

ПРИМЕНЕНИЕ КОРПУСНЫХ ТЕХНОЛОГИЙ В ОБУЧЕНИИ АНГЛИЙСКОМУ ЯЗЫКУ СТУДЕНТОВ ЯЗЫКОВЫХ НАПРАВЛЕНИЙ ПОДГОТОВКИ

Золотов П.Ю.

Тамбовский государственный университет имени Г.Р. Державина
pitirim93@gmail.com

Корпусная лингвистика, раздел языкознания, занимающийся созданием и использованием электронных текстовых корпусов, зародившись во второй половине прошлого века, продолжает свое развитие в качестве инструмента для других разделов лингвистики. Корпусные технологии, как одно из направлений информационно-коммуникационных технологий (ИКТ) определяются как совокупность средств и методов обработки и анализа данных электронных лингвистических корпусов, значительно облегчили лингвистам процедуру сбора материала для исследований, одновременно с этим повысив качество выборки, а также со временем предоставив более широкий выбор жанрового разнообразия материала. В данной же статье будет рассмотрен лингводидактический потенциал использования корпусных технологий применительно к студентам языковых направлений подготовки.

Начнем с обзора электронных лингвистических корпусов английского языка, подходящих для обучения английскому как иностранному. В процессе этого обучения рекомендуется использовать корпуса открытого доступа, то есть свободно доступные онлайн или требующие только регистрации. Существует группа лингвистических корпусов, созданная М. Дэвисом, профессором Университета имени Бригама Янга. В эту группу входят Британский национальный корпус (BNC), Корпус современного американского английского языка (COCA), Корпус «Time Magazine» (TMC), и Корпус исторического американского английского (COHA) [1].

Британский национальный корпус – это одноязычный корпус, представляющий британский английский язык конца 20-го века. Корпус включает в себя коллекцию примеров письменного и разговорного британского английского из различных жанров, источников и тематических областей. Общее количество текстов составляет около 100 миллионов слов.

Корпус современного американского английского – самый большой и единственный свободно распространяемый корпус американского варианта английского языка. Он содержит 220 тысяч текстов, 520 миллионов слов за период с 1990 по 2017 год. Корпус иллюстрирует самые разные жанры: устную речь, представленную записями более 150 телевизионных и радиопрограмм; раздел художественной литературы, включающий в себя рассказы, романы и пьесы из литературных журналов для взрослых и детей, сценарии фильмов; популярные журналы, газеты; академические журналы делятся на классы в соответствии с классификацией библиотеки Конгресса.

Корпус журнала «Time» состоит из более чем 275 000 статей из архива журнала «TIME» и содержит более 100 миллионов слов текстов, написанных с 1923 по 2006 год.

Корпус исторического американского английского – это исторический корпус из 400 миллионов слов, в том числе 100 000 текстов с 1810 по 2009 год. Он дает возможность проследить изменения языка за десятилетия. Корпус содержит тексты из цифровой библиотеки «Создание Америки», в том числе 267 монографий и более 100 000 статей в журналах 19-го века по образованию, психологии, американской истории, социологии, религии, науке и технике, а также тексты различных жанров из архивов универсальной цифровой библиотеки «Project Gutenberg».

Далее следует отметить, что подход, основанный на использовании в лингводидактике данных, полученных из корпусов в зарубежной литературе, был назван «data-driven learning», что есть «обучение на основе данных». Этот термин был введен Тимом Джонсом в 1990 году, чтобы описать индуктивный подход, при котором учащиеся считаются своего рода «детективами», анализирующими определенные языковые структуры с помощью корпусных технологий на целевом языке для формирования обобщений и отслеживания закономерностей [2]. Обучение на основе данных корпуса заключается в использовании инструментов и методов корпусной лингвистики в педагогических целях. Используя этот подход, учителя могут предоставить учащимся прямой доступ к корпусу, чтобы они могли исследовать результаты поиска в корпусе для изучения закономерностей.

Среди преимуществ подхода обучения на основе данных корпуса можно назвать привнесение аутентичного материала в процесс освоения иностранного языка аудиторией. Кроме этого, в числе достоинств использования корпусных технологий состоит некоторый элемент

самостоятельного открытия нового, который делает обучение более увлекательным, помимо мотивации учащихся. Считается, что основанные на корпусе действия повышают самостоятельность учащегося, поскольку обучающихся учат наблюдать за языком и делать обобщения, а не зависеть в этом плане только от учителя. Поэтому использование корпусов в лингводидактике может быть эффективным при изучении языковых аспектов, таких как грамматика, дискурс (организация речи и письма), лексика, произношение (описывает акценты и исследует интонацию) и многие другие.

