

КОГНИТИВНЫЕ И КОРПУСНЫЕ ТЕХНОЛОГИИ СЛАВЯНСКОЙ ФИЛОЛОГИИ

Светлана Самуиловна Дикарева

кандидат филологических наук,
доцент кафедры межъязыковых коммуникаций и журналистики
руководитель научно-исследовательской группы
по прикладной когнитивистике
Факультет славянской филологии и журналистики
Таврическая академия
Крымский федеральный университет имени В. И. Вернадского
Симферополь, Республика Крым
E-mail: dikareva@crimea.edu



Предварительные замечания

Цифровое информационное пространство представляет собой новую среду для филологических исследований славянских языков в открытом академическом диалоге, направленном на взаимопонимание и сотрудничество. В академической коммуникации принципиальное значение имеет открытый доступ к тем знаниям, которые представлены в глобальной сети Интернет.

Когнитивные и корпусные технологии

Как представляется, методология современной славистики строится на основе интегрированных когнитивных и корпусных технологий. Интегрированная парадигма открывает новые перспективы для сотрудничества и взаимопонимания славистов [2]. Результаты исследования внедрены в учебный процесс на факультете славянской филологии и журналистики Таврической академии КФУ имени В. И. Вернадского (учебные дисциплины «Квантитативная лингвистика и информационные технологии», «Национальный корпус русского языка как исследовательский и образовательный ресурс»).

Лингвистический корпус, представленный как информационно-справочная система, представляет собой эффективный и полезный инструмент, который успешно используют не только лингвисты, но и все, кому интересно знать, как функционируют языки в разных социально-коммуникативных ситуациях [1], [3], [5], [6].

По объему и охвату жанров текстов роль лидеров принадлежит Национальным лингвистическим корпусам. Впервые термин «Национальный корпус» появился в названии Британского национального корпуса (British National Corpus, BNC), созданного в 1990-е годы в Великобритании специалистами-лексикографами.

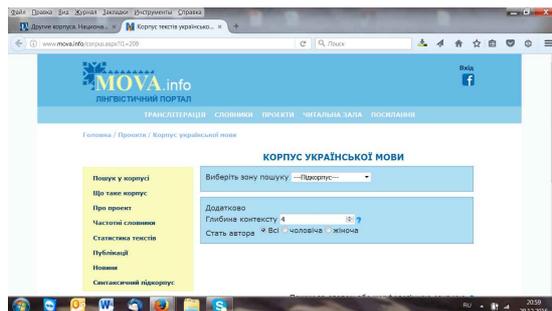
Национальным корпусом называют большой (не менее 100 млн. словоупотреблений) персонифицированный и репрезентативный массив текстов, характеризующий язык данной страны в целом. В славянских языках первым национальным корпусом был Чешский национальный корпус (Český národní korpus), содержащий коллекцию текстов на чешском литературном языке разных периодов (на октябрь 2016 года включает более 800 млн. словоупотреблений). Электронная коллекция открытого Национального корпуса русского языка на май 2016 года составляла более 600 млн.

Юрий Владимирович Рождественский, русский учёный-филолог, философ и культуролог, указывал: «Универсальная теория словесности получила название *«общая филология»*, т.к. она рассматривает опыт не одной речевой культуры, в частности русской, а всех речевых культур мира и кодифицирует речевые произведения как итог сравнения разных культур. В соответствии с данными общей филологии основные вехи стиля речи и стиля жизни связаны с созданием новых материалов речи и соответствующих им новых орудий речи» [4, с. 123].

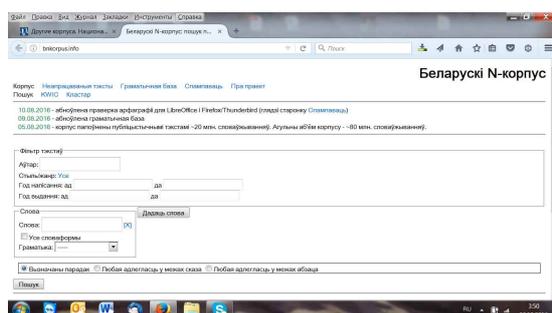
Сегодня стиль речи, стиль жизни, новые материалы речи и новые орудия речи славянской словесности обусловлены стремительным развитием корпусных ресурсов, созданных славистами разных стран и открытым в глобальном информационном пространстве Интернета. Всё большую популярность приобретают масштабные научно-исследовательские проекты, получившие название «краудсорсинга».

