

## Список использованных источников

1. Костомаров, В. Г. Выступление на круглом столе «Языковая политика в современной России» 16 апреля 2002 года / В. Г. Костомаров // Журнал «Мир русского слова». – № 2. – 2002. – С. 24.

2. Лебединский, С. И. Методика преподавания русского языка как иностранного : учеб. пособие / С. И. Лебединский, Л. Ф. Гербик. – Минск : БГЭУ, 2011. – 309 с.

*In this article questions of coordination of subject-teachers and specialists in Russian philology in the course of training of foreign citizens in higher education institutions of Republic of Belarus are considered. Special attention is paid to work on this direction in technical college.*

УДК 378.147:811.161.1'243-057.875:81

## КОРПУС ТЕКСТОВ РУССКОГО ЯЗЫКА В УЧЕБНОЙ ДЕЯТЕЛЬНОСТИ ИНОСТРАННЫХ СТУДЕНТОВ-ФИЛОЛОГОВ

**И. Г. Гомонова**

*В статье анализируется место корпусных технологий и корпусных методов в учебной деятельности иностранных студентов-филологов. Корпус текстов рассматривается как эффективный инструмент для решения разнообразных дидактических и исследовательских задач, стоящих перед студентом, изучающим русский язык как иностранный.*

В качестве неотъемлемого компонента учебной и исследовательской деятельности студента-филолога выступают в настоящее время компьютерные (информационные) технологии, которые помогают усовершенствовать и ускорить процесс получения, обработки и передачи информации. Информационные технологии в лингвистике – «это совокупность законов, методов и средств получения, хранения, передачи, распространения, преобразования информации о языке и законах его функционирования с помощью компьютеров» [1, с. 8].

В полной мере необходимость осваивать и использовать компьютерные технологии распространяется и на процесс филологического образования иностранных студентов.

Иностранные студенты-филологи, которые обучаются по направлению «Русская филология (компьютерное обеспечение)», знакомятся с электронными лингвистическими ресурсами в рамках

дисциплин «Введение в компьютерную филологию», «Инженерия знаний», «Методы автоматической обработки текста», «Корпусная лингвистика», «Формализация языка в экспертных системах» (общее количество часов, отведенных на аудиторные занятия и самостоятельную работу, – 414). Для иностранных студентов-филологов, выбравших направление «Русская филология (русский язык как иностранный)», с этой целью в учебные планы включена дисциплина «Компьютерные технологии в преподавании русского языка как иностранного» (общее количество часов – 64), дополнить которую могут дисциплины, входящие в компонент учреждения высшего образования.

С полной уверенностью можно утверждать, что независимо от направления специальности иностранному студенту, получающему филологическое образование, необходимо знакомство с компьютерными ресурсами, относящимися к сфере корпусной лингвистики.

Работа с лингвистическими корпусами, т. е. с массивами текстов, собранными в соответствии с определенными принципами, представленными в электронном виде, унифицированными, размеченными по определенному стандарту, обеспеченными специализированной поисковой системой, в настоящее время представляет собой один из основных методов лингвистических исследований, при помощи которого могут решаться самые разные задачи. Значительное количество корпусов создается и уже создано для многих языков мира.

Корпусная русистика начала развиваться в 1980-е годы, пережила некоторый спад в 1990-е годы и в настоящее время активно развивается и уже достигла существенных результатов, большинство из которых находится в свободном доступе [2; 3].

Изучая дисциплину «Корпусная лингвистика», иностранные студенты-русисты знакомятся с характерными особенностями различных доступных on-line корпусов языка специальности, но, несомненно, основным ресурсом, на базе которого студенты осваивают принципы работы корпуса текстов и вырабатывают у себя навыки использования корпусных методов в исследовании и преподавании языка, является Национальный корпус русского языка (НКРЯ) [4] – крупнейший и самый востребованный в настоящее время ресурс, содержащий тексты на русском языке. Объем более 600 млн словоупотреблений, богатая разметка, включающая метаинформацию и лингвистическую (морфологическую, семантическую, синтаксическую и др.) информацию, достойная техническая база и другие преимущества НКРЯ позволяют успешно использовать данный ресурс для решения самых разнообразных дидактических и исследовательских задач.

Освоив базовый терминологический аппарат корпусной лингвистики, познакомившись теоретически с типами лингвистических задач, решаемых с применением корпусных технологий, студенты приступают к практической работе с корпусом текстов. В случае с иностранными студентами такая деятельность заключается преимущественно в выполнении ими заданий, подготовленных преподавателем.

