

УДК 378

АНГЛОЯЗЫЧНОЕ НАУЧНОЕ ПИСЬМО: ЗАТРУДНЕНИЯ НАЧИНАЮЩИХ РУССКОЯЗЫЧНЫХ АВТОРОВ¹

Е.И. Шпит (Томск, Россия)

В.Н. Куровский (Томск, Россия)

Аннотация

Постановка проблемы. В статье рассматривается проблема многочисленных несоответствий научных текстов на английском языке, создаваемых начинающими русскоязычными авторами, нормам и ожиданиям международного научного сообщества.

Цель статьи – предложить концепцию применения инструментов компьютерного текстового анализа в педагогических целях для выявления актуальных потребностей обучающихся в области написания научного текста на английском языке.

Методологию исследования составляют сравнительный дискурс-анализ корпуса рукописей научных статей студентов магистратуры и аспирантуры российского технического вуза и корпуса опубликованных статей международных авторов. Оба корпуса представляют одну инженерную дисциплину и созданы специально для целей исследования. Анализ проводился с применением инструмента компьютерной лингвистики Gramulator.

Результаты. Количественный и качественный анализ дифференцирующих n-грамм в Gramulator показал, что письмо начинающих русскоязычных исследователей в значительной степени отличается от письма опытных международных авторов. Приведенные примеры указывают на низкое лексическое разнообразие; недостаточно развитые навыки применения грамматических явлений, отсутствующих в русском языке; склонность избегать употребления предикативных придаточных предложений и средств хеджирования; склонность к чрезмерному употреблению слов и формулировок, типичных для русского научного стиля, и к чрезмерному уточнению.

Заключение. Предлагаемая авторами концепция призвана придать процессу обучения письму более студентоцентрированный характер. Выявление лингвистических проблем в письме начинающих авторов на фоне письма целевого дискурсивного сообщества позволяет повысить осведомленность студентов о различиях в риторическом выборе в двух языках, скорректировать стратегические навыки такого выбора при составлении собственного текста на английском языке и тем самым приблизиться к нормам целевого дискурсивного сообщества.

Ключевые слова: научный дискурс, научное письмо, дискурс-аналитический подход, сравнительный корпусный анализ, дискурс-анализ, дискурсивное сообщество, Gramulator.

Шпит Елена Ирисметовна – аспирант кафедры педагогики и психологии образования факультета психолого-педагогического и специального образования института психологии и педагогики, Томский государственный педагогический университет; доцент кафедры иностранных языков, Томский государственный университет систем управления и радиоэлектроники; ORCID-ID: 0000-0003-2568-224X; e-mail: forester_2007@mail.ru

Куровский Василий Николаевич – доктор педагогических наук, профессор, заведующий лабораторией развития образования Научно-образовательного центра теории образования, Томский государственный педагогический университет; e-mail: v.kuroskii@yandex.ru

Благодарность. Авторы выражают огромную признательность разработчику Gramulator, теоретику и практику академического письма Филипу МакКарти (Philip McCarthy) за оказанную помощь в овладении инструментом и обучении навыкам интерпретирования результатов.

¹ Исследование выполнено за счет гранта Российского научного фонда, проект № 22-29-01331.

Постановка проблемы. Необходимость публиковать результаты своих исследований в международных рецензируемых журналах на английском языке становится все более актуальной в современном глобализованном мире. Уже на этапе магистратуры перед студентами ставится задача пробовать свои силы в написании научной статьи или тезисов для международных журналов и конференций на английском языке. На уровне аспирантуры наличие англоязычных статей становится почти обязательным условием итоговой аттестации. Однако статистика по показателям количества публикаций российских авторов в базах данных Scopus и Web of Science, а также по уровню цитируемости статей, написанных российскими авторами, все еще удручает [Бородик и др., 2021; Мохначева, Цветкова, 2021]. Низкие уровни российского участия и цитируемости могут подразумевать многочисленные проблемы в области международной публикационной активности наших ученых. Среди них недостаточно развитые навыки научной письменной коммуникации на английском языке, которые могут привести к множеству сопутствующих проблем, например высокому уровню неуверенности в написании научного текста для высокорейтинговых журналов или нежеланию связываться с довольно жесткой системой международного рецензирования. Соответственно, прослеживается востребованность обучения навыкам написания научных текстов на английском языке именно с учетом их дальнейшей публикации в международных рецензируемых изданиях. Если мы хотим добиться повышения показателей международной публикационной активности российских ученых, необходимо ввести в университетское образование систематизированное обучение академическому письму как «самостоятельной отрасли знания и дисциплине, имеющей хорошо разработанные теоретико-методологические основы» [Короткина, 2018а, с. 65].

Формирование академического письма (АП) как отрасли знания и дисциплины начинается в России с 2010-х гг. Активизируются дис-

куссии о терминологии, содержании и месте АП на страницах журналов (например, «Высшее образование в России»). Переводятся и вводятся в обучение международные пособия по обучению написанию научных статей на английском языке. Разрабатываются собственные учебные и учебно-методические пособия, а также массовые открытые онлайн-курсы по обучению разным аспектам АП. Большим шагом в формировании организационно-методологических основ академического письма как дисциплины стало создание Консорциума центров письма, который объединил различные университетские центры, офисы и лаборатории письма [Базанова, Короткина, 2017]. Объединяя экспертов в области АП, распространяя международный опыт с учетом российской реальности, аккумулируя опыт российских практиков и совершенствуя технологии обучения, Консорциум создает «концептуальное поле» АП в России [Там же, с. 53]. Усилиями участников консорциума создается научно-методическая основа АП как самостоятельной дисциплины.

Под «академическим письмом» («academic writing») подразумевается тип письма, который характерен для работ, создаваемых в академической среде (включая обучение в университете) и для академической, т.е. научной, среды [Hyland, 2011]. Эпистемологической основой АП являются методы научного убеждения, и уже на этапе обучения в университете студент погружается в письменные работы (эссе, рефераты, курсовые и т.д.), которые должны обладать характеристиками научного стиля. Термин «научное письмо», который мы будем использовать в данной статье, несколько отличается от термина «академическое письмо» более узкой направленностью знаний, умений и навыков (ЗУН), а именно нацеленностью сообщить научному сообществу результаты собственной исследовательской деятельности. В этом смысле «научное письмо» больше соотносится с английским термином «research writing» или «writing for research publication purposes» («письмо для научно-публикационных целей») [Короткина, 2018б; Flowerdew, 2013; Hyland, 2007; Kwan, 2010], т.е.

в данной работе под термином «научное письмо» мы будем подразумевать навыки написания научного текста с учетом стилистических (соблюдение норм научного стиля языка письма), жанровых (соблюдение норм собственно-научных жанров) и дискурсивных (соблюдение норм и традиций дискурсивного сообщества) конвенций.

Термин «научный дискурс» является синонимичным термину «academic discourse»; оба относятся к образу мышления и употребления языка, который существует в научном сообществе [Hyland, 2011], или к совокупности текстов, объединенных общей тематической соотнесенностью [Чернявская, 2018]. Карасик [2000] под научным дискурсом подразумевал разновидность институционального дискурса, предполагающую отношения между представителями научной общественности, которые могут не знать друг друга, но должны общаться в соответствии с нормами данного социума. Целью научного дискурса является развитие научного знания, которое достигается путем решения ряда задач, включая когнитивно-коммуникативные (достижение истины путем обсуждения проблемы, ведения экспериментальных изысканий, изучения имеющейся литературы и т.д.) и социально-коммуникативные (распространение научных знаний путем публикации результатов, чтения лекций, проведения семинаров и т.д.) [Комарова, 2012]. Наличие экстралингвистического компонента (научное сообщество, институциональность, распространение знаний, социальная ситуация и т.д.) отличает научный дискурс от научного функционального стиля, который в большей мере касается общих черт стиля, отличающих его от других функциональных стилей. Соответственно, изучение различных лингвистических явлений дискурса происходит параллельно с их соотнесенностью с экстралингвистическими характеристиками дискурса и называется «дискурс-анализом» [Bhatia et al., 2008; Hyland, 2011; Макаров, 2003].

Есть множество методов дискурс-анализа, начиная с изучения жанровых характеристик и заканчивая анализом влияния выбираемых дис-

курсивных формул на человека (примеры можно найти: [Hyland, 2011, Чернявская, 2018; Макаров, 2003]). Например, англоязычное научное письмо, создаваемое русскоязычными авторами, рассматривалось с точки зрения ясности и недвусмысленности выражения [Terenin, 2013], удобочитаемости [Добрынина, 2019] или лексических и грамматических ошибок [Добрынина, 2017; 2018]. Более подробно данная тема будет рассмотрена в обзоре научной литературы.

Свой вклад в совершенствование методов и технологий обучения англоязычному научному письму в России мы видим в расширении возможностей «дискурс-аналитического подхода» [Fairclough, 2003; Hyland, 2018; Huckin, 2003; Reid, 2018]. Этот подход также называют риторическим, индуктивным, «down-top» и т.д. Сущность подхода заключается в изучении практик применения риторических средств для достижения определенных прагматических целей, а также корректировке тех, что считаются несоответствующими нормам дискурсивного сообщества. Поэтому данный подход, с одной стороны, рассматривает «текст как социальное событие, анализ которого связывает воедино социальный контекст коммуникации, дискурсивные практики и лингвистические механизмы их реализации» [Шейпак, 2020, с. 95], с другой – письмо как способ социализации в дискурсивное сообщество [Connor, 2002]. Довольно часто этот подход основывается на сопоставительной риторике (контрастивной или межкультурной риторике – *contrastive or intercultural rhetoric*). Еще в 1966 г. Р. Каплан [Kaplan, 1966], родоначальник контрастивной риторики, обнаружил, что общая риторика (приемы и способы построения речи) письма имеет свои культурные особенности. Например, письмо российских авторов, по мнению Каплана, отличается длинными, тяжеловесными предложениями и наличием отступлений, которые не относятся к непосредственной линии аргументации. Однако как подход к изучению лингвистических различий в речи двух и более групп людей, контрастивная риторика получила распространение в 1990-е гг. (обзор в: [Connor, 2002]).

Предлагаемая в данном исследовании концепция основана на сопоставительном анализе письма, проводимого с помощью инструмента компьютерной лингвистики Gramulator [McCarthy et al., 2012b]. Этот инструмент позволяет провести статистический анализ двух корпусов текста и извлечь наиболее часто используемые n-граммы (сочетания слов), дифференцирующие каждый корпус. Кроме того, функционал инструмента позволяет изучить использование этих n-грамм в текстах и проверить гипотезы о возможных причинах выбора тех или иных риторических средств. Создаваемая картина применения способов выражения мысли может очертить спектр коммуникативных проблем и тем самым способствовать модификации содержания обучения письму. Сочетание количественного и качественного анализов, реализуемых в данной технологии, позволяет сфокусировать дискурс-аналитический подход в обучении письму на трудностях, присущих определенной категории авторов. Это придает процессу обучения письму студент-центрированный характер, при котором учитываются не только отнесенность письма к целевому дискурсивному сообществу, но и языковые и коммуникативные потребности конкретной группы обучающихся.

Актуальность исследования. Поскольку на данном этапе уровень научно-методической разработки вопросов обучения англоязычной письменной научной коммуникации с учетом культурных особенностей русскоязычных авторов недостаточно развит, необходимо использовать различные способы расширения теоретико-методологической базы дисциплины, в том числе инструментарий компьютерной лингвистики. Кроме того, методика формирования способности «анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия» (УК-5, ФГОС ВО 3++)² в содержании обучения письму довольно слабо представлена конкретными примерами такого культурного разно-

образия. Инструменты компьютерного текстового анализа, например Gramulator, способны пролить свет на определенные проблемы в письме, отражая уровень сформированности компетенции. До настоящего времени никто из известных нам российских авторов не рассматривал данный инструмент для педагогических целей, в чем и состоит новизна данного исследования.

