

МИНИСТЕРСТВО НА ОБРАЗОВАНИЕТО И НАУКАТА НА РЕПУБЛИКА БЪЛГАРИЯ
КИРИЛО-МЕТОДИЕВСКИ НАУЧЕН ЦЕНТЪР ПРИ БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ
ИЖЕВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМ. М. Т. КАЛАШНИКОВА
НАУЧНОЕ СООБЩЕСТВО “ПИСЬМЕННОЕ НАСЛЕДИЕ”
DIGITAL MEDIEVALIST SCHOLARLY COMMUNITY
ФОНДАЦИЯ „УСТОЙЧИВО РАЗВИТИЕ НА БЪЛГАРИЯ“

**Писменото наследство
и информационните технологии**

El’Manuscript–2014

Материали от V международна научна конференция
Варна, 15–20 септември 2014 г.

София · Ижевск
2014

Сборникът е издаден с финансовата подкрепа на Министерството на образованието и науката на Република България по процедура за подкрепа на международни научни форуми.

Отговорни редактори: проф. дфн В. А. Баранов
 доц. д-р В. Желязкова
 д-р А. М. Лаврентъев

Редактори: Нели Ганчева, Веселка Желязкова (български текст)
 О. В. Зуга, В. А. Баранов (руски текст)
 Кевин Хокинс (Kevin Hawkins) (английски текст)

Писменото наследство и информационните технологии [Текст] : материали от V международна науч. конф. (Варна, 15–20 септември 2014 г.) / отг. ред. В. А. Баранов, В. Желязкова, А. М. Лаврентъев. — София ; Ижевск, 2014. — 448 с.

Сборникът съдържа материали от конференция, посветена на разработването и създаването на съвременни средства за съхраняване, описване, обработка, анализ и публикуване на ръкописни и старопечатни книжовни паметници и исторически извори, а също и на въпросите за подготвянето на електронни ресурси в областта на хуманитаристиката и тяхното използване в научните изследвания и преподаването.

© Кирило-Методиевски научен център — БАН, 2014
© Ижевский государственный технический университет
им. М. Т. Калашникова, 2014
© Авторски колектив, 2014
© Лилия Тошкова — графичен дизайн на корицата, 2014

ISBN 978–954–9787–25–2

Электронный корпус региональных текстов “Жизненный круг”: проблемы типологии¹

Ю. Н. Драчева, П. Н. Задумина

Электронный корпус, региональный ключевой культурный текст, региональная картина мира, концепт, диалектная языковая картина мира

The Circle of Life Electronic Corpus of Regional Texts: Typology Problems

Yulia Dracheva, Polina Zadumina

This article deals with typology problems arising from an electronic corpus of regional texts. Using the *Circle of Life* corpus, the authors describe how to analyze the functioning of concepts in key regional cultural texts. The studying of key regional cultural texts in the form of corpora reveals basic concepts of a local language's picture of the world and provides a key to comprehending traditional culture, such as that of the Russian North.

Одной из самых динамично развивающихся сфер прикладного языкознания является корпусная лингвистика. Представление языковых фактов в виде электронного собрания текстов (собственно корпуса), снабженного поисковыми инструментами, метаразметкой, морфологической разметкой, системой аннотирования является неисчерпаемым источником материала для учебных и исследовательских проектов. К электронным корпусам относят известные национальные корпуса (например, Британский и Американский корпуса английского языка, Национальный корпус русского языка), электронные библиотеки (например, Проект “Гутенберг”, сайт “Библиотека Максима Мошкова”), электронные архивы современных газет (например, архивы *MailOnline*, *New York Times*), архивы с оцифрованными старыми газетами и возможностью поиска (например, *British Newspaper Archive*) и т. д. Объединяет же такие типологически разные проекты то, что основной единицей корпуса является *текст*. Типология электронных корпусов основывается на функциональном принципе: исследовательские корпуса, иллюстративные, мониторинговые (динамические), статические (авторские), мультимедийные и корпуса параллельных текстов.

Электронная лексикография в регионалистику, в частности в русскую диалектологию, пришла достаточно давно и имеет ряд успешно реализуемых проектов:

¹ Статья подготовлена при финансовой поддержке Российского государственного научного фонда (РГНФ) (конкурс поддержки молодых ученых 2014 года, проект № 14–34–01263 “Мультимедийный корпус вологодских текстов “Жизненный круг”).

диалектные фонохрестоматии (например, тамбовская, вологодская), электронные словари различных типов на основе картотек говоров (например, Электронный словарь строения слов вологодских говоров) и корпуса. Можно сказать, что современная регионалистика в России переживает бум создания региональных/диалектных электронных корпусов. Самый разработанный региональный электронный проект в этой области — это Мультимедийный Саратовский диалектологический текстовый корпус при Центре изучения народно-речевой культуры имени профессора Л. И. Баранниковой, целью которого является построение модели традиционного сельского общения на диалекте, для чего осуществляется наполнение каждого подкорпуса (на основе отдельного говора) текстовым материалом, отражающим типы диалектной речи (речь бытовую, фольклорную, обрядовую), различные формы речи (диалог, полилог, монолог), разнообразную тематику сельского общения, социальную дифференциацию носителей говора (пол, возраст, профессия, уровень образования) [Крючкова, Гольдин 2011].

