

К ВОПРОСУ О КОРПУСНОМ ПОДХОДЕ В ОБУЧЕНИИ РУССКОМУ ЯЗЫКУ КАК ИНОСТРАННОМУ

Яцевич Е.А.

Белорусский государственный университет

Беларусь, Минск

Статья посвящена описанию сущности корпусного подхода в лингвистике, причем акцент сделан именно на лингводидактическом потенциале данного направления. В статье обосновывается важность создания методически ориентированного корпуса языковых учебных текстов и учет полученной на его основе информации при разработке учебно-методических материалов для обучения студентов иностранному языку.

Ключевые слова: корпусный подход, методически ориентированный корпус языковых текстов, обучение иностранному языку, обучение русскому языку как иностранному.

Высокий уровень развития компьютерных технологий и их доступность широкому кругу пользователей не могут не оказывать влияние на методы проведения лингвистических исследований, а также на методики работы при преподавании различных дисциплин, в частности, при изучении иностранного языка.

Как отмечает Л.В. Рычкова, «важнейшим инструментом преподавателя РКИ могут и должны стать полнотекстовые базы данных, дающие возможность наблюдать за функционированием единиц языка в реальных контекстах» [1].

Метод анализа текстов на базе корпусов относится к сфере корпусной лингвистики, которая оформилась в качестве отдельной научной отрасли к началу девяностых годов двадцатого века. По определению В.П. Захарова, «корпусная лингвистика – раздел компьютерной лингвистики, занимающийся разработкой общих принципов построения и использования лингвистических корпусов (корпусов текстов) с применением компьютерных технологий. Под лингвистическим, или языковым, корпусом текстов понимается большой, представленный в машиночитаемом виде, унифицированный, структурированный, размеченный, филологически компетентный массив языковых данных, предназначенный для решения конкретных лингвистических задач» [2, с. 7].

Корпусный менеджер позволяет работать с текстовой и лингвистической информацией, получить статистические данные в соответствии с поставленной лингвистической задачей, построить при необходимости конкорданс, который представляет собой перечень всех словоупотреблений для конкретного слова в рамках корпуса со ссылкой на источник. Корпусы также позволяют оптимизировать лексикографическую работу над словарями различных типов. Корпусы помогают увеличить эффективность работы с языковыми и речевыми данными, в силу того что «корпус – это уменьшенная модель языка или подъязыка» [2, с. 18].

Среди самых известных и разработанных иноязычных корпусов на данный момент следует назвать Британский национальный корпус (British National Corpus – BNC), лингвистический Банк английского языка (Bank of English), Корпус современного американского английского (Corpus of Contemporary American English – COCA), Лимерик-Белфаст корпус устного академического английского языка LIBELCASE (LIBEL Corpus of Spoken Academic English) и др. Среди русскоязычных корпусов следует назвать Национальный корпус русского языка (НКРЯ), Национальный корпус русского литературного языка (НКРЛЯ), Русский интернет-корпус (I-RU) и т.д. Отдельно следует назвать русскоязычные корпусы, имеющие лингводидактическую направленность: учебный подкорпус НКРЯ, Хельсинкский аннотированный корпус (ХАНКО), а также в рамках направления Learner Corpora создаются корпусы для анализа языка и речи инофонов, например корпус Russian Language Learner Corpus of Academic Writing – RULEC – (корпус академического письма изучающих русский язык, который создан в Портландском университете), пропедевтический корпус, создаваемый в Гродненском университете имени Янки Купалы.

В самом общем виде все задачи корпусной лингвистики можно условно разделить на две группы: к первой могут быть отнесены те, которые направлены на формирование корпуса, его разметку и разработку корпусных менеджеров, ко второй – те, которые сконцентрированы на проведении экспериментальных исследований на базе корпуса (корпусов).