Обзор научной литературы. В силу многогранности понятия «дискурс» в настоящее время в лингвистике существует множество подходов к определению этого понятия и критериев, различающих типы дискурса [Макаров, 2003; Чернявская, 2018]. Тем не менее общим для всех является то, что дискурс объединяет лингвистическую и экстралингвистическую компоненты [Ван Дейк, 1998; Карасик, 2000; Gee, 2005; Widdowson 2007]. Как уже упоминалось выше, под «научным дискурсом» мы понимаем разновидность институционального дискурса, по В.И. Карасику [2000]. Принадлежность одному институту знаний определяет необходимость всех участников научного сообщества «говорить на одном языке», чтобы быть понятыми и принятыми этим сообществом. Во многом эта способность определяется умением использовать единую риторику дискурса, т.е. единые методы научного убеждения [Короткина, 2018а].

Сравнительный дискурс-анализ в данной работе во многом опирается на анализ специфических характеристик письма с точки зрения «научного функционального стиля». Под научным функциональным стилем мы подразумеваем разновидность литературного языка, связанного с научной деятельностью и основанного на «общности формы общественного сознания» [Чернявская, 2018, с. 117]. Однако дискурс, помимо «облигаторных моделей языковой деятельности» функционального стиля, включает еще и «содержательно-смысловую общность текстов» дисциплины или сообщества [Там же]. Совокупно это предполагает, что исследователь, желающий быть частью конкретного научного сообщества, должен знать и воспроизводить как универсальные для всех научных текстов модели языковой деятельности, так и коммуника-

² Федеральный государственный образовательный стандарт высшего образования для магистратуры и специалитета [Электронный ресурс]. URL: <https://fgosvo.ru/fgosvo/index/25> (дата обращения: 30.07.2022)

тивные речевые нормы и конвенции этого сообщества. Примеры последних могут быть обнаружены в результате сравнительного дискурсивного анализа, в данном случае выполненного инструментом компьютерной лингвистики.

Сравнительный дискурсивный анализ позволяет выявить отличительные особенности набора текстов учащихся и набора текстов целевого сообщества, предполагая, что эти особенности сигнализируют об отклонениях или несоответствиях желаемым нормам и ожиданиям. В этой связи стоит сказать, что многие исследователи отмечают существенные отличия в научном дискурсе разных дисциплин, выступающих как единое тематическое сообщество [Becher, 1987, 1994; Connor, 2002; Hyland, 2011]. Отсюда и понятие «дискурсивное сообщество», т.е. сообщество людей, объединенное общими социальными целями, общими средствами обмена знаниями и информацией, общими нормами и традициями дискурса и т.д. [Swales, 1990].

Исследование научного дискурса, как правило, основывается на анализе научных статей и/или аннотаций к научным статьям. Первые выступают в качестве основной жанровой единицы в коммуникативном и прагматическом планах научной деятельности и отвечают всем признакам научной прозы [Колесникова, Ридная, 2016]. Вторые являются «лицом» научной статьи, потому что информируют читателя о ее содержании и убеждают его прочесть статью целиком [Hyland, 2000]. Кроме того, аннотации представляют собой легкодоступный законченный текст научного стиля, который удобен для сравнительного анализа [McCarthy et al., 2007]. У российских исследователей довольно высок интерес к аннотациям, что можно также объяснить значительными расхождениями в их построении на русском и английском языках [Добрынина 2017; 2018; 2019; Тивьяева, Кузнецова, 2020; Толстова, 2020; Шпит, Куровский, 2020; Zanina, 2017]. Тем не менее очевидно, что полные тексты статей могут дать гораздо больше пищи для размышлений об уровне сформированности навыков письменной научной коммуникации [Кочетова, Ильинова, 2020;

Kochetova, Kononova, 2018; Shpit, McCarthy, 2022; Terenin, 2013; Zanina, 2016].

Прежде чем рассматривать отличительные признаки научного стиля в двух языках (английском и русском), остановимся на их общих характеристиках. Главными отличительными чертами научного функционального стиля в любом языке являются отвлеченно-обобщенность, подчеркнутая логичность и терминологичность [Колесникова, 2002]. При этом также важны объективность, точность, ясность и некатегоричность [Кожина, 1968; Колесникова, 2002; 2010; Котурова, 2009; Чернявская, 2006], отражающие надкультурность научной коммуникации. Кроме того, можно обозначить и общие экстралингвистические признаки научного дискурса: общая для дискурсивного сообщества картина мира, участники коммуникации, социальная ситуация, ценности и установки участников сообщества и др. [Ван Дейк, 1998]. Очевидно, что для исследователя, стремящегося опубликовать свои результаты в престижном международном журнале, недостаточно знать только характеристики научного стиля. Необходимо также учитывать нормы и ожидания дискурсивного сообщества, которые лингвистически проявляются в применении дискурсивных формул сообщества.

Нормы научной коммуникации в русском языке имеют первостепенную важность, поскольку авторы из России, прежде всего, вовлечены в устную и письменную практику научной коммуникации на родном языке. Среди риторических норм выделяют, например, большое количество существительных и прилагательных (и связанную с ними активность родительного падежа), склонность к сложным предложениям с большим количеством вложенных конструкций (причастных/деепричастных оборотов, придаточных предложений и длинных определительных фраз), применение устойчивых лексических и фразеологических единиц, а также избегание использования местоимений *я* и *мы* [Прядильникова, 2016; Трофимова, 2004]. О.А. Лаптева [1995] отмечает широкое употребление наречно-предикативных слов, кратких страдательных причастий прошедшего времени и воз-

вратных форм на -ся, -сь, а также редуцирование голоса автора. Она также подчеркивает тенденции к повышению значимости существительных через номинализацию глаголов, к упрощению синтаксической структуры предложения через применение цепочек родительных падежей существительных для ограничения объема понятия, а также применение предложно-падежных оборотов. Многолетняя практика написания научных текстов на русском языке в соответствии с указанными нормами, их постоянное пополнение и неизбежное многократное обращение к ним молодых исследователей приводят к укоренению таких образцов речи в памяти читателей и авторов новых текстов.

Нормы научной коммуникации в английском языке в настоящее время во многом опираются на английский язык как *lingua franca*, т.е. как единый языковой код для осуществления обмена научными знаниями в международной среде [Bhabha, 2004; Bhatia, 2002; Hyland, 2011; Tardy, 2004]. Англосаксонские традиции письма, а именно строгость, лаконичность и информативность, вполне отвечают духу информационной эры и глобализованной реальности. К. Хайленд [Hyland, 2011] отмечает следующие характерные черты письма англоязычных авторов: оно имеет более выраженную структуру и цели; авторы часто повторяют основные утверждения исследования; текст включает много свежих ссылок; риторика текста избегает отступлений от основной мысли; авторы избегают прямых утверждений и стремятся их смягчить, используя хеджирующие средства; письмо имеет множество соединительных элементов, чтобы ясно показать связь между частями абзаца или текста. Некоторые рекомендации по формулированию идей в научном тексте можно найти в обучающих курсах для авторов не-носителей английского языка. Например, композиционные советы от Springer Nature³ включают учет ожиданий читателя увидеть определенного типа ин-

формацию в определенном разделе статьи, части раздела, абзаце и даже части предложения. С точки зрения риторики разработчики курса, например, рекомендуют использовать короткие предложения (20–25 слов, «одна мысль на одно предложение»); где возможно, использовать активный залог вместо пассивного; быть как можно более лаконичным и избегать использования ненужных и общих слов. При этом важным является соблюсти баланс между чрезмерной академичностью языка и просторечием и донести основные результаты своего исследования не только до целевой, предметной аудитории, но и до более широкой аудитории из смежных дисциплин [Graff, 2000]. Много рекомендаций дается на сайте American Psychological Association⁴, начиная с формальных требований, утвержденных для стиля APA, и заканчивая общими советами по механике текста. И наконец, каждое издательство публикует требования к рукописям, в которых даны рекомендации, касающиеся не только оформления текста, но и языка изложения. Хотя, предполагается, что авторы читают подобные рекомендации перед подачей рукописи в печать, довольно большой процент рукописей возвращается (например: [Belcher, 2007; McKercher et al., 2007]). Интересно, но замечания по фактическим лингвистическим ошибкам при возврате рукописей составляют меньший процент, чем замечания, касающиеся риторики (например, отсутствие ясности письма, сложность понимания логики изложения).

Общий **сопоставительный дискурс-анализ** традиций и норм в международном и русском научном письме в общественных и гуманитарных науках был проведен И.Б. Короткиной [2018]. Автор описала основные отличия по 10 аспектам, начиная с традиций в построении и содержании разных разделов статьи и заканчивая общим стилем и языком письма. По мнению исследователя, стилистически международное письмо отличается последовательностью,

³ Springer Nature. Author tutorials. Writing in English: краткий обучающий курс для авторов. URL: <https://www.springer-nature.com/gp/authors/campaigns/writing-in-english> (дата обращения: 28.05.2022).

⁴ APA style. Style and Grammar Guidelines: руководство по написанию научных текстов на английском языке в стиле APA. URL: <https://apastyle.apa.org/style-grammar-guidelines/abbreviations/definition> (дата обращения: 28.05.2022).

ясностью и убедительностью аргументации, выстроенной с учетом широкой по специализации аудитории и различных точек зрения. Письмо российских авторов, в свою очередь, отличается многословием и излишним академизмом (перегруженностью терминологией и формальными фразами). С точки зрения языка международные тексты компактны и удобочитаемы, с ограниченным количеством номинализации и пассивных конструкций. Тогда как письмо от русскоязычных авторов зачастую трудно воспринимается, изобилует номинализацией, безличность создает двусмысленность и синтаксическую сложность. Если вспомнить упомянутые ранее языковые конструкции, рекомендуемые к использованию в научном письме на русском языке, можно с легкостью представить себе возможные переводы таких конструкций на английский язык, выполненные с помощью словарей и переводчиков, и понять, что имеет в виду И.Б. Короткина под излишней номинализацией, многословием, двусмысленностью и синтаксической сложностью.

Приводимые далее примеры исследований российских ученых раскрывают отличия двух типов письма как проявления различных компонентов иноязычной письменной научной компетенции: лингвистического, дискурсивного, стратегического, социокультурного и т.д. [Кудряшова, 2007; Соловова, 2003; Хуторской, 2003]. При сравнении образцов применения местоимения *we* в академическом тексте [Kochetova, Kononova, 2018] было обнаружено, что русскоязычные авторы склонны прибегать к инклюзивному *we* для того, чтобы убедить читателя разделить или принять точку зрения автора, тогда как англоязычные авторы используют инклюзивное *we* для проявления солидарности на основе общих знаний, убеждений и мнений (прагматический компонент). Результаты исследования устойчивых словосочетаний в научном дискурсе [Kochetova, Ильинова, 2020] показали, что русскоязычные авторы применяют значительно больше фраз с абстрактными существительными и фраз, нетипичных для англоязычного корпуса; и значительно меньше

фраз с последующими придаточными или глагольными конструкциями, а также элементов хеджирования. Все эти наблюдения авторы связывают с явлениями языкового перехода (*code-switching*) и языкового смешения (*code-mixing*) (стратегический и социокультурный компоненты). Результаты изучения способов хеджирования в англоязычных и русскоязычных статьях по менеджменту и маркетингу [Zanina, 2016] показали, что авторы русскоязычных статей гораздо реже прибегают к таким средствам, чем их англоязычные коллеги, что отражает недостаточно развитые социокультурные компетенции студентов. Анализ опубликованных аннотаций по прикладной лингвистике [Толстова, 2020] показал, что тексты, написанные русскоязычными авторами, меньше по объему, имеют более сокращенную структуру и отличаются меньшим лексическим разнообразием (лингвистический и дискурсивный компоненты). Жанровые отличия аннотаций были изучены на примере лингвистики [Тивьяева, Кузнецова, 2020] и менеджмента [Zanina, 2017] и доказывают, что в разных дисциплинах элементы содержания аннотаций заметно отличаются (дискурсивный компонент). В дополнение, исследование Н.К. Рябцевой [2018] позволило обозначить основные отличия между заголовками научных статей, написанных русскоязычными и международными авторами, и сформулировать основные принципы перевода названия статей на английский язык (социокультурный компонент).

