

Сусанна АНГУРЯН
Армянский государственный
педагогический университет

КОРПУСНАЯ ЛИНГВИСТИКА В МЕТОДИКЕ ПРЕПОДАВАНИЯ ЯЗЫКОВ

В статье говорится о возможности использования в практике преподавания языков лингвистических данных различных корпусов, что позволит максимально приблизить учебный процесс к реальной действительности и значительно повысить его эффективность. Использование корпусных технологий приобретает особую актуальность при обучении продуктивным видам речевой деятельности, а также позволяет лучше организовать самостоятельную творческую работу учащихся, привить им навыки работы с современными информационными ресурсами, научить использовать их в практической деятельности.

Ключевые слова: лингвистические корпусы, корпусные технологии, Википедия, работа с текстом, самостоятельная работа, иностранные языки, продуктивные виды речевой деятельности

Одним из самых доступных и эффективных способов использования телекоммуникативных средств при изучении иностранных языков является обращение к данным корпусной лингвистики. Они, к тому же, позволяют менять объем информации и подбирать типы заданий в расчете на каждого учащегося.

Термин *корпусная лингвистика* как наименование раздела языкознания, который занимается разработкой, фиксацией и использованием совокупности текстов (лингвистических корпусов), был создан и введен в употребление в середине XX века. Первым большим компьютерным корпусом считается Брауновский корпус (*Brown Corpus*, BC), созданный в 1960-е годы в Университете Брауна и содержащий 500 текстов (по 2 тыс. слов) на английском языке. Он постепенно расширялся, вскоре уже включал около 1 млн. текстов и послужил эталоном для создания корпусов на других языках. Очевидно, что созданные специалистами корпусы лингвистических данных не могли не привлечь к себе внимания методистов, то есть не рассматриваться с точки зрения полезности в процессе преподавания соответствующих иностранных языков. Методики преподавания языка с использованием корпуса отличаются обращением к реальным современным источникам и текстам, то есть к таким материалам, с которыми учащиеся сталкиваются (или будут сталкиваться) в повседневной жизни, в профессиональной сфере. Возможность быстрого получения с помощью корпусных данных сведений о правильности, вариативности или частотности

того или иного употребления позволяет значительно экономить время, максимально приблизить учебный процесс к реальной действительности.

Наличие в Интернет-пространстве большого количества текстов в электронной форме существенно облегчило задачу создания больших корпусов размером в сотни миллионов слов, но не ликвидировало множества уже определенных создателями проблем, таких как отбор текстов, снятие возникающих при этом проблем с авторскими правами, приведение всех текстов в единообразную форму, распределение корпусных данных по темам, стилям, грамматическим составляющим и т.д.

Общие, обширные, большие по объему лингвистические корпусы существуют (или разрабатываются) для многих языков (греческого, немецкого, чешского, финского, армянского, китайского, японского, болгарского и др. языков). Например, Национальный корпус русского языка на сегодняшний день является одним из самых больших и содержит более 300 млн. образцов текстовых словоупотреблений.

В настоящее время создаются более конкретные по цели использования, а потому ограниченные по объему корпусы, то есть современный лингвистический корпус – это совокупность текстов, собранных в соответствии с определёнными принципами (уровень владения языком и тип пользователя, стилистическая принадлежность, определенные лексико-грамматические характеристики слов и т.д.) и обеспеченных соответствующей поисковой системой. Таковы существующие в наше время лингвистические корпусы разного типа: двуязычные, тематические, стилистически ограниченные, жанровые, авторские, учебные и т.д.

Целесообразность создания таких дифференцированных текстовых корпусов объясняется исследователями многими причинами, среди которых наиболее важными являются следующие: достаточная для каждого пользователя обширность и убедительность фрагмента любого типа; необходимость представления вариантов реальных контекстов реализации конкретных лингвистических фактов; возможность быстро проанализировать представленные таким образом языковые факты (сравнить, количественно сопоставить и т.д.); возможность многократного использования однажды созданного корпуса для решения различных лингвистических задач.

Очевидно, что созданные специалистами корпусы лингвистических данных не могли не привлечь к себе внимания методистов, то есть не рассматриваться с точки зрения полезности в процессе преподавания соответствующих иностранных языков. Методики преподавания языка с использованием корпуса отличаются обращением к реальным современным источникам и текстам, то есть к таким материалам, с которыми учащиеся сталкиваются (или будут сталкиваться) в повседневной жизни, в профессиональной сфере. Возможность быстрого получения с помощью корпусных данных сведений о правильности, вариативности или частотности

того или иного употребления позволяет значительно сэкономить время, максимально приблизить учебный процесс к реальной действительности.

