

ИНСТРУМЕНТЫ КОРПУСНОГО АНАЛИЗА В ОБУЧЕНИИ ИНОСТРАННОМУ ЯЗЫКУ

Рассматриваются инструменты корпусного анализа, обосновывается целесообразность привлечения их к обучению иностранному языку, раскрывается лингвометодический потенциал корпусного инструментария при формировании лингвистической компетенции, описываются пути интеграции корпусных исследований в процесс обучения. Приводятся примеры корпусных заданий. Используются материалы англоязычных лингвистических компьютерных корпусов.

Ключевые слова: корпусная лингвистика; корпусная лингвостатистика; корпусные инструменты лингвистического анализа; корпусная компетенция; конкорданс; коллокации; семантическая просодия; лексическая плотность; статистическая наглядность; английский язык.

Введение

Сегодня использование корпусов в лингводидактике еще только складывается. Однако целесообразность содружества методики и лингвостатистики отмечалась многими авторами, начиная от И.В. Рахманова [1] и заканчивая сегодняшними корпусными лингводидактами (Н.Б. Гвишиани [2], R. Carter [3], D. Biber, R. Reppen [4, 5] и др.). Совершенно разумно, что достижения и наработки корпусной лингвистики, появление которой явилось значительным событием в истории науки, не может и не должно оставаться в рамках одной науки лингвистики. Одним из первых результатов интеграции корпусных методов стало изменение технологий создания словарей, вслед за которыми последовали значительные трансформации учебных материалов. Ни одно из уважаемых изданий учебной литературы на сегодняшний день уже не может обходиться без корпусов при отборе содержания учебника, устранения рассогласованности между содержанием обучения и объектами контроля, а также анализа ошибок обучаемых для непрерывного уточнения содержания обучения. Корпус как эффективная справочная система стал одним из обязательных форматов представления языковых данных в совокупности с более традиционными грамматиками и словарями. Более того, очевидно, что традиционный методический принцип наглядности под воздействием корпусных грамматик сегодня превращается в *принцип статистической наглядности*. Визуализация лингвостатистических результатов в виде гистограмм, графиков и словарных областей стала неотъемлемой частью не только словарей, но и современных учебников английского языка. Таким образом, вслед за корпусной трансформацией в практике создания словарей существенно изменилась технология создания лингвометодических материалов, а также наметились новые лингвостатистические тенденции в способах предъявления языковой информации в учебной литературе.

Вместе с тем практика применения корпусных инструментов в аудиторной работе при обучении иностранному языку не приживается так же стремительно. Поэтому целью данной статьи является популяризация идеи корпусного обучения, информирование профессиональной аудитории о потенциале корпусного инструментария при обучении иностранному язы-

ку. В разделе 1 статьи последовательно рассматриваются инструменты корпусного анализа и релевантные им вопросы на основе уже имеющегося опыта их применения. Обоснованность отрефлексированных нами элементов работы с опорой на корпус подтверждается результатами опытно-экспериментального исследования, которое проводилось в группах студентов уровня владения языком B2+ по шкале Совета Европы. Следует отметить, что применение корпусного подхода в обучении не ограничивается указанным уровнем и во многом зависит от желания преподавателя приобретать корпусную компетенцию и включать в процесс обучения задания с опорой на корпус. Как корпусные инструменты, так и основанные на них задания вариативны и обладают существенной гибкостью в плане их применения в обучении. Иллюстративный материал разделов 1 и 2 дает более полное представление о категориях обучающихся и типах корпусных заданий, которые могут обогатить как аудиторную практику, так и самостоятельную работу студентов. В заключении обсуждается необходимость развития корпусной компетенции у обучаемых, а также ее совершенствования у преподавателей, которым, мы надеемся, будет полезен как наш собственный опыт, так и реферативный анализ методической литературы по обучению с опорой на корпус.

1. Корпусные инструменты

Мы рассмотрим те инструменты корпусного анализа, пришедшие к нам из корпусной лингвистики, которые могут стать основой для создания лингводидактических материалов. Они могут быть полезны преподавателям иностранного языка, но в первую очередь статья адресована тем, кто обучает иностранному языку будущих лингвистов и преподавателей, а также ведет профессионально направленное обучение иностранному языку.

1.1. Лингвистическая наглядность конкорданса.

К корпусным инструментам анализа в рамках обучения с помощью баз данных (Data-driven approach) традиционно относится *генерирование строк конкорданса*, которые представляют собой «нарезку текстов корпуса с рассматриваемой лексической единицей (ЛЕ)» [6. С. 13]. Т. Джонс, стоявший у истоков обучения с опорой на конкорданс, полагал, что обучаемый