Преимущества студенческого корпуса. Анализ опубликованных англоязычных текстов, написанных авторами, для которых родным языком является русский, несомненно, дает бесценный материал для обновления и пополнения содержания обучения. Однако мы считаем эти тексты «чистым» вариантом письма, т.е. они либо написаны автором с достаточно высоким уровнем английского языка и большим опытом написания академических текстов, либо прошли довольно серьезную проверку посредниками по проверке грамотности текста (*literacy brokers*) [Lillis, Curry, 2006]. На наш взгляд, гораздо больше инфор-

мации для целей формирования навыка можно получить из текстов, создаваемых самими авторами на начальном этапе работы над рукописью, или текстов студентов в процессе их обучения навыкам письменной научной коммуникации [Paquot, 2005; Flowerdew, 2001]. С другой стороны, мы допускаем, что такие тексты могут содержать грамматические, орфографические и пунктуационные ошибки, способные несколько исказить результаты статистического анализа. Тем не менее мы выбираем этот путь и рассматриваем письмо начинающих авторов в рамках студенческого (учебного, ученического) корпуса.

Под «студенческим корпусом» понимается корпус работ, написанных начинающими авторами. Эти корпуса также могут быть общими, как, например, Международный корпус для изучающих английский язык (International Corpus of Learner English, ICLE)⁵, и специальными, как, например, корпус студенческих текстов в представлениях будущих научных исследований (research proposals) [Smirnova, 2019; Smirnova, Strinyuk, 2020], корпус студенческих рукописей научных статей [Шпит, Куровский, 2021] и корпус студенческих описаний дипломных проектов по менеджменту [Виноградова и др., 2020]. Исследование [Smirnova, 2019] обнаружило гораздо меньшее обращение русскоязычных студентов к анафорическим выражениям и проблемы с их применением (повторяемость слов/фраз, несогласование по числу и др.). Другое исследование этого же корпуса [Smirnova, Strinyuk, 2020] показало, что студенты гораздо реже и менее корректно применяют средства хеджирования. Исследование [Шпит, Куровский, 2020] аналогично [Толстова, 2020] обнаружило ограниченное разнообразие лексического наполнения аннотаций к научным статьям. Помимо этого, было замечено, что письмо российских начинающих авторов в техниче-

ской области отличается чрезмерным использованием пассивного залога и предложных фраз, а также более длинными предложениями и более сложным синтаксисом. В работе [Виноградова и др., 2020] было обнаружено, что студенческие тексты уступают профессиональным по клаузуальной и фразовой сложности, особенно с точки зрения их разнообразия. Однако предпочтение *of*-фраз в письме русскоязычных студентов значительно выше [Виноградова и др., 2020; Шпит, Куровский, 2020]. Сравнительный корпусный анализ в [Shpit, McCarthy, 2022] выявил множество расхождений в письме начинающих русскоязычных инженеров-исследователей и опытных международных авторов, среди которых значительно большее количество предложных и именных фраз, значительно меньшее количество глаголов и глагольных фраз.

Рассмотренные выше исследования доказывают, что существует множество различий в риторике научной прозы авторов из России и представителей международного научного сообщества. Причем в большинстве упомянутые исследования базировались на ограниченной категории авторов, это позволяет утверждать, что обнаруженные признаки выявлены относительно соответствующей категории дискурсивного сообщества (чаще всего дисциплинарного). Следовательно, обучение письму должно включать теоретический и практический материал, ориентированный на формирование у студентов осведомленности о риторических нормах, принятых в соответствующем дисциплинарном сообществе, и развитие навыков применения этих риторических норм. У студентов должно складываться четкое понимание того, что существуют значительные различия в том, каким образом формулируются высказывания в научных текстах для русского читателя и для международного. При этом применение стратегий и норм научного стиля родного языка в текстах на английском языке неминуемо приведет к значительным отклонениям от традиций и ожиданий международных дискурсивных сообществ и, возможно, нарушениям в ясности и удобочитаемости текста в целом.

⁵ International Corpus of Learner English / S. Granger, M. Dupont, F. Meunier [et al.]; Presses universitaires de Louvain, Version 3, 2020. 277 p. URL: https://www.new.enl.uth.gr/flt/wp-content/uploads/sites/8/2021/02/ICLEv3_manual.pdf (дата обращения: 07.03.2022).

Методологию данного исследования составляет применение инструмента компьютерного текстового анализа Gramulator, который позволяет получить не только количественные, но и качественные результаты [McCarthy et al., 2012b]. Подход к текстовому анализу в Gramulator основывается на комбинировании двух методов: сравнительного корпусного анализа (Contrastive Corpus Analysis – CCA) и машинного определения дифференциалов (Machine Differential Diagnosis – MDD). CCA рассматривается как концептуальная схема подхода, а MDD – как практическое применение этого подхода. Совместно эти методы создают выходной результат инструмента – наборы n-грамм. Данное исследование опирается на дифференцирующие n-граммы (differentials), т.е. сочетания слов, типичные для одного корпуса и нетипичные для другого. При анализе дифференциалов были выявлены гендерные различия в новостном дискурсе [Wen et al., 2013], различия между правдивым и лживым дискурсом [McCarthy et al., 2012a], а также различия между научными дискурсами корейских ученых и их американских коллег [Min, McCarthy, 2010; 2013]. Помимо n-грамм, инструмент имеет множество других функций текстового анализа, а также встроенные одновыборочный тест Стьюдента (t-test) и точный тест Фишера (Fisher’s exact test), которые позволяют быстро определять статистическую значимость результатов.

Специальные корпуса создавались в интересах одной из кафедр радиотехнического факультета Томского университета систем управления и радиоэлектроники (ТУСУР). Начинающие авторы кафедры – это студенты магистратуры и аспирантуры, которые активно участвуют в реализации научных грантов РФФИ и РНФ и создают статьи для публикации в международных журналах и материалах конференций. Все студенты специализируются в области электромагнитной совместимости (ЭМС). Большинство из них владеют английским языком на уровне A2 – B1 (по шкале системы CEFR). В процессе обучения в университете никто из участников не проходил целенаправленной подготовки по научно-

му письму. Однако не исключена возможность самостоятельного овладения данным навыком, например при прохождении открытых онлайн-курсов, чтении специальных учебных пособий и рекомендаций для авторов от различных изданий или через активное чтение англоязычных статей по специальности.

Для целей исследования были созданы два корпуса: студенческий (СК) и экспертный (ЭК). Отличительными особенностями обоих корпусов являются их дисциплинарное и жанровое единство. Это означает, что все тексты относятся к ЭМС и представляют собой научные статьи. Однако студенческие тексты – это рукописи статей, написанные до подачи в международные журналы, а тексты экспертов, т.е. международных авторов, это опубликованные научные статьи по ЭМС. Студенческий корпус содержит 80 текстов от 23 авторов; некоторые студенты являются авторами нескольких работ, написанных в разное время. Общий объем корпуса составляет 142 954 словоупотреблений. Все тексты были написаны с 2018 до 2021 г. (начало). Средний объем (M) текстов – 1840 слов (стандартное отклонение (SD) 904). Тексты были собраны специалистом-лингвистом с кафедры иностранных языков университета, вычитывающим студенческие рукописи до подачи в журналы. Экспертный корпус включает 94 текста от 83 авторов (M = 3175 слов; SD = 1406). Общий объем составляет 295 704 словоупотребления. 55 текстов (59 %) написаны авторами, аффилированными с учреждениями англоязычных стран (США – 38, Великобритания – 10, Канада – 5, Австралия – 2). Остальные 39 статей (41 %) написаны авторами из 17 стран мира (например, Германия, Южная Корея, Швеция, Тайвань). Статьи русскоязычных авторов исключались. Превалирующее число текстов от авторов из англоязычных стран позволяет предположить, что данный дискурс, вероятнее всего, является прототипичным и грамматически корректным. Между тем достаточное количество текстов от авторов других стран мира обеспечивает вариативность языка, которая характерна для глобализованного научного сообщества. Все тексты от международных

авторов были собраны с помощью консультантов профилирующей кафедры. Хотя средний объем экспертного текста почти в два раза больше, в данном исследовании это не имеет критического значения, поскольку используемые статистические методы не зависят от размера текста, а берут в расчет соотношение количества измеряемых единиц к объему текста.

Чтобы инструмент Gramulator мог аккуратно обработать данные, тексты обоих корпусов были «очищены» от заголовков, подзаголовков, информации об авторах, благодарностей, списка литературы, списка условных обозначений, формул, таблиц, рисунков и приложений. Также были удалены опечатки, явные орфографические, грамматические и пунктуационные ошибки, кроме тех, удаление которых привело бы к нарушению авторского стиля. Мы стремились максимально сохранить язык автора, однако удалили то, что могло бы помешать компьютерной обработке. Для проведения анализа в Gramulator тексты были переведены в текстовый (txt) формат.

С помощью Gramulator корпуса были проанализированы на предмет выявления дифференциалов. Для извлечения дифференциалов была использована функция Process. По мере необходимости выполнялись одновыборочные тесты Стьюдента и точные тесты Фишера для определения статистической значимости результатов.

Результаты. На первом этапе был проведен сравнительный корпусный анализ, в результате которого было получено 1397 дифференцирующих биграмм для корпуса международных авторов и 595 для студенческого корпуса. Отличительной особенностью дифференциалов студенческого корпуса является то, что большинство из них представляют собой словосочетания, в которых оба слова или одно из слов относятся к специализированной лексике (например, *the decomposition, ultrashort pulse, protect against*). Это можно объяснить тем, что все тексты написаны студентами одной кафедры, исследовательская деятельность которой сконцентрирована на определенных направлениях. Напротив, корпус международных авто-

ров представляет набор текстов от ученых разных исследовательских центров, работающих в разных областях ЭМС. Такое внутривидовое разнообразие объясняет тот факт, что дифференциалы международного корпуса отличаются более общим лингвистическим характером и соотносительностью с устойчивыми речевыми образцами, характерными для академической речи (например, *such as, that is, in table, over the*). Учитывая то, что дифференциалы – это n-граммы, которые типичны для одного корпуса, но не типичны для другого, можно сделать вывод, что студенческий корпус не отличается активным применением устойчивых речевых образцов, характерных для данного дискурсивного сообщества. Далее были проанализированы несколько биграмм из первых 25 для каждого корпуса. Ограниченный набор биграмм прежде всего обусловлен размерами статьи. Кроме того, мы считаем, что самые первые, максимально часто используемые биграммы будут достаточно информативными для создания общей картины основных различий между дискурсами. Однако мы не исключаем, что анализ остальных позволил бы выявить другие значимые характеристики письма. Рассмотрение биграмм из экспертного корпуса позволит выявить риторические образцы, которые характерны для данного дисциплинарного сообщества, но недостаточно представлены в студенческом дискурсе.

Дифференциалы корпуса экспертов. Самой первой биграммой корпуса экспертов является *such as* (ЭК: 230 раз в 72 текстах, 76,60 %; СК: 12 раз в 9 текстах, 11,25 %; $p < .001$). Мы проверили оба корпуса на другие лингвистические средства, используемые для приведения примеров. Было обнаружено, что эксперты также значительно чаще применяют фразы *like* и *for instance* ($p < .001$), в то время как студенты чаще всего прибегают к использованию фразы *for example* (табл. 1). Данный пример может свидетельствовать об ограниченности дискурсивных навыков, проявляющейся в низком разнообразии языковых средств для связывания смыслового содержания.

