный стиль текста** — один из пяти функциональных стилей речи (также выделяются публицистический, официально-деловой, разговорный и художественный), связанный со сферой науки и производства и выполняющий информативную и доказательную функции. Для научного стиля характерны такие свойства, как отвлеченность и обобщенность, подчеркнутая логичность и смысловая точность, терминованность и ясность, объективность изложения и стандартность, лаконичность и безличность. Подобный стиль характерен для статей в научных журналах, диссертаций, монографий и докладов, а также учебников и учебно-методических пособий [10, с. 131–135]. Научный текст иногда признается разновидностью **моделированного текста** – особого типа текста с устоявшейся стандартной структурой и некими регламентирующими параметрами или блоками [9, с. 107].

**Научная статья** – один из видов научной публикации как основного результата исследования, который содержит краткий, но достаточный для понимания отчет о проведенном исследовании и объективное обсуждение его значимости. В статье выделяются такие части, как: заглавие, аннотация, ключевые слова, введение, методы исследования, основные результаты и их обсуждение, заключение и список источников [7, с. 47; 8, с. 109–110; 9, с. 107]. В настоящее время подавляющая часть статей публикуется на английском языке в соответствующих научных журналах; подобные статьи, как правило, структурированы по требованиям мирового научного сообщества [7, с. 48].

**Ключевые слова** – наиболее важные и существенные для понимания текста слова, которые служат так называемыми «вехами» или «операторами» в процессе восприятия читателем текста для последующего его членения на смысловые группы и их обобщения, а также итогового понимания содержимого [12, с. 100]. Набор подобных слов позволяет выявить тематику текста и обеспечить его быстрое обнаружение по данным КС в поисковой системе или в базе данных научных работ; в большинстве случаев именно от набора ключевых слов зависит, будет ли отображаться в результатах поиска текст [5, с. 155; 11, с. 279; 13, с. 82]. Как правило, ключевые слова идут непосредственно перед текстом статьи, после аннотации, но могут быть представлены и в самой аннотации для введения читателей в курс дела [13, с. 82]. Они представляют собой одно- или многокомпонентные лексические группы (т. е. под ключевыми словами могут подразумеваться как слова, так и словосочетания) [15, с. 76]. При публикации статьи в отечественном научном журнале иногда обязательным для требования является перевод заглавия, аннотации и ключевых слов на английский язык, хотя в современном мире даже при этих мерах цитируемость англоязычных научных журналов остается очень высокой и недостижимой для журналов на других языках [7, с. 48].

Для КС характерны следующие свойства и качества:

- это наиболее частотные наименования, которые обозначают признак предмета, состояние или действие;
- они представлены значимой лексикой, обобщены по тематике, являются стилистически нейтральными;
- они связаны друг с другом пересечением значений;

- их количество варьируется в пределах от 5 до 15 или от 8 до 10 слов, что соответствует объему оперативной памяти человека;
- их повторение не может быть чрезмерно частым, чтобы текст не был исключен из результатов поиска;
- набор определяет контексты максимально предсказуемых слов [4, с. 86].

Слова, которые заявлены в качестве ключевых к научной статье, но не встречаются в ее тексте, называются *тегами*. Их роль сводится к характеристике текста статьи и указанию на его отличие от текстов других публикаций [1, с. 36]. Отношения между тегами и КС можно охарактеризовать как «общее – частное» [2, с. 1–2].

## 2. Анализ ключевых слов в научных статьях

В рамках исследования были проанализированы 12 статей из указанных 4 англоязычных научно-исторических журналов (в скобках приводится количество предваряющих статью ключевых слов). Представим имена авторов и названия статей на языке оригинала.

### **The Journal of Medieval History**

- 1) Randall Todd Pippenger. Lives on hold: the Dampierre family, captivity and the crusades in thirteenth-century Champagne (8 КС)
- 2) Gerd Althoff, Iben Fonnesberg-Schmidt & William Kynan-Wilson. Framing papal communication in the central Middle Ages (5 КС).
- 3) Kim Overlaet. The ‘joyous entry’ of Archduke Maximilian into Antwerp (13 January 1478): an analysis of a ‘most elegant and dignified’ dialogue (5 КС).

