



DOI: 10.18721/JHSS.10407

УДК: 378.02:37.016

## **ПОТЕНЦИАЛ КОРПУСНОЙ ЛИНГВИСТИКИ В ПОДГОТОВКЕ СПЕЦИАЛИСТОВ В ОБЛАСТИ КОМПЬЮТЕРНОЙ ЛИНГВОДИДАКТИКИ**

**А.В. Дмитриев, М.С. Коган**

Санкт-Петербургский политехнический университет Петра Великого,  
Санкт-Петербург, Российская Федерация

В последние годы наблюдается заметное повышение роли корпусной лингвистики (КЛ) в мире. Число доступных корпусов исчисляется тысячами, позволяя решать самые разные исследовательские задачи в области лингвистики, культурологии, социологии, истории, когнитивной лингвистики и других гуманитарных наук. Несмотря на то, что КЛ оказывает существенное влияние на методику обучения иностранным языкам, анализ литературы показывает, что потенциал КЛ не в полной мере находит отражение в программах подготовки преподавателей иностранных языков и переводчиков. В статье дается обзор наиболее значимых современных проектов в области КЛ, анализируется роль и место КЛ в подготовке будущих специалистов в области компьютерной лингводидактики – магистерской программы в рамках направления «Лингвистика», открытой в 2016г в СПбПУ. Анализ рабочих программ дисциплин (РПД), изучаемых в рамках магистерской программы «Компьютерная лингводидактика», показал, что большинство из них содержит либо специальные модули, направленные на освоение и применение подходов КЛ, либо позволяют интегрировать подходы КЛ при изучении определенных тем. Утверждается, что для успешной работы с корпусами студенты должны овладеть корпусной грамотностью, которая предполагает решение поисковых и исследовательских задач разной сложности: от знакомства с устройством корпуса до создания собственных подкорпусов и решения эвристических задач. Выделяются три этапа освоения больших корпусов, приводятся примеры заданий для каждого этапа на материале Национального корпуса русского языка (НКРЯ). В заключении обсуждаются дальнейшие направления исследований.

**Ключевые слова:** корпусная лингвистика, компьютерная лингводидактика, лингвистические корпусы, методология DDL, рабочая программа дисциплины, Национальный корпус русского языка, комплекс поисковых заданий в НКРЯ

**Ссылка при цитировании:** Дмитриев А.В., Коган М.С. Потенциал корпусной лингвистики в подготовке специалистов в области компьютерной лингводидактики // Научно-технические ведомости СПбГПУ. Гуманитарные и общественные науки. 2019. Т. 10, № 4. С. 69–85. DOI: 10.18721/JHSS.10407

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

## **THE POTENTIAL OF CORPUS LINGUISTICS IN TRAINING FOREIGN LANGUAGE TEACHERS MAJORING IN COMPUTER ASSISTED LANGUAGE TEACHING**

**A.V. Dmitrijev, M.S. Kogan**

Peter the Great St. Petersburg Polytechnic University,  
St. Petersburg, Russian Federation

In modern world practice, Corpus Linguistics (CL) is gaining serious positions. The number of available corpora accounts several thousands which allows conducting various research in linguistics, cultural studies, sociology, history, cognitive linguistics and other humanities. The paper reviews some of the most substantial ones. Although CL has a remarkable impact on language teaching the literature overview shows that its potential is only modestly used in curricula of future language teachers and translators. The paper analyzes the role and place of CL in the curriculum of the Computer Assist-

ed Language Teaching (CALT) Master program launched at SPbPU in 2016. The analysis of working programs of disciplines included into the CALT Master Program has revealed that most of them either contain special modules aiming at acquisition of and applying CL methods in research and teaching or allow integrating CL approaches in studying certain topics. The paper states that students have to acquire corpus literacy which means solving corpus search and research tasks of different complexity ranging from getting familiar with the corpus structure to compiling own subcorpora and dealing with heuristic type tasks. Three stages of students' familiarization with big corpora are distinguished. The recommended tasks for each stage for the Russian National corpus developed by authors are provided. Further research is discussed.

**Keywords:** Corpus Linguistics, linguistic corpora, DDL approach, working syllabuses, the Russian National Corpus, search tasks for the RNC

**Citation:** A.V. Dmitrijev, M.S. Kogan, The Potential of Corpus Linguistics in Training Foreign Language Teachers Majoring in Computer Assisted Language Teaching, St. Petersburg State Polytechnical University Journal. Humanities and Social Sciences, 10 (4) (2019) 69–85. DOI: 10.18721/JHSS.10407

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

### Введение. Постановка проблемы

Корпусная лингвистика (далее КЛ) – это новое направление в лингвистике, которое существует чуть более полувека и занимается созданием и использованием корпусов для решения различных языковых (и не только) проблем. Сегодня корпуса стали неотъемлемой частью лингвистики, одним из ее краеугольных камней.

В настоящей статье рассматривается потенциал современной КЛ в подготовке специалистов в области компьютерной лингводидактики, обучающихся по магистерской программе «Компьютерная лингводидактика» Высшей школы лингводидактики и перевода Санкт-Петербургского политехнического университета Петра Великого.

Овладение методами КЛ способствует развитию и лингвистической/иностранной компетенции, и ИКТ компетенции, что очень важно для наших выпускников, так как предполагается, что после окончания магистратуры они смогут преподавать иностранные языки для специальных целей (ЯСЦ/LSP) по некоторым техническим, а также экономическим и гуманитарным направлениям с широким применением ИКТ.

В рамках данной статьи мы планируем обсудить роль КЛ в современном мире, исследовать ее место КЛ в программе подготовки магистров по Программе *Компьютерная лингводидактика*, обсудить способы повышения эффективности/усиления дидактического компонента использования корпусных технологий в учебном процессе.

### Роль и место КЛ в современном мире

После появления корпусов вся лингвистическая наука стала другой, поскольку все ее

достижения теперь необходимо интерпретировать в контексте «корпусной эпохи» (термин В.А. Плунгяна). По меткому замечанию В.А. Плунгяна, «корпус – это не только мощный инструмент исследования языка, но и новая идеология, ориентирующая исследователя на текст как главный объект теоретической рефлексии» [1, с. 14]. Другими словами, вся современная лингвистика должна быть лингвистикой корпусов: «корпус вернул лингвистике ее настоящий полноправный объект»<sup>1</sup>, корпус позволяет понять, каков язык на самом деле, а не каким мы хотим, чтобы он был. «Теперь для овладения языком человеку нужны не две, а три вещи: словарь, грамматика и корпус текстов данного языка. Потому что и словарь, и грамматика, в общем-то, бесполезны вне этого живого пространства, где язык, собственно, и функционирует» [ibid].

В настоящее время мы являемся свидетелями глобального покрытия КЛ – как содержательно, так и методически – различных профессиональных и научных зон в области гуманитарных исследований – истории, социологии, литературоведения и не только. Вместе с тем, КЛ уже выходит за пределы и гуманитарных наук: так, недавно британские ученые на основе анализа корпуса газетных текстов периода 1800–2014 годов попытались пролить свет на историческую засуху и дефицит воды в Великобритании и в остальном мире в течение последних 200 лет [2, с. 212]. Все это дает нам основания утверждать, что современная КЛ – это не только и не столько дисциплина, сколько особая *культуромика науки* (science

<sup>1</sup> Плунгян В.А. Лекция: Почему современная лингвистика должна быть лингвистикой корпусов. Прочитана 1 октября 2009 года. URL: <https://polit.ru/article/2009/10/23/corpus/>.

culturomics), что может определяться как «исследование культуры человечества, направление его развития во времени посредством количественного анализа слов и фраз в очень больших объемах оцифрованных текстов» [3, 4].

Классическим примером внедрения технологии корпусов в современную жизнь является голосовой помощник Яндекс Алиса, который изучал различные текстовые корпуса, перенимая у них тот или иной стиль разговора. Методология корпуса стала частью – или даже предвестником – всех современных технологий, на основе которых создаются системы искусственного интеллекта, таких как машинное обучение и нейронные сети.

Сегодня КЛ характеризуется не только тысячами корпусов, в некоторых из которых содержатся десятки миллиардов слов, но и международными проектами, которые могут использовать лингвисты, изучающие разные языки. КЛ становится областью исследования и создания метакорпусов (корпуса корпусов):

- открытая БД KELLY (списки частотных слов для 9 языков и 72 языковых пар для изучающих иностранный язык и переводчиков) [5, с. 155];

- ICC (International Comparable Corpus) (создание началось в 2017 году; конечная цель проекта – внести вклад в сопоставительные исследования английского и других языков с использованием сопоставимых наборов данных разговорных, письменных и, возможно, электронных регистров) [2, с. 131];

- специальный портал Wortschatz в Лейпцигском университете (393 корпуса, 252 языков);

- система Sketch Engine corpus (588 корпусов, 94 языка) [6];

- корпус Aganea в Университете Комениуса в Братиславе (66 корпусов и 31 язык) и др. [7].

Все эти тенденции, безусловно, необходимо принимать во внимание, говоря о месте и роли КЛ и в образовательном пространстве в том числе.