Многие технологии, которые применяются исследователями в области корпусной лингвистики, являются новациями с практической точки зрения, а не с теоретической: стремительное развитие ИТ-сферы, а также увеличение мощности и скорости современных компьютеров, привело к трансформации техник и приемов работы с лингвистическими данными, повысив эффективность проводимых исследований. «Понятие корпуса является продолжением традиционных картотек, с которыми всегда работали лингвисты. В XX веке эти картотеки стали компьютерными и общедоступными» [2, с. 33].

К основным характеристикам корпусов относятся репрезентативность, сбалансированность, полнота, наличие разметки.

Необходимо подчеркнуть, что репрезентативность корпуса понимается как пропорциональное и достаточное представление текстов различной жанрово-стилевой и временной отнесенности и т.п. для отражения объективной картины речевой деятельности независимо от намерений исследователя (группы исследователей), создающих корпус. В.П. Захаров отмечает, что «корпус выступает как некоторый объект, призванный послужить моделью некоторой внешней по отношению к нему реальности. Именно репрезентативность корпуса определяет достоверность полученных на его материале результатов» [2, с. 18].

Разметка (аннотация) корпуса – «это введенная автоматически или вручную лингвистическая или метатекстовая информация обо всех выбранных единицах корпуса: тексте, предложении, текстоформе, морфеме, звуке и т. д.» [3].

Типология корпусов различна в силу многоаспектности: например,

различение корпусов на основе универсальности выражается в выделение корпусов универсальных, которые отражают все языковые и речевые факты, и корпусов для специальной цели, которые строятся для наблюдения некоторого факта культурного либо лингвистического характера; по типу же языковых данных корпусы делятся на письменные, устные и смешанные и т.д. Методология создания любого корпуса находится в прямой зависимости от его типа.

При создании корпуса выделяют основные этапы. Например, В.П. Захаров [2, с. 18] приводит следующий перечень. Первый – отбор текстов в соответствии с четко определенными критериями, второй – представление текстов в электронном виде, третий – предварительная обработка текстов (корректурa по необходимости) и их анализ (создание экстралингвистических и библиографических описаний), четвертый – конвертирование и графематический анализ, пятый – разметка текста, шестой – проверка автоматической разметки (корректурa по необходимости), седьмой (заключительный) – работа корпусного менеджера, который осуществляет быстрый многоаспектный поиск и статистическую обработку, восьмой – обеспечение доступа к корпусу (по необходимости) и девятый – создание документов, отражающих различные аспекты создания и использования корпуса.

Среди авторов работ, посвященных описаниям корпусных исследований, а также отражающих применение корпусов в преподавании, следует назвать Д. Байбера, Л.Н. Беляеву [4], Н.Б. Гвишиани, В.П. Захарова [2], С. Кондрад, М. Копотева [3], В.А. Плунгяна [5], В.В. Рыкова, Л.В. Рычкову [1, 7], Е.П. Соснину [6], Л. Флауердью и др.

Для характеристики использования корпусных технологий непосредственно при преподавании иностранных языков инофонам Л.В. Рыčkова приводит цитату Г. Астона, который одним из первых начал использовать данные Британского национального корпуса после его создания и который четко обозначил два практических направления при работе с корпусом текстов в целях обучения иностранному языку, «в первом, которое могло бы быть названо подходом типа COBUILD, корпус, прежде всего, представляет собой ресурс, который может использовать методист и учитель для создания справочников, разработки учебных программ и материалов. Во втором, которое можно было бы назвать обучением, основанным на данных, корпус, прежде всего, представляет собой ресурс для учащегося» [7]. Следует подчеркнуть, что описанные направления предельно точно соответствуют векторам развития корпусной лингводидактики.

Среди базовых методов работы с корпусом при лингводидактическом анализе следует назвать генерирование конкорданса, подсчет абсолютной частотности (т.е. сколько раз конкретный объект содержится в имеющемся наборе объектов), анализ ключевых слов как лингвостатистическую основу для отбора тематического лексического ядра словарного запаса обучаемого, поиск коллокаций (т.е. слов, встречающихся со статистически значимой частотностью в окружении изучаемого слова), коллигаций (т.е. более длинные сочетания слов с их вероятными грамматическими предпочтениями, помогают определить

способность либо неспособность сопоставляться с определенными единицами или выступать в определенной синтаксической функции), создание лексико-грамматического профиля слов и другие.