### **The Journal of Pacific History**

- 4) Lars Eckstein & Anja Schwarz. The Making of Tupaia’s Map: A Story of the Extent and Mastery of Polynesian Navigation, Competing Systems of Wayfinding on James Cook’s Endeavour, and the Invention of an Ingenious Cartographic System (8 КС)
- 5) Vicki Luker & Jane Buckingham. Histories of Leprosy: Subjectivities, Community and Pacific Worlds (6 КС)
- 6) Maaike Derksen. Local Intermediaries? The Missionising and Governing of Colonial Subjects in South Dutch New Guinea, 1920–42 (5 КС)

### **Cultural and Social History**

- 7) Jon Stobart & Mark Rothery. Fashion, Heritage and Family. New and Old in the Georgian Country House (5 КС)
- 8) Irene Gammel. Lacing Up the Gloves. Women, Boxing and Modernity (5 КС)
- 9) Claire Langhamer. Love, Selfhood and Authenticity in Post-War Britain (5 КС)

### **The International Journal of the History of Sport**

- 10) Dario Brentin & Dejan Zec. From the Concept of the Communist ‘New Man’ to Nationalist Hoolliganism: Research Perspectives on Sport in Socialist Yugoslavia (5 КС)
- 11) Marcus P. Chu. Post-Handover Hong Kong’s International Sporting Bids: A Win-Less-Lose-More Journey (5 КС)
- 12) Paul Bowman: Making Martial Arts History Matter (5 КС)

В совокупности в 12 статьях используется 67 предваряющих ключевых слов и словосочетаний, встречающихся в текстах всего 2192 раза. Наиболее частотное из них – ключевое слово *Tupaia* из статьи авторов Eckstein & Schwartz (279 раз). В среднем одно ключевое слово встречается в тексте 182,7 раз, что объясняется не только его высокой частотностью, но и частотностью иных слов (количество упоминаний в одной статье превышает показатель 100 употреблений; сюда относятся слова: *leprosy* (147 раз, статья Luker & Buckingham), *goeroes* (146 раз, статья Derksen), *boxing* (129 раз, статья Gammel), *love* (210 раз, статья Langhamer) и *martial arts* (114 раз, статья Bowman). Среднее количество предваряющих ключевых слов на одну статью составляет 5,6, что, в целом, соответствует свойствам КС и объясняется объемом оперативной памяти человека. Из 67 ключевых слов только 5 (около 7,5 %) можно охарактеризовать как теги, поскольку они, хотя и характеризуют текст статьи и позволяют ее отделить от иных научных публикаций, но в тексте не встречаются ни разу.

В семи были найдены 13 слов и словосочетаний (БТ), близких по теме малым ключевым словам и тегам; у 7 подобных БТ обнаружены лексические отношения с тегами, а у 6 – с МКС. В тех статьях, где присутствовали БТ, они соотносились либо только с МКС, либо с тегами; при этом имели место случаи, когда для обнаруженных МКС (статьи авторов Althoff et al., Eckstein & Schwarz и Brentin & Zec) не было соответствующих БТ. В Таблице представлены данные об употреблении клю-

ЧЕВЫХ СЛОВ ВО ВСЕХ 12 СТАТЬЯХ.