Магистральные направления в создании инновационной образовательной среды с привлечением баз данных и других цифровых ресурсов были заложены в научно-педагогической школе Р.Г. Пиотровского и развиваются Л.Н. Беляевой, Е.С. Полат, П.В. Сысоевым, С.В. Титовой и др. Вопросами выделения профессионального терминологического слоя лексики с помощью методов статистической оценки небольших корпусов текстов и статистикой речи занимались М.А. Марусенко, Т.А. Лаздинь

и другие представители Санкт-Петербургской школы математической лингвистики (И.В. Азарова, А.О. Гребенников, О.Н. Гринбаум, Л.Н. Засорина, В.П. Захаров, Г.Я. Мартыненко, А.С. Герд и др.). Безусловно, это находит свое отражение и в заметном повышении роли КЛ в подготовке лингвистов-исследователей, преподавателей иностранных языков и переводчиков.

Существенно обогащая и расширяя возможности методики обучения иностранным языкам, в частности профессионально-ориентированного обучения [8, 9], не оставаясь в стороне от его психолого-педагогического и когнитивного потенциала [10–14], КЛ стремительно развивается в сеть междисциплинарных направлений, одним из которых сегодня является корпусная педагогика, или *корпусная лингводидактика*. В ее разработку вносят свой вклад исследователи отечественных и зарубежных школ: Н.Б. Гвишиани, Е.Е. Голубкова, Т.Б. Назарова, Т.А. Клепикова, С.О. Савчук, Т.В. Толстова, О.В. Нагель, В.П. Захаров, В.В. Рыков, Р.П. Мильруд, В.А. Плунгян, Д. Байбер, А. Болтон, С. Браун, Т. Джонс, Р. Реппен, Т. Кобб, Д. Кристал, М. Льюс, Р. Картер, М. Маккарти, Т. МакЭнери, К. Триббл, М. Скотт, К. Лакман, С. Конрад, М. Хоуи, Л. Флаурдюю, С. Йохансон, Дж. Лич, М. Дейвис и др.

Все большую популярность получают методически-ориентированные корпуса. Примерами могут служить:

- ресурс LINGVATORIUM, созданный коллективом центра лингвистических исследований им. А.А. Худякова в Санкт-Петербургском государственном экономическом университете;

- корпус инженерных текстов на базе Томского политехнического университета (Н.Н. Шаламова, А.Ю. Фильченко);

- мультимедийный корпус устных текстов и упражнений ELISA (English Language Interview Corpus as a Second-Language Application), созданный в университете г. Тюбинген, Германия ([http://www.uni-tuebingen.de/elisa\\_index.html](http://www.uni-tuebingen.de/elisa_index.html));

- корпус ошибок обучаемых, собранный в РГПУ им. А.И. Герцена (М.В. Колина, Е.А. Николаева);

- корпус русских учебных текстов (КРУТ), разработанный командой лингвистов и программистов Школы лингвистики ВШЭ (корпус представляет собрание студенческих ВКР, и одним из достоинств является возможность поиска по типам студенческих ошибок с указанием причин их возникновения) и др.

В свое время У. Рёмер, называя такие

корпуса специализированными, определила их в три основные группы: LSP (language for specific purpose) корпуса; ученические корпуса; параллельные корпуса и корпуса переводов [15].

Поскольку корпуса активно и многопланово интегрируются в учебный процесс, как синхронный, так и диахронный, все насущнее становится методология составления упражнений на основе корпусных технологий, которые довольно быстро решают проблему дефицита заданий и существенно увеличивают скорость подбора примеров, что позволяет составлять задания по мере необходимости, в соответствии с возникшей идеей или проблемой.

Исследователи обращают внимание не только на то, какие упражнения/задания можно разработать с помощью корпуса (см., например, [16]), но и на какой языковой уровень необходимо разрабатывать, и как это делать, какой метод выбирать для конкретной целевой аудитории и каких ошибок разработчику следует избегать при составлении подобных заданий [17-20].

Ввиду того, что следование методологии составления упражнений на основе корпуса непосредственно связано с выбором наиболее эффективного метода обучения, то в качестве одного из возможных решений в современной КЛ предлагается образовательный формат DDL (data-driven learning), предложенный еще в 1991 году Т. Джонсом. Он предложил студентам выступить в роли исследователей, которые с помощью компьютерных технологий и корпусов должны были проводить свои собственные языковые мини-исследования [21]. В конечном счете, основная задача студента «научиться учиться» (learn how to learn) [21, с. 19], что достигается, в том числе, с помощью корпусных заданий, которые «позволяют учащемуся примерить на себя роль экспериментатора, который проводит свое уникальное исследование, а не компилирует чужие идеи» [21, с. 14].

Результаты недавно опубликованного мета-анализа убедительно свидетельствуют о том, что DDL является эффективным методом обучения ИЯ в различной учебной обстановке, для различных категорий учащихся, преследующих разные языковые цели [24]. Перспективность такого подхода была обоснована и отечественными исследователями [23–25]. Корпусное обучение на основе Web-дидактики является одним из причин отказа от традиционно объяснительного подхода в пользу когнитивно-коммуникативных практик, то есть

переход к «обучению через исследование» [26, с. 159], а также от парадигмы «учащийся-исследователь» (learner-as-researcher) к новой парадигме «учащийся-путешественник» (learner-as-traveller) [26]. Перспективным представляется также сочетание DDL подхода и CLIL технологии (content-language integrated learning).

Необходимо отметить, что современные ИКТ, на наш взгляд, недостаточно внедрены в профессиональную деятельность лингвиста-преподавателя, что также признают и авторы упомянутого мета-анализа, замечая, что, несмотря на свой потенциал и эффективность, использование подходов КЛ для изучения ИЯ не является повсеместно/широко применяемым методом в обучении/изучении ИЯ [27]. Недостаточная разработанность методов применения корпусной лингвистики в обучении ИЯ характеризует в большей степени отечественную науку. Введя запросы на русском (*корпусная лингвистика*) и английском (*corpus linguistics*) языках на платформах Elibrary и Google Scholar, можно заметить, что количество публикуемых работ за рубежом будет более чем в 4 раза (Elibrary) и более чем в 70 раз (Google Scholar) больше, чем русскоязычных отечественных исследований.

#### **Методология применения КЛ в лингводидактике. Инструментарий**

Целесообразность создания и смысл использования корпусов определяется следующими предпосылками:

1. С конца XX века словари и учебники грамматики создаются на основе корпусов.
2. КЛ используется в теоретической лингвистике с целью сбора необходимых для анализа определенного явления данных их естественном контексте.
3. На основе кейсов проводится машинное обучение для различных областей прикладной лингвистики и информационных технологий.
4. Корпусы могут быть полезны для многих других задач, связанных с языком. Созданный и подготовленный корпус может использоваться повторно разными исследователями и для разных целей, что является одним из критериев устойчивости электронных образовательных ресурсов.

За время существования корпусных технологий КЛ обогатилась собственным корпусным инструментарием педагогически-ориентированного анализа. Данный инструментарий позволяет исследователю анализировать лексические чанки, гапаксы, находить и различать коллокации и коллигации, лексические пучки,

или кластеры, освоить стратегию хеджинга, типичную для современного научного дискурса как в естественно-научной области [28], так и в гуманитарной сфере и т.д. [29].

В конечном счете это позволяет:

- по любому слову построить конкорданс (список всех употреблений данного слова в контексте со ссылками на источник);

- получить данные о частоте словоформ, лексем, грамматических категорий, проследить изменение частот и контекстов в различные периоды времени, получить данные о совместной встречаемости лексических единиц и т.д.;

- изучать динамику процессов изменения лексического состава языка, проводить анализ лексико-грамматических характеристик в разных жанрах и у разных авторов, и т.д.;

- использовать корпус как источник и инструмент многоаспектных лексикографических работ по подготовке разнообразных исторических и современных словарей.

Помимо уже существующих и широко апробированных программ-конкордансеров (наибольшей популярностью сегодня пользуются Concordance, MicroConcord, MonoCorc, TACT, TACTWeb, SARA), разрабатываются и активно интегрируются в обучение такие новые инструменты анализа, как корпусные менеджеры (например, Bonito CQP, DDC, WebCorp, Xaira), которые по своему функционалу стоят на порядок выше конкордансеров [30]. Обзор существующих программ можно найти в [31–33].

Своевременно говорить и о праве на существование *корпусного подхода*, или *корпусной методологии* как совокупности методов лингвистического исследования, основанного на корпусах текстов, ориентированного на прикладное изучение языка, его функционирование в реальных текстовых средах, что важно не только для преподавания LSP, но и для правильной организации работы в формате DDL для будущих специалистов в области компьютерной лингводидактики [34–36]. «Корпусный метод помогает отобрать микрокосм (the microcosm), о котором говорил еще Г. Пальмер, который подразделял учебный словарь на две большие группы: «строго отобранный материал, который он называл микрокосм (the microcosm), и стихийный» [37, с. 115].

#### **Результаты. Анализ РПД ООП «Компьютерная лингводидактика»**

В 2016 году на кафедре лингвистики и межкультурной коммуникации Санкт-Петербургского политехнического университета Петра Великого была разработана

ООП «Компьютерная лингводидактика» подготовки магистров лингвистики по заочной форме обучения. Программа получила лицензию, и в 2016 году был проведен первый набор студентов, заинтересованных в перспективах данного направления. Центральными курсами в ООП «Компьютерная лингводидактика» являются: «Компьютерная лингводидактика», «Введение в автоматизированную обработку текста. Лингвистические базы данных», «Информационная культура педагога», «Разработка образовательных электронных ресурсов», «Методика дистанционного обучения иностранным языкам», «Мобильные технологии в обучении иностранным языкам». Именно они содержат весь тот потенциал, который реализует профессиональные компетенции не только по педагогике, но и по прикладной и компьютерной лингвистике и должен позволить достичь высокого уровня сформированности ИКТ-компетенции в соответствии с отечественными и зарубежными требованиями [38].