Так, для первичного знакомства с корпусом, выработки умения создавать пользовательский подкорпус студенты-иностранцы выполняют задания типа «Определите, сколько романов, авторы которых – женщины, а время и место действия – Россия постсоветского периода, представлено в НКРЯ. Определите, сколько раз в этих произведениях употребляется прилагательное *медвежий*. Приведите контексты, включающие фразеологизмы с компонентом *медвежий* (например, *медвежья услуга*)». В процессе выполнения подобных заданий делается акцент и на широких возможностях метаразметки корпуса, и на обязательном использовании количественных методов при обращении к корпусным данным (отметим, что предпочтение количественных методов – один из отличительных признаков корпусной лингвистики при сравнении ее с традиционной лингвистикой).

Выполнение заданий типа «Определите, в каком художественном произведении впервые (по данным НКРЯ) употребляется слово *инфузория*» позволяет иностранным студентам познакомиться с дополнительными настройками корпуса и выработать навык упорядочивать контексты по дате создания или в обратном порядке в ходе самостоятельной работы с корпусом.

Знакомство иностранных студентов с морфологической разметкой НКРЯ предполагает выполнение заданий типа «Изучите морфологическую разметку НКРЯ и назовите, какие категории глагола в ней представлены. Сделайте запрос на формы: а) 1 лица множественного числа сложного будущего времени; б) 2 лица единственного числа повелительного наклонения; в) женского рода сослагательного наклонения. Приведите по одному примеру». Работая с морфологической разметкой НКРЯ, обращаем внимание студентов на отсутствие в ней сложных форм и на специфику формирования в корпусе запроса на такие формы.

Задание типа «Приведите по одному примеру сложной формы сравнительной и превосходной степени имени прилагательного, оформив соответствующий запрос» в случае работы с подготовленной аудиторией можно усложнить следующим образом: «Приведите два контекста с ненормативной (аналитической) формой степени сравнения прилагательного, оформив и описав соответствующий запрос». Как

известно, корпус призван адекватно отражать «реальную жизнь» языка, которая включает в себя и ненормативные проявления. Задание, сформулированное таким образом, «провоцирует» корпус на выдачу контекстов, содержащих ненормативные формы. Анализ таких ошибок, представленных в реальных (а не искусственно созданных) контекстах, способствует их предупреждению в речи студентов, изучающих русский язык как иностранный: *Во-первых, у всех людей, независимо от расы и пола, одинаково небольшие (относительно других зубов) клыки. Это позволяет думать, что у всех наших прямых предков, вплоть до самых древнейших, тоже были относительно небольшие клыки* (Б. Стариков. Восхождение к предкам // «Знание-сила», 2013); *«В-4» по внешнему виду напоминая стандартный движок, был в два раза мощнее и еще более надежнее* (И. Бояшов. Танкист, или «Белый тигр» (2008)). Подобные запросы-«провокации» могут быть направлены и на выявление контекстов с ошибками другого типа: *И даже во фразеологическом словаре русского языка можно найти более тридцати выражений со словом НОС. Но чтобы умело пользоваться этими выражениями, следует соблюдать следующие правила: 1. Если вы повесили нос (то есть огорчились), то зарубите себе на носу, и хоть кровь из носу выполняйте правило первое: выше нос!* (Носария // «Трамвай», 1990); *Мы сформировали специальную структуру в форме автономного государственного учреждения под названием «Агентство регионального маркетинга»* (Д. Михалев. Новосибирская область вплотную займется улучшением имиджа региона // РИА Новости, 2008.09.23) – речевая ошибка в виде тавтологии; *Не известно, на каком языке к нам обращаются с просьбой **оплатить за антенну*** («Не могу работать без клея... Дайте стержней!»), – о чем пишут уральцы на пейджеры телекомпаний // Новый регион 2, 2008.02.15); *Кроме того, сейчас сложилась определенная культура, особенно свойственная для молодежи* (Д. Менделеева, Ю. Ратанов. Психотерапевт: Что прячут счастливые аккаунты соцсетей? (2015.11.06)) – синтаксическая ошибка в организации словосочетаний со связью управления и т. п.

Знакомство иностранных студентов с семантической разметкой НКРЯ происходит в процессе выполнения заданий типа «Оформив запрос “наречие, обозначающее большую длительность + глагол речи”, приведите количественную характеристику конкорданса и два-три контекста, в которых субъектом заданного действия является не человек». Напр.: *В углу комнаты, на чемоданах (позже на тумбочке) находился приёмник и постоянно каркал, играл, вещал и пел* (Э. Лимонов. У нас была Великая Эпоха (1987)); *Желтые страницы, похоже, немного ввали. Или хитрили* (В. Пелевин. Любовь к трем цукербринам (2014)).