Таблица

Данные о частотности КС в статьях

Автор	Упоминаний КС	Теги / МКС	Наиболее частотные КС и БТ	Частотность
Pippenger	195 / 201	0 / 2	<i>family</i> (54) <i>crusades / crusade</i> (37) <i>Champagne</i> (35)	24,4 / 22,3
Althoff et al.	47 / 76	0 / 1	<i>Pope</i> (29) <i>Rome</i> (15) <i>papal communication</i> (14)	9,4 / 12,7
Overlaet	135	0	<i>Antwerp</i> (50) <i>Maximilian of Austria / Maximilian</i> (48) <i>joyous entries / joyous entry</i> (19)	27
Eckstein & Schwarz	347 / 382	2 / 1	<i>Tupaia</i> (279) <i>wayfinding</i> (27) <i>Polynesian</i> (26)	43,4 / 34,7
Luker & Buckingham	269	0	<i>leprosy</i> (147) <i>Pacific</i> (56) <i>Makogai</i> (33)	44,8
Derksen	238	0	<i>goeroes</i> (146) <i>Dutch New Guinea</i> (59) <i>local intermediaries</i> (15)	47,6
Stobart & Rothery	110	0	<i>consumption</i> (31) <i>elite</i> (27) <i>country house</i> (18)	22
Gammel	212 / 215	1 / 0	<i>boxing</i> (129) <i>women</i> (57) <i>modernity</i> (11)	42,4 / 30
Langhamer	278	0	<i>love</i> (210) <i>post-war</i> (43) <i>authenticity</i> (11)	55,6
Brentin & Zec	75 / 83	0 / 2	<i>Yugoslavia</i> (41) <i>physical culture</i> (19) <i>socialism</i> (11)	15 / 11,9
Chu	105 / 115	0 / 2	<i>Hong Kong</i> (78) <i>China</i> (18) <i>2009 East Asian Games</i> (9)	21 / 16,4
Bowman	181 / 201	2 / 0	<i>martial arts</i> (114) <i>history</i> (62) <i>tradition</i> (17)	36,2 / 28,7

В Таблице определено общее количество упоминаний КС в статье, среднее количество упоминаний

наний одного КС на статью (в формате «без учета БТ / с учетом БТ» при наличии БТ в статье); количество тегов и МКС; а также представлены три наиболее часто употребляемых ключевых слова для каждой статьи. При равенстве упоминаний в таблицу включалось то слово, которое ранее было представлено в списке предваряющих ключевых слов. Жирным выделены предваряющие статью КС. Для удобства вычислений числовые параметры были округлены до десятых.

Из представленных табличных данных можно вывести закономерность, заключающуюся в следующем: при игнорировании БТ и их невключении в список ключевых слов средняя частотность употребления одного КС в целом по статье возрастает. Корреляция между частотностью при включении и игнорировании БТ в каждом из журналов подтвердила это предположение, хотя и не везде параметр корреляции оказался одинаковым. Наиболее высок он для *The International Journal of History of Sport* (0,999), чуть меньше – для журналов *Cultural and Social History* (0,986) и *Journal of Medieval History* (0,983). Наименьший параметр 0,880 для журнала *The Journal of Pacific History* частично объясняется присутствием двух тегов (при восьми КС) в статье авторства Eckstein & Schwarz и трех БТ, им соответствующих. Более высокая частотность употребления КС и БТ вместе взятых по сравнению с частотностью только КС встречается в статье Г. Альтхоффа и его соавторов (Althoff et al.), и причиной этой более высокой частотности является БТ, превышающее по числу упоминаний любое ключевое слово (29 раз).

### 3. Особенности соотношения выделенных ключевых слов и содержания статей

В текстах анализируемых статей можно отметить следующие особенности употребления ключевых слов (как обычных слов, так и МКС и тегов), позволяющие раскрыть содержание статей.

В статье Рэндалла Пиппенджера (Pirpenger) наиболее часто употребляемыми КС являются *family*, *crusades* (также форма единственного числа *crusade*) и *Champagne*, что позволяет сформировать представление о содержании статьи. При этом в списке КС нет имени главного действующего лица, жизнеописание которого представлено в статье – рыцаря Ренара II де Дампьера, но есть два МКС *regency* (ему соответствует более часто употребляемый гипоним *regent*) и *Dampierre-le-Château*.