Конечно, можно предположить, что однажды все региональные корпуса войдут в систему Национального корпуса русского языка (в котором есть диалектный подкорпус [Качинская 2009]), но этому препятствуют и отличия в подходах к созданию корпусов, и различающееся программное обеспечение, и вопрос авторских прав. Проблема также и в том, что национальные корпуса предлагают систему функционирования языка хотя и во всех его вариантах, но как единой гомогенной системы. В настоящее же время считается доказанным, что региональные варианты языка отличаются от литературного языка не только на уровне фонетических, лексических и грамматических структур (то есть того, что материально выражено и может быть включено в общий корпус, проиндексировано и т. д.), но и на уровне когнитивных.

В нашем проекте мы ориентируемся на понятие *регионального ключевого культурного текста* и доказываем возможность его использования для обоснования проекта корпуса именно региональных текстов: рассказов диалектоносителей, фольклорных записей, статей в местных СМИ, произведений вологодских писателей и поэтов.

Когнитивная лингвистика активно развивается за рубежом (Ch. J. Fillmore, L. Talmy, G. Lakoff, W. Croft, D. A. Cruse и др.) и в России (Ю. С. Степанов, В. И. Карасик, В. В. Красных, С. Г. Воркачев, Е. С. Кубрякова, И. А. Стернин, З. Д. Попова, В. Н. Телия, Н. Д. Арутюнова, Т. И. Вендина, К. И. Демидова, В. Е. Гольдин и др.) и предлагает различные подходы к анализу региональной концептосферы. Наш проект направлен на комплексное рассмотрение в функциональном, когнитивном и коммуникативном аспектах базовых концептов деятельности севернорусского крестьянства — <труд> и <игра>. Эти концепты представляют собой реализацию “жизненного круга” человека и воплощение того круга смыслов, который, по В. фон Гумбольдту, очерчен вокруг каждого народа.

Исследование концепта сквозь призму региональных культурных текстов позволяет нам говорить о концепции многофункционального корпуса. Подкорпус устной диалектной речи является динамическим, исследовательским и иллюстративным. Подкорпус литературный объединяет статическую и иллюстративную функции с некоторой исследовательской составляющей. Также учитываются статико-динамический иллюстративный фольклорный подкорпус и динамический исследовательский газетный подкорпус. Мультимедиазация является основным требованием ко всей базе корпуса. Подкорпус значимых исследовательских работ, выполненных на местном языковом материале, также имеет прикладное значение.

Таким образом, на основе анализа записей диалектной речи, полученных в результате диалектологических экспедиций в районы Вологодской области, строится когнитивная модель различных видов деятельности севернорусского крестьянина. В материалах периодической печати (районные газеты “Северная Новь”, “Звезда”, “Призыв”) анализируются особенности пропаганды общественно-значимых идей, а также отражение в статьях культурно значимой материально-бытовой и деятельностной сфер жизни. Исследование текстов писателей-вологоджан позволит смоделировать художественно-эстетический образ труда, праздников, игры в индивидуально-авторской картине мира (В. И. Белов, Н. А. Клюев, Н. И. Рубцов, А. Я. Яшин, О. А. Фокина, С. В. Викулов, И. Д. Полуянов, В. Ф. Тендряков и др.). Материалы устного народного творчества (пословицы, поговорки, приметы и пр.) позволяют дополнить описание концептологической составляющей деятельностной сферы жизни человека.

Проект “Мультимедийный корпус вологодских текстов “Жизненный круг” рассчитан на три года. В первый год выполняется работа по теоретическому обоснованию корпуса, по сбору и первичной обработке языкового материала (в устной и письменной форме фиксации), решению вопроса программного обеспечения (оболочки), наполнению текстовой базы (подкорпус диалектной речи и литературный подкорпус). Второй год планируется посвятить анализу уже собранных материалов, первичной разметке (токенизация, лемматизация, морфологический анализ) и созданию газетного подкорпуса. В третий год работы над проектом требуется редактировать и дополнять базу, совершенствовать информационно-справочный и поисковой аппарат корпуса.

Научная ценность проекта заключается в выявлении региональных черт общерусских концептов и комплексном описании локальной языковой картины мира; практическая же значимость проекта определяется созданием электронного корпуса, включающего в себя тексты, аудио- и видеоматериалы, запечатлевшие различные этапы “жизненного круга” человека в отдельном регионе. Электронный корпус потенциально является самой действенной формой сохранения и распространения культурно-исторического наследия региона, а также обеспечения преемственности традиционных ценностей.

Литература

- Качинская 2009 — Качинская И. Б. Корпус диалектных текстов в Национальном корпусе русского языка: состояние и перспективы // Лексический атлас русских народных говоров (Материалы и исследования): 2009. СПб.: Наука, 2009. С. 57–68.
- Крючкова, Гольдин 2011 — Крючкова О. Ю., Гольдин В. Е. Корпус русской диалектной речи: концепция и параметры оценки // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной Международной конференции “Диалог” (Бекасово, 25–29 мая 2011 г.). Вып. 10 (17). М.: Изд-во РГГУ, 2011. С. 359–367.