Но потенциал технологий корпусной лингвистики в решении задач, связанных с обучением иностранному языку инофонов, особенно при формировании профессиональной компетенции, увеличится в случае создания специальных, тематических, методически ориентированных на определенную предметную область корпусов. В качестве примеров можно привести Хельсинкский аннотированный корпус (ХАНКО), корпус инженерных текстов, созданный в Томском политехническом университете; ресурс ЛИНГВАТОРИУМ, созданный в Санкт-Петербургском государственном экономическом университете, ориентированный на преподавателей английского языка и т.д.

Соснина Е.П. подчеркивает, что «корпусный подход, или метод лингвистического исследования, основанный на корпусах текстов, ориентирован на прикладное изучение языка, его функционирование в реальных средах и текстах, что важно для преподавания языка» [6].

Литература

1. Рычкова, Л.В. Корпусные технологии в преподавании русского языка как иностранного / Л.В. Рычкова, С.Н. Киеня // Документы [Электронный ресурс]. – 2000-2015. – Режим доступа : <http://rudocs.exdat.com/docs/index-263394.html>. – Дата доступа : 21.01.2020.
2. Захаров, В.П. Корпусная лингвистика: учебник для студентов гуманитарных вузов / В.П. Захаров, С.Ю. Богданова. – Иркутск: ИГЛУ, 2011. – 161 с.
3. Копотев, М.В. Введение в корпусную лингвистику / М.В. Копотев // Электронная библиотека bookz.ru [Электронный ресурс]. – 2003-2017. – Режим доступа : http://bookz.ru/authors/mihail-kopotev/vvedenie_529/1-vvedenie_529.html. – Дата доступа : 20.01.2020.
4. Беляева, Л.Н. Корпусные технологии извлечения терминологии в задачах переводной лексикографии / Л.Н. Беляева // Структурная и прикладная лингвистика. Вып. 10: межвуз. сб. / под ред. А. С. Герда. – СПб.: Изд-во С.-Петербург. ун-та, 2014. – С. 169-181.
5. Плунгян, В.А. О проекте корпусной грамматики русского языка / В.А. Плунгян, Н.М. Стойнова // Русский язык за рубежом, № 3 (256), 2016. – С. 12–19.
6. Соснина, Е.П. Корпусная лингвистика и корпусный подход в обучении иностранному языку / Е.П. Соснина // Прикладная лингвистика [Электронный ресурс]. – 2008-2017. – Режим доступа : http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_linguistics_language_teaching/. – Дата доступа : 21.01.2021.
7. Рычкова, Л.В. Корпусные технологии в обучении русскому языку в условиях близкородственного двуязычия / Л.В. Рычкова // Научная библиотека учреждения образования «Витебский государственный университет имени П.М.Машерова» [Электронный ресурс]. – 2007-2017. – Режим доступа : <http://lib.vsu.by/xmlui/bitstream/handle/123456789/3676/Рычкова%20Л.В..pdf?sequence=1&isAllowed=y>. – Дата доступа : 21.01.2021.

TO THE QUESTION OF THE CORPUS APPROACH IN TEACHING THE RUSSIAN LANGUAGE AS A FOREIGN LANGUAGE

Jacevich K.A.

Belarusian State University

Belarus, Minsk

The article is devoted to the description of the essence of the corpus approach in linguistics, and the emphasis is made on the linguodidactic potential of this direction. The article substantiates the importance of creating a methodically oriented corpus of language educational texts and taking into account the information obtained on its basis in the development of educational and methodological materials for teaching students a foreign language.

Keywords: corpus approach, methodically oriented corpus of language texts, teaching a foreign language, teaching Russian as a foreign language.