



УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЪЕДИНЕНИЕ ПО НАПРАВЛЕНИЯМ ПЕДАГОГИЧЕСКОГО ОБРАЗОВАНИЯ
НАУЧНАЯ ШКОЛА «ПРИКЛАДНЫЕ ИССЛЕДОВАНИЯ ЯЗЫКА И РЕЧИ» ШКОЛА Р. Г. ПИОТРОВСКОГО
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А. И. ГЕРЦЕНА

ПРИКЛАДНАЯ ЛИНГВИСТИКА В НАУКЕ И ОБРАЗОВАНИИ

СБОРНИК ТРУДОВ
VII МЕЖДУНАРОДНОЙ НАУЧНОЙ КОНФЕРЕНЦИИ

10–12 апреля 2014 года
Санкт-Петербург

КОМПЬЮТЕРНЫЕ ЯЗЫКОВЫЕ РЕСУРСЫ В ИССЛЕДОВАНИИ РЕГИОНАЛЬНЫХ ОТЛИЧИЙ УЗУАЛЬНЫХ НОРМ РЕЧИ*

В статье рассматриваются виды компьютерных языковых ресурсов, комплексное использование которых актуально в исследовании региональных вариантов узуальных норм русской речи. Предлагается методика корпуса-эталона, основанная на сочетании возможностей основного модуля Национального корпуса русского языка и комплекса специализированных языковых ресурсов, последовательно создаваемых на аутентичном языковом материале региональных СМИ Гродненщины.

Ключевые слова: компьютерные языковые ресурсы, узус русской речи, корпус-эталон, иллюстративный корпус, язык СМИ.

L. V. Rychkova

COMPUTER LANGUAGE RESOURCES RELEVANT FOR THE INVESTIGATION OF THE RUSSIAN REGIONAL NORMS DIFFERENCES

The article deals with the variety of computer language resources relevant for the investigation of the Russian regional usage norms. The reference corpus research approach is proposed based on using both the possibilities of the main unit of the National corpus of Russian and a set of specialized language resources consequently created on the authentic language material of the Grodno region mass media.

Key words: computer language resources, Russian speech usage, reference corpus, illustrative corpus, language of mass media.

Антропоцентричность, признаваемая современной лингвистикой как фундаментальное свойство естественного языка, актуализировала рассмотрение проблем, которые ранее языкознанием не рассматривались вовсе либо находились на периферии лингвистического знания. К таким проблемам можно отнести исследования различных форм речи и речевой коммуникации, включая проблемы взаимодействия языков, проблему «культурных коннотаций» и обусловленных ими социокультурных конструктов в языке, проблемы этноязыковой идентичности личности, когнитивной семантики и многие другие, решение которых стало возможным в связи с развитием компьютерной техники, а также технологий

© РЫЧКОВА Людмила Васильевна. Гродненский государственный университет им. Я. Купалы, Гродно, Беларусь. E-mail: l.v.rychkova@mail.ru, rychkova@grsu.by

* Подготовлено в рамках проекта, реализуемого при поддержке Белорусского республиканского фонда фундаментальных исследований (договор № Г13Р-050 от 16.04.2013 г.).

создания и использования репрезентативных компьютерных языковых ресурсов, формирующих адекватную экспериментально-доказательную базу лингвистических исследований. Особая роль отводится созданию и использованию (в том числе уже существующих) языковых ресурсов в решении комплексной задачи исследования региональных отличий узуальных норм речи, выявления динамики нормы, находящей отражение не только в устной речи, но и в таком актуальном языковом материале, как язык СМИ.

Известно, что функционирование языка в окружении других языков, в условиях языковых контактов может приводить к возникновению вариантов узуса речи — региональных и национальных. Данная проблема имеет несколько аспектов: собственно лингвистический, социолингвистический, психолингвистический, педагогический и др. Все эти направления развиваются в мировой лингвистике. Соответственно, активно создаются корпуса вариантов того или иного языка, которые используются не только в качестве исследовательских ресурсов, но и находят непосредственное применение в педагогической практике, прежде всего, для лингводидактических целей. Так, например, наряду с Британским национальным корпусом существует свободно доступный он-лайн Корпус американского варианта современного английского языка. Широкую известность получил Международный корпус английского языка, постоянно пополняемый как с точки зрения языкового материала, так и с точки зрения все новых корпусов национальных вариантов английского языка.

Понятие национального варианта, в силу вполне очевидных причин, вплоть до последнего времени не было актуальным для русского языка. Тем не менее не существовало и лингвистических корпусов его региональных вариантов, несмотря на то что изучение русского языка как средства межнационального общения было начато еще в СССР. Так, например, в 1985 г. при участии и под общей редакцией А. Е. Михневича была издана коллективная монография «Русский язык в Белоруссии», посвященная изучению состояния русского языка в республике. Именно А. Е. Михневичем применительно к белорусскому варианту русского языка был впервые предложен термин «нациолект», подчеркивающий, с его точки зрения, «во-первых, реальность и относительную самостоятельность обозначаемого им явления; во-вторых, онтологически и концептуально соотносимый с другими «лект»-образованиями (диалект, социолект, идиолект, нациолект) и в силу этого, в-третьих, достаточно определенно указывающий на невозможность превращения данного явления в национальный вариант русского литературного языка» [2, с. 11–12]. Отметим, что такой подход не исключает, а, напротив, предполагает региональное варьирование узуса русской речи.