Среди корпусных технологий в первую очередь стоит выделить конкорданс. Этот инструмент предполагает использование программного обеспечения корпусов для поиска появления определенного слова или фразы в контексте. В первую очередь этот инструмент очень полезен для изучения лексики. Результаты поиска в конкордансе могут быть использованы для того, чтобы студенты могли самостоятельно найти значение слова. При этом контекст может подсказать им, что искомое слово может иметь разные значения. На начальных этапах стоит использовать заранее отобранные преподавателем примеры, чтобы студентам было легче найти информацию. Можно использовать распечатанные результаты поиска в конкордансе. Конкорданс предлагает ознакомиться с реальным использованием слов в языке. На основе материалов, извлеченных из корпусов, были даже созданы учебники для обучения английскому языку, например, Touchstone (автор: Майкл МакКартти).

Осваивая иностранный язык очень важно изучить не только лексику саму по себе, но и сочетаемость слов. К примеру, зная слова на английском смотреть – to look и телевизор – a TV, мы, казалось бы, можем легко их соединить в словосочетание смотреть телевизор – to look a TV. Однако мы знаем, что это неверно, правильно будет – to watch a TV. Такую сочетаемость слов, то есть коллокации, мы можем легко проверить с помощью корпусов. В Таблице 1 приведено сравнение коллокаций английских глаголов hear и listen, где показаны их первые пять самых частотных употребления.

Таблица 1.

Сравнение коллокаций глаголов hear и listen

	word	w1	w2	w1/w2	score
1	surprised	100	0	200.0	83.4
2	delighted	75	0	150.0	62.6
3	prayer	47	0	94.0	39.2
4	relieved	37	0	74.0	30.9
5	mercy	33	0	66.0	27.5
	word	w2	w1	w2/w1	score
1	attentively	19	0	38.0	91.1
2	carefully	118	4	29.5	70.7
3	ter	9	0	18.0	43.1
4	willing	36	3	12.0	28.8
5	willingness	11	1	11.0	26.4

В обучении грамматике английского языка также есть место корпусным технологиям. Здесь в первую очередь следует сказать о коллигации, то есть грамматической сочетаемости слов. Примером коллигации может служить управление глаголов. В конкордансе мы можем проверить, требует ли определенный глагол после себя какой-либо предлог или нет, а также, можем ли мы использовать после этого глагола другой глагол в форме инфинитива, или в форме герундия. В корпусе мы также можем изучать реальное использование временных конструкций [3]. Фразовые глаголы, являющиеся всегда сложным элементом для освоения, также могут быть изучены и исследованы с помощью корпусных технологий.

Существуют лингвистические корпуса, в которых записана устная речь с приложенной ее фонетической транскрипцией. Подобные источники помогут обучающимся студентам языковых направлений подготовки на занятиях по практической и теоретической фонетике. Такие корпуса предоставляют примеры живого потока речи, где можно услышать различные варианты и диалекты английского языка.

С помощью корпусных технологий можно также обучать правилам построения речевых высказываний с точки зрения дискурса. То, как соблюсти когерентности текста и какие средства

когезии использовать можно найти в корпусах. Поиск в конкордансе поможет понять в каких ситуациях лучше использовать конкретное средство когезии.

В последнее время корпусные технологии стали применяться в области формирования прагматических навыков студентов. Прагматика изучает то, как язык используется в общении, то как мы используем язык и какими лингвистическими средствами выражаем свои мысли. Некоторые аспекты прагматики также подлежат обучению с помощью корпусных технологий, например, использование вежливого обращения к собеседнику.

Создание «виртуального корпуса» в Британском национальном корпусе – это корпус, созданный из базы данных текстов. Можно создать корпус либо по ключевым словам в текстах (например, тексты со словами «инвестиции», «баскетбол» или «биология»), либо по типу информации о текстах (например, дата, название или источник), либо по комбинации ключевой и текстовой информации.

Затем можно редактировать виртуальные корпуса, выполнять поиск в определенном виртуальном корпусе, сравнивать частоту слова, фразы или грамматической конструкции в разных виртуальных корпусах, а также создавать «списки ключевых слов» на основе текстов в своем виртуальном корпусе.

Таким образом, роль корпусных технологий в обучении студентов языковых направлений подготовки может быть очень значительной. Дидактический потенциал корпусных технологий охватывает все уровни языка, делая основной упор на лексику и грамматику, однако, не упуская такие аспекты, как фонетика, дискурс и даже прагматика. Корпусные технологии являются инновационной технологией, заслуживающей внимания и предлагающей большие возможности для преподавателей и студентов вузов.

Литература

1. Osipova, E.S. Corpus Linguistics Technology In Teaching English As A Foreign Language. 2018. P. 273-283.
2. Johns, Tim. Should you be persuaded: two examples of data driven learning. ELR Journal (New Series). 1991, P. 1-16.
3. Сысоев П.В. Лингвистический корпус в методике обучения иностранным языкам // Язык и культура. 2010. Вып. 1(9). С. 99-111.