Таблица 1

Использование фраз для приведения примеров

Table 1

Use of phrases for giving examples

N-граммы	СК (N, %)	ЭК (N, %)
such as	9 (11,25 %)	72 (76,60 %)***
like	6 (7,5 %)	37 (39,36 %)***
for instance	0	14 (14,89 %)***
as an example	4 (5 %)	11 (11,70 %)
for example	49 (61,25 %)	49 (52,13 %)

Примечание. N – количество текстов, % – их процент в корпусе.

Второй наиболее частой биграммой ЭК является фраза *as shown* (ЭК: 249 раз в 60 текстах, 63,83 %; СК: 1 раз, 1,25 %; $p < .001$). Наряду с *as illustrated* и *as depicted* эти фразы используются для привлечения внимания читателей к данным в таблице или на рисунке. В таком виде (*'as + причастие 2'*) эти фразы довольно редко используются в СК. Это наводит на мысль о некоторых жанровых различиях в дискурсе. Конструкция *'as + причастие 2'* является вводной конструкцией, которая дополняет основную информацию в предложении. Широкое употребление такой конструкции указывает на ее распространенность, т.е. прототипичность, в экспертном письме. Практически полное

отсутствие таких конструкций в студенческом письме может характеризовать жанровую особенность русского научного стиля в обращении с визуальными данными или отсутствие опыта в обращении к визуальной информации у студентов, поэтому требует дальнейших исследований. Кроме того, студенческие тексты имеют значительно меньше таких глаголов, как *demonstrate*, *illustrate*, *depict*, и их производных (табл. 2). В своих текстах студенты чаще всего полагаются на глагол *show* и его производные (*shows*, *showed*, *shown*, *showing*), что может свидетельствовать об ограниченности лексического разнообразия для передачи одной и той же мысли.

Таблица 2

Использование глаголов для обращения к визуальной информации

Table 2

Use of verbs for addressing the visuals

Униграммы	СК (N, %)	ЭК (N, %)
demonstrate (+ производные)	15 (20 %)	54 (57,45 %)***
illustrate (+ производные)	4 (5 %)	38 (40,43 %)***
depict (+ производные)	1 (1,25 %)	31 (32,98 %)***
show (+ производные)	79 (98,75 %)	90 (95,75 %)

Примечание. N – количество текстов, % – их процент в корпусе.

Биграмма *that is* является третьей, самой распространенной в экспертном письме (ЭК: 140 раз в 59 текстах, 62,77 %; СК: 9 раз в 9 текстах, 11,25 %; $p < .001$). Подобно *that are/were* и *that can/could*, *that is/was* вводит определяющее придаточное предложение (*restrictive clause*, *that*) и как таковое несколько отличается от неопределяющего придаточного (non-

restrictive clause, *'запятая + which'*), которое вводится *which is/was*, *which are/were*, *which can/could*. Как видно из табл. 3, студенты значительно реже используют словосочетания *'that + be/can'* в настоящем и прошедшем времени и *'which + can'* в прошедшем. Такие результаты могут происходить от слабого понимания различий между указанными двумя типами при-

даточных предложений в английском языке. В русском языке нет различий между этими видами придаточных предложений; и в том и в другом случае ставится 'запятая + который/

которая/которые/что'. Студенты, возможно, не обладая достаточными языковыми навыками, переносят синтаксические нормы из родного языка в английский текст.

Таблица 3

Биграммы, начинающие придаточные предложения

Table 3

Bigrams for relative clauses

Биграммы	СК (N, %)	ЭК (N, %)
that is	9 (11,25 %)	59 (62,77 %)***
that can	6 (7,5 %)	45 (47,87 %)***
that was	0	14 (14,89 %)***
that could	0	14 (14,89 %)***
which could	0	12 (12,77 %)***
that are	19 (23,75 %)	47 (50 %)***
that were	0	11 (11,70 %)***
which is	53 (66,25 %)	71 (75,53 %)***
which can	25 (31,25 %)	36 (38,30 %)***
which was	9 (11,25 %)	14 (14,89 %)***
which are	32 (40 %)	39 (41,49 %)***
which were	4 (5 %)	4 (4,26 %)***

Примечание. N – количество текстов, % – их процент в корпусе.

Еще одна биграмма, связанная с синтаксисом, это фраза *is that* (ЭК: 73 раза в 39 текстах, 41,49 %; СК: 7 раз в 7 текстах, 8,75 %; $p < .001$). Эта фраза обычно применяется для построения сложноподчиненного предложения с предикативным придаточным, которое объясняет свойство, прис-ваиваемое подлежащему/предмету (например, *limitation, advantage, benefit, requirement*).

Например:

The benefit of the approach **is that** *it is possible to protect equipment without additional embedded structures*. – Преимущество данного подхода **состоит в том, что есть возможность** защитить оборудование без добавления встроенных структур.

В студенческом научном письме такая идея чаще всего выглядит так:

The benefit of the approach **is the possibility of protection** without additional embedded structures. – Преимуществом данного подхода **является возможность** защиты оборудования без добавления встроенных структур.

То есть традиции русского научного стиля, а именно применение расширенной номиналь-

ной структуры вместо придаточного предложения, неопытными авторами часто переносятся в англоязычный письменный дискурс [Добрынина, 2018; 2019; Korotkina, 2018]. Забегая вперед, отметим, что биграмма *possibility of* является одной из самых частых в студенческом корпусе (СК: 72 раза в 34 текстах, 42,50 %; ЭК: 10 раз в 8 текстах, 8,51 %; $p < .001$).

Еще одной значимой биграммой, которая дифференцирует экспертное письмо от студенческого, является *could be* (ЭК: 122 раза в 42 текстах, 44, 68 %; СК: 1 раз, 1,25 %; $p < .001$). Мы изучили другие модальные глаголы, глаголы и наречия, которые используются для хеджирования, и обнаружили, что студенты значительно реже прибегают к применению этих языковых средств. Хеджирование – умение быть менее категоричным в утверждениях, демонстрировать вежливость – является общепринятой нормой в международном научном письме. Самым распространенным средством хеджирования в обоих корпусах является модальный глагол *can* (ЭК: 1375 раз в 94 текстах, 100 %; СК: 573 раза в 76 текстах, 95 %; $p = .043$). Кроме того,

оба корпуса отличаются частым использованием глаголов *seem* и *assume* и их производных (табл. 4). Однако использование других средств недооценено студентами. Такая ситуация мо-

жет быть вызвана слабым пониманием студентами важности хеджирования, т.е. ограниченностью социокультурных навыков, что согласуется с [Smirnova, 2019; Zanina, 2016].

Таблица 4

Средства хеджирования

Table 4

Hedging devices

Глаголы и наречия	СК (N, %)	ЭК (N, %)
may	23 (28,75 %)	71 (75,53 %)***
could	4 (5 %)	59 (62,77 %)***
indicate/indicates	7 (8,75 %)	33 (35,11 %)***
might	2 (2,5 %)	29 (30,85 %)***
likely	0	20 (21,27 %)***
suggest/suggests	5 (6,25 %)	22 (23,4 %)**
tend/tends	3 (3,75 %)	16 (17,02 %)**
possibly	0	10 (10,63 %)**
can	76 (95 %)	94 (100 %)*
imply/implies	4 (5 %)	14 (14,89 %)*
unlikely	0	7 (7,45 %)*
seem/seems	10 (12,5 %)	15 (15,96 %)
probably	2 (2,5 %)	7 (7,45 %)

Примечание. N – количество текстов, % – их процент в корпусе.

Дифференциалы студенческого корпуса. Как уже говорилось выше, одной из самых распространенных биграмм СК является фраза *possibility of* (СК: 72 раза в 34 текстах, 42,5 %; ЭК: 10 раз в 8 текстах, 8,51 %; $p < .001$). Так как *possibility of* предполагает после себя именную или герундиальную фразу, можно предположить, что это вариант номинализации придаточного предложения (как было упомянуто выше). Альтернативными вариантами в этом случае могут быть предложения с *possible that/possible to, allow/enable/provide/make*

possible или *can/could/may/might*. Как видно из табл. 6, биграмма *possible that* является нетипичной для обоих корпусов. Что касается глаголов, то студенты значительно чаще используют *make possible* ($p < .001$) и *allow* ($p = .02$) и реже – *enable* ($p < .001$) и *provide* ($p = .017$). Как упоминалось выше, опытные авторы довольно часто прибегают к применению модальных глаголов при большем их разнообразии. Можно сделать вывод, что студенты обладают гораздо менее разнообразным лексическим и синтаксическим арсеналом.

Таблица 5

Способы выражения 'возможности'

Table 5

Expressing 'possibility'

N-граммы 'возможности'	СК (N, %)	ЭК (N, %)
make possible	21 (26,25 %)***	3 (3,19 %)
enable (+ производные)	2 (2,5 %)	25 (26,6 %)***
provide (+ производные)	50 (62,5 %)	75 (79,79 %)*
allow (+ производные)	56 (70 %)	49 (52,13 %)*
possible to	48 (60 %)**	33 (35,11 %)
possible that	0	4 (4,26 %)

Примечание. N – количество текстов, % – их процент в корпусе.

Следующей значимой биграммой в студенческом письме является фраза *the obtained* (СК: 73 раза в 36 текстах, 45 %; ЭК: 12 раз в 6 текстах, 6, 38 %; $p < .001$). *The obtained*, вполне вероятно, является частью триграмм *the obtained results*, *the obtained waveforms*, *the obtained expressions* и других. Частое использование таких фраз может означать, что студенты придают особое значение уточнению предмета повествования (табл. 7). Напротив, опытные авторы из всего перечисленного часто прибегают только к таким фразам, как *under test* ($p < .001$) и *the proposed* ($p = .718$). Употребление начинающими авторами явных механизмов ссылки и уточ-

нения можно объяснить их стремлением проявить сильную связность (когезию) между элементами повествования. Однако это также может означать, что они недооценивают основополагающее значение артикля *the* в английском языке; этот артикль сам по себе является указанием на что-то упомянутое выше, и дополнительные механизмы ссылки могут быть избыточными. Прослеживается низкий уровень знаний об этом явлении в английской грамматике, что может сигнализировать о возможной недостаточной развитости навыков применения и других языковых явлений, отсутствующих в русской грамматике.

Таблица 6

Механизмы ссылки/уточнения

Table 6

Referencing and specifying devices

Биграммы ссылки	СК (N, %)	ЭК (N, %)
<i>the obtained</i>	36 (45 %)***	6 (6,38 %)
<i>under study</i>	20 (25 %)***	5 (5,32 %)
<i>the investigated</i>	16 (20 %)***	3 (3,19 %)
<i>under investigation</i>	10 (12,5 %)***	1 (1,06 %)
<i>the developed</i>	12 (15 %)*	5 (5,32 %)
<i>under test</i>	3 (3,75 %)	22 (23,4 %)***
<i>the proposed</i>	17 (21,25 %)	23 (24,47 %)

Примечание. N – количество текстов, % – их процент в корпусе.

Биграммы *is advisable* и *advisable to* используются практически в третьей части студенческих текстов (*is advisable* – СК: 36 раз в 26 текстах, 32,5 %; ЭК: ни разу; $p < .001$; *advisable to* – СК: 37 раз в 27 текстах, 33,75 %; ЭК: ни разу; $p < .001$) и очень часто образуют триграмму *is advisable to*. Эта фраза также ни разу не встречается в экспертном письме (СК: 35 раз в 25 текстах, 31,25 %; ЭК: ни разу; $p < .001$). Согласно смыслу *is advisable to* мы можем предположить, что начинающие авторы стремятся выразить совет, рекомендацию или целесообразность чего-либо. Аналогичным смыслом обладают и другие слова, например *desirable*, *appropriate*, *recommended*, и модальный глагол *should*. Как видно из табл. 8, студенты имеют представление о различных способах выражения желательности чего-либо, но применяют их гораздо реже, чем опытные авторы.