В статье Герда Альтхоффа (Althoff et al.) наиболее часто упоминаемым является не КС, а БТ *Pope*, встречающееся 29 раз (соответствует ключевому словосочетанию *Medieval papacy*) при том, что среди ключевых слов есть три словосочетания с прилагательным *papal*, образованным от данного существительного. Среди ключевых слов нет конкретных имен римских пап (например, *Gregory I*, *Innocent II* или *Alexander III*), об общении которых с монархами идет речь в статье. Тем не менее, словосочетание *papal communication*, которое является третьим по частотности среди КС и БТ, позволяет понять содержание статьи.

В статье К. Оверлет (Overlaet) наиболее часто употребляемыми являются три КС: *Antwerp* (50 раз), название города Антверпен (встречаются также исторические варианты *Antwerpen* или *Nantwerpen*); *Maximilian* (48 раз), имя эрцгерцога австрийского Максимилиана I (в тексте встречается как более полный титул *Maximilian of Austria*, так и просто имя) и *joyous entry* (19 раз с учетом формы множественного числа), название события, которое буквально переводится как «радостный въезд». Описание значения этого ключевого слова раскрывается в тексте (первый торжественный мирный визит монарха в один из городов). В наборе ключевых слов не встречаются теги.

В статье Л. Экштейна (Eckstein et al.) наиболее употребительным (279 раз) стало слово *Tupaia*, имя полинезийского проводника, автора карты островов Океании, которой пользовалась экспедиция Джеймса Кука. При этом в ключевые слова не входит словосочетание *Tupaia's map*, характеризующее собственно карту. К тегам в статье относятся словосочетания *star navigation* (из БТ встречается только гипоним *star course*), называющее ориентирование в пространстве у моряков, и *Indigenous knowledges and ontologies*, собирательный термин, связанный с культурой коренных народов (к БТ для этого тега относятся гипонимы *Polynesian* и *South Pacific*, характеризующие регион, где совершал плавание Джеймс Кук). Два упомянутых гипонима не были включены в список КС, поскольку авторы статьи считали достаточным размещение материала в журнале, название которого свидетельствует о его специализации – история стран Тихоокеанского региона.

В статье В. Люкер (Luker et al.) чаще других использовано слово, заявленное в заголовке *leprosy* (147 раз), переводящееся как «проказа» и именуемое описываемое явление. На втором месте по частотности, однако, идет ключевое слово *Pacific* (56 раз), что, скорее всего, дополнительно акцентирует внимание читателей на регионе, история которого рассматривается. Названия островов *Makogai* (архипелаг Фиджи) и *Molokai* (Гавайские острова) позволяют сделать вывод о том, где именно была эпидемия проказы и где применяли меры по борьбе против этого заболевания.

В статье М. Дерксен (Derksen) из ключевых слов лидером по частотности является слово *goeroes* (146 раз), которое в аннотации переводится как «учитель, наставник» (речь идет о подготовленных католическими священниками учителей, набранных из числа коренного населения Полинезии). Второе из КС по числу упоминаний *Dutch New Guinea* (59 раз) позволяет определить географическую территорию, история которой освещается в данной статье, – это современная Папуа-Новая Гвинея. Находящиеся на третьем и четвертом местах *local intermediaries* и *catholic mission* называют явление – миссионерская деятельность Нидерландов на данной территории. Также в аннотации разъясняется содержание ключевого словосочетания *dual colonisation*: развитие территории не только усилиями колониальных властей, но и местных общественных и культурных деятелей (в том числе и вышеозначенных *goeroes*).

В статье Дж. Стобарта (Stobart et al.) набор КС не позволяет определить время и место анализируемых событий: на это намекает только прилагательное *Georgian* в заголовке, которое не является ключевым словом, но дает указания на эпоху правления британского короля Георга III. Тем не менее, предваряющие статью ключевые слова *consumption*, *elite*, *material culture* указывают на то, что статья посвящена быту домов Англии конца XVIII века (то есть материальной культуре) и образу жизни их владельцев – британской аристократии.