должен совершать «открытия» (discovery learning) в изучаемом языке, а преподаватель создает условия, в которых эти открытия могут состояться. Таким образом, была предложена модель обучения, опирающаяся на лингвистическую наглядность конкорданса. Обучаемый получает доступ к фактам состоявшихся, отработанных речевым сообществом употреблений и выступает в роли исследователя [7, 8]. Прием вертикального чтения [3], или сканирования, конкорданса позволяет обучаемому составить лексико-грамматический профиль (ЛГК) ЛЕ, т.е. выявить наиболее вероятное грамматическое оформление слова и его типичную сочетаемость. Этот процесс также представляет собой конденсированное чтение, так как обучаемый фактически прочитывает много текстов одновременно, подмечая важные для собственной речепродукции особенности употребления лексических единиц в наиболее вероятностном грамматическом окружении, т.е. самостоятельно изучает коллокационно-коллизионный профиль слова [9]. Кроме того, студент примеряет на себя роль исследователя, использует корпусные инструменты для развития навыков корпусного анализа. Пионеры корпусного подхода не указывали, что существует минимально возможный уровень, с которого можно начинать корпусное обучение. Вместе с тем упражнения на сортировку способны выполнять студенты и уровня владения языком A1–A2. Например, рассортировать строки конкорданса по заданным параметрам с целью усвоения лексико-грамматических особенностей употребления: на строки с существительным *influence (on)* и строки с глаголом *influence*.

1.2. Профилактика интерференции и фосилизации ошибок. Изучая строки сочетаемости, студент также получает профилактику возможной интерференции. Н.Б. Гвишиани, на основании опыта обучения студентов-лингвистов с опорой на корпус, отмечает, что при контрастивном анализе речепродукции носителей языка и учащихся частотными оказываются ошибки, вызванные интерференцией родного и изучаемого языков. Так, по данным международного корпуса («английский как иностранный») речи учащихся (The International Corpus of Learner English, ICLE), национально-специфической ошибкой русско-говорящих студентов является употребление *possibility to do something* ('возможность делать что-либо'). Если же обратиться к британскому компоненту корпуса (ICE-GB), то самым частым распространением выражения окажется *possibility of*, *possibility that*, *possibility for*, в то время как вариант *possibility to do* отсутствует вообще. На основе изучения 10–15 строк конкорданса обучаемый может самостоятельно прийти к выводу о том, что продиктованное схемой родного языка употребление *possibility to* следует исключить [2. С. 24]. В некоторых случаях ошибки и заблуждения обучаемых настолько укоренились в «русском» английском: *to explain me; sorry me; to influence on; in the nearest future; highest education; the humanitarian department; you have two variants* и т.д., – что требуется существенная доказательная база для их осознания. Так как закрепившиеся ошибки встречаются практически на всех уровнях обучения, то и от-

бор корпусных примеров может проводиться с учетом уровня сложности.

1.3. Расширенный контекст в корпусе. Еще одна функция корпусного инструментария – опция *расширенный контекст*, которая существует в каждом корпусе и дает доступ к нескольким предложениям исходного текста. Расширенный контекст способствует сохранению включенности ситуации, сообщает сведения о времени и обстоятельствах создания текста, авторе и источнике публикации. В профессиональной литературе высказывается достаточно радикальная точка зрения о том, что корпус, в особенности большой диверсифицированный корпус (например, БНК – Британский национальный корпус), является «музеем» текстов, в котором включенность ситуации потеряна. Текст оказывается оторванным от коммуникативной ситуации, поскольку «реальность текст не сопровождает» [10. Р. 32–33]. Обозначенная функция расширенного контекста помогает представить ситуацию, в которой данное употребление слова состоялось, и в какой-то мере снять обозначенную трудность. Однако, возможно, полностью снять такие трудности может небольшой специализированный корпус, составление которого имеет целый ряд преимуществ [11].

1.4. К вопросу о преимуществах компиляции «малого» корпуса. В этой связи следует упомянуть возможность самостоятельного составления усилиями команды преподавателей небольшого корпуса, ориентированного на нужды будущих специалистов. Известно, что сегодня для таких языков, как английский, русский и другие созданы большие диверсифицированные и репрезентативные аннотированные корпусы. Вместе с тем ни один корпус не в состоянии эффективно служить всем целям сразу. Поэтому создание на базе университетов специализированных так называемых домашних (*home-made*) языковых корпусов, ориентированных на нужды факультета, кафедры или университета, становится повседневной университетской практикой. Подобные университетские проекты могут ставить перед собой задачи по компиляции, например, таких профессионально ориентированных лингвистических баз данных, как корпус дискурса преподавателей в аудитории, корпус ошибок обучаемых, корпус изучаемой предметной области и т.д. При сегодняшнем развитии программного обеспечения по управлению лингвистическими базами данных (например, корпусменеджер WordSmith Tools) создание малого корпуса становится как технически реализуемым, так и оправданным с точки зрения затрат усилий университетской команды специалистов. Примерами таких «малых» университетских корпусов могут служить мультимедийный корпус устных текстов и упражнений ELISA (English Language Interview Corpus as a Second-Language Application [12]), созданный в университете г. Тюбинген, Германия; корпус ошибок обучаемых, собранный в РГПУ им. А.И. Герцена [13]; ресурс LINGUATORIUM, созданный коллективом центра лингвистических исследований им. А.А. Худякова в Санкт-Петербургском государственном экономическом университете; корпус инженерных текстов на базе Томского политехнического университета [14]; корпус

профессионально-направленных текстов, созданный нами для экспериментального обучения [6], и др.