Это может означать, что у студентов, возможно, не было необходимости всесторонне рассуждать по поводу предмета исследования или их исследовательский и/или лингвистический опыт не позволил им сделать это. Кроме того, студенты избегают активного применения модального глагола *should*, что может свидетельствовать об их стремлении выдержать официоз русского научного стиля.

Еще одной распространенной биграммой студенческого письма является фраза *dependence of* (СК: 44 раза в 27 текстах, 33,75 %; ЭК: 4 раза в 2 текстах, 2,13 %; $p < .001$). Эта биграмма нередко является частью триграммы *the dependence of*, частота которой также значительно выше в студенческих текстах (табл. 9). Помимо терминов (например, *frequency dependence*), эта фраза применяется, чтобы описать, как два

Таблица 7

Выражение совета и целесообразности

Table 7

Devices for giving advice

Униграммы	СК (N, %)	ЭК (N, %)
advisable	30 (37,5 %)***	0
should	26 (32,5 %)	68 (72,34 %)***
appropriate	10 (12,5 %)	32 (34,04 %)**
recommended	4 (5 %)	18 (19,15 %)**
desirable	6 (7,5 %)	14 (14,89 %)

Примечание. N – количество текстов, % – их процент в корпусе.

и более предмета взаимосвязаны, и как таковая должна быть типичной для всех научных текстов. Вопрос в том, что эксперты используют для этих целей. Мы предполагаем, что биграмма *dependence of* является синонимичной *relationship between* и *function of* в контексте дисциплин технической направленности. Поэтому мы проверили наличие этих фраз в обоих корпусах и обнаружили, что опытные международные авторы используют их значительно чаще. Тогда как *dependence of* применяется ими исключительно в терминологии (*frequency dependence*,

temperature dependence). Это свидетельствует о том, что студенты не видят семантической разницы между словом 'зависимость' в двух случаях и поэтому используют знакомое им слово 'dependence' оба раза. Мы также выявили, что начинающие авторы значительно чаще используют существительное *dependence(s)*, тогда как опытные авторы намного чаще используют прилагательные (*in*)*dependent* и глагол *depend* (и его производные). Такие результаты, к сожалению, также говорят о довольно ограниченном разнообразии риторических средств у студентов.

Таблица 8

N-граммы со смыслом 'зависеть/зависимый/зависимость'

Table 8

N-grams meaning 'relationship/dependence'

N-граммы	СК (N, %)	ЭК (N, %)
dependence(s)	39 (48,75 %)***	10 (10,64 %)
dependence of	27 (33,75 %)***	2 (2,13 %)
the dependence of	16 (20 %)***	1 (1,06 %)
(in) <i>dependent</i>	6 (7,5 %)	38 (40,43 %)***
function of	6 (7,5 %)	36 (38,3 %)***
relationship between	1 (1,25 %)	12 (12,77 %)***
depend (+ производные)	30 (37,5 %)	54 (57,45 %)*

Примечание. N – количество текстов, % – их процент в корпусе.

Таким образом, проведенный анализ позволил обнаружить множество расхождений в риторическом выборе начинающих русскоязычных авторов, что в итоге может означать, что их англоязычные тексты расходятся с нормами и ожиданиями международного дискурсивного сообщества. Множество примеров говорит о недостаточной сформированности лингвистических навыков (например, по грамматическим явлениям,

отсутствующим в русском языке). Проблемы с языковыми навыками тесно связаны и с другими навыками, анализ которых показал не столько их отсутствие, сколько ограниченность. Если предположить, что студенческие рукописи являются автоматизированным переводом предварительно составленных русских текстов, то можно предположить, что авторы не видят различий в стилистических нормах в двух языках и/или

не подключают (или не имеют) навыки планирования, обдумывания и перефразирования сформулированных мыслей в формат, принятый в международном письме, перед их переводом. Поэтому их риторический выбор можно рассматривать как языковую интерференцию [Добрынина, 2018; 2019; Кочетова, Ильинова, 2020; Korotkina, 2018; Terenin, 2013].

Данное исследование имеет ряд **ограничений**. Во-первых, статистический анализ в любом инструменте позволяет проанализировать письмо с точки зрения лингвистических средств, использованных в нем, поэтому исключается прямое измерение его экстралингвистических характеристик. Кроме того, в таких инструментах невозможно измерить развитость навыков самоконтроля, интеллектуальную активность, мотивированность и т.д., которые являются составной частью стратегических навыков. Следовательно, невозможно в полной мере и с полной уверенностью говорить об уровне сформированности различных навыков коммуникативной компетенции. Однако индуктивный анализ лингвистического выбора может создать некий образ основных побуждений и ЗУН авторов, приведших к такому выбору.

Во-вторых, компьютерный анализ строился на текстах из одной технической дисциплины и с применением студенческих работ одной кафедры вуза. С уверенностью можно сказать, что корпусный анализ более широкого спектра образцов письма или текстов из других дисциплин может дать иные результаты (n-граммы). Кроме того, мы рассмотрели только несколько n-грамм; анализ большего количества фраз может расширить спектр конкретных лингвистических проблем и их возможные экстралингвистические причины. Однако в качестве проверки концепции выбранные примеры показали, что их анализ позволяет раскрыть проблемы, касающиеся разных компонентов письменной компетенции.

В-третьих, тексты из студенческого корпуса принадлежат авторам, которые не проходили никакого формального обучения англоязычному академическому письму. Поэтому сопоставительный корпусный анализ мог выявить особенности

письма, которые неактуальны для тех, кто имеет хорошие языковые навыки и ясное представление об англоязычном научном стиле. Однако, поскольку практика обучения научному письму до сих пор не является обязательной в российских вузах, мы можем предположить, что социокультурные, дисциплинарные и индивидуальные факторы, лежащие в основе выявленных особенностей, могут рассматриваться общими для большинства русскоязычных авторов инженерных направлений. Кроме того, для любой категории авторов будет полезным понимание того, как их письмо соотносится с письмом целевого международного дискурсивного сообщества.

И последнее, анализ в Gramulator не предполагает автоматизированную оценку индивидуальной работы и поэтому не может применяться в ежедневной практике обучения письму. Однако этот инструмент можно вполне применять для выявления динамики развития компетенции группы учащихся на основе сопоставительного анализа с их ранними работами. Кроме того, можно применять возможности Conco dancer, чтобы выявить частоту употребления конкретных слов и словосочетаний.

Педагогическая значимость. Полученные в ходе исследования результаты позволяют утверждать, что применение инструментов компьютерной лингвистики может расширить арсенал инструкторов письма, заинтересованных в повышении осведомленности о наиболее значимых различиях между письмом студентов и их целевым дискурсивным сообществом. Четкое понимание этих различий может придать большую осмысленность процессу учения и обучения научному письму, поскольку обнаруженные биграммы служат маркерами определенных лексико-грамматических средств, которые авторы выбирают при создании своих текстов на английском языке. Эти маркеры позволяют диагностировать возможные коммуникативные проблемы анализируемой группы, а понимание их истоков позволяет скорректировать содержание и методы обучения навыкам написания научных текстов на английском языке. Суть корректировки сводится к следующим трем

направлениям: *совершенствование лингвистических навыков, формирование когнитивно-коммуникативных навыков и развитие навыков критического анализа.*

1. Поскольку уровень владения иностранным языком лежит в основе всей письменной коммуникативной компетенции, необходимо постоянно *совершенствовать лингвистические навыки.* Применяя традиционные (объяснительно-иллюстративные и репродуктивные) и современные (поисково-исследовательские, проблемно-деятельностные, развивающие и т.д.) методы, подходы и технологии, необходимо стремиться к формированию у студентов багажа лексики с акцентом на разнообразии и синонимичности, а также знаний и умений применять различные синтаксические структуры. Кроме того, необходимо фокусировать внимание на явлениях, отсутствующих в русском языке (например, артикли, герундии, существительные в препозитивном положении, определяющие/неопределяющие придаточные предложения) и добиваться умения оперировать этими знаниями при формулировании мысли на английском языке. В качестве положительных примеров необходимо использовать образцы риторического выбора из менторских текстов (опубликованные научные статьи на английском языке, демонстрирующие высокий уровень языковой грамотности, разнообразия и соответствия дискурсивному сообществу). Для анализа некорректного риторического выбора можно брать примеры из студенческих текстов.

2. *Формирование когнитивно-коммуникативных навыков* должно происходить параллельно с совершенствованием языковых навыков, когда каждый пример языкового материала сопровождается разъяснением его значимости с точки зрения соотносительности с нормами и ожиданиями дискурсивного сообщества. Применяя объяснительно-иллюстративные и поисково-исследовательские методы, необходимо добиваться понимания студентами существенных социокультурных различий между научным стилем изложения в русском языке и английском языке международного научного общения

(например, хеджирование, излишняя номинализация). Помимо этого, в процессе обучения необходимо активно применять информационные ресурсы для работы над англоязычным текстом, тем самым повышая уровень информационной грамотности учащихся.

3. Особое внимание необходимо уделять развитию личностных качеств, а именно *навыков критического мышления.* Используя различные приемы и методы анализа и самоанализа, контроля и самоконтроля, оценки и самооценки (например, при двойном переводе, взаимном рецензировании, пред- и постпереводческой обработке текста, а также при анализе собственной работы с учетом изученных явлений), можно значительно усилить стратегические навыки и совместно с информационной грамотностью сформировать навыки автономии при работе с иноязычным текстом. Эти умения могут значительно повысить уровень мотивации и уверенности в собственных возможностях создавать научные тексты на английском языке и публиковать их в международных рецензируемых журналах. Последнее (уровень уверенности в собственных возможностях) хорошо проявилось в результатах анкетирования участников формирующего эксперимента. До начала курса только 2 человека из 24 оценили свою уверенность от 8 до 10 баллов, тогда как после прохождения курса таких было 16 человек.

Можно заметить, что данное исследование не только подтвердило изложенные в теоретическом обосновании выводы российских теоретиков и практиков АП, но и предоставило конкретные примеры самых распространенных особенностей в письме начинающих авторов инженерных направлений. Причем эти особенности представляют собой самые частотные реальные примеры, извлеченные из письма студентов, а не потенциальные проблемные явления, взятые из сторонних источников и затем проверенные в имеющемся корпусе. Соглашаясь со всеми упомянутыми исследователями, хотелось бы повториться, что научному письму необходимо обучать; это теоретическая и одновременно практическая деятельность, которая требует длительного,

поэтапного освоения и постоянной практики. Такая деятельность должна включать не только универсальные знания, умения и навыки, формируемые традиционным подходом к обучению АП и характерные для любой страны и любой дисциплины, но и те ЗУН, которые обусловлены социокультурными и дисциплинарными факторами, характерными для определенной категории обучающихся и формирующиеся при реализации дискурс-аналитического подхода.

Заключение. В данном исследовании представлена концепция применения инструмента компьютерной лингвистики для реализации дискурс-аналитического подхода в обучении навыкам письменной научной коммуникации на ан-

глийском языке. Предлагаемая технология позволяет обнаружить и проанализировать конкретные примеры лингвистического выбора определенной категории учащихся в сопоставлении с целевым дискурсивным сообществом. Применение этих знаний помогает скорректировать процесс обучения АП, придавая ему большую осознанность при выборе риторических средств и формулировании высказываний на английском языке, а также большую сфокусированность на реальных языковых и коммуникативных потребностях учащихся. В сочетании с обучением универсальным ЗУН этот подход придает большую осмысленность и определенность процессу обучения научному письму на английском языке.

