

**ШКОЛА-СЕМИНАР «НАЦИОНАЛЬНЫЙ КОРПУС  
РУССКОГО ЯЗЫКА И ПРЕПОДАВАНИЕ»  
(Москва, 20 – 23 октября 2008 г.)**

С 20 по 23 октября 2008 г. Институт русского языка им. В.В. Виноградова АН РФ совместно с Высшей школой экономики проводил школу-семинар «Национальный корпус русского языка и преподавание». В семинаре приняли участие 53 участника из вузов, школ и исследовательских институтов России, Белоруссии, Киргизии, Казахстана, Эстонии, Молдавии.

В первый день участники семинара прослушали лекции ученых из института языкознания РАН, института русского языка им. В.В. Виноградова РАН, МГУ и ВШЭ. В последующие дни занятия проводились в компьютерных классах ВШЭ, где под руководством преподавателей МГУ и ВШЭ решались поисковые задачи.

В результате обучения участники семинара познакомились с целевой информационно-поисковой системой, дающей возможность получить самый разнообразный и максимально представительный иллюстративный материал о самых разных фактах русского языка.

Национальный корпус русского языка (НКРЯ) был подготовлен в рамках программы РАН «Филология и информатика» специалистами из нескольких институтов РАН при участии университетов Москвы, Санкт-Петербурга, Воронежа и Ижевска и с апреля 2004 г. открыт на общедоступном сайте компании «Яндекс» <http://ruscorpora.ru>

С появлением НКРЯ русский язык вошёл в число тех языков мира, для научного описания которых имеются не только полные академические словари и академические грамматики, но и собственный полный «академический корпус». Главной особенностью корпуса является наличие особой дополнительной информации о свойствах входящих в его состав текстов, т.е. разметки. В НКРЯ в настоящее время используется четыре типа разметки: метатекстовая, морфологическая, семантическая и акцентная. Система разметки постоянно совершенствуется.

Национальный корпус предполагает отражение языка во всем многообразии жанров, стилей, территориальных и социальных вариантов и т.п. Корпус содержит сбалансированный состав текстов, где отражены все типы письменных и устных текстов, существующих на данном языке в данное время (художественные разных жанров, публицистические, учебные, научные, деловые, разговорные, диалектные и т.п.). Эти тексты представлены пропорционально их доле в языке. НКРЯ по объёму является самым большим национальным корпусом: планируемый объём – 200 млн словоупотреблений, из которых 100 млн представляют XVIII, XIX и первую половину XX в., а другие 100 млн – современный русский язык (с 50-х гг. XX в. до настоящего времени). В настоящее время объём корпуса превысил отметку в 150 млн (ср.: объём Британского национального корпуса – 100 млн словоупотреблений).

НКРЯ предназначен прежде всего для обеспечения научных исследований лексической и грамматической структуры языка. С появлением корпуса ограничений на объём анализируемого материала и скорость поиска информации по существу нет, т.е. в распоряжении очень быстро оказываются огромные массивы текстов самого разного типа.

В НКРЯ включены: корпус современных текстов, древнерусский корпус, корпус параллельных текстов, корпус поэтических текстов, образовательный корпус, устный корпус, диалектный корпус. Подробная инструкция по правилам пользования НКРЯ расположена на главной странице сайта.

Освоение НКРЯ в вузах и школах только начинается. Преподавателям вузов он поможет быстро найти материал и для исследовательской деятельности, и для занятий со студентами. Необходимо осваивать этот новый ресурс и внедрять в научную и учебную деятельность.

**Н.В. Изотова**