Результаты анализа показали, что большинство дисциплин содержат темы, предполагающие либо использование знаний обучающихся о КЛ в проведении лингвистических исследований и в обучении ИЯ (например, «Научный дискурс», «Практикум по культуре общения первого иностранного языка», «Разработка образовательных электронных ресурсов», «Информационная культура педагога», «Методика дистанционного обучения иностранным языкам»), либо углубленное изучение ими теоретических основ КЛ и их практическое применение (например, «Введение в автоматизированную обработку текста. Лингвистические базы данных»). Иными словами, данные анализа позволяют говорить о КЛ как о дисциплине, укрепляющей междисциплинарные связи (МДС) в ООП «Компьютерная лингводидактика», что позволяет говорить о центрообразующем потенциале КЛ в подготовке специалистов данного профиля (см. табл. 1). Междисциплинарной модели обучения иностранному языку и ее интегрированию в современное образовательное пространство посвящен ряд работ наших коллег [39–42].

В табл. 1 не содержится информации о дисциплинах, в рамках которых КЛ используется «факультативно»: например, в курсе «Практика перевода английского языка (технический профиль)», «Практикум по культуре речевого общения первого ИЯ», «Практикум по культуре речевого общения второго ИЯ», «Общее языкознание и теория лингвистических учений», «Научный дискурс». Публикации по использованию DDL в

Таблица 1

**Результаты анализа РПД ООП 45.04.02\_04 Компьютерная лингводидактика  
Corpus Linguistics related topics in working programs of the CALT Master Program**

Название дисциплины	Название разделов, касающихся КЛ	Трудоемкость освоения раздела, ач
Компьютерная лингводидактика	Тема 10. Электронные ресурсы: классификация и структура;	19
	Тема 21. Применение интернет-ресурсов в лингводидактике. Разновидности лингвистических интернет-ресурсов и их особенности.	18
Введение в автоматизированную обработку текста. Лингвистические базы данных	Модуль 4. Корпусные технологии на практике	36
Информационно-коммуникационные технологии в лингвистических исследованиях	Тема: Базы данных и лингвистические информационные ресурсы.	22
НИР (рассредоточенная)	3 семестр	254
Темы ВКР, утвержденные Приказом СПбПУ, защита которых состоится в январе 2020 г.	1) Разработка дидактических ресурсов по дисциплине «Английский язык для специалистов социально-экономической сферы» с использованием подходов КЛ; 2) Обучение фразеологической лексике с использованием подходов КЛ (направление «Лингвистика»)	

преподавании немецкого [43, 44], французского [45] и испанского [46] языков и наши собственные исследования [24, 47] говорят о целесообразности использования этого подхода в «Практикуме по культуре речевого общения второго ИЯ», которыми в ВШЛиП являются немецкий, испанский и французский. Интерес, который проявляют исследователи к привлечению корпусных методов к дискурсивному анализу [см., например, 48, 49, 50] позволяет предположить/заключить, что методы КЛ также могут быть успешно интегрированы в изучение дисциплины «Научный дискурс».

Далее мы более подробно остановимся на анализе нашего опыта интегрирования Национального корпуса русского языка (НКРЯ) в программу подготовки специалистов в области компьютерной лингводидактики.

**Способы интегрирования НКРЯ в программу подготовки специалистов в области компьютерной лингводидактики**

Практическая работа с корпусом, а именно поиск единиц по языковым уровням — от морфемного до синтаксического, — четко коррелирует с теми знаниями, умениями и навыками, которые студенты получают в магистерском курсе «Общее языкознание и история лингвистических учений», где одно из цен-

тральных мест занимает учение об уровневой организации языковой системы и позволяет учащимся, с одной стороны, успешно справляться с поисковыми заданиями в корпусе, а с другой, увидеть теоретическую дисциплину базового блока в качественно новом ракурсе — в свете структурной лингвистики, являющейся, фундаментом, точкой отсчета для всех прикладных лингвистических исследований XX — начала XXI веков.

В настоящее время имеется ряд исследований, в которых авторы приводят разные форматы упражнений на основе НКРЯ, начиная с фонологического уровня и заканчивая поиском контекстуальных связей и окружений на синтаксическом уровне [36, 51, 52]. Расширяя имеющийся опыт, мы стремимся адаптировать существующую методологию для обучения студентов в области компьютерной лингводидактики. НКРЯ является универсальным лингвистическим инструментарием, позволяющим проводить поиск по текстам самой разной тематической направленности, временному охвату, жанрам и т.д.

Перейдем к рассмотрению составленных нами заданий, предназначенных для пользователей с разным уровнем знакомства с корпусом НКРЯ (начальный, средний и продвинутый). Нами был разработан модульный план,

состоящий из ступеней последовательного освоения навыков работы с корпусом — от знания функционально-технических возможностей поисковой системы корпуса и основополагающей терминологии до умения самостоятельно составить комплекс упражнений на основе такого корпуса. План разработан на материале Национального корпуса русского языка.

Методологически наш план «держится» на идее о том, что для выполнения лингвистических исследований на основе корпусов особое значение имеет умение рационально пользоваться корпусом. Это умение называется *корпусной грамотностью*, к числу теоретических вопросов которой относятся следующие: «1) знание базовых понятий, методов и технологий корпусной лингвистики; 2) изучение и обобщение основных способов и этапов создания корпуса; 3) типология корпусов; 4) паспортизация текста; 5) характеристика различных видов внешней и внутренней разметки текста; 6) Интернет как корпус и др. Важным практическим навыком является умение корректно задать параметры поиска в Национальном корпусе русского языка с обращением к другим корпусным ресурсам русского языка в открытом пространстве Интернет» [26, с. 160]. Примерный план изучения НКРЯ представлен в табл. 2.

### Обсуждение

На *начальном этапе* студенты знакомятся с устройством поисковой системы НКРЯ и «корпусными» терминами, которые необходимо знать для последующей работы с корпусом, например, *корпус, репрезентативность, грамматическая разметка, семантическая разметка, лемма, лемматизация, токенизация, паспорт леммы, грамматическая омонимия, шум* и т.д. Здесь мы учим студентов овладевать техническими возможностями работы с корпусом: пользоваться грамматическими и семантическими признаками в лексико-грамматическом поиске; добавлять и удалять поисковые строки; правильно вводить словосочетания в поисковые строки, а также целые фразы; отличать словоформу от леммы, знать, где прочитать описание леммы и как «растянуть» контекст до полного; пользоваться опцией «Настройки», выдаваемой вместе с результатами, форматом KWIC, а также «Дополнительными признаками» в корпусе; знать разные форматы выдачи результатов, с возможностями отфильтровать контексты со снятой и неснятой омонимией; уметь пользоваться всеми форматами сохранения результатов. Хорошим по-

мощником на этом этапе является Инструкция пользователя Национального корпуса русского языка<sup>2</sup>.

Наиболее показательные упражнения на эти навыки можно продемонстрировать следующими конкретными примерами, которые стандартно начинаются с фразы «Найдите в основном корпусе...»:

1. *прилагательное «самодовольный» в основном корпусе, прочитайте паспорт леммы и полный контекст с данной словоформой;*
2. *сочетания слова «бодрый» с животными мужского рода; отсортируйте по алфавиту;*
3. *все слова, заканчивающиеся на -сть, но не на -исть, и обозначающие поведение и хорошие поступки человека;*
4. *последовательность трех однородных сказуемых мужского рода прошедшего времени, начинающихся с приставки об-; полученные результаты скачайте в формате Excel;*
5. *только те тексты, где есть случаи употребления леммы «задоринка» в качестве независимого существительного, не в составе фразеологизма;*
6. *словоформу «уж» только со снятой омонимией.*

Следует отметить, что уже на ознакомительном этапе возможны небольшие открытия в области семантики, которая осваивается позже. Так, в одном из заданий предлагалось найти всех животных мужского рода, стоящих с прилагательным *бодрый* и представить в формате выдачи KWIC. Помимо бодрых коней, псов, жуков, петухов в результатах студенты увидели *бодрого Гоголя*. Растянув контекст, студенты убедились в том, что речь идет об известном русском писателе, однако усомнились в правильности заданных критериев поиска в семантической и грамматической разметке, поскольку в параметрах были отмечены только животные. Те, кто догадался заглянуть в толковые словари на [www.slovari.ru](http://www.slovari.ru)<sup>3</sup> с удивлением обнаружили, что гоголь — это подвид уток. Поскольку практически все контексты были с неснятой омонимией, то автор «Мертвых душ» оказался среди бодрых животных.

Другим не менее интересным результатом оказались контексты со свободной леммой *задоринка*. Задание заключалось в нахождении примеров, где эта словоформа была бы

<sup>2</sup> Национальный корпус русского языка. Инструкция пользователя // Российская академия наук. Институт русского языка им. В. В. Виноградова. Москва. 84с. URL: <http://ruscorp.org/new/instruction-main.pdf>

<sup>3</sup> Ресурс, содержащий более 20 томов словарей русского языка.