Методисты, занимающиеся изучением проблемы повышения эффективности изучения языка с применением корпусных технологий отмечают, что в этом случае происходит сдвиг акцента с дедуктивного подхода в преподавании к индуктивному, а это значит, что непосредственная опора на эмпирические данные кардинально меняет функции всех составляющих, в том числе и участников, учебного процесса:

- преподавателя, который становится координатором, куратором;
- студента, который учится анализировать встретившиеся употребления;
- учебной грамматики, которая теряет излишнюю абстрактность, схематизм, оторванность от современности /Leech, 2001: 227/.

Все эти положения полностью соответствуют современным дидактическим концепциям, теоретическим методическим разработкам, направленным на обоснование необходимости применения новых методов и приемов для повышения эффективности преподавания всех дисциплин, в частности иностранных языков.

Таким образом, в центре обучения языку с использованием данных определенного корпуса оказывается учащийся, который сам находит и выбирает примеры, сам их анализирует, сам делает на основании своего анализа выводы и т.д. Новая дидактическая парадигма затрагивает роль преподавателя в меньшей степени, чем учащегося. Преподаватель должен не столько руководить, сколько помогать учащемуся в изучении иностранного языка, направлять его, указывать на новые ресурсы, и корпус оказывается для этого идеальным и перспективным инструментом.

Наиболее популярным достижением корпусной лингвистики является создание в 2001 году сайта Wikipedia (основатель - Джимми Уэйлс) – свободной энциклопедии, которая представляет собрание текстов, характеризующих практически все предметы, явления, факты, события прошлого и настоящего, известных стран и малочисленных народов и т.д. Одним из основных принципов пополнения материалов сайта является тот, согласно которому любой желающий может не только добавить свою статью, но и обоснованно (со ссылками на авторитетные работы в данной области) исправить чужую. Сейчас, по данным Интернета, аудитория сайта Wikipedia составляет около 450 млн. посещений в месяц, хотя сам сайт не является коммерческим /www.vsozial.ru/.

Исследователи отмечают много достоинств этой свободной энциклопедии, однако главными из них являются три:

- она действительно общедоступна, поскольку из неё может черпать информацию любой желающий, у кого есть выход в Интернет;
- она универсальна, поскольку здесь есть практически всё и обо всём (здесь уже более 17 млн. статей);

- она многоязычна: перевод созданных на английском языке статей Википедии осуществляется на более, чем 250 языков, среди которых есть и армянский.

Материалы Википедии постоянно обновляются, поэтому она считается и одним из самых быстрых СМИ. По скорости обновления новостей Википедия обошла такие мировые сайты, как YouTube и MySpace.

Значение материалов этого сайта в познавательной деятельности человека трудно переоценить, но не менее (если не более) он важен и для правильной и продуктивной организации учебного времени. Ведь, по данным Интернета, “свободная энциклопедия” была создана как дополнительный, доступный, бесплатный, образовательный проект, предназначенный огромному количеству читателей: школьников и студентов, родителей и преподавателей, представителей разных народов и религий и т.д. В этой связи хочется отметить значение составляющих слово Wikipedia частей: Wiki - в переводе с гавайского языка означает «быстрый» и pedia - в переводе с древнегреческого означает «обучение». И если значение слова *энциклопедия* – «обучение в (полном) круге (наук)», то значение слова *википедия* можно определить как «быстрое обучение» /www.ru.wictionary.org/.

Такая работа особенно продуктивна при прохождении определенных лексических единиц. Материалы Corpus-a могут, например, использоваться таким образом. Из пройденного по учебной программе (учебнику) текста выделяются такие слова, которые представляются интересными с точки зрения их определения в Википедии /Сысоев, 2010/. Сильному студенту можно предложить найти в Википедии соответствующую статью и представить ее суть в аудитории так, чтобы было понятно и интересно всем студентам (в случае необходимости это может сделать и преподаватель). Затем этот уже знакомый в общих чертах текст раздается всей группе и тем же самым студентом (или преподавателем) прорабатывается более подробно – задаются вопросы по каждому предложению, объясняются незнакомые слова с переводом и т.д. Такая работа способствует выполнению следующих задач:

- отработка определенной лексико-грамматической конструкции;
- выявление родо-видовых отношений предметов и способов их языкового выражения;
- формирование у всех студентов интереса к работе в Интернете и удовлетворение их познавательных интересов.