Библиографический список

1. Базанова Е.М., Короткина И.Б. Российский консорциум центров письма // Высшее образование в России. 2017. № 4 (211). С. 50–57.
2. Бородик К.А., Дикусар К.С., Богатов В.В. Тренды публикационной активности российских исследователей за период 2016–2020 гг. по данным международных баз научного цитирования Web of Science Core Collection и Scopus // Управление наукой и наукометрия. 2021. Т. 16, № 4. С. 571–595. DOI: 10.33873/2686-6706.2021.16-4.571-595
3. Ван Дейк Т.А. К определению дискурса // Ideology: A Multidisciplinary Approach. Лондон: Сэйдж пабликэйшнс, 1998. 384 с. URL: <http://psyberlink.flogiston.ru/internet/bits/vandijk2.htm> (дата обращения: 01.08.2022).
4. Виноградова О.И., Смирнова Е.А., Виклова А.В., Пантелеева И.М. Синтаксическая сложность академического текста: корпусное исследование работ русскоязычных студентов, изучающих английский язык в сравнении с текстами профессиональных авторов // Вестник РГГУ. Сер.: Литературоведение. Языкознание. Культурология. 2020. № 7. С. 107–129. DOI: 10.28995/2686-7249-2020-7-107-129
5. Добрынина О.Л. Академическое письмо для публикационных целей: стилистические погрешности // Высшее образование в России. 2019. Т. 28, № 10. С. 38–49. DOI: 10.31992/0869-3617-2019-28-10-38-49
6. Добрынина О.Л. Грамматические ошибки в англоязычном академическом письме: причины появления и стратегии коррекции // Высшее образование в России. 2017. № 8/9. С. 100–107.
7. Добрынина О.Л. Проблемы англоязычного академического письма: лексические ошибки, причины их появления и стратегии коррекции // Высшее образование в России. 2018. Т. 27, № 10. С. 75–83. DOI: 10.31992/0869-3617-2018-27-10-75-83
8. Карасик В.И. О типах дискурса // Языковая личность: институциональный и персональный дискурс. Волгоград: Перемена, 2000. 220 с.
9. Кожина М.Н. К основаниям функциональной стилистики. Пермь: Изд-во ПГУ, 1968. 251с.
10. Колесникова Н.И., Ридная Ю.В. Жанровая модель научной статьи на русском и английском языках // Высшее образование в России. 2016. № 6 (202). С. 98–105.
10. Колесникова Н.И. От конспекта к диссертации: учеб. пособие по развитию навыков письменной речи. М.: Флинта: Наука, 2002. 288 с.

12. Колесникова Н.И. Что важно знать о языке и стиле научных текстов? Статья первая // Высшее образование в России. 2010. № 3. С. 130–137. URL: <https://cyberleninka.ru/article/n/chto-vazhno-znat-o-yazyke-i-stile-nauchnyh-tekstov/viewer> (дата обращения: 01.08.2022).
13. Комарова Ю.А. Функционально-лингвистическое описание англоязычного научного дискурса // Вестник СПбГУ. Язык и литература. 2012. № 4. С. 78–85.
14. Короткина И.Б. Академическая грамотность и методы глобальной научной коммуникации // Научный редактор и издатель. 2017. № 2 (1). С. 08–13. DOI: 10.24069/2542-0267-2017-1-8-13
15. Короткина И.Б. Академическое письмо: необходимость междисциплинарных исследований // Высшее образование в России. 2018а. Т. 27, № 10. С. 64–74. DOI: 10.31992/0869-3617-2018-27-10-64-74
16. Короткина И.Б. Английский язык для научно-публикационных целей как новое направление педагогических исследований // Отечественная и зарубежная педагогика. 2018б. Т. 1, № 4 (52). С. 115–130. URL: <https://cyberleninka.ru/article/n/angliyskiy-yazyk-dlya-nauchno-publikatsionnyh-tseley-kak-novoe-napravlenie-pedagogicheskikh-issledovaniy/viewer> (дата обращения: 01.08.2022).
17. Котюрова М.П. Тенденция к формированию новых явлений в научном стиле речи // Известия Южного федерального университета. Филологические науки. 2009. № 1. С. 77–86. URL: <https://philol-journal.sfedu.ru/index.php/sfuphilol/article/view/167> (дата обращения: 01.08.2022).
18. Кочетова Л.А., Ильинова Е.Ю. Англоязычный академический дискурс в транслингвальной ситуации: корпусно-ориентированное изучение лексических маркеров // Вестник Волгоградского государственного университета. Языкознание. 2020. Т. 19, № 5. С. 25–37. DOI: 10.15688/jvolsu2.2020.5.3
19. Кудряшова О.В. Компоненты коммуникативной компетенции при обучении письменной речи // Вестник ЮУрГУ. Сер.: Лингвистика. 2007. № 15 (87). URL: <https://cyberleninka.ru/article/n/komponenty-kommunikativnoy-kompetentsii-pri-obuchenii-pismennoy-rechi> (дата обращения: 01.08.2022).
20. Лаптева О.А. Как пишут ученые // Русская речь. 1995. № 2. С. 55–62.
21. Макаров М.Л. Основы теории дискурса. М.: ИТДГК «Гнозис», 2003. 280 с. URL: chrome-extension://efaidnbnmnibpcjpcglclefindmkaj/http://yanko.lib.ru/books/cultur/makarov-osnovu_teorii_diskursa-8l.pdf (дата обращения: 01.08.2022).
22. Мохначева Ю.В., Цветкова В.А. Динамика развития российского сегмента научных публикаций (по данным Web of Science Core Collection и Scopus) // Научные и технические библиотеки. 2021. № 6. С. 15–28. DOI: 10.33186/1027-3689-2021-6-15-28
23. Прядильникова Н.В. Практическая функциональная стилистика русского языка: учеб. пособие. Самара: Изд-во Самарского ун-та, 2016. Ч. 2. 164 с.
24. Рябцева Н.К. Название как доминантный компонент научного текста: русско-английские межъязыковые «несоответствия» // Вестник Волгоградского государственного университета. Языкознание. 2018. Т. 17, № 2. С. 33–43. DOI: 10.15688/jvolsu2.2018.2.4
25. Соловова Е.Н. ELT NEWS @ VIEWS // The Russian magazine for English language teachers. 2003. No. 2 (27). Динтернал International Book Distribution. С. 2–11.
26. Тивьяева И.В., Кузнецова Д.Л. Компаративный анализ структурно-содержательной организации аннотаций к русскоязычным и англоязычным научным статьям // Актуальные проблемы филологии и педагогической лингвистики. 2020. № 3. С. 139–152. DOI: 10.29025/2079-6021-2020-3-139-152
27. Толстова Т.В. Корпусный анализ жанровых конвенций в российских и зарубежных аннотациях к научным статьям // Вестник Московского университета. Сер. 19: Лингвистика и межкультурная коммуникация. 2020. № 1. С. 118–129.
28. Трофимова Г.К. Русский язык и культура речи: курс лекций. М.: Флинта, Наука, 2004. 160 с.

29. Хуторской А.В. Ключевые компетенции как компонент личностно-ориентированной парадигмы образования // Народное образование. 2003. № 2. С. 58–64.
30. Чернявская В.Е. Интерпретация научного текста. М.: КомКнига, 2006. 128 с.
31. Чернявская В.Е. Лингвистика текста. Лингвистика дискурса: учеб. пособие. М.: ЛЕНАНД, 2018. 200 с.
32. Шейпак С.А. Академическое письмо: критика традиционного подхода // Высшее образование в России. 2020. Т. 29, № 2. С. 92–103. DOI: 10.31992/0869-3617-2020-29-2-92-103
33. Шпит Е.И., Куровский В.Н. Изучение англоязычного академического письма инструментами компьютерной лингвистики // Высшее образование в России. 2020. Т. 29, № 7. С. 89–103. DOI: 10.31992/0869-3617-2020-29-7-89-103
34. Шпит Е.И., Куровский В.Н. Специализированный корпус для обучения академическому письму: создание и применение // Профессиональное образование: проблемы и достижения: матер. X Междунар. науч.-практ. конф., Томск, 17 декабря 2020 г. Томск: Том. гос. пед. ун-т, 2021. С. 161–166.
35. Becher T. Disciplinary discourse // Studies in Higher Education. 1987. Vol. 12 (3). P. 261–274. DOI: 10.1080/03075078712331378052
36. Becher T. The significance of disciplinary differences // Studies in Higher Education. 1994. Vol. 19 (2). P. 151–161. DOI: 10.1080/03075079412331382007
37. Belcher D.D. Seeking acceptance in an English-only research world // Journal of Second Language Writing. 2007. Vol. 16 (1). P. 1–22.
38. Bhabha H.K. The location of culture (Routledge Classics). 2nd ed. London: Routledge, 2004. 440 p.
39. Bhatia V., Flowerdew J., Jones R. Advances in discourse studies. Routledge Taylor and Francis Group. London and New York, 2008. 274 p.
40. Bhatia V.K. A generic view of academic discourse. Academic discourse. Edited by J. Flowerdew. London: Longman, 2002. P. 21–39.
41. Connor U. New directions in contrastive rhetoric // TESOL Quarterly. 2002. Vol. 36 (4). P. 493–510. DOI: 10.2307/3588238
42. Fairclough N. Analysing discourse: Textual analysis for social research. London: Routledge, 2003. 270 p.
43. Flowerdew J. English for research publication purposes. The handbook of English for specific purposes. Edited by B. Paltridge, S. Starfield. Malden, MA: Wiley-Blackwell, 2013. P. 301–321.
44. Flowerdew L. The exploitation of small learner corpora in EAP materials design: In M. Ghadessy & R. Roseberry (Eds.) Small Corpus Studies and ELT. Amsterdam & Philadelphia: Benjamins, 2001. P. 363–379.
45. Gee J.P. An introduction to discourse analysis: Theory and method. Routledge. Taylor and Francis Group. London and New York, 2005. 248 p.
46. Graff G. Scholars and sound bites: The myth of academic difficulty // PMLA (Publications of the Modern Language Association of America). 2000. Vol. 115 (5). P. 1041–1052.
47. Huckin T.N. Specificity in LSP // Ibérica, Revista de la Asociación Europea de Lenguas para Fines Específicos. 2003. No. 5. P. 3–18.
48. Hyland K. Academic discourse. The Bloomsbury Companion to Discourse Analysis. Edited by K. Hyland, B. Paltridge. London: Continuum, 2011. P. 171–184.
49. Hyland K. Disciplinary discourses: Social interactions in academic writing. Harlow, UK: Pearson Education Limited, Longman, 2000. 211 p.
50. Hyland K. Sympathy for the devil? A defence of EAP // Language Teaching. 2018. No. 3. P. 383–399. DOI: 10.1017/S0261444-818000101