Таблица 2

**План изучения НКРЯ**  
**The plan for the study of RNC**

Этап/уровень освоения НКРЯ	Цель этапа	Примеры заданий
Начальный	Знакомство с устройством НКРЯ, освоение «корпусной» терминологии	Поисковые задания, начинающиеся с фразы «Найдите в основном корпусе...» 1) прилагательное «самодовольный» в основном корпусе, прочитайте паспорт леммы и полный контекст с данной словоформой; 2) сочетания слова «бодрый» с животными мужского рода; отсортируйте по алфавиту;
Средний	поиск единиц по языковым уровням – от морфемного до синтаксического; освоение структурно-семантической организации текста	- задайте подкорпус X и найдите в нем ...; - задайте грамматический признак X и проверьте; - выберите среди семантических признаков категорию X и проанализируйте.
Продвинутый	последовательный анализ языковых реалий, создание собственных подкорпусов, применение эвристического подхода	1) Пользуясь поэтическим корпусом, докажите, что в XVIII веке слово «приличен» могло иметь ударение на последнем слоге: «приличЕн». Приведите этот пример (эти примеры). 2) Когда в Корпусе появляются первые примеры глагола «реагировать»? Приведите первые два примера. Что они значат? Когда появляются первые примеры глагола «реагировать» с какой-либо приставкой? Приведите первые два примера.

не в составе фразеологизма. С первого взгляда, суть задания – в умении пользоваться оператором ‘-’ для отсечения предлога *без* и существительного *сучок* (в ответ на запрос по поисковому слову *задоринка*, корпус выдает 155 вхождений леммы в 143 документах).

При анализе полученных результатов студенты смогли убедиться, что лемма *задоринка* действительно может свободно функционировать в контекстах, но со значением ‘радость’, ‘задор’:

1. *Серьезные все как один, без задоринки и смешинки. – Давай в арку, - без задоринки и смешинки предложил третий <...> Они брели шеренгой в полтротуара, с задоринкой, мимо озабоченно-неотложных граждан [Александр Силаев. Армия Гутэнтака (2007)].*

2. *С нее, мелованной финскими бумажниками, с идеального отпечатка, смотрел на татарина его сомик, совсем как живой, и было в его взгляде что-то утешающее, даже задоринка некая проглядывала [Дмитрий Липскеров. Последний сон разума (1999)].*

При этом студентам необходимо разъяснить, что этимология этого слова никак не мотивирована значением ‘радость’. Поскольку фразеологизм *без сучка, без задоринки* означает ‘сделать что-то идеально’, то рассматривать

словоформу *задоринка* следует исключительно в семантическом поле значений словоформы *сучок*, что и подтверждается толковыми словарями: *задорина* – ‘шероховатость на гладкой поверхности’<sup>4</sup>. Именно в таком значении лемма выступает в следующих найденных контекстах в НКРЯ, и при этом также не в составе фразеологизма:

1. *Воля бы Дементьева, он всю повесть мою сострогал бы под гладь, не осталось бы ни задоринки [А.И. Солженицын. Бодался теленок с дубом (1967-1974)].*

2. *Пора научиться лелеять свои несовершенства и задоринки-сучки! [Игорь Мартынов. Святая пустота // «Столица», 1997.07.01].*

3. *Мчались мы на вихрях-конях по своей великой равнине, а он навстречу нам выдул стужу-метель и так засорил глаза, что в самых невинных сучках-задоринках нам примерещились бесы-вредители [Михаил Эпштейн. Ленин-Сталин // «Родник», 1989].*

4. *Я, конечно, ни в чем не была виновата, ибо если бы на моем партийном пути была бы хоть малейшая задоринка, то в обстановке 1937 года и учитывая работу, которую я тогда выполняла (зав. отделом школ и вузов Сталинского райкома*

<sup>4</sup> Словарь <http://slovari.ru/search.aspx?s=0&p=3068>



г. Киева), то я бы прошла не как жена, а по самостоятельному делу [Обращения родственников репрессированных командиров Красной Армии к руководителям страны (1937-1960)].

Мы неоднократно подмечали, что для студентов, особенно не-лингвистов, подобные функционально-семантические «открытия» становятся стимулом, «корпусным проводником» к более глубокому изучению языковых изменений.

*На следующем этапе* начинается содержательная работа с корпусом — а именно по поиску единиц по языковым уровням — от морфемного до синтаксического. Такой подход важен потому, что он четко коррелирует с теми знаниями, умениями и навыками, которые студенты получили в курсе «Общее языкознание и история лингвистических учений», где одно из центральных мест занимает учение об уровневой организации языковой системы. Работая с различными контекстами, представленными в корпусе, студенты получают необходимые навыки работы со структурно-семантической организацией текста, что подразумевает не просто изолированный анализ конкретных единиц, но умение сформулировать необходимые выводы на основе внутритекстовых глубинных связей, а также выявленных коллокаций.

Цель данного этапа работы с НКРЯ — научить студентов осуществлять поиск в имеющихся в НКРЯ подкорпусах (газетном, параллельном, акцентологическом, устном, синтаксически размеченном и т.д.). Для этого необходимо задавать некоторый вектор движения, формирующий преимущественно пассивные навыки работы, но определенным образом воспитывающим системность и последовательность поиска. Примеры формулировок такого вектора представлены в табл.2. Вектор (или, как называют его авторы «опоры») уже методически обоснован [53]. Мы полагаем, что на примере таких «опор» у обучающегося формируются и так называемые *полезные траектории движения* [26, с. 161].

Благодаря поиску в подкорпусах, учащиеся могут подбирать примеры на материале, который близок их интересам. Опция, позволяющая формировать подкорпуса отдельных периодов, позволяет продемонстрировать, что язык меняется даже на протяжении вполне обозримого периода времени, — факт, который далеко не всегда очевиден и понятен учащимся.

Помимо использования уже имеющихся подкорпусов, студенты получают задания по созданию собственных подкорпусов, решая не столько технические и исполнительские,

сколько творческие задачи, которые формируют совокупность навыков для последующего создания уже своих авторских упражнений на основе корпусных технологий, для интеграции их в самостоятельно разработанные авторские курсы.

Задания на основе корпуса позволяют наглядно показать и проверить на практике тезис о том, что язык развивается. Составление авторского подкорпуса с последующим поиском в нем должно опираться на самую логику языковой системы, или ее срез в определенный временной интервал. Например, можно задать подкорпус художественных произведений первой половины XIX века, но искать в нем те языковые реалии, которые появились существенно позже (или наоборот, уже исчезли к началу XIX века), или грамматические формы слов, которые отсутствовали в языке того времени, неправильно и бессмысленно. Если для магистрантов с базовой лингвистической подготовкой, поступивших в магистратуру после получения диплома бакалавра лингвистики такие выводы кажутся очевидными благодаря пройденным курсам по языкознанию, лексикологии, типологии, теоретической грамматике, то для магистрантов, не имеющих базового лингвистического образования, это не так просто осознать и практически реализовать.

По нашему мнению, задания следующего характера позволяют лучше понять смысл такого поиска: *Создайте подкорпус русской поэзии первой половины XIX века (1800-1850). Посмотрите, употребляется ли в этом подкорпусе слово «кофе» в среднем роде. Сравните результаты и составьте таблицу частотности употребления форм мужского и среднего рода.* Студенту ставится задача на сопоставление результатов поиска с современным употреблением этой лексемы. С одной стороны, очевидно, что форма среднего рода в языке того времени будет «перекрываться» формой мужского рода. С другой стороны, студенту предлагается проанализировать встречаемость формы среднего рода и объяснить, почему именно в тех или иных контекстах она присутствует в языке «пушкинской эпохи». Иными словами, на таких заданиях магистранты учатся оценивать каждое языковое явление не дискретно, не как констатацию факта, а как неотъемлемую часть динамики языка, эволюции языковой системы, и учим объяснять найденные результаты с опорой на эти знания.

При знакомстве освоении корпусных технологий, на наш взгляд, необходимы небольшие лекционные вставки о теории языка и

его свойствах, где-то и в чем-то повторяющие пройденный студентами материал в курсе «Общее языкознание и история лингвистических учений», а где-то вводящие качественно новую информацию, необходимую для эффективной работы в корпусе/проведения корпусных исследований.

Здесь можно привести следующие примеры заданий:

1. *Найдите в подкорпусе текстов документальной прозы середины XX века предлог «ради»: отберите все контексты, где он используется в постпозиции по отношению к существительному, составьте частотный список таких существительных.*

2. *Проанализируйте предлог «по-над» на предмет его функциональности в художественных текстах второй половины XX века.*

3. *Найдите в основном корпусе примеры сочетания союза «как» с названиями животных. Выпишите один пример, созданный на основе коннотаций, и один пример сравнения, созданного на примере авторских ассоциаций.*

4. *Проанализируйте слово «покойник» в художественных текстах XIX и XX веков. Выявите, как изменилась семантика данной единицы на протяжении 200 лет.*

5. *Найдите в публицистических текстах первой половины XX века слова «ворота» и «врата», сравните их контекстуальные встречаемости и проследите, в каком значении употребляется каждое из них.*

6. *Проанализируйте лексику «бандформирование» в газетном подкорпусе текстов периода 1990-2000 года с целью выявления ее коллокаций / сочетаемости с прилагательными и глаголами.*

7. *Изучите раздел про дательно-субъектное отношение. Обратите внимание, какие в этой конструкции признаки у зависимого слова. Задайте поиск примеров с этим отношением.*

8. *Найдите коррелятивы «Кто..., тот...» Например: Кто историю изучает, тому известно..., Кто хорошо ходил на занятия, тот быстро получил зачет... и пр.*

9. *Создайте подкорпус: Язык рекламы. Найдите сочетания «незаменимый» + существительное. Отберите идиоматические выражения.*

При работе с синтаксическим уровнем важно научить / обратить внимание студентов на необходимость задействовать несколько полей в поисковой системе, причем расставить нужные параметры следует также исходя из логики самой искомой синтаксической структуры. Так, в одном из заданий требовалось найти контексты по следующей модели: 'до + сущ.