1.5. Обзор корпусной статистики в применении к обучению. Сегодня, когда филологически компетентная текстовая база данных обладает разметкой и корпусом-менеджером (управляющей программой), к основным методам работы с корпусами текстов могут быть причислены все виды корпусной лингвостатистики: подсчет абсолютной частотности слов, выделение наиболее частотных содержательных (в отличие от служебных) слов по отдельным жанрам корпуса (включая профессионально релевантные), диаграмма частотности слова по различным жанрам корпуса, определение ключевых слов, выделение кластеров (статистически значимых цепочек слов, расположенных контактно), корпусный поиск сочетаемости (коллокатов, которые могут располагаться дистанционно). Легко заметить, что многие из этих функций основаны на статистических показателях корпуса. Следует подчеркнуть, что до прихода компьютерного анализа больших массивов лингвистических данных основным индикатором верности и естественности высказывания служила интуиция носителей языка. Сегодня и у не-носителей появилась возможность оценить аутентичность речи как письменной, так устной с помощью лингвостатистических показателей.

1.6. Семантическая просодия и корпусная статистика. Многочисленные корпусные примеры дают возможность изучения еще более тонких особенностей употребления. Так, внимание корпусных лингвистов привлекло явление *семантической просодии*

(semantic prosody), которая понимается как вероятное употребление слова в определенных контекстах, положительных или отрицательных [2. С. 24]. Например, отмечается, что глагол *cause* ('являться причиной') в более чем 90% случаев (автором изучались 250 вхождений в корпусе объемом в 1 млн словоупотреблений и 38 тыс. вхождений в корпусе объемом в 120 млн словоупотреблений) встречается с негативно коннотированными словами *cancer, crisis, delay* [15]. По корпусу LOB (The Lancaster-Oslo/Bergen Corpus, LOB corpus) глагол сочетается с негативными словами в 80% употреблений, с нейтральными и позитивными – в 18% и в 2% употреблений соответственно [15].

В нашем исследовании кластер *something of a* выступает в препозиции: (i) к отрицательно коннотированным существительным – в 49% случаев; (ii) к нейтрально коннотированным существительным – в 39% случаев; (iii) к положительно коннотированным существительным – в 12% случаев. Если сопоставить процентное отношение отрицательных и положительных коннотаций существительных коллокатов, то явственна тенденция сочетания кластера *something of a* с отрицательными словами: 80 против 20%. Самые иллюстративные результаты представлены в таблице. Такое тяготение к отрицательным контекстам связывается с функцией кластера *something of a* как средства выражения коммуникативной стратегии смягчения негативности и категоричности высказывания, характерной для носителей английской лингвокультуры [16].

Корпусная статистика сочетаемости кластера *something of a* с негативно окрашенными существительными

Слово	Количество вхождений в БНК с кластером <i>something of a</i>	Общее количество вхождений слова в БНК	Процентное соотношение вхождений с кластером	Коэффициент взаимного притяжения (MI) в БНК
SHOCK	29	4 163	0,70	7,60
PROBLEM	22	28 251	0,08	4,44
MISNOMER	19	93	20,43	12,47
DISAPPOINTMENT	16	1 451	1,10	8,26
PARADOX	11	651	1,69	8,87
CULT	10	878	1,14	8,31
DILEMMA	9	1 062	0,85	7,88

Разнообразная статистика корпуса имеет существенный объяснительный потенциал и может использоваться для оценки аутентичности высказывания самим преподавателем, а также для постановки задач исследования для обучаемых. (см. пример в разделе 2).

В отличие от *cause* и *something of a*, слово *provide* тяготеет к положительным контекстам в сочетаниях со словами *care, help, relief, support* [15]. В «докорпусную» эпоху данное явление не изучалось с возможностью количественного подтверждения на аутентичных материалах. Ранжирование коллокаций по силе притяжения и выявление семантической просодии были использованы нами при составлении лексически-направленных заданий.

1.7. Поиск коллокаций по корпусу. Говоря о развернутом поиске коллокаций, или сочетаемости, сле-

дует обозначить некоторые особенности корпусного поиска. При автоматическом поиске коллокаций задается как искомая часть речи коллоката, так и «расстояние» (измеряемое в словах) между изучаемым словом и его коллокатом. Кроме того, при составлении корпусного запроса можно выбирать *тип контекста*, или *жанр*, – например, отдать предпочтение профессионально ориентированным, устным или письменным жанрам. Еще одна замечательная особенность корпусного поиска заключается в том, что при поиске коллокатов в корпусе есть возможность задать *зону поиска*, например: от 1 до 5 слов, слева или справа от изучаемого слова. При поиске сочетающихся прилагательных, например, к существительному *stigma* мы получим частотное словосочетание *a huge stigma*. Если же расширить поиск до 2–4 слов в окрестности от изуч-

емого, то есть вероятность обнаружить так называемый чанк (chunk – англ. ‘кусок’, chunking – образование типичных, в разной степени устойчивых цепочек слов), например (*there is*) *a huge stigma attached to...* (‘существует презрительное / негативное отношение к...’) или цепочку (*it*) *takes a bit of getting used to* (‘к этому надо привыкнуть’). С точки зрения когнитивных теорий именно способность оперировать такими частотными, отработанными речевым сообществом (в том числе и профессиональным) чанками и составляет основу беглости речи носителей языка. Выявление подобных статистических особенностей речи является, на наш взгляд, важным профессионально релевантным навыком специалистов языка и образования, которые ведут обучение иностранному языку специалистов разных направлений.