51. Hyland K. Writing in the academy: Reputation, education and knowledge. London: Institute of Education Press, 2007. 28 p.
52. Kaplan R.B. Cultural thought patterns in intercultural education. *Language Learning*, 1966. 648 p. DOI: 10.1111/j.1467-1770.1966.tb00804.x
53. Kochetova L.A., Kononova I.V. Corpus-based contrastive study of discursive strategy of construing interpersonal relations in English language academic discourse // *J. Sib. Fed. Univ. Humanit. soc. sci.* 2018. P. 1–11. DOI: 10.17516/1997-1370-0353
54. Korotkina I.B. Russian scholarly publications in anglophone academic discourse: The clash of tyrannosaurs // *Integration of Education*. 2018. Vol. 22 (2). P. 311–323. DOI: 10.15507/1991-9468.091.022.201802.311-323
55. Kwan B.S.C. An investigation of instruction in research publishing in doctoral programs: The Hong Kong case // *Higher Education*. 2010. Vol. 59, No. 1. P. 55–68.
56. Lillis T., Curry M.J. Professional academic writing by multilingual scholars: Interactions with literacy brokers in the production of English-medium texts // *Written Communication*. 2006. Vol. 23 (1). P. 3–35. DOI: 10.1177/0741088305283754
57. McCarthy P., Lehenbauer B., Hall C., Duran N., Fujiwara Y., McNamara D. A coh-metrix analysis of discourse variation in the texts of Japanese, American, and British Scientists // *Foreign Languages for Specific Purposes*. 2007. Vol. 6. URL: https://www.researchgate.net/publication/242322281_A_Coh-Metrix_Analysis_of_Discourse_Variation_in_the_Texts_of_Japanese_American_and_British_Scientists (access date: 01.08.2022).
58. McCarthy P.M., Duran N.D., Booker L.M. The devil is in the details: New directions in deception analysis. In: *Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference, FLAIRS-25*. 2012a. P. 188–195.
59. McCarthy P.M., Watanabe S., Lamkin T.A. The Gramulator: A tool to identify differential linguistic features of correlative text types. In: P.M. McCarthy and C. Boonthum Denecke (Eds.) «Applied Natural Language Processing: Identification, Investigation and Resolution». 2012b. P. 312–333. Hershey, PA: IGI Global. DOI: 10.4018/978-1-60960-741-8.ch018
60. McKercher B., Law R., Weber K., Song H., Hsu C. Why referees reject manuscripts // *Journal of Hospitality & Tourism Research*. 2007. Vol. 31 (4). P. 455–470.
61. Min H.S., McCarthy P.M. Contrastive corpus analysis on the writing of American and Korean scientists using Gramulator // *Language Research*. 2013. Vol. 49 (3). P. 681–702.
62. Min H.S., McCarthy P.M. Identifying varieties in the discourse of American and Korean scientists: A contrastive corpus analysis using the Gramulator. In: *Proceedings of the 23rd International Florida Artificial Intelligence Research Society Conference, FLAIRS-23*. 2010. P. 247–252.
63. Paquot M. Towards a productively-oriented academic word list. In: Walinski J.; Kredens K.; Gozdz-Rozkowski S. *Practical Applications in Language and Computers*, 2005. Peter Lang: Frankfurt and Main, 2007. P. 127–140. URL: <http://hdl.handle.net/2078.1/76045>
64. Reid N. Getting published in international journals: Writing strategies for European social scientists. Professional Publications Press, 2018. 309 p.
65. Shpit E.I., McCarthy P.M. Addressing discourse differences in the writing of Russian engineering students and international researchers // *Language Teaching Research*. July. 2022. DOI: 10.1177/13621688221109809
66. Smirnova E., Strinyuk S. Hedges in Russian EAP writing: A corpus-based study of research papers in management // *Journal of English as a Lingua Franca*. 2020. Vol. 9 (1). P. 81–101. DOI: 10.1515/jelf-2020-2033
67. Smirnova E.A. Referential coherence of academic texts: a Corpus-Based analysis of L2 research papers in management // *Journal of Language and Education*. 2019. Vol. 5 (4). P. 112–127. DOI: 10.17323/jle.2019.9688

68. Swales J. Genre analysis: English in academic and research settings. Boston: Cambridge UP, 1990. 274 p.
69. Tardy C. The role of English in scientific communication: lingua franca or Tyrannosaurus rex? // Journal of English for Academic Purposes. 2004. Vol. 3 (3). P. 247–269. DOI: 10.1016/j.jeap.2003.10.001
70. Terenin A. Clear Writing as a problem of Russian learners of English // Middle-East Journal of Scientific Research. 2013. Vol. 16, No. 11. P. 1512–1519. DOI: 10.5829/idosi.mejsr.2013.16.11.12039
71. Wen X.A., McCarthy P.M., Strain A.C. Gramulator analysis of gendered language in cable news reportage. In: Proceedings of the 26th International Florida Artificial Intelligence Research Society Conference, FLAIRS. 2013. P. 243–248.
72. Widdowson H.G. Discourse Analysis. Oxford University Press, 2007. 152 p.
73. Zanina E. Move structure of research article abstracts on management: Contrastive study (the case of English and Russian) // Journal of Language and Education. 2017. Vol. 3 (2). P. 63–72. DOI: 10.17323/2411-7390-2017-3-2-63-72
74. Zanina E. Strategic hedging: a comparative study of methods, results and Discussion (and Conclusion) // Journal of Language and Education. 2016. Vol. 2 (2). P. 52–60. DOI: 10.17323/2411-7390-2016-2-2-52-60

SCIENCE WRITING IN ENGLISH: DIFFICULTIES FOR NOVICE RUSSIAN AUTHORS

E.I. Shpit (Tomsk, Russia)

V.N. Kurovskii (Tomsk, Russia)

Abstract

Statement of the problem. The paper discusses multiple deviations of science texts in the English language written by Russian novice writers from the norms and expectations of the international discourse community.

The purpose of the article is to propose a concept of utilizing computer linguistics tools of automated text analysis for pedagogical purposes to identify real needs of students in science writing in English.

The methodology is based on the contrastive discourse-analysis of two corpora: students' research paper manuscripts and published research papers by international expert writers. Both corpora belong to one engineering discipline and were compiled specifically for the research purposes. The contrastive corpus analysis was conducted by means of a computer linguistics tool Gramulator.

Research results. The quantitative and qualitative analysis of differential bigrams in Gramulator demonstrated that the writing by novice Russian authors significantly differs from the expert writing. The considered bigrams indicated lower lexical diversity; insufficient skills in using grammar that is absent in the Russian language; tendency to underuse predicative clauses and hedging devices; tendency to overuse words and phrases typical of Russian scientific style, as well as referencing and specifying words and phrases.

Conclusions. The proposed concept is aimed to make the academic writing process more student-centered. The identification of common linguistic problems in the novice researchers' writing may raise students' awareness of the differences in the rhetorical choice in the two languages, improve strategic skills in making the proper choice when writing their own texts, and, thus, approach to the norms and expectations of the target discourse community.

Keywords: *science discourse, science writing, discourse-analysis approach, contrastive corpus analysis, discourse-analysis, discourse community, Gramulator.*

Shpit Elena I. – PhD Candidate, Department of Pedagogy and Psychology of Education, Faculty of Psychological, Pedagogical and Special Education, Institute of Psychology and Pedagogy, Tomsk State Pedagogical University (Tomsk, Russia); Associate Professor, Department of Foreign Languages, Tomsk State University of Control Systems and Radio Electronics (Tomsk, Russia); ORCID-ID: 0000-0003-2568-224X; e-mail: forester_2007@mail.ru

Kurovskii Vasily N. – DSc (Pedagogy), Professor, Department of Social Pedagogy, Head of Education Development Laboratory, Research and Education Centre of Education Theory, Tomsk State Pedagogical University (Tomsk, Russia); e-mail: v.kurovskii@yandex.ru

References

1. Bazanova E.M., Korotkina I.B. Russian writing centers consortium // *Vyshee obrazovanie v Rossii* (Higher Education in Russia). 2017. No. 4. P. 50–57.
2. Borodik K.A., Dikusar K.S., Bogatov V.V. Publication activity trends among Russian researchers in 2016–2020, based on international scientific Citation databases: Web of Science Core Collection and Scopus // *Upravlenie naukoj i naukometriya* (Science Governance and Scientometrics). 2021. Vol. 16 (4). P. 571–595. DOI: 10.33873/2686-6706.2021.16-4.571-595
3. Van Deik T. To the definition of discourse. Ideology: A Multidisciplinary Approach. 1998. URL: <http://psyberlink.flogiston.ru/internet/bits/vandijk2.htm> (access date: 01.08.2022).
4. Vinogradova O.I., Smirnova E.A., Viklova A.V., Panteleeva I.M. Syntactic complexity of academic text: A corpus study of written production by learners of English with Russian L1 in comparison with expert texts of English authors // *Vestnik RGGU. Seriya: Literaturovedenie. Yazykoznanie. Kulturologiya* (RSUH Bulletin. Series: Literary Theory. Linguistics. Cultural Studies). 2020. No. 7. P. 107–129. DOI: 10.28995/2686-7249-2020-7-107-129

5. Dobrynina O.L. Grammar errors in academic writing in English: Causes and strategies of correction // *Vysshee obrazovanie v Rossii (Higher Education in Russia)*. 2017. No. 8/9. P. 100–107.
6. Dobrynina O.L. Propaedeutics of errors in abstracts of papers written in Russian // *Vysshee obrazovanie v Rossii (Higher Education in Russia)*. 2015. No. 2. P. 42–50.
7. Dobrynina O.L. Problems of academic writing: Lexical errors, their causes and correction strategies // *Vysshee obrazovanie v Rossii (Higher Education in Russia)*. Vol. 27 (10). P. 75–83. DOI: 10.31992/0869-3617-2018-27-10-75-83
8. Karasik V.I. About the types of discourse. In: *Linguistic personality: institutional and personal discourse*. Volgograd, 2000. 220 p.
9. Kozhina M.N. To the foundations of functional stylistics. Perm: PSU publ., 1968. 251 p.
10. Kolesnikova N.I. What is important to know about the language and style of academic texts? (Paper 1) // *Vysshee obrazovanie v Rossii (Higher Education in Russia)*. 2010. No. 3. P. 130–137. URL: <https://cyberleninka.ru/article/n/chto-vazhno-znat-o-yazyke-i-stile-nauchnyh-tekstov/viewer> (access date: 01.08.2022).
11. Kolesnikova N.I. From the summary to the thesis: A study guide for developing writing skills. Moscow: Flinta, 2002. 288 p.
12. Kolesnikova N.I., Ridnaya Yu.V. The genre model of a scientific paper in the Russian and English languages // *Vysshee obrazovanie v Rossii (Higher Education in Russia)*. 2016. No. 6. P. 98–105.
13. Komarova Yu.A. Functional-stylistic characterization of the science discourse // *Vestnik SPbGU. Yazyk i literatura (Bulletin of Saint-Petersburg University. Language and Literature)*. 2012. No. 4. P. 78–85.
14. Korotkina I.B. Academic Literacy and Methods of Global Scientific Communication // *Nauchnyy redaktor i izdatel (Science Editor and Publisher)*. 2017. Vol. 2 (1). P. 08–13 (In Russ., abstract in Eng.). DOI: 10.24069/2542-0267-2017-1-8-13
15. Korotkina I.B. (b) Academic writing in Russia: The urge for interdisciplinary studies // *Vysshee obrazovanie v Rossii (Higher Education in Russia)*. 2018a. Vol. 27 (10). P. 64–74. DOI: 10.31992/0869-3617-2018-27-10-64-74
16. Korotkina I.B. (a) English for research publication purposes as a new field of educational studies // *Otechestvennaya i zarubezhnaya pedagogika (National and Foreign Pedagogy)*. 2018b. Vol. 1, No. 4 (52). P. 115–130.
17. Kotyurova M.P. Tendency to new phenomena developing in scientific style of speech // *Izvestiya Yuzhnogo federal'nogo universiteta. Filologicheskie nauki (News of Southern Federal University. Philology)*. 2009. No. 1. P. 77–86. URL: <https://philol-journal.sfedu.ru/index.php/sfuphilol/article/view/167> (access date: 01.08.2022).
18. Kochetova L.A., Ilyinova E.Yu. English academic discourse in translanguistic context: Corpus-based study of lexical markers // *Vestnik Volgogradskogo gosudarstvennogo universiteta. Yazykoznanie (Bulletin of Volgograd State University. Linguistics)*. 2020. Vol. 19 (5). P. 25–37. DOI: 10.15688/jvolsu2.2020.5.3
19. Kudryashova O.V. The components of communicative competence in teaching writing speech // *Vestnik YuUrGU. Seriya: Lingvistika (South Ural State University Bulletin. Linguistics)*. 2007. No. 15 (87). URL: <https://cyberleninka.ru/article/n/komponenty-kommunikativnoy-kompetentsii-pri-obuchenii-pismennoy-rechi> (access date: 01.08.2022).
20. Lapteva O.A. How scientists write // *Russkaya rech (Russian Speech)*. 1995. No. 2. P. 55–62.
21. Makarov M.L. The foundations of discourse theory. Moscow: Gnozis, 2003. 280 p. URL: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/http://yanko.lib.ru/books/cultur/makarov-osnovu_teorii_diskursa-8l.pdf (access date: 01.08.2022).
22. Mokhnacheva Yu.V., Tsvetkova V.A. Dynamics of the Russian segment of science publications (based on Web of Science Core Collection and Scopus data) // *Nauchnye i tekhnicheskie biblioteki (Scientific and Technical Libraries)*. 2021. No. 6. P. 15–28. DOI: 10.33186/1027-3689-2021-6-15-28