В статье И. Гаммель (Gammel) пара ключевых слов *woman* (57) и *boxing* (129), которые чаще других упоминаются в тексте, сразу же позволяет сделать акцент на теме – истории стремительного развития женского бокса, начавшегося, согласно аннотации, еще в 1920-е гг. Намного реже упоминаются в тексте лексемы *gender* и *modernity*, которые акцентируют внимание на изменении социальных гендерных ролей. Роль тега в статье играет *self-representation*, которому соответствуют всего два БТ в виде синонимов *representing* и *self-realization*; но который позволяет понять основную мысль статьи – появление новой возможности для самореализации женщин.

Статья К. Лангхамера (Langhamer) связана как с историей британского послевоенного общества, так и с литературой, о чем свидетельствуют выдержки из письменных текстов. Все ключевые слова представлены собственно в заголовке статьи – это *love* (210), *post-war* (43), *authenticity* (11), *self-hood* (9) и *Britain* (5).

В статье Д. Brentина (Brentin et al.) наличествуют два МКС – это *sport history* и *historiography* (по 2 упоминания), однако только для первого удалось найти более конкретные БТ в виде названий спортивных сообществ *Crvena Zvezda* и *Dinamo Zagreb*, об извечном противостоянии которых и знаменитой потасовке между их болельщиками на загребском стадионе «Максимир» 1990 г. говорится в самом начале статьи. Остальные ключевые слова употребляются намного чаще и помогают сформулировать тему публикации – развитие физкультуры и спорта в социалистической Югославии. Включение лексемы *socialism* позволяет подчеркнуть тот факт, что речь идет о СФРЮ (Социалистической Федеративной Республике Югославии), а не о довоенном Королевстве Югославии и не об образовавшейся позднее Союзной Республике Югославия.

В статье М. Чу (Marcus P. Chu) ключевые слова *Hong Kong* и *China* в контексте не только указывают на страны, о спортивных организациях которых идет речь, но и свидетельствуют об их сотрудничестве в плане подготовки Гонконга к принятию мероприятий, заявки на которые он подавал: *the 2006 Asian Games bid*, *the 2009 East Asian Games bid* и *the 2011 IOC Session bid* (последние два — МКС, им соответствуют более гиперонимы *2009 East Asian Games* и *2011 International Olympic Committee Session*, описывающие собственно события, а не процедуры подачи заявки на их проведение). Ключевые слова подобраны так, чтобы описывать список принятых Гонконгом мер, но не их результаты (результаты описаны устойчивым выражением *Win-Less-Lose-More*, характеризующим результаты как неудачные).

В статье П. Боумэна (Bowman) наиболее частотным КС является *martial arts* (114 раз), которое вместе с другим КС *history* (62 раза) может входить в состав единого выражения *martial arts history*, формулирующего тему статьи. Остальные три слова каких-либо значимых уточнений не вносят, хотя два из них (*invented tradition* и *allochronism*) являются тегами с соответствующими БТ *tradition* (15 раз, гипероним) и *allochronic* (3 раза, образованное от него прилагательное).

Если сопоставить результаты этого исследования с результатами предыдущего, проведенного нами на базе русскоязычных статей из научно-исторических журналов [6], то можно отметить одинаковую роль ключевых слов как в русскоязычных, так и англоязычных научных статьях; схожими также оказались высокие показатели корреляции между средней частотностью употребления ключевых слов и отнесения к ним БТ (от 90 %). Однако есть одно существенное различие: в англоязычных

статьях в плане раскрытия темы статьи отдается предпочтение прежде всего обычным ключевым словам, а не тегам. В англоязычных статьях теги составляют не более 7,5 %, в то время как в русскоязычных статьях эта доля может достигать 15 % [6, с. 70].

### Заключение

Результаты проведенного исследования позволяют сформулировать несколько значимых для дальнейших исследований суждений.

В научной статье ключевые слова не только помогают читателю определить тематику текста и интерпретировать его содержание, но и позволяют отобразить статью в результатах поиска по КС в соответствующей поисковой системе, причем в зависимости от выбора ключевых слов текст может как появиться на первых страницах результатов поиска, так и не появиться вообще.