Приведем конкретный пример. В учебнике встречается слово **каноз**. Оно, если даже понятно студентам, по сути, представляет большой познавательный интерес. Один из сильных студентов получает задание найти в Интернете и соответствующим образом представить другим студентам это слово. В Википедии статья о каноз, снабженная многочисленными фотографиями, выглядит следующим образом:

A **canoe** is a lightweight narrow boat, typically pointed at both ends and open on top, propelled by one or more seated or kneeling paddlers facing the direction of travel using a single-bladed paddle. In International Canoe Federation nomenclature used in some European countries such as the United Kingdom the term canoe refers to kayaks, while canoes are called *Canadian canoes*.

Canoes are used for racing, whitewater canoeing, touring and camping, free-style, and general recreation. The intended use of the canoe dictates its hull shape and length and construction material.

Constructed between 8200 and 7600 BC, and found in the Netherlands, the Pesse canoe may be the oldest known canoe. Excavations in Denmark reveal the use of dugouts and paddles during the Ertebølle period, (ca 5300 BC – 3950 BC).

Sailing canoes (see Canoe sailing) are propelled by means of a variety of sailing rigs. Common classes of modern sailing canoes include the 5 m² and the International 10 m² Sailing canoes. The latter is otherwise known as the International Canoe, and is one of the fastest and oldest competitively sailed boat classes in the western world. The log canoe of the Chesapeake Bay is in the modern sense not a canoe at all, though it evolved through the enlargement of dugout canoes.

Простая формулировка слова «каное» может выглядеть так: «A **canoe** is a small narrow boat, typically human-powered, which is normally open on top, but can be decked over (i.e. covered, similar to a kayak) and which was constructed somewhere between 8200 and 7600 BC». Она способствует не только пониманию слова «каное», но и пониманию слова «кайак».

Далее, другому (тоже сильному) студенту предлагается устно (синхронно) перевести данную формулировку на армянский язык. Затем студентам раздается печатный вариант этого текста и проводится работа по расширению представления об этом предмете: чтение текста, лексико-грамматические упражнения, т.е. объяснение и закрепление некоторых (2-3) лексико-грамматических единиц, таких как слова bow, stern, paddler и др., конструкции to be powered by, *can* be fitted with и т.д.

Возможен и другой вариант работы. Из пройденного текста преподаватель выбирает ряд употребительных слов, затем дает задание нескольким студентам найти в Corpus-е примеры употребления (по 4-5) каждого из этих слов и представить их остальным. Эти примеры могут быть проанализированы студентами группы (в аудитории или дома) по конкретному заданию (сравните формулировки, данные в Интернете и в каком-нибудь печатном словаре...; сформулируйте значение, в котором использовано данное слово в вашем тексте...; определите разницу в сочетательных возможностях...; составьте свои примеры...; используя новый лексико-грамматический материал, составьте небольшой текст (диалог)...; переведите на армянский язык... и т. д.) и таким образом при неоднократном повторении, усвоены.

տեխնոլոգիաների կիրառումը առանձնահատուկ նշանակություն է ստանում խոսքային գործունեության արտադրողական տեսակների ուսուցման առումով, քանի որ այն հնարավորություն է ընձեռում ավելի լավ կազմակերպել սովորողների ինքնուրույն ստեղծագործական աշխատանքը, սերմանել նրանց մեջ ժամանակակից տեղեկատվական ռեսուրսներով աշխատանքային հմտություններ, սովորեցնել կիրառել դրանք գործնականում:

Քանալի բառեր. Լեզվաբանական կորպուսներ, կորպուսային տեխնոլոգիաներ, Վիկիպեդիա, աշխատանք տեքստերի հետ, ինքնուրույն աշխատանք, օտար լեզուներ, խոսքային գործունեության արտադրողական տեսակներ

S. ANGURYAN – *Corpus Linguistics in the Methods of Teaching Languages.* – The paper discusses some methods of teaching languages based on Corpus Linguistics. The application of Corpus Linguistics or corpora (structured sets of electronically stored samples of ‘real world’ texts) to the language teaching practice aims at bringing the learning process as much to reality as possible, greatly enhancing its efficiency. Using the text-corpus methods/technology is of particular relevance in teaching productive skills (speaking and writing); it is a feasible tool to better organize students’ independent creative work, to instill in them the ability to work with modern information resources, to instruct how to apply the latter effectively in practice.

Key words: Corpus Linguistics, Corpus technology, Wikipedia, work with texts, independent work, foreign languages, productive skills