род. + глаг. + числит.' (например, *до зарплаты осталось пятьдесят рублей*):

На этом полезно разъяснить термин «шум». В принципе с этим явлением студенты уже познакомились, когда делали упражнение на *бодрый* с животными мужского рода, и им попались контексты с писателем Гоголем. Это грамматическая омонимия, и студенты этот термин усвоили. Но многие исследователи называют ее также шумом, который, в частности, встретился в результатах:

*От аэродрома в Тушино до Приморья почти десять тысяч километров* [Андрей Островский. Воспитать врага // «Владивосток», 2003.09.29].

Наречие *почти* корпус интерпретировал как повелительное наклонение леммы *почитать*. В других результатах попало большое количество словосочетаний, содержащих глагол *оставалось (осталось)*.

Поскольку наши студенты активно работают с английским языком, то занятия предполагают также использование возможностей параллельного корпуса в НКРЯ. НКРЯ уникален тем, что содержит параллельные корпуса по 11 парам языков, включая пары английский – русский и немецкий – русский, в которые входят оригинальные тексты как на русском языке с переводом на ИЯ, так и на иностранном языке с переводом на русский.

Эти задания призваны сформировать у студентов дополнительные навыки работы с иноязычным текстом, глубже проанализировать авторские варианты перевода, дать им оценку и сравнить со своими вариантами перевода. Например, в таких заданиях, как «*Какие английские выражения соответствуют русскому выражению «клевать носом»?»* и «*Задайте подкорпус текстов, переведенных с английского на русский. Выясните, какие английские выражения переводятся на русский наречием «виновато»»*, студенту предлагается не просто найти нужную информацию, но и проанализировать ее с точки зрения адекватности и валидности перевода. Учащийся должен уметь оценить и перспективы использования этого материала в педагогической деятельности, особенно связанной с преподаванием русского языка как иностранного. Такие задания лучше раскрывают МДС, на основе которых нагляднее формируются педагогические и ИКТ компетенции – что, еще раз подчеркнем, обуславливает специфику ООП «Компьютерная лингводидактика» и отличает ее от других программ.

Недостаток времени не всегда позволяет вывести студентов на качественно новый уровень работы с корпусом: выполнение заданий, требующих вдумчивого и последовательного

анализа языковых реалий. Здесь предполагается, что студент овладел инструментарием корпуса и может свободно проводить поиск по семантическим и грамматическим параметрам, разбираясь в контекстах и сравнивая их между собой. Суть заданий заключается не просто в том, представить какие-то результаты поиска, а в том, чтобы доказать, что эти результаты действительно соответствуют состоянию языковой системы в определенный промежуток времени – как по статистике поиска, так и по содержанию контекстуальных связей с учетом конкретной эпохи.

Кроме того, в формулировке заданий не говорится, какой именно подкорпус должен создать студент: ему необходимо самому решить, где ему нужно будет искать. Возможно, что студенту придется «перепробовать» несколько вариантов созданных им подкорпусов, но в конечном итоге, он должен выбрать и продемонстрировать кратчайший путь к решению проблемы. Например:

1. *Пользуясь поэтическим корпусом, докажете, что в XVIII веке слово «приличен» могло иметь ударение на последнем слоге: «приличЕн». Приведите этот пример (эти примеры).*

2. *Из социальной сети Facebook нередко приходит сообщение: «Интересуюсь о Вас на Facebook. Здравствуйте, Natasha! Я Вас приглашаю посмотреть мою личную страницу на Facebook...». Воспользовавшись корпусом, оцените, насколько конструкция «интересоваться о ком-либо» распространена, в каких текстах она встречается. Сравните ее частотность с частотностью конструкции «интересоваться кем-либо».*

3. *Сегодня можно сказать «Банк не имеет право отказать в открытии счёта» и «Банк не имеет права отказать в открытии счёта». А какой из этих вариантов является более старым? Задание. Найдите самые ранние и поздние примеры употребления винительного падежа в этой конструкции («не иметь право») и родительного («не иметь права»).*

4. *Когда в Корпусе появляются первые примеры глагола «реагировать»? Приведите первые два примера. Что они значат? Когда появляются первые примеры глагола «реагировать» с какой-либо приставкой? Приведите первые два примера.*

Заключительный этап освоения НКРЯ / работы с НКРЯ предполагает реализацию всех полученных знаний и умений в процессе выполнения заданий на этапах, рассмотренных выше. На этом этапе перед студентом могут поставлены задачи методического характера, а именно – создание своих собственных

упражнений на разные темы, что способствует дальнейшему развитию педагогических компетенций, сформированных при изучении дисциплин профиля «Компьютерная лингводидактика».

### Заключение

Проведенное нами исследование отражает вопросы, связанные с ролью КЛ в подготовке специалистов в области Компьютерной лингводидактики. Анализ РПД дисциплин показал, что три дисциплины содержат модули, посвященные КЛ. Обращение к методам КЛ происходит также в рамках рассредоточенной научно-исследовательской работы (НИР) магистрантов и некоторых выпускных квалификационных работ. В ряде дисциплин, таких как Практикум по культуре речевого общения первого / второго иностранного языка, Практика перевода английского языка (технический / гуманитарный профиль). Практикум по межкультурной коммуникации немецкого / французского / испанского языка, Научный курс подходы КЛ могут использоваться как дополнительные при рассмотрении определенных тем этих курсов.

Выполнение поисковых заданий в корпусе позволяет в новом ракурсе посмотреть на некоторые темы / положения, изученные в теоретических дисциплинах, таких как «Общее языкознание и история лингвистических учений». С другой стороны, исследования показывают, что интенсивного семестрового курса, направленного на освоение подходов КЛ в обучении ИЯ для будущих преподавателей ИЯ, недостаточно для уверенного овладения необходимыми техническими, корпусно-лингвистическими и педагогическими знаниями, связанными с использованием подхода DDL в обучении ИЯ. А. Ленко-Шиманска считает, что для этого необходимо, чтобы студенты выполняли поисковые и исследовательские задания в корпусах при изучении разных дисциплин [54].

Профессионально организованная опора на ресурсы такого масштаба открывает перспективы создания учебных пособий совершенно нового типа [55]: современных и минимально адаптированных, которые могут быть комплексно интегрированы в разные образовательные платформы.

Логично продолжить исследовать некоторые теоретические аспекты, например, роль КЛ в программе подготовки компьютерных лингводидактов с точки зрения междисциплинарной парадигмы.

Все это позволяет заключить, что для более эффективного использования подходов КЛ в разных дисциплинах программы *Компьютерная лингводидактика* целесообразно разработать единое пособие, содержащее конкретные задания на обращение к корпусам в рамках разных дисциплин Программы «Компьютерная лингводидактика» и рекомендации для преподавателей. Обращение к разным корпусам способствует формированию лингвисти-

ческой, компьютерной, исследовательской компетенции студентов. Наличие единого пособия позволит оптимизировать процесс формирования указанных компетенций в рамках разных дисциплин.

«Объединив усилия, корпусная лингвистика и лингводидактика смогут найти адекватное и пропорциональное отражение безграничной стихии языка в ограниченном по объему и доступном для усвоения учебном пособии» [54, с. 115].