Среднее количество ключевых слов в статье из англоязычного научно-исторического журнала составляет от 5 до 6, что соответствует и оперативной памяти человека, и требованиям журналов. Тегами в англоязычных статьях являются около 7,5 % ключевых слов: не будучи использованы в тексте статьи (за исключением списка предваряющих КС), они фиксируют уникальные характеристики текста, позволяющие выделить его в тематически близких публикациях.

Тематически близкие слова и словосочетания, которые состоят в определенных лексических отношениях с тегами или мало употребляемыми ключевыми словами, в англоязычных статьях встречаются не так часто, поскольку ключевые слова специально подобраны так, чтобы раскрывать содержание статьи. В некоторых случаях содержание КС дополнительно раскрывается в аннотацию.

Высокие показатели среднего количества упоминаний ключевых слов по отношению к одной статье или всему материалу исследования объясняются наличием КС с огромным количеством упоминаний в конкретной статье (279 раз – максимальный показатель для одного ключевого слова). Частотность КС в заголовке, аннотации и тексте статьи почти всегда возрастает, если при определении списка КС полностью игнорируются любые близкие к ключевым слова.

### ЛИТЕРАТУРА

1. Абрамов Е. Г. Подбор ключевых слов для научной статьи // Научная периодика: проблемы и решения. 2011. № 2. С. 35–40.
2. Афонин С. А., Бахтин А. В. Об одном методе автоматического построения гиперонимов с помощью внешней поисковой системы // III Всероссийская конференция с международным участием «Знания – Онтологии – Теории» (ЗОНТ-2011), Новосибирск, 3-5 октября 2011. С. 1–10.
3. Бусыгина О. А. Структурно-семантический анализ ключевых слов к научной статье (на материале «Вестника ЧГУ») // Проблемы порождения и восприятия речи: материалы XI выездной школы-семинара (7-8 декабря 2012 г., Череповец) / отв. ред. Р.Л. Смулаковская. Череповец: ЧГУ, 2013. С. 5–11.
4. Ванюшкин А. С., Гращенко Л. А. Методы и алгоритмы извлечения ключевых слов // Новые информационные технологии в автоматизированных системах. 2016. № 19. С. 85–93.
5. Гринева М., Гринева М., Лизоркин Д. Анализ текстовых документов для извлечения тематически сгруппированных ключевых терминов // Труды Института системного программирования РАН. 2009. С. 155–165.
6. Екимов М. А. Особенности подбора ключевых слов в русскоязычных научных статьях (на материале журналов по истории из списка ВАК) // Вестник магистратуры. 2020. № 1–5. С. 65–71.
7. Иваницкая Е. В. Статья как элемент научной коммуникации: особенности формирования научного авторитета в международном сообществе // Научный редактор и издатель. 2018. № 3 (1-2). С. 44–51.
8. Казакова А. Г. Научная статья как результат деятельности исследователя // Культура и образование. 2013. № 1 (10). С. 109–115.
9. Камшилова О. Н. Малые формы научного текста: ключевые слова и аннотация (информационный аспект) // Известия Российского государственного педагогического университета им. А.И. Герцена. Филология. 2013. № 156. С. 107–117.
10. Колесникова Н. И. Что важно знать о языке и стиле научных текстов // Высшее образование в России. В помощь аспиранту. 2010. № 3. С. 130–137.
11. Москвитина Т. Н. Ключевые слова и их функции в научном тексте // Известия ЧГПУ. Филология и искусствоведение. 2009. № 11. С. 277–283.
12. Осипова И. А. Модельное представление процессов понимания текста на основе ключевых слов // Вестник Оренбургского государственного университета. 2012. № 11 (147). С. 100–105.
13. Родригес В. Как написать эффективные заглавие статьи и авторское резюме и выбрать подходящие ключевые слова // Научный редактор и издатель. 2018. № 3 (1-2). С. 81–84.