В последнее десятилетие в изменившихся геополитических условиях в России стала восприниматься как актуальная проблема функционирования русского языка за пределами метрополии. Функционирование языка в иноязычном окружении имеет свои особенности, которые при определенных условиях могут закрепляться в узуальной норме, приспосабливаясь к функционированию в разных условиях, например, в речи билингвов. Существует проблема кодификации норм при описании функционирования языков за пределами метрополий, в частности русского языка на территории Республики Беларусь.

В решении вышеназванных проблем особую роль могут и должны сыграть «региональные» корпуса русского языка. Именно такой корпус на материале русскоязычных СМИ Гродненщины создается в настоящее время в рамках международного проекта, реализуемого учеными Гродненского государственного университета имени Янки Купалы (Беларусь) и Института русского языка имени В. В. Виноградова Российской академии наук. Создаваемый корпус носит статус иллюстративного и представляет собой первый, пилотный, вариант регионального лингвистического (размеченного) корпуса русского языка [3]. Предполагается, что этот корпус положит начало формированию «в составе НКРЯ нового модуля, представляющего региональные варианты литературного стандарта» [1, с. 352].

Привлечение текстов СМИ в качестве базового языкового материала для создания первого регионального корпуса русского языка обусловлено несколькими факторами: во-первых, данная сфера занимает сегодня главенствующее положение в системе этнической и межэтнической коммуникации, оказывая влияние на формирование стереотипов публичного речевого поведения и социокультурных конструктов; во-вторых, в данных текстах функционирует не только общеупотребительная лексика, но и лексика специальная, в частности, именно тексты СМИ отражают формирование языков для социокультурных целей; в-третьих, динамичность языка СМИ (отмечается многими исследователями) позволяет выявить семантический потенциал лексических единиц и специфику его реализации в региональном узусе, обусловленную аксиологическими и иными социокультурными факторами.

Можно предположить, что специфические особенности, обнаруженные в СМИ Гродненщины, будут заключаться не столько в использовании белорусских, польских и литовских лексических вкраплений (Гродненская область непосредственно граничит как с Польшей, так и с Литвой), сколько в наличии прежде всего семантических, сочетаемостных и стилистических отклонений и различий в частотных распределениях словоформ и лексем. Для выявления таких отклонений наиболее эффективно могут быть использованы лингвостатистические методы, применяемые на большом массиве текстов. Выявление и системное описание такой специфики

может быть достоверным и эффективным лишь на основе использования корпусных технологий, для чего предлагается абсолютно новый для такого рода исследований подход, основанный на применении корпуса-эталона, построенного «на паритетной основе» на материале основного модуля Национального корпуса русского языка. Данный корпус является общепризнанным авторитетным ресурсом, обладающим свойствами полноты и репрезентативности, необходимыми для построения корпуса-эталона с целью использования последнего для выявления специфики узуса русского языка в условиях Гродненского приграничного полиэтничного региона.

Сравнение корпусных данных позволит выявить общие и контрастно-специфические черты изменений в русском языке под воздействием других языков, а также стабильные и нестабильные участки языковой системы. Интерпретация результатов анализа потребует привлечения данных социолингвистики, психолингвистики и лингвокультурологии. Способствовать выявлению семантических парадигм, отобранных для исследования лексических единиц, будет создание системы таких взаимосвязанных компьютерных языковых ресурсов, как словари-конкордансы и дифференциальные частотные словоуказатели.

Список литературы

1. Кустова Г. И., Савчук С. О. Изучение лексико-семантической и социокультурной специфики русской речи на территории Республики Беларусь (на материале текстов СМИ) // Труды международной конференции «Корпусная лингвистика — 2013». СПб.: С.-Петербургский гос. университет, Филологический факультет, 2013. С. 344–352.
2. Михневич А. Е. и др. Русский язык в Белоруссии. Мн.: Наука и техника, 1985. 272 с.
3. Рычкова Л. В. и др. Принципы создания газетного корпуса СМИ Гродненщины // Компьютерная лингвистика: научное направление и учебная дисциплина: сб. науч. ст. Вып. 2. Гомель: ГГУ им. Ф.Скорины, 2012. С. 88–92.

References

1. Kustova G. I., Savchuk S. O. Izuchenie leksiko-semanticheskoy i sotsiokul'turnoy spetsifiki russkoy rechi na territorii Respubliki Belarus' (na materiale tekstov SMI) // Trudy mezhdunarodnoy konferentsii «Korpusnaya lingvistika — 2013». SPb.: S.-Peterburgskiy gos. universitet, Filologicheskiy fakul'tet, 2013. S. 344–352.
2. Mikhnevich A. E. i dr. Russkiy yazyk v Belorussii. Mn.: Nauka i tekhnika, 1985. 272 s.
3. Rychkova L. V. i dr. Printsipy sozdaniya gazetnogo korpusa SMI Grodnenshchiny // Komp'yuternaya lingvistika: nauchnoe napravlenie i uchebnaya distsiplina: sb. nauch. st. Vyp. 2. Gomel': GGU im. F. Skoriny, 2012. S. 88–92.