## СПИСОК ЛИТЕРАТУРЫ

1. Плуныян В.А. Корпус как инструмент и как идеология: О некоторых уроках современной корпусной лингвистики // Русский язык в научном освещении. 2008. № 2 (16). С. 7–20.
2. Corpus linguistics. International conference. 25–28 July 2017, University of Birmingham. Abstract book ed. S. Hunston and M. Nasseri 2017. URL: <http://paulslals.org.uk/ccr/CL2017ExtendedAbstracts.pdf>.
3. Michel J.-B., Shen Y.K., Aiden A.P., Veres A., Gray M.K., Pickett J.P., Hoiberg D., Clancy D., Norvig P., Orwant J., Pinker S., Nowak M.A., Aiden E.L. Quantitative analysis of culture using millions of digitized books // Science. No 331(6014). 2011. Pp. 176–182.
4. Масевич А.Ц., Захаров В.П. Методы корпусной лингвистики в исторических и культурологических исследованиях // Труды XIX Международной объединенной научной конференции «Компьютерная лингвистика и вычислительные онтологии». 2016. С. 24–43.
5. Kilgarriff A., Charalabopoulou F., Gavrilidou M., Johannessen J.B., Khalil S., Kokkinakis S.J., Lew R., Sharoff S., Vádlapudi R., Volodina E. Corpus-based vocabulary lists for language learners for nine languages Language Resources and Evaluation. 2014. Vol. 48 (1). Pp. 121–163.
6. Kilgarriff A., Baisa V., Bušta J., Jakubíček M., Kovář V., Michelfeit J., Rychlý P., Suchomel V. The Sketch Engine: ten 2019s on. // Lexicography. ASIALEX. 2014. Vol.1. No1. Pp. 7–36. DOI: 10.1007/s40607-014-0009-9
7. Benko V. Aranea: Yet Another Family of (Comparable) Web Corpora // In: P. Sojka et al. (Eds.) Text, Speech and Dialogue. TSD 2014. Lecture Notes in Computer Science. 2014. Vol. 8655. Springer, Cham. Pp. 247–254.
8. The Handbook of English for Specific Purposes / Ed. by B. Paltridge and Sue Starfield. Oxford: Wiley-Blackwell. 2013.
9. Алексеева Л.Е. Методика обучения профессионально ориентированному иностранному языку: Курс лекций: Метод. пособие. СПб.: Филол. ф-т СПбГУ, 2007.
10. Huhta M., Vogt K., Johnson E., Tulkki H. Needs Analysis for Language Course Design: A holistic approach to ESP / Ed. by D. Holl. Cambridge: Cambridge Univ. Press. 2013.
11. Bartlett F.C. Remembering: A Study in Experimental Social Psychology. Cambridge: Cambridge Univ. Press, 1932.
12. Выготский Л.С. Педагогическая психология. М.: Педагогика-Пресс, 1999.
13. Залевская А.А. Психолингвистические исследования. Слово. Текст: Избр. труды. М.: Гнозис, 2005.
14. Скребцова Т.Г. Когнитивная лингвистика: Курс лекций. СПб.: Филол. ф-т СПбГУ, 2011.
15. Römer U. Corpora and language teaching // Origin and history of corpus linguistics – corpus linguistics vis-à-vis other disciplines (ed. A. Lüdeling M. Kytö). 2009. Pp. 112–131. Berlin, Germany: Walter de Gruyter GmbH & Co. KG, 10785
16. Раицкая Л.К. Дидактические возможности корпусных интернет-технологий в преподавании иностранного языка в высшей школе // Вестник Московского государственного областного университета. Серия: Педагогика. 2009. № 4. С. 123–127.
17. Sinclair J. Reading Concordances. An Introduction. London: Pearson Longman, 2003.
18. Shaw E.M. Teaching Vocabulary Through Data-driven Learning, 2011. URL: <https://www.semanticscholar.org/paper/Teaching-Vocabulary-Through-Data-Driven-LearningShaw/d1758aa87bcf9211957f1f5738d4d39bd38c2f72>
19. Коган М.С. Пути интегрирования подходов корпусной лингвистики в программу подготовки переводчиков // VII международная научно-практическая конференция «Перевод. Язык. Культура». 2016. С. 215–219.
20. Коган М.С., Куликова Е.В. Корпус как ресурс для изучения лексико-грамматических аспектов английского языка для специальных целей // IX международная научно-практическая конференция «Перевод. Язык. Культура». 2018. С. 97–103.
21. Johns T. Should you be persuaded – Two samples

of data-driven learning materials // In T. Johns & P. King (Eds.) *Classroom Concordancing*. Birmingham University [ELR Journal, 4]. 1991. Pp. 1-16.

22. **Boulton A., Cobb T.** Corpus use in language learning: A meta-analysis // *Language Learning*. 2017. Vol. 67(2), Pp.348–393.

23. **Агафонова Л.И.** Некоторые вопросы использования корпусных технологий как фактора повышения качества обучения иностранным языкам // *Известия Российского государственного педагогического университета имени А.И. Герцена*. СПб., 2009. № 87. С. 80-88.

24. **Садовникова О.Э.** Прямое и косвенное использование корпусов в зарубежной лингводидактике // *Научно-педагогический журнал Восточной Сибири Magister Dixit*. 2013. № 2. С. 152–161.

25. **Kogan M., Kolotaeva A., Yaroshevich A., Zakharov V.** On prospects of using the DDL approach in GSP course // *The European Proceedings of Social & Behavioural Sciences EpSBS*. 2018. Pp. 1763–1775.

26. **Дикарева С.С.** Корпусные технологии в режиме диалога «преподаватель — студент-исследователь» // *Труды международной конференции «Корпусная лингвистика — 2011»*, 27–29 июня 2011 г., Санкт-Петербург. С.-Петербургский гос. университет. Филологический факультет. 2011. С. 157–162.

27. **Cobb T., Boulton A.** Classroom applications of corpus analysis // In (Eds) *Cambridge Handbook of English Corpus Linguistics* (ed. D. Biber & R. Reppen). 2015. Pp. 478-497. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139764377.027. 2015

28. **Hyland K.** Writing without conviction? Hedging in science research articles, *Applied Linguistics*. 1996. Vol. 17(4). Pp. 433-454.

29. **Safronenkova E.L.** Hedging vs tolerance in presenting the scientific result in research articles (based on English research articles of the humanities field), *St. Petersburg State Polytechnical University Journal. Humanities and Social Sciences*, 2019. Vol. 10, № 3, Pp. 51-57. DOI: 10.18721/JHSS.10305

30. **Захаров В.П., Богданова С.Ю.** Корпусная лингвистика: Учебник — 2-е изд. СПб: СПбГУ. РИО. Филологический факультет, 2013.

31. **Anthony L.** A critical look at software tools in corpus linguistics // *Linguistic Research*, 2013. Vol. 30, No 2, Pp. 141–161.

32. **Brezina V.** Collocation graphs and networks: Selected applications // *Lexical collocation analysis: Advances and applications / ed. P. Cantos-Gómez & M. Almela-Sánchez*. Cham, Switzerland: Springer International. 2018. Pp. 59–83.

33. **Коган М.С., Ярошевич А.М., Колотаева А.Ю., Захаров В.П., Шрот-Вихерт З., Тильманс А.** К проблеме создания списка высококачественных слов и выражений немецкого языка для

специальных целей // *Компьютерная лингвистика и вычислительные онтологии*. Выпуск 2 (Труды XXI Международной объединенной конференции «Интернет и современное общество, IMS-2018, Санкт-Петербург, 31 мая – 2 июня 2018 г. Сборник научных статей). СПб: Университет ИТМО, 2018. С. 44–55.

34. **Соснина Е.П.** Корпусная лингвистика и корпусный подход в обучении иностранному языку // *Прикладная лингвистика: официальный сайт Ульяновского государственного технического университета*. URL: [http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus\\_linguistics\\_language\\_teaching/](http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_linguistics_language_teaching/)

35. **Богданова С.Ю.** Возможности корпусной методологии в решении лингвистических задач // *Вестник ИГЛУ*. № 2 (18). 2012. С.47–50.

36. **Добрушина Н.Р.** Корпусная методика обучения русскому языку [Corpus-based approach in teaching the Russian language] // *Национальный корпус русского языка*. 2006–2008. Новые результаты и перспективы. СПб.: Нестор-История. 2009. С. 338–351.

37. **Горина О.Г.** К вопросу о корпусном отборе ключевых лексических единиц // *Stephanos*. 2017. № 1(21). С. 111–117.

38. **Титова С.В., Самойленко О.Ю.** Структура информационно-коммуникационной компетенции преподавателя вуза // *Вестник Тамбовского университета*. Серия: Гуманитарные науки. 2017. Т. 22. №3 (167). С. 39–48.

39. **Иовлева В.И., Попова Н.В.** Междисциплинарная модель обучения иностранному языку как отражение интегративных тенденций современного образовательного пространства // *Актуальные проблемы языкознания*. 2015. Т. 1. С. 129–130.

40. **Халыпина Л.П., Попова Н.В., Кузнецова О.В.** Междисциплинарное проектирование как средство реализации интегрированного предметно-языкового обучения в вузе // *Научно-технические ведомости СПбГПУ. Гуманитарные и общественные науки*. 2017. Т. 8, № 3. С. 145–152. DOI: 10.18721/JHSS.8314

41. **Акопова М.А., Попова Н.В.** Междисциплинарные связи в высшей школе. Санкт-Петербург. 2014.

42. **Попова Н.В., Коган М.С., Вдовина Е.К.** Предметно-языковое интегрированное обучение (CLIL) как методология актуализации междисциплинарных связей в техническом вузе // *Вестник Тамбовского университета*. Серия: Гуманитарные науки. 2018. Т. 23, № 173 С.29–42.