14. Сахарный Л. В., Штерн А. С. Набор ключевых слов как тип текста // ...*Слово отвечает*: памяти А.С. Штерна и Л. В. Сахарного. Пермь: Изд-во Пермского ун-та, 2006. С. 50–59.
15. Шереметьева С. О., Осминин П. Г. Методы и модели автоматического извлечения ключевых слов // Вестник Южно-Уральского государственного университета. Лингвистика. 2015. № 1. Т. 12. С. 76–81.
16. Arjukhina E. G., Grudeva E. V., Churilina L. N. On secondary text as a structural and semantic unity // Вопросы когнитивной лингвистики. 2017. № 2 (51). С. 101–106.

*M. A. Ekimov (Cherepovets, Russia)*

**PECULIARITIES OF KEYWORDS CHOOSING FOR ENGLISH LANGUAGE SCIENCE ARTICLES (BASING ON DATA OF HISTORY JOURNALS INCLUDED IN SCOPUS BASE<sup>2</sup>)**

**Abstract.** This article considers the peculiarities of keywords choosing and mentioning in English-language science articles of world history from history journals, which are indexed in Scopus international abstract and citation base. It presents such aspects as choosing keywords to approach the chosen theme of science text and trend of mentioning the keywords directly in the text of article. The research material includes 12 different articles from 4 English-language history journals, The Journal of Medieval History, The Journal of Pacific History, Cultural and Social History, The International Journal of History of Sport. The keywords were studied from the perspectives of their total amount, structure and frequency of references in a text. A definition of tag as one of keyword types is present in this paper; a tag is a word or word collocation which is included in the list of keywords but is never mentioned in headlines, abstracts or texts. The study has showed that the average amount of keywords in articles from English-language history journal varies from 5 to 6 and not only meets the journal requirements, but also corresponds to human short term memory. As a rule, the keywords have very high frequency of mentioning inside the article text. Another research made on Russian-language material using the same methods showed that tags are used less frequently in the articles from English-language humanities journals than the related Russian-language articles.

**Keywords:** keyword, tag, frequency of mention, science text, science article, history journal, English-language journal

**REFERENCES**

1. Abramov E. G. Podbor klyuchevykh slov dlya nauchnoi stat'i, Nauchnaya periodika: problemy i resheniya, 2011, no. 2, pp. 35–40.
2. Afonin S. A., Bakhtin A. V. Ob odnom metode avtomaticheskogo postroeniya giperonimov s pomoshch'yu vneshnei poiskovoi sistemy, III Vserossiiskaya konferentsiya s mezhdunarodnym uchastiem «Znaniya – Ontologii — Teorii» (ZONT-2011), Novosibirsk, 3–5 oktyabrya 2011, pp. 1–10.
3. Busygina O. A. Strukturno-semanticheskii analiz klyuchevykh slov k nauchnoi stat'e (na materiale «Vestnika ChGU»), Problemy porozhdeniya i vospriyatiya rechi : materialy XI vyezdnoi shkoly-seminara (7-8 dekabrya 2012 g., Cherepovets) / Otv. red. R.L. Smulakovskaya, Cherepovets, ChGU, 2013, pp. 5–11.
4. Vanyushkin A. S., Grashchenko L. A. Metody i algoritmy izvlecheniya klyuchevykh slov, Novye informatsionnye tekhnologii v avtomatizirovannykh sistemakh, 2016, no.19, pp. 85–93.
5. Grineva M., Grinev M., Lizorkin D. Analiz tekstovykh dokumentov dlya izvlecheniya tematicheskikh sgruppovannykh klyuchevykh terminov, Trudy Instituta sistemnogo programmirovaniya RAN, 2009, pp. 155–165.
6. Ekimov M. A. Osobennosti podbora klyuchevykh slov v russkoyazychnykh nauchnykh stat'yakh (na materiale zhurnalov po istorii iz spiska VAK), Vestnik magistratury, 2020, no. 1–5, pp. 65–71.
7. Ivanitskaya E. V. Stat'ya kak element nauchnoi kommunikatsii: osobennosti formirovaniya nauchnogo avtoriteta v mezhdunarodnom soobshchestve, *Nauchnyi redaktor i izdatel'* [Science Editor and Publisher], 2018, no. 3 (1-2), pp. 44–51.
8. Kazakova A. G. Nauchnaya stat'ya kak rezul'tat deyatelnosti issledovatelya, *Kul'tura i obrazovanie* [Culture and education], 2013. no. 1 (10), pp. 109–115.
9. Kamshilova O. N. Malye formy nauchnogo teksta: klyuchevye slova i annotatsiya (informatsionnyi aspekt), *Izvestiya Rossiiskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gertsena. Filologiya* [Izvestia: Herzen University Journal of Humanities & Sciences], 2013, no. 56, pp. 107–117.
10. Kolesnikova N. I. Chto vazhno znat' o yazyke i stile nauchnykh tekstov, Vysshee obrazovanie v Rossii. V pomoshch' aspirantu. 2010, no.3, pp. 130–137.
11. Moskvitina T. N. Klyuchevye slova i ikh funktsii v nauchnom tekste, *Izvestiya ChGPU, Filologiya i iskusstvovedenie* [South Ural State Humanitarian Pedagogical University Bulletin], 2009, no. 11, pp. 277–283.