Как уже указывалось, возможность поиска статистически значимой комбинации слов особенно важна для не-носителей английского языка и способствует формированию корректной сочетаемости. Корпусные примеры действуют как профилактика интерференции, если у говорящего нет упроченного в прошлом опыта употребления коллокаций, узально ограниченных словосочетаний [17]. Важными для корпусного исследования являются трудные с точки зрения употребления слова случаи «деформации идиом», «оживления» их компонентов, намеренное обыгрывание или мета-метафора (*his eggs were in too many baskets*) [18]. Такие употребления частотны, например, в политическом дискурсе как средство убеждения. Как преподаватель, так и студент могут научиться проводить корпусное исследование по дистрибуции изучаемого идиоматического выражения в разных жанрах корпуса, отобрать лишь релевантные в контексте исследования употребления. В зависимости от целей исследования внимание может быть сосредоточено на диалектных, регистровых, стилистических особенностях функционирования такой фразеологической единицы.

Самостоятельные корпусные исследования помогают оказывать одновременное воздействие на все анализаторы обучаемого, способствуют усвоению, ведут к идиоматичной и правильной речепродукции. Вместе с тем такой корпусный анализ развивает исследовательские навыки будущих специалистов, лингвистов и преподавателей иностранного языка.

1.8. Классификация заданий по приемам корпусного инструментария. В нашей работе был создан банк лексически направленных заданий и предложена классификация по приемам используемого корпусного инструментария: (i) задания с использованием строк сочетаемости; (ii) задания с использованием функции поиска коллокаций; (iii) задания с использованием статистически выделяемых ключевых слов и других видов статистической наглядности корпуса; (iv) комбинированные задания, в которых используются и корпусные, и традиционные приемы. Также был введен прием «решение исследовательской задачи», под которым понимается самостоятельное исследование студента с использованием корпусных инструментов [6]. Так, студенту-лингвисту может быть поручено исследовать, действительно ли в устном академическом дискурсе, как и в письменном, не исполь-

зуются идиомы и фразовые глаголы; подобрать примеры их использования, если такие существуют. Вопрос, например, который должен решить будущий преподаватель иностранного языка с помощью корпусного исследования, заключается в следующем: правильно ли то, что использование идиом и фразовых глаголов при оценке устной части высказывания в рамках обучения академическому английскому языку приравнивается к ошибке? Проделав небольшое корпусное исследование, будущий профессионал образования сможет верифицировать свои догадки с опорой на аутентичные массивы лингвистических данных.

1.9. Лексическая плотность. К функциям корпусного инструментария можно также отнести элемент корпусной статистики под названием *соотношение количества словоформ к количеству лексем* (type / token ratio). Этот показатель является индикатором лексической плотности [19] и вариативности текста и может использоваться при отборе содержательных текстов для аннотирования.

Кроме того, с помощью показателя лексической плотности внимание будущего специалиста по английскому языку может быть привлечено к существенным регистровым особенностям письменной и устной речи. Так, по данным последних корпусных грамматик английского языка, самыми лексически плотными являются новостные сообщения, в то время как устные обиходно-бытовые диалоги представляют собой самые лексически неплотные высказывания [20]. В противоположность выверенным и отредактированным текстам новостных статей, обиходно-бытовой диалог между участниками разворачивается в «живом эфире», когда отсутствие временных ресурсов исключает редактирование грамматических средств и планирование, корректировка сказанного происходит в последующих репликах, а произнесенное высказывание нельзя удалить. Предметы, окружающие собеседников, могут быть вовлечены в ситуацию дейктическими средствами или жестами. В среднем по корпусу LSWE (Longman Corpus of Spoken and Written English) лексические (содержательные) слова составляют 50% от всех слов корпуса. Типичная дистрибуция содержательных слов по устным (разговорным) и письменным (новостным) источникам составляет 41 и 63% соответственно; количество служебных слов в письменном дискурсе превышает их число в устном: 44 и 37% соответственно, а включения (*erm, right, yeah, please*) наблюдаются только в устной речи – 15% [Там же].

Следует отметить, что такая корпусная статистика стала непременным атрибутом грамматик, словарей и учебников английского языка, создаваемых для обычных студентов, изучающих иностранный язык. Излишне говорить, что навыки анализа такой корпусной статистики, безусловно, имеют большое значение для формирования профессиональной компетентности будущих языковых специалистов.

Как будет показано в следующем разделе, приоритеты в пользу выбора того или иного корпусного инструмента могут меняться сообразно потребностям обучаемых, целям, задачам обучения. В этой связи необходимо также отметить лингводидактический

потенциал *переводного корпуса* (тексты оригиналов и их переводов), позволяющий уделять пристальное внимание межъязыковым соответствиям. Анализ иллюстративного материала при сопоставлении текстов оригинала и перевода при обучении художественному переводу [18] представляет собой еще одно использование корпусного инструментария, который может служить основой достоверности выводов при исследовании трудностей перевода.