43. **Vyatkina N.** Data-driven learning for beginners: The case of German verb-preposition collocations // *ReCALL*. 2016. Vol. 28, No.2. Pp. 207-226. DOI: 10.1017/S0958344015000269

44. **Vyatkina N.** Data-driven learning of collocations: Learner performance, Proficiency, and Perceptions // *Language Learning & Technology*. 2016. Vol. 20, No.3. Pp. 159–179. <http://llt.msu.edu/issues/october2016/vyatkina.pdf>
45. **O'Sullivan Í, Chambers A.** Learners' writing skills in French: Corpus consultation and learner evaluation // *Journal of Second Language Writing*. 2006. 15(1), 49–68. DOI: 10.1016/j.jslw.2006.01.002
46. **Lavid J.** Contrastive patterns of mental transitivity in English and Spanish: A student-centred corpus-based study // In E. Hidalgo, L. Quereda & J. Santana (Eds.) *Corpora in the foreign language classroom*. (pp. 237-252). Amsterdam: Rodopi. 2007.
47. **Kogan M., Yaroshevich A., Ni O.** Corpus-Based Teaching German Compound Nouns and Lexical Bundles for Improving Academic Writing Skills of Engineering Students// *Lidil*. 2018. № 58. Pp. 1–30. URL :<http://journals.openedition.org/lidil/5438>
48. **Baker P.** *Using Corpora in Discourse Analysis*. London: Continuum. 2006.
49. **Hyland K.** *Corpus Informed Discourse Analysis: The Case of Academic Engagement* // *Academic Writing: At the Interface of Corpus and Discourse* (ed. M. Charles, D. Pecorari, S. Hunston). London, U.K.: Continuum International Publishing Group, pp. 110–128. ISBN 9781847064363
50. **Чернявская В.Е.** Дискурсивный анализ и корпусные методы: необходимое доказательное звено? объяснительные возможности качественного и количественного подходов // *Вопросы когнитивной лингвистики*. 2018. № 2 (55). С. 31–37.
51. **Добрушина Н.Р.** Как использовать Национальный корпус русского языка в образовании? // *Национальный корпус русского языка: 2003–2005*. М.: Индрик, 2005. С. 308-329.
52. **Слободян Е.А.** Обучение методике составления упражнений с помощью Национального корпуса русского языка // *Гуманитарные исследования в Восточной Сибири и на Дальнем Востоке*. 2010. №1.
53. **Ольховская А.И., Парамонова М.К.** Корпус в преподавании русского языка и литературы // *Rhema*. Рема. 2019. № 1. С. 93–123.
54. **Leńko-Szymańska A.** Training teachers in data-driven learning: Tackling the challenge // *Language Learning & Technology*. 2017. Vol. 21 (3) Pp. 217–241.
55. **Сибирцева В.Г., Хоменко А.Ю., Баранова Ю.Н.** Национальный корпус русского языка как основа новаторских электронных учебников // *Образовательные технологии и общество*. 2013. Т. 16. № 3. С. 508-521.

*Статья поступила в редакцию 17.10.2019*

## REFERENCES

- [1] **V.A. Plungyan**, Korpus kak instrument i kak ideologiya: O nekotorykh urokakh sovremennoy korpusnoy lingvistiki [Corpus as a tool and ideology: Some lessons of modern corpus linguistics]// *Russkii yazyk v nauchnom osveshchenii*. 2 (16) (2008) 7–20.
- [2] *Corpus linguistics. International conference. 25-28 July 2017, University of Birmingham. Abstract booked*. S. Hunston and M. Nasser. 2017. URL: <http://paulslals.org.uk/ccr/CL2017ExtendedAbstracts.pdf>
- [3] **J.-B. Michel, Y.K. Shen, A.P. Aiden, A. Veres, M.K. Gray, J.P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M.A. Nowak, E.L. Aiden**, Quantitative analysis of culture using millions of digitized books, *Science*. 331 (6014) (2011) 176–182.
- [4] **A.Ts. Masevich, V.P. Zakharov**, Metody korpusnoy lingvistiki v istoricheskikh i kulturologicheskikh issledovaniyakh [Corpus linguistics methods in historical and culturological studies], *Trudy XIX Mezhdunarodnoy obyedinennoy nauchnoy konferentsii «Kompyuternaya lingvistika i vychislitelnyye ontologii»*. 2016. S. 24–43.
- [5] **A. Kilgarriff, F. Charalabopoulou, M. Gavrilidou, J.B. Johannessen, S. Khalil, S.J. Kokkinakis, R. Lew, S. Sharoff, R. Vadlapudi, E. Volodina**, Corpus-based vocabulary lists for language learners for nine languages *Language Resources and Evaluation*. 48 (1) (2014) 121–163.
- [6] **A. Kilgarriff, V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, V. Suchomel**, The Sketch Engine: ten years on, *Lexicography*. ASIALEX. 1(1) (2014) 7–36. DOI: 10.1007/s40607-014-0009-9
- [7] **V. Benko**, Aranea: Yet Another Family of (Comparable) Web Corpora, In: P. Sojka et al. (Eds.) *Text, Speech and Dialogue*. TSD 2014. Lecture Notes in Computer Science. Springer, Cham. 8655 (2014) 247–254.
- [8] *The Handbook of English for Specific Purposes* / Ed. by B. Paltridge and Sue Starfield. Oxford: Wiley-Blackwell. 2013.
- [9] **L.Ye. Alekseyeva**, Metodika obucheniya professionalno oriyentirovannomu inostrannomu yazyku: Kurs lektsiy: Metod. posobiye. [Methodology of teaching professionally-oriented foreign language: A course of lectures: Manual]. SPb.: Filol. f-t SPbGU, 2007.
- [10] **M. Huhta, K. Vogt, E. Johnson, H. Tulkki**, Needs Analysis for Language Course Design: A holistic approach to ESP / Ed. by D. Holl. Cambridge: Cambridge Univ. Press. 2013.

- [11] **F.C. Bartlett**, *Remembering: A Study in Experimental Social Psychology*. Cambridge: Cambridge Univ. Press, 1932.
- [12] **L.S. Vygotskiy**, *Pedagogicheskaya psikhologiya* [Pedagogical psychology]. M.: Pedagogika-Press, 1999.
- [13] **A.A. Zalevskaya**, *Psikholingvisticheskiye issledovaniya. Slovo. Tekst: Izbr. trudy* [Psycholinguistic studies. Word. Text: Selecta], M.: Gnozis, 2005.
- [14] **T.G. Skrebtsova**, *Kognitivnaya lingvistika: Kurs lektsiy* [Cognitive linguistics: A course of lectures]. SPb.: Filol. f-t SPbGU, 2011.
- [15] **U. Römer**, *Corpora and language teaching, Origin and history of corpus linguistics – corpus linguistics vis-à-vis other disciplines* (ed. A. Lüdeling M. Kytö). 2009. R. 112–131. Berlin, Germany: Walter de Gruyter GmbH & Co. KG, 10785.
- [16] **L.K. Raitskaya**, *Didakticheskiye vozmozhnosti korpusnykh internet-tekhnologiy v prepodavanii inostrannogo yazyka v vysshey shkole* [Didactic uses of corpus-related technologies in teaching a foreign language in higher school], *Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta. Seriya: Pedagogika*. 4 (2009) 123–127.
- [17] **J. Sinclair**, *Reading Concordances. An Introduction*. London: Pearson Longman, 2003.
- [18] **E.M. Shaw**, *Teaching Vocabulary Through Data-driven Learning*, 2011. URL: <https://www.semanticscholar.org/paper/Teaching-Vocabulary-Through-Data-Driven-LearningShaw/d1758aa87bcf9211957f1f5738d4d39bd38c2f72>
- [19] **M.S. Kogan**, *Puti integrirovaniya podkhodov korpusnoy lingvistiki v programmu podgotovki perevodchikov* [Ways of integrating approaches of corpus linguistics into the translators training programme], VII mezhhdunarodnaya nauchno-prakticheskaya konferentsiya «Perevod. Yazyk. Kultura». 2016. Pp. 215–219.
- [20] **M.S. Kogan, Ye.V. Kulikova**, *Korpus kak resurs dlya izucheniya leksiko-grammaticheskikh aspektov angliyskogo yazyka dlya spetsialnykh tseley* [Corpus as a resource for lexico-grammar aspect of ESP], IX mezhhdunarodnaya nauchno-prakticheskaya konferentsiya «Perevod. Yazyk. Kultura». 2018. Pp.97–103.
- [21] **T. Johns**, *Should you be persuaded – Two samples of data-driven learning materials*, In T. Johns & P. King (Eds.) *Classroom Concordancing*. Birmingham University [ELR Journal, 4]. 1991. Pp. 1–16.
- [22] **A. Boulton, T. Cobb**, *Corpus use in language learning: A meta-analysis*, *Language Learning*. 67 (2) (2017) 348–393.
- [23] **L.I. Agafonova**, *Nekotoryye voprosy ispolzovaniya korpusnykh tekhnologiy kak faktora povysheniya kachestva obucheniya inostrannym yazykam* [Some issues of corpus technologies as a quality development factor for teaching foreign languages ], *Izvestiya Rossiyskogo gosudarstvennogo pedagogicheskogo universiteta imeni A.I. Gertsena*. SPb. 87 (2009) 80–88.
- [24] **O.E. Sadovnikova**, *Pryamoye i kosvennoye ispolzovaniye korpusov v zarubezhnoy lingvodidaktike* [Direct and indirect corpora applications foreign language pedagogy], *Nauchno-pedagogicheskyy zhurnal Vostochnoy Sibiri Magister Dixit*. 2 (2013) 152–161.
- [25] **M. Kogan, A. Kolotaeva, A. Yaroshevich, V. Zakharov**, *On prospects of using the DDL approach in GSP course*, *The European Proceedings of Social&Behavioural Sciences EpSBS*. 2018. Pp. 1763–1775.
- [26] **S.C. Dikareva**, *Korpusnyye tekhnologii v rezhime dialoga «prepodavatel – student-issledovatel»* [Corpus technologies in conversational mode “teacher – student-researcher”], *Trudy mezhdunarodnoy konferentsii «Korpusnaya lingvistika – 2011»*, 27–29 iyunya 2011 g., Sankt-Peterburg. S.-Peterburgskiy gos. universitet. Filologicheskyy fakultet. 2011. Pp. 157–162.
- [27] **T. Cobb, A. Boulton**, *Classroom applications of corpus analysis*, In (Eds) *Cambridge Handbook of English Corpus Linguistics* (p/red. D. Biber& R. Reppen). 2015. Pp. 478–497. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139764377.027. 2015
- [28] **K. Hyland**, *Writing without conviction? Hedging in science research articles*, *Applied Linguistics*. 17 (4) (1996) 433–454.
- [29] **E.L. Safronenkova**, *Hedging vs tolerance in presenting the scientific result in research articles (based on English research articles of the humanities field)*, *St. Petersburg State Polytechnical University Journal. Humanities and Social Sciences*, 10 (3) (2019) 51–57. DOI: 10.18721/JHSS.10305
- [30] **V.P. Zakharov, S.Yu. Bogdanova**, *Korpusnaya lingvistika : Uchebnik – 2-ye izd.* [Corpus linguistics: textbook, 2-nd edition] SPb: SPbGU. RIO. Filologicheskyy fakultet, 2013.
- [31] **L. Anthony**, *A critical look at software tools in corpus linguistics*, *Linguistic Research*, 30 (2) (2013) 141–161.
- [32] **V. Brezina**, *Collocation graphs and networks: Selected applications*, *Lexical collocation analysis: Advances and applications*, ed. P. Cantos-Gómez & M. Almela-Sánchez. Cham, Switzerland: Springer International. 2018. Pp. 59–83.
- [33] **M.S. Kogan, A.M. Yaroshevich, A.Yu. Kolotayeva, V.P. Zakharov, Z. Shrot-Vikhter, A. Tilmans**, *K probleme sozdaniya spiska vysokochastotnykh slov i vyrazheniy nemetskogo yazyka dlya spetsialnykh tseley* [On problem of creating a German engineering academic words and formulas list], *Kompyuternaya lingvistika i vychislitelnyye*

