

Средства поиска и представления лингвистических данных в электронном издании как основа для изучения лексики и семантики древнерусского языка (на материале "Повести временных лет")

Р. А. Аникина

Ижевский государственный технический университет, Ижевск, Россия

1. На сегодняшний день при высоком развитии информационных технологий актуальным является создание в сети Интернет электронных коллекций, которые дают возможность хранения, обработки и изучения древнейших и средневековых русских и славянских текстов. Наличие подобных электронных изданий позволяет разносторонне исследовать древние рукописные памятники. Но при работе с таким материалом возникает ряд проблем, обусловленных наличием в сети Интернет лишь фотокопий или переводных текстов, что, в свою очередь, не позволяет полноценно исследовать интересующий текст [Баранов 2006].

В настоящее время уже существует большое количество порталов, где пользователь может увидеть рукописи или старопечатные книги, но не имеет возможностей сделать, например, выборку лингвистических единиц.

2. Информационно-аналитическая система «Манускрипт» содержит коллекции древнерусских текстов, а также информацию о разработках в области их электронной публикации и о методах создания полнотекстовых баз данных, содержащих сложные по структуре и составу рукописные памятники (URL: <http://manuscripts.ru/>).

Обработка материала ведется с помощью пользовательских модулей сайта (модуль выборок и запросов; специализированный редактор Olded; модуль грамматических словарей; морфологический анализатор), что дает не только возможность читать текст в виде, близком к оригиналу, но и возможность быстрой выборки искомого языкового материала.

3. Наше исследование ориентировано на изучение лексических единиц с корнем *добр-* в тексте «Повести временных лет» (далее – ПВЛ) по Лаврентьевской летописи (далее – ЛЛ). Источником исследования является электронное издание «Повести временных лет» по ЛЛ, сверенное с факсимильным изданием ПВЛ и опубликованное на портале «Манускрипт».

Исследуемые единицы были выбраны с помощью запросных форм портала, позволяющих познакомиться с контекстами, указателями и осуществить выборку данных.

Исследованием ПВЛ занимались И. П. Еремин, Д. С. Лихачев, О. В. Творогов, В. В. Колесов, Е. В. Душечкина, Г. Ф. Ковалев, Н. П. Некрасов, Г. А. Хабургаев, А. А. Шахматов и другие ученые [Еремин 1966; Лихачев 1979; Творогов 1981; Творогов 2001; Колесов 2001; Душечкина 1973; Ковалев 1980; Некрасов 1896, 1897; Хабургаев 1979; Шахматов 1940]. Язык ПВЛ интенсивно изучался в отечественной науке XIX века П. А. Лавровским, М. А. Колосовым, Н. П. Некрасовым [Лавровский 1852; Колосов 1872; Некрасов 1896].

Более обстоятельно исследуется лексический