2. Примеры корпусных заданий и учебных исследований

Приведем ряд примеров отдельных заданий, а также небольших самостоятельных исследований типа «решение исследовательской задачи», которые были апробированы в группах студентов – лингвистов, регионоведов и переводчиков II и IV курсов

обучения. Обучение проводилось с привлечением материалов нашего собственного корпуса, составленного из текстов, профессионально релевантных для данных направлений, и лингвистической базы БНК, из которой также отбирались контексты профессиональной направленности.

2.1. Пример задания с опорой на конкорданс. Это задание направлено на усвоение лексической единицы, которая трудна с узульской точки зрения, и рекомендовано студентам начиная с уровня владения языком В2. Задания на сортировку конкорданса, как указывалось ранее, могут включаться на самых ранних уровнях обучения. В данном задании обучаемому предлагается трансформировать информацию, представленную в конкордансе, в коллокационно-коллизионный профиль слова. Задание направлено на развитие коллокационной компетентности. Представлен фрагмент конкорданса существительного *stigma*:

On the one hand, it seems that it is the illegality and stigma attached to heroin use which cause users their problems more than the drug itself.
feel that why the hell shouldn't they, y'know. It's all the stigma attached to it and all the things that people have to do to obtain i
men defined as 'exchanging information'. Clearly, there is a stigma attached to gossip among English males, an unwritten rule to...
Besides all these difficulties, the bankrupt also faces problems with the social stigma still attached to going bankrupt. In the public eye the bankrupt is seen as
we know how to get public funding. There is a stigma against Muslims, so your name is important.
migrants from former colonies carry an additional stigma that affects their political attitudes.
illness and hypochondria must shoulder some of the blame in such situations, because the stigma attached to these disorders owes much to doctors' negative attitudes. Patients who mistakenly

Ожидаемый результат состоит в том, что в полном конкордансе студенты отметят наиболее частотные коллокации с прилагательными *certain* (*certain amount of*), *huge*, *additional* и выделят другие наиболее существенные черты грамматики слова. С точки зрения грамматически возможных вариантов существительного *stigma* привлекают внимание сочетание с глагольной формой в постпозиции и явственная тенденция наполнения этой коллизионной схемы глаголом *attached to* (38 вхождений из 264 в БНК), реже – с *associated with*. Оборот *there is* – заметное синтаксическое предпочтение сочетания (*certain / huge / certain amount of*) *stigma attached to*. Закрепление лексической единицы происходит в дальнейших заданиях на составление обучаемыми собственных предложений, поиск примеров с изучаемой единицей в текстах СМИ, коммуникативных заданиях.

2.2. Пример задания «решение исследовательской задачи». Как показало наше исследование, корпусная статистика по употребимости слова или словосочетания в разных регистрах и контекстах корпуса обладает значительным объяснительным потенциалом. Данное исследовательское задание нацелено на изучение регистрационных особенностей и было апробировано в аудиториях студентов-регионоведов, лингвистов и переводчиков.

Задание формулируется следующим образом: «В материалах BBC в синхронном переводе цитаты российского политика глагол *exclude* (*We do not “exclude” the possibility of joining the UN sanctions*) дается в кавычках как не совсем соответствующий контексту. Во всех последующих цитирований высказывания источ-

ники британских СМИ отдают предпочтение глаголу *rule out*, несмотря на то что в словаре (Longman) эти два глагола определяются как синонимы. Используя разбивку корпуса БНС по жанрам, исследуйте данные единицы и найдите объяснение таким фактам употребления». Ожидаемый результат состоит в том, что, по данным корпуса, коллокация *rule out the possibility* встречается преимущественно в новостных, газетных текстах, а *exclude the possibility* частотна лишь в текстах академического жанра. Логичный вывод состоит в том, что узус глагола латинского происхождения *exclude* ограничен академическими контекстами и употребление *exclude* в новостном контексте не соответствует ожиданиям языкового сообщества.

Задания с использованием корпусной статистической наглядности могут комбинироваться с другими традиционными заданиями на закрепление лексических единиц. Приемы корпусного обучения лучше варьировать с коммуникативными заданиями, т.е. выходом в речевую практику.

Ограничением в использовании корпусного инструментария на занятиях может быть несформированность корпусной компетенции преподавателя, недостаточная техническая оснащенность учебного заведения и нехватка времени, отводимого на изучение дисциплины.

Обоснованность предлагаемых заданий подтверждается результатами экспериментального обучения. Апробация заданий с опорой на корпусные инструменты, преимущественно нацеленная на совершенствование лексических навыков, проводилась в группах студентов – лингвистов, регионоведов и переводчиков. По

результатам проверки уровня сформированности лексических навыков средний балл в экспериментальных группах по сравнению с контрольными вырос на 6–8%. Среди положительных моментов применения корпусного подхода были отмечены: рост ряда специфических умений (инструментальных, эвристических, умений учиться), высокая доля самостоятельности при решении возникающих затруднений, более высокая степень вовлеченности студентов в учебный процесс. Расчет достоверности результатов, или неслучайности приращения, производился на основании вычисления критерия Крамера–Уэлча и уточнялся по критерию Вилкоксона–Манна–Уитни [21].