<sup>2</sup> This work was supported by the Russian Foundation for Basic Research (project No. 20-012-00290 “Oral and written narrative unit as a secondary text: features of a creating process by different categories of the Russian language native speakers”).



12. Osipova I. A. Model'noe predstavlenie protsessov ponimaniya teksta na osnove klyuchevykh slov, *Vestnik Orenburgskogo gosudarstvennogo universiteta* [Vestnik of the Orenburg State University], 2012, no. 11 (147), pp. 100–105.
  13. Rodrigues V. Kak napisat' effektivnye zaglavie stat'i i avtorskoe rezyume i vybrat' podkhodyashchie klyuchevye slova, *Nauchnyi redaktor i izdatel'* [Science Editor and Publisher], 2018, no. 3 (1-2), pp. 81–84.
  14. Sakharnyi L. V., Shtern A. S. Nabor klyuchevykh slov kak tip teksta, ... Slovo otzovetsya: pamyati A. S. Shtern i L. V. Sakharnogo, Perm', Izd-vo Permskogo un-ta, 2006, pp. 50–59.
  15. Sheremet'eva S. O., Osminin P. G. Metody i modeli avtomaticheskogo izvlecheniya klyuchevykh slov, *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Lingvistika* [Bulletin of South Ural State University, Series «Linguistics»], 2015, no.1, vol. 12, pp. 76–81.
  16. Arjukhina E. G., Grudeva E. V., Churilina L. N. On secondary text as a structural and semantic unity, *Voprosy Kognitivnoy Lingvistiki* [Issues of Cognitive Linguistics], 2017, no. 2 (51), pp. 101–106.
- 

Екимов М. А. Особенности подбора ключевых слов в англоязычных статьях (на примере исторических журналов, входящих в Scopus) // Гуманитарно-педагогические исследования. 2020. Т. 4. № 2. С. 47–55.

Ekimov M.A. Peculiarities of Keywords Choosing for English Language Science Articles (Basing on Data of History Journals Included in Scopus Base), *Gumanitarno-pedagogicheskie issledovaniya* [Humanitarian and pedagogical Research], 2020, vol. 4, no 2, pp.47-55

Дата поступления статьи – 02.04.2020; печ. л.

#### *Сведения об авторе*

**Екимов Марк Анатольевич** – магистрант кафедры отечественной филологии и прикладных коммуникаций Гуманитарного института ФГБОУ ВО «Череповецкий государственный университет», г. Череповец, Россия; markus9339@mail.ru

#### **Author:**

**Mark A. Ekimov**, Masters student, Department of Russian Philology and Applied Communications, Humanities Institute, Cherepovets State University (ChSU), Cherepovets, Russia; [markus9339@mail.ru](mailto:markus9339@mail.ru)

---