ontologii. Vypusk 2 (Trudy XXI Mezhdunarodnoy obyedinennoy konferentsii «Internet i sovremennoye obshchestvo, IMS-2018, Sankt-Peterburg, 31 maya - 2 iyunya 2018. Sbornik nauchnykh statey). SPb: Universitet ITMO, 2018. Pp. 44–55.

[34] **Ye.P. Sosnina**, Korpusnaya lingvistika i korpusnyy podkhod v obuchenii inostrannomu yazyku [Corpus linguistics and corpus-based approach in foreign language teaching], *Prikladnaya lingvistika: ofitsialnyy sayt Ulyanovskogo gosudarstvennogo tekhnicheskogo universiteta*. URL: [http://ling.ulstu.ru/linguistics/resourses/literature/articles/corpus\\_linguistics\\_language\\_teaching/](http://ling.ulstu.ru/linguistics/resourses/literature/articles/corpus_linguistics_language_teaching/)

[35] **S.Yu. Bogdanova**, Vozmozhnosti korpusnoy metodologii v reshenii lingvisticheskikh zadach [Possibilities of Corpus methodology in solving linguistic problems], *Vestnik IGLU*. 2 (18) (2012) 47–50.

[36] **N.R. Dobrushina**, Korpusnaya metodika obucheniya russkomu yazyku [Corpus-based approach in teaching the Russian language], *Natsionalnyy korpus russkogo yazyka. 2006–2008. Novyye rezultaty i perspektivy*. SPb.: Nestor-Istoriya. 2009. Pp. 338–351.

[37] **O.G. Gorina**, K voprosu o korpusnom otbore klyuchevykh leksicheskikh yedinit [Towards corpus-informed computing of key words], *Stephanos*. 1 (21) (2017) 111–117.

[38] **S.V. Titova, O.Yu. Samoilenko**, Struktura informatsionno-kommunikatsionnoy kompetentsii prepodavatelya vuza [Structure of higher educational institution lecturer's information and communication competence], *Vestnik Tambovskogo universiteta*. Seriya: Gumanitarnyye nauki. 22 (3) (167) (2017) 39–48.

[39] **V.I. Iovleva, N.V. Popova**, Mezhdistsiplinarnaya model obucheniya inostrannomu yazyku kak otrazheniye integrativnykh tendentsiy sovremennogo obrazovatel'nogo prostranstva [Interdisciplinary model of foreign language teaching as reflection of integrating trends in modern educational environment], *Aktualnyye problemy yazykoznaneya*. 1 (2015) 129–130.

[40] **L.P. Khalyapina, N.V. Popova, O.V. Kuznetsova**, Interdisciplinary design as a means of implementing content and language integrated learning in higher education, *St. Petersburg State Polytechnical University Journal. Humanities and Social Sciences*, 8 (3) (2017) 145–152. DOI: 10.18721/JHSS.8314

[41] **M.A. Akopova, N.V. Popova**, Mezhdistsiplinarnyye svyazi v vysshey shkole [Interdisciplinary links in higher education]. Sankt-Peterburg. 2014.

[42] **N.V. Popova, M.S. Kogan, Ye.K. Vdovina**, Predmetno-yazykovoye integrirovannoye obucheniye (CLIL) kak metodologiya aktualizatsii mezhdistsiplinarnykh svyazey v tekhnicheskoye vuze

[Content and language integrated learning (CLIL) as actualization methodology of interdisciplinary links in technical university], *Vestnik Tambovskogo universiteta*. Seriya: Gumanitarnyye nauki. 23 (173) (2018) 29–42.

[43] **N. Vyatkina**, Data-driven learning for beginners: The case of German verb-preposition collocations, *ReCALL*. 28 (2) (2016) 207–226. DOI:10.1017/S0958344015000269

[44] **N. Vyatkina**, Data-driven learning of collocations: Learner performance, Proficiency, and Perceptions, *Language Learning & Technology*. 20 (3) (2016) 159–179. URL: <http://llt.msu.edu/issues/october2016/vyatkina.pdf>

[45] **Í. O'Sullivan, A. Chambers**, Learners' writing skills in French: Corpus consultation and learner evaluation, *Journal of Second Language Writing*. 15 (1) (2006) 49–68. DOI: 10.1016/j.jslw.2006.01.002

[46] **J. Lavid**, Contrastive patterns of mental transitivity in English and Spanish: A student-centred corpus-based study, In E. Hidalgo, L. Quereda & J. Santana (Eds.) *Corpora in the foreign language classroom*. (pp. 237–252). Amsterdam: Rodopi. 2007.

[47] **M. Kogan, A. Yaroshevich, O. Ni**, Corpus-Based Teaching German Compound Nouns and Lexical Bundles for Improving Academic Writing Skills of Engineering Students, *Lidil*. 58 (2018) 1–30. URL: <http://journals.openedition.org/lidil/5438>

[48] **P. Baker**, *Using Corpora in Discourse Analysis*. London: Continuum. 2006.

[49] **K. Hyland**, *Corpus Informed Discourse Analysis: The Case of Academic Engagement, Academic Writing: At the Interface of Corpus and Discourse* (p/r M. Charles, D. Pecorari, S. Hunston). London, U.K.: Continuum International Publishing Group, pp. 110–128. ISBN 9781847064363

[50] **Chernyavskaya V.Ye.** Diskursivnyy analiz i korpusnyye metody: neobkhodimoye dokazatel'noye zveno? obyasnitel'nyye vozmozhnosti kachestvennogo i kolichestvennogo podkhodov [Discourse analysis and corpus approaches: a missing evidence-based link? Towards qualitative and quantitative approaches in language studies.], *Voprosy kognitivnoy lingvistiki*. 2 (55) (2018) 31–37.

[51] **N.R. Dobrushina**, Kak ispolzovat Natsionalnyy korpus russkogo yazyka v obrazovanii? [How should the Russian National Corpus be used in teaching?], *Natsionalnyy korpus russkogo yazyka: 2003–2005*. M.: Indrik, 2005. Pp. 308–329.

[52] **Ye.A. Slobodyan**, Obucheniye metodike sostavleniya uprazhneniy s pomoshchyu Natsionalnogo korpusa russkogo yazyka // Gumanitarnyye issledovaniya v Vostochnoy Sibiri i na Dalnem Vostoke. 1 (2010).

[53] **A.I. Olkhovskaya, M.K. Paramonova**, Korpus v prepodavanii russkogo yazyka i literatury [Corpus in



teaching the Russian language and literature], Rhema. Rema. 1 (2019) 93–123.

[54] **A. Leńko-Szymańska**, Training teachers in data-driven learning: Tackling the challenge, Language Learning & Technology. 21 (3) (2017) 217–241.

[55] **V.G. Sibirtseva, A.Yu. Khomenko, Yu.N. Baranova**, Natsionalnyy korpus russkogo yazyka kak osnova novatorskikh elektronnykh uchebnikov [The Russian National Corpus as the basis for innovative electronic textbooks], Obrazovatelnyye tekhnologii i obshchestvo. 16 (3) (2013) 508–521.

*Received 17.10.2019*

#### **СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS**

**Дмитриев Александр Владиславович**

**Dmitrijev Alexander V.**

E-mail: avd84@list.ru

**Коган Марина Самуиловна**

**Kogan Marina S.**

E-mail: m\_kogan@inbox.ru