Ниже предлагаются адреса и изображения интерфейсов лингвистических корпусов славянских языков. Более полный перечень корпусов славянских языков размещён на сайте НКРЯ <http://ruscorpora.ru>, раздел главного меню «Другие корпуса».

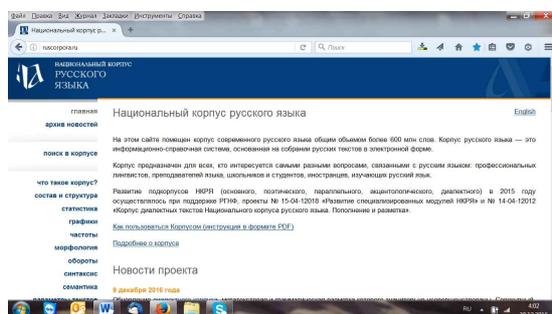
Корпусы славянских языков



Изображение 1. Корпус текстов украинского языка
<http://www.mova.info/corpus.aspx?l1=209>



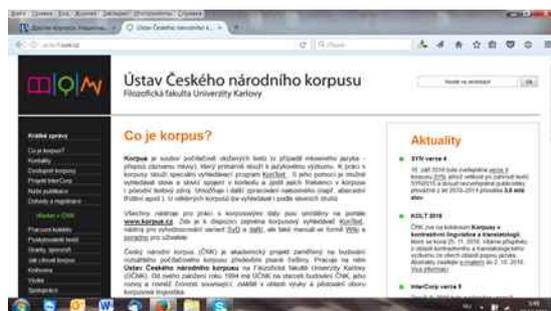
Изображение 2. Корпус белорусского языка
<http://bnkorporus.info/>



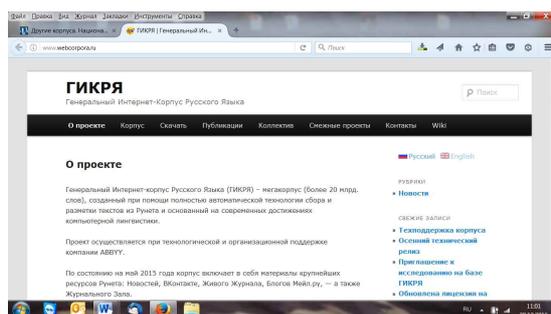
Изображение 3. Национальный корпус русского языка
<http://ruscorpora.ru/index.html>



Изображение 4. Тюбингемские корпуса русских текстов
<http://www.lingexp.uni-tuebingen.de/sfb441/b1/en/korpora.html>



Изображение 5. Чешский национальный корпус
<http://ucnk.ff.cuni.cz/>



Изображение 6. Генеральный интернет-корпус русского языка
<http://www.webcorpora.ru/>

ЛИТЕРАТУРА

1. Гиндин С.И. О культурных корнях корпусной лингвистики и её возможных импликациях для теоретического и прикладного языковедения /С.И.Гиндин // Труды международной конференции «Корпусная лингвистика – 2015». – СПб.: С.-Петербургский гос.университет, Филологический факультет, 2015. – С. 170–180 .
2. Дикарева С. С. Когнитивные и корпусные технологии в единой парадигме прикладной лингвистики / С. С. Дикарева //Ученые записки Таврического национального университета им. В. И. Вернадского, серия «Филология. Социальные науки». – Том 25 [64]. –№ 2. ч.2., – Симферополь. 2012. – С. 3 – 7.
3. Плунгян В. А. Почему современная лингвистика должна быть лингвистикой корпусов» [Электронный аудиоресурс] Публичные лекции на Полит.ру URL: <http://www.polit.ru/lectures/2009/10/23/corpus/html> Доступ свободный
4. Рождественский Ю.В. Принципы современной риторики / Ю.В. Рождественский; под ред. В.И. Аннушкина. – 4-е изд. испр. – М.: Флинта: Наука, 2005. – 176 с.
5. Савчук С.О. Лексико-семантические особенности текстов региональных СМИ по данным корпусного обследования/ С. О. Савчук//Труды международной конференции «Корпусная лингвистика – 2015». – СПб.: С.-Петербургский гос. университет, Филологический факультет, 2015. – С. 398–406.
6. Труды международной конференции «Корпусная лингвистика – 2015». – СПб.: С.-Петербургский гос. университет, Филологический факультет, 2015.– 466 с.