Заключение

Необходимо подчеркнуть, что для того, чтобы пользоваться возможностями корпуса, как студенту, так и преподавателю необходимо ознакомиться с правилами корпусного запроса и освоить элементарные навыки работы с такой базой данных. Последовательное рассмотрение возможностей корпусной лингвистики в лингводидактическом контексте позволяет убедительно продемонстрировать ее потенциал для развития ключевых иноязычных компетенций. Усилия по освоению возможностей корпуса позволяют и преподавателю, и студенту эффективно пользоваться корпусом как большой, аутентичной справочной системой и сформировать навыки автономного исследователя.

Выдвигая идею развития корпусной компетенции, мы, прежде всего, видим необходимость ее развития при обучении лингвистов и будущих преподавателей иностранного языка, так как обращение к корпусам

текстов и корпусным технологиям как к средству поддержки обучения иностранному языку является одним из актуальных направлений совершенствования существующих методик. В своей убежденности мы опираемся на существенный потенциал корпусного обучения.

Такие характеристики языкового компьютерного корпуса, как достоверность и аутентичность, репрезентация различных регистровых особенностей, лингвистическая и статистическая наглядность, а также разнообразные поисковые возможности, рассмотренные в лингвометодическом контексте, позволяют эффективно развивать ключевые иноязычные компетенции.

Приведенный в разделе 1.4 опыт составления малого профессионально направленного корпуса предполагает работу не только одного преподавателя, но и коллектива преподавателей в содружестве со специалистами-предметниками. Целесообразность компиляции корпуса предметной области также состоит в использовании корпусных процедур для отбора более точного содержания обучения. В нашей практике преподавания профессионально ориентированного иностранного языка идея создания небольшого корпуса доказала свою состоятельность и универсальность, т.е. применимость при обучении иностранному языку для различных специальностей. Разработанный нами курс, посвященный лингвометодическому потенциалу корпусных исследований, включен в программу магистратуры по тестологии СПбГУ. Мы искренне благодарим руководителя программы профессора И.Ю. Павловскую за еще одну возможность реализации наших идей при подготовке магистров-тестологов.

ЛИТЕРАТУРА

1. Рахманов И.В., Гез Н.И., Зимняя И.А. и др. Основные направления в методике преподавания иностранных языков в XIX–XX вв. / под ред. И.В. Рахманова. М. : Педагогика, 1972. 317 с.
2. Гвишиани Н.Б. Практикум по корпусной лингвистике : учеб. пособие по английскому языку. М. : Вышш. шк., 2008. 191 с.
3. O'Keeffe A., McCarthy M., Carter R. From Corpus to Classroom : Language Use and Language Teaching. Cambridge : Cambridge Univ. Press, 2007. 315 p.
4. Biber D., Conrad S., Reppen R. Corpus Linguistics : Investigating Language Structure and Use. Cambridge : Cambridge Univ. Press, 1998. 300 p.
5. Reppen R. Using Corpora in the Language Classroom. New Edition. S. I. : Cambridge Univ. Press, 2010. 118 p.
6. Горина О.Г. Использование технологий корпусной лингвистики для развития лексических навыков студентов-регионоведов : автореф. дис. ... канд. пед. наук. М., 2014. 24 с.
7. Johns T. Should You Be Persuaded – Two Samples of Data-driven Learning Materials // Johns T., King Ph. (eds.) Classroom Concordancing : ELR Journal. 1991. Vol. 4. P. 1–16. URL: http://www.lexically.net/wordsmith/_corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf (дата обращения: 1.12.2015).
8. Горина О.Г. Психологические основы обучения иностранному языку специальности с опорой на языковой корпус // Вестник ЛГУ им. А.С. Пушкина. СПб., 2014. Т. 7, № 1. С. 172–179.
9. 12th Teaching and Language Corpora Conference. URL: <https://www.uni-giessen.de/faculties/f05/engl/ling/tale> (дата обращения: 30.02.2017).
10. Brawn S. Designing and exploiting small multimedia corpora for autonomous learning and teaching // Hidalgo E., Quereda L., Santana J. (eds) Corpora in the Foreign Language Classroom : Selected Papers from the Sixth International Conference on Teaching and Language Corpora (TaLc 6). Amsterdam : Rodopi, 2007. P. 32–33.
11. Горина О.Г. Определение словарного запаса обучаемых с помощью корпуса // Вестник ЛГУ им. А.С. Пушкина. СПб., 2013. Т. 7, № 2. С. 201–209.
12. ELISA – English Language Interview Corpus as a Second-Language Learning Application // Xmars. URL: http://www.xmarks.com/s/site/www.uni-tuebingen.de/elisa/html/elisa_index.html (дата обращения: 30.02.2017).
13. Камшилова О.Н., Колина М.В., Николаева Е.А. Разработка корпуса текстов петербургских школьников: задачи и перспективы // Прикладная лингвистика в науке и образовании. СПб., 2008. С. 92–98.
14. Шаламова Н.Н., Фильченко А.Ю. Корпусная лингвистика и ее использование в профильно-ориентированном преподавании иностранных языков. Томск : ТПУ, 2004.
15. Stubbs M. Corpus evidence for norms of lexical collocation // Cook G., Seidlhofer B. (eds.) Principle and Practice in Applied Linguistics : Studies in Honour of H.G. Widdowson. Oxford : Oxford Univ. Press, 1995. P. 245–56.
16. Gorina O.G., Khrabrova V.E. Linguistic Hedging as a Communicative Strategy (Within the Scope of Corpus Studies). Vestnik NSU. Series: Linguistics and Intercultural Communication, 2017. Vol. 15, № 3. P. 44–53. (In Russ.)
17. Горина О.Г. Использование технологий корпусной лингвистики для развития лексических навыков студентов-регионоведов : дис. ... канд. пед. наук. М., 2014. 325 с.