23. Pryadilnikova N.V. Practical functional stylistics of the Russian language. A study guide. Samara: Samara University, 2016. Part 2. 164 p.
24. Ryabtseva N.K. Academic paper titles and their dominating patterns: a Russian-English perspective // Vestnik Volgogradskogo gosudarstvennogo universiteta. Yazykoznanie (Bulletin of Volgograd State University. Linguistics). 2018. Vol. 17 (2). P. 33–43. DOI: 10.15688/jvolsu2.2018.2.4
25. Solovova E.N. ELT NEWS @ VIEWS // The Russian magazine for English language teachers. 2003. No. 2 (27). Dinternal International Book Distribution. P. 2–11.
26. Tivyaeva I.V., Kuznetsova D.L. Comparative analysis of abstract content and structure in Russian and English research papers // Aktualnye problemy filologii i pedagogicheskoy lingvistiki (Current Issues in Philology and Pedagogical Linguistics). 2020. No. 3. P. 139–152. DOI: 10.29025/2079-6021-2020-3-139-152
27. Tolstova T. Corpus analysis of genre conventions in Russian and international research article abstracts // Vestnik Moskovskogo universiteta. Ser. 19: Lingvistika i mezhkulturnaya kommunikatsiya (Moscow University Bulletin. Linguistics and Intercultural Communication). 2020. No. 1. P. 118–129.
28. Trofimova G.K. Russian language and culture of speech: lectures. Moscow: Flinta, Nauka, 2004. 160 p. URL: <https://biblioclub.ru/index.php?page=book&id=56264> (access date: 01.08.2022).
29. Khutorskiy A.V. Key competences as a component of student-centered paradigm in education // Narodnoe obrazovanie (National Education). 2003. No. 2. P. 58–64.
30. Chernyavskaya V.Ye. Interpreting a science text. Moscow: KomKniga, 2006. 128 p.
31. Chernyavskaya V.Ye. Text linguistics. Discourse linguistics: A study guide. Moscow: LENAND, 2018. 200 p.
32. Sheypak S.A. Academic writing: Critics of the traditional EAP approach // Vyshee obrazovanie v Rossii (Higher Education in Russia). 2020. Vol. 29 (2). P. 92–103. DOI: 10.31992/0869-3617-2020-29-2-92-103
33. Shpit E.I., Kurovskii V.N. Analysing academic texts with computational linguistics tools // Vyshee obrazovanie v Rossii (Higher Education in Russia). 2020. Vol. 29 (7). P. 89–103. DOI: 10.31992/0869-3617-2019-29-7-89-103
34. Shpit E.I., Kurovskii V.N. Specialist corpus for academic writing instruction: compilation and application. In: Proceedings of the 10th International scientific-practical conference, Tomsk, December 17, 2020. Tomsk: TSPU, 2021. P. 161–166.
35. Becher T. Disciplinary discourse // Studies in Higher Education. 1987. Vol. 12 (3). P. 261–274. DOI: 10.1080/03075078712331378052
36. Becher T. The significance of disciplinary differences // Studies in Higher Education. 1994. Vol. 19 (2). P. 151–161. DOI: 10.1080/03075079412331382007
37. Belcher D.D. Seeking acceptance in an English-only research world // Journal of Second Language Writing. 2007. Vol. 16 (1). P. 1–22.
38. Bhabha H.K. The location of culture (Routledge Classics). 2nd ed. London: Routledge, 2004. 440 p.
39. Bhatia V.K. A generic view of academic discourse. Academic discourse. Edited by J. Flowerdew. London: Longman, 2002. P. 21–39.
40. Bhatia V., Flowerdew J., Jones R. Advances in discourse studies. Routledge Taylor and Francis Group. London and New York, 2008. 274 p.
41. Connor U. New directions in contrastive rhetoric // TESOL Quarterly. 2002. Vol. 36 (4). P. 493–510. DOI: 10.2307/3588238
42. Fairclough N. Analysing discourse: Textual analysis for social research. London: Routledge, 2003. 270 p.
43. Flowerdew J. English for research publication purposes. The handbook of English for specific purposes. Ed. by B. Paltridge, S. Starfield. Malden, MA: Wiley-Blackwell, 2013. P. 301–321.

44. Flowerdew L. The exploitation of small learner corpora in EAP materials design: In M. Ghadessy & R. Roseberry (Eds.) *Small Corpus Studies and ELT*. Amsterdam & Philadelphia: Benjamins, 2001. P. 363–379.
45. Gee J.P. *An introduction to discourse analysis: Theory and method*. Routledge. Taylor and Francis Group. London and New York, 2005. 248 p.
46. Graff G. Scholars and sound bites: The myth of academic difficulty // *PMLA* (Publications of the Modern Language Association of America). 2000. Vol. 115 (5). P. 1041–1052.
47. Huckin T.N. Specificity in LSP // *Ibérica, Revista de la Asociación Europea de Lenguas para Fines Específicos*. 2003. No. 5. P. 3–18.
48. Hyland K. *Academic discourse. The Bloomsbury Companion to Discourse Analysis*. Edited by K. Hyland, B. Paltridge. London: Continuum, 2011. P. 171–184.
49. Hyland K. *Disciplinary discourses: Social interactions in academic writing*. Harlow, UK: Pearson Education Limited, Longman, 2000. 211 p.
50. Hyland K. Sympathy for the devil? A defence of EAP // *Language Teaching*. 2018. No. 3. P. 383–399. DOI: 10.1017/S0261444-818000101
51. Hyland K. *Writing in the academy: Reputation, education and knowledge*. London: Institute of Education Press, 2007. 28 p.
52. Kaplan R.B. Cultural thought patterns in intercultural education. *Language Learning*, 1966. 648 p. DOI: 10.1111/j.1467-1770.1966.tb00804.x
53. Kochetova L.A., Kononova I.V. Corpus-based contrastive study of discursive strategy of construing interpersonal relations in English language academic discourse // *J. Sib. Fed. Univ. Humanit. soc. sci.* 2018. P. 1–11. DOI: 10.17516/1997-1370-0353
54. Korotkina I.B. Russian scholarly publications in anglophone academic discourse: The clash of tyrannosaurs // *Integration of Education*. 2018. Vol. 22 (2). P. 311–323. DOI: 10.15507/1991-9468.091.022.201802.311-323
55. Kwan B.S.C. An investigation of instruction in research publishing in doctoral programs: The Hong Kong case // *Higher Education*. 2010. Vol. 59, No. 1. P. 55–68.
56. Lillis T., Curry M.J. Professional academic writing by multilingual scholars: Interactions with literacy brokers in the production of English-medium texts // *Written Communication*. 2006. Vol. 23 (1). P. 3–35. DOI: 10.1177/0741088305283754
57. McCarthy P., Lehenbauer B., Hall C., Duran N., Fujiwara Y., McNamara D. A coh-matrix analysis of discourse variation in the texts of Japanese, American, and British Scientists // *Foreign Languages for Specific Purposes*. 2007. Vol. 6. URL: https://www.researchgate.net/publication/242322281_A_Coh-Matrix_Analysis_of_Discourse_Variation_in_the_Texts_of_Japanese_American_and_British_Scientists (access date: 01.08.2022).
58. McCarthy P.M., Duran N.D., Booker L.M. The devil is in the details: New directions in deception analysis. In: *Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference, FLAIRS-25*. 2012a. P. 188–195.
59. McCarthy P.M., Watanabe S., Lamkin T.A. The Gramulator: A tool to identify differential linguistic features of correlative text types. In: P.M. McCarthy and C. Boonthum Denecke (Eds.) *“Applied Natural Language Processing: Identification, Investigation and Resolution”*. 2012b. P. 312–333. Hershey, PA: IGI Global. DOI: 10.4018/978-1-60960-741-8.ch018
60. McKercher B., Law R., Weber K., Song H., Hsu C. Why referees reject manuscripts // *Journal of Hospitality & Tourism Research*. 2007. Vol. 31 (4). P. 455–470.
61. Min H.S., McCarthy P.M. Contrastive corpus analysis on the writing of American and Korean scientists using Gramulator // *Language Research*. 2013. Vol. 49 (3). P. 681–702.

62. Min H.S., McCarthy P.M. Identifying varieties in the discourse of American and Korean scientists: A contrastive corpus analysis using the Gramulator. In: Proceedings of the 23rd International Florida Artificial Intelligence Research Society Conference, FLAIRS-23. 2010. P. 247–252.
63. Paquot M. Towards a productively-oriented academic word list. In: Walinski J.; Kredens K.; Gozdz-Rozkowski S. Practical Applications in Language and Computers, 2005. Peter Lang: Frankfurt and Main, 2007. P. 127–140. URL: <http://hdl.handle.net/2078.1/76045>
64. Reid N. Getting published in international journals: Writing strategies for European social scientists. Professional Publications Press, 2018. 309 p.
65. Shpit E.I., McCarthy P.M. Addressing discourse differences in the writing of Russian engineering students and international researchers // Language Teaching Research. July. 2022. DOI: 10.1177/13621688221109809
66. Smirnova E., Strinyuk S. Hedges in Russian EAP writing: A corpus-based study of research papers in management // Journal of English as a Lingua Franca. 2020. Vol. 9 (1). P. 81–101. DOI: 10.1515/jelf-2020-2033
67. Smirnova E.A. Referential coherence of academic texts: a Corpus-Based analysis of L2 research papers in management // Journal of Language and Education. 2019. Vol. 5 (4). P. 112–127. DOI: 10.17323/jle.2019.9688
68. Swales J. Genre analysis: English in academic and research settings. Boston: Cambridge UP, 1990. 274 p.
69. Tardy C. The role of English in scientific communication: lingua franca or Tyrannosaurus rex? // Journal of English for Academic Purposes. 2004. Vol. 3 (3). P. 247–269. DOI: 10.1016/j.jeap.2003.10.001
70. Terenin A. Clear Writing as a problem of Russian learners of English // Middle-East Journal of Scientific Research. 2013. Vol. 16. No. 11. P. 1512–1519. DOI: 10.5829/idosi.mejsr.2013.16.11.12039
71. Wen X.A., McCarthy P.M., Strain A.C. Gramulator analysis of gendered language in cable news reportage. In: Proceedings of the 26th International Florida Artificial Intelligence Research Society Conference, FLAIRS. 2013. P. 243–248.
72. Widdowson H.G. Discourse Analysis. Oxford University Press, 2007. 152 p.
73. Zanina E. Move structure of research article abstracts on management: Contrastive study (the case of English and Russian) // Journal of Language and Education. 2017. Vol. 3 (2). P. 63–72. DOI: 10.17323/2411-7390-2017-3-2-63-72
74. Zanina E. Strategic hedging: a comparative study of methods, results and Discussion (and Conclusion) // Journal of Language and Education. 2016. Vol. 2 (2). P. 52–60. DOI: 10.17323/2411-7390-2016-2-2-52-60