Выполнение подобных заданий требует анализа большого количества контекстов, включающих заданные сочетания слов, и способствует пополнению лексического запаса иностранных студентов путем усвоения лексических единиц, относящихся к определенной тематической группе. Кроме того, контексты такого рода, отобранные студентами на занятиях по корпусной лингвистике, могут служить в качестве языкового материала на занятиях по русскому языку как иностранному.

Как известно, основной проблемой лингвистической разметки является широко распространенная в русском языке омонимия. Любой семантический запрос в корпусе приводит к большому количеству «шума», уменьшить который (а в некоторых случаях и вовсе исключить) позволяют так называемые фильтры, которые могут носить семантический, грамматический или семантико-грамматический характер. Данному аспекту работы с разметкой корпуса посвящены задания типа «Опишите фильтр, позволяющий получить контексты с существительным *море* преимущественно в количественном значении. Приведите два контекста в качестве примера». Выполняя это задание, студенты создают достаточно простой грамматический фильтр «*море* + сущ. в родительном падеже». Результатом такого запроса будут преимущественно метафорические сочетания со значением большого количества кого-л. или чего-л.: *Когда я научился читать в школе, когда я понял, что могу этим миром управлять, когда я пришел в библиотеку и увидел море книг – дома тоже была этажерка с книгами, случайными – этот мир меня уже ни на минуту не отпускал* (Е. Тюшина, В. Краснов. «Облака» Владимира Краснова // «Новгородские ведомости», 2013); *Разве мало вам того моря крови, того моря слёз, что мы пролили за войну?* (В. Астафьев. Обертон (1995–1996)).

Навыки комплексного использования лингвистической разметки разного вида развивают задания типа «Определите, какое цветное прилагательное активно используется в простой форме превосходной степени. Приведите контекст из художественного произведения».

Специфику разных корпусов НКРЯ (Газетного, Поэтического, Параллельного, Устного и др.) студенты анализируют, выполняя задания типа:

– «В Газетном корпусе оформите запрос, чтобы исследовать сочетаемость прилагательного *липовый*. Приведите два контекста, в которых это прилагательное употребляется в переносном значении» (результатом будут многочисленные контексты с сочетаниями типа *липовые документы / справки / сертификаты / дипломы / диссертации / информация / вузы / ученые* и т. п.);

– «Создайте подкорпус поэтических текстов с каламбурной рифмой, приведите его количественную характеристику. Приведите контекст,

в котором каламбурная рифма формируется с участием местоимения *мой* (И. п. м. р.)» (для выполнения данного задания в пользовательском подкорпусе, включающем 108 документов, студенты делают запрос на местоимение *мой* в заданной форме и в зоне рифмовки и находят контекст с омонимической рифмой *мой – мой: Пожаром жги и морем мой, Ты поцелуев смыть не сможешь И никогда не уничтожишь Сознания, что в веках ты – мой* (М. А. Кузмин. «Целованные мною руки...», 1912));

– «Определите, как на английский язык переводится русское обращение *братцы*. Приведите три разных варианта» (задание выполняется в Параллельном корпусе; студенты задают направление перевода, понимая, что русско-английский и англо-русский подкорпус – это разные компоненты корпуса; в случае неуверенного владения языком переводных текстов студенты пользуются функцией перехода на ЯндексПереводчик; результатом выполнения задания будут разные варианты перевода данного обращения – *brothers, friends, lads, mates, everybody*, в том числе его отсутствие в переводном тексте) и др.

Таким образом, учебные задания, выполняемые иностранными студентами на базе корпуса текстов, разнообразны и привлекательны в плане расширения лингвистического кругозора, активизации уже имеющихся знаний и, конечно, развития новых умений и навыков, оптимизирующих процесс изучения и исследования неродного языка и придающих этому процессу творческий характер.

#### Список использованных источников

1. Зубов, А. В. Информационные технологии в лингвистике : учеб. пособие для студ. лингв. фак-тов высш. учеб. заведений / А. В. Зубов, И. И. Зубова. – М. : Издательский центр «Академия», 2004. – 208 с.
2. Резникова, Т. И. Лингвистически аннотированные корпуса русского языка (обзор общедоступных ресурсов) / Т. И. Резникова, М. В. Копотев // Национальный корпус русского языка: 2003–2005. Результаты и перспективы. – М. : Индрик, 2005. – С. 31–61.
3. Резникова, Т. И. Славянская корпусная лингвистика: современное состояние ресурсов / Т. И. Резникова // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. – СПб. : Нестор-История, 2009. – С. 402–461.
4. Национальный корпус русского языка [Электронный ресурс]. – Режим доступа : <http://www.ruscorpora.ru>. – Дата доступа : 02.02.2019.

*The article analyzes the place of corpus technologies and corpus methods in the educational activities of foreign students. The corpus of texts is considered as an effective tool for solving a variety of didactic and research tasks facing a student who studies Russian as a foreign language.*