18. Тер-Минасова С.Г. Синтагматика речи : онтология и эвристика : Общая и английская синтагматика составных номинативных групп. 2-е изд. М. : Либроком, 2009. 200 с.
19. Гвишиани Н.Б. Введение в контрастивную лексикологию (Англо-русские межъязыковые соответствия). М. : Либроком, 2010. 288 с.
20. Biber D., Johansson S., Leech G. et al. Longman Grammar of Spoken and Written English. S. I. : Pearson Education Limited, 1999. 1204 p.
21. Новиков Д.А. Статистические методы в педагогических исследованиях : Типовые случаи. М. : МЗ-Пресс, 2004. 67 с. URL: <http://www.mtas.ru/uploads/stat.zip> (дата обращения: 10.02.2017).

Статья представлена научной редакцией «Филология» 26 апреля 2018 г.

CORPUS RESEARCH TOOLS IN L2 TEACHING

Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal, 2018, 435, 187–194.

DOI: 10.17223/15617793/435/24

Olga G. Gorina, Higher School of Economics (St. Petersburg, Russian Federation). E-mail: gorina@bk.ru.

Keywords: corpus linguistics; corpus linguistic statistics; corpus tools of linguistic analysis; corpus competence; concordance; collocation; semantic prosody; lexical density; visualization of language patterns; English language.

The aim of the article is to inform professional readership of corpus analysis potential for L2 teaching purposes, which is based on the author's own implementation experience of corpus-based activities in L2 classroom. Thus, the article is divided into four sections, including Introduction (1), Corpus Tools (2), Examples of Classroom Use (3) and Conclusion (4). Introduction outlines the recent corpus-driven changes in attitudes to language statistics, which are reflected in corpus-informed text books. Section Two, which has nine subsections, deals with corpus tools and notions of corpus analysis (concordance, collocation and colligation search, corpus statistics, semantic prosody etc.) in L2 teaching context. In particular, the author discusses condensed reading, concordance vertical scanning for lexico-grammatical profiling and other teaching tools to develop L2 linguistic competence. These are later supported (Section 3) by some corpus-oriented classroom activities with possible teaching outcomes outlined. Some experience-based comments are also given regarding the language level of students that could benefit from corpus data analysis. Based on the research results, Conclusion elaborates on the idea of corpus competence as well as the necessity of corpus tools to be used by both language teaching professionals and students. Theoretically, as was initially suggested by data-driven teaching pioneers, not only the researcher, but also the learner should be given the chance of studying language through corpus. The article advocates that corpus tools for collocation search together with colligation (or probable grammar structures) detection are powerful means to develop both language and research skills. In addition to corpus-based activities and theoretical grounding behind that, the author also shared her own experience on compiling a corpus of professional discourse. Both the idea and the practicality of a small university-made corpus are evaluated. A brief comparison of a diversified corpus (such as the British National Corpus) versus a “home”-made corpus is provided. The research also draws attention to a term *chunk of language*, which has been adopted by western teaching methodology, and is being considered in the article through frequency – probability (corpus) terms. It is suggested that a chunk of language bigger than a collocation lends itself to being discovered through the combination of various corpus tools. These frequent language chunks (such as *there is certain stigma attached to...*) account for a large part of native speaker's vocabulary and fluency. They are believed to be stored in memory in great amounts and retrieved virtually undivided. Chunks could be subject to minor colligational corrections while speaking. The author believes that frequent language chunks could be discovered by language learners as an educational research activity under the guidance of a language instructor. As has already been mentioned, the article provides some research activity examples in Section 3. Thus, the article will equip the reader with a clear understanding of corpus linguistic potential in the foreign language classroom as well as with the capacity and confidence to engage in corpus analysis. It might be particularly beneficial for non-native speakers of English who happen to teach English in ESP context, since the author believes that corpus research ends the monopoly of language intuition, which in today's world is being gradually replaced by corpus statistics.

REFERENCES

1. Rakhmanov, I.V. et al. (1972) *Osnovnye napravleniya v metodike prepodavaniya inostrannyykh yazykov v XIX–XX vv.* [The main directions in the methodology of teaching foreign languages in the nineteenth and twentieth centuries]. Moscow: Pedagogika.
2. Gvishiani, N.B. (2008) *Praktikum po korpusnoy lingvistike: ucheb. posobie po anglyiskomu yazyku* [Workshop on corpus linguistics: studies. An English manual]. Moscow: Vyssh. shk.
3. O'Keeffe, A., McCarthy, M. & Carter, R. (2007) *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge Univ. Press.
4. Biber, D., Conrad, S. & Reppen, R. (1998) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge Univ. Press.
5. Reppen, R. (2010) *Using Corpora in the Language Classroom*. New Edition. Cambridge: Cambridge Univ. Press.
6. Gorina, O.G. (2014) *Ispol'zovanie tekhnologii korpusnoy lingvistiki dlya razvitiya leksicheskikh nabykov studentov-regionovedov* [Using the technology of corpus linguistics for the development of lexical skills of regional students]. Abstract of Pedagogy Cand. Dis. Moscow.
7. Johns, T. (1991) Should You Be Persuaded – Two Samples of Data-driven Learning Materials. *Classroom Concordancing: ELR Journal*. 4. pp. 1–16. [Online] Available from: http://www.lexically.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf. (Accessed: 1.12.2015).
8. Gorina, O.G. (2014) Psychological grounds of using corpora in the ESP classroom. *Vestnik LGU im. A.S. Pushkina*. 7(1). pp. 172–179. (In Russian).
9. Giessen University. (2016) *12th Teaching and Language Corpora Conference*. [Online] Available from: <https://www.uni-giessen.de/faculties/f05/engl/ling/talc>. (Accessed: 30.02.2017).
10. Brawn, S. (2007) Designing and exploiting small multimedia corpora for autonomous learning and teaching. In: Hidalgo, E., Quereda, L. & Santana, J. (eds) *Corpora in the Foreign Language Classroom: Selected Papers from the Sixth International Conference on Teaching and Language Corpora* (TaLC 6). Amsterdam: Rodopi.
11. Gorina, O.G. (2013) Using Corpus Evidence to Compile Learner's Vocabulary. *Vestnik LGU im. A.S. Pushkina*. 7(2). pp. 201–209. (In Russian).
12. Uni-tuebingen.de. (n.d.) *ELISA – English Language Interview Corpus as a Second-Language Learning Application*. [Online] Available from: http://www.uni-tuebingen.de/elisa/html/elisa_index.html. (Accessed: 30.02.2017).
13. Kamshilova, O.N., Kolina, M.V. & Nikolaeva, E.A. (2008) Razrabotka korpusa tekstov petersburgskikh shkol'nikov: zadachi i perspektivy [Development of a corpus of texts of Petersburg schoolchildren: tasks and prospects]. In: *Prikladnaya lingvistika v nauke i obrazovanii* [Applied linguistics in science and education]. St. Petersburg: [s.n.]. pp. 92–98.

14. Shalamova, N.N. & Fil'chenko, A.Yu. (2004) *Korpusnaya lingvistika i ee ispol'zovanie v profil'no-orientirovannom prepodavanii inostrannykh yazykov* [Corpus linguistics and its use in the profile-oriented teaching of foreign languages]. Tomsk: TPU.
15. Stubbs, M. (1995) Corpus evidence for norms of lexical collocation. In: Cook, G. & Seidlhofer, B. (eds) *Principle and Practice in Applied Linguistics: Studies in Honour of H.G. Widdowson*. Oxford: Oxford Univ. Press.
16. Gorina, O.G. & Khrabrova, V.E. (2017) Linguistic Hedging as a Communicative Strategy (Within the Scope of Corpus Studies). *Vestnik NSU. Series: Linguistics and Intercultural Communication*. 15(3). pp. 44–53. (In Russian).
17. Gorina, O.G. (2014) *Ispol'zovanie tekhnologiy korpusnoy lingvistiki dlya razvitiya leksicheskikh nayk studentov-regionovedov* [Using the technology of corpus linguistics for the development of lexical skills of students majoring in regional studies]. Pedagogy Cand. Dis. Moscow.
18. Ter-Minasova, S.G. (2009) *Sintagmatika rechi: ontologiya i evristika: Obshchaya i angliyskaya sintagmatika sostavnykh nominativnykh grupp* [Speech syntagmatics: ontology and heuristics: General and English syntagmatics of composite nominative groups]. 2nd ed. Moscow: Librokom.
19. Gvishiani, N.B. (2010) *Vvedenie v kontrastivnyu leksikologiyu: (Anglo-russkie mezh'yazykovye sootvetstviya)* [Introduction to contrastive lexicology: (English-Russian interlanguage correspondence)]. Moscow: Librokom.
20. Biber, D., Johansson, S., Leech, G. et al. (1999) *Longman Grammar of Spoken and Written English*. Pearson Education Limited.
21. Novikov, D.A. (2004) *Statisticheskie metody v pedagogicheskikh issledovaniyah: Tipovye sluchai* [Statistical methods in pedagogical research: Typical cases]. Moscow: MZ-Press. [Online] Available from: <http://www.mtas.ru/uploads/stat.zip>. (Accessed: 10.02.2017).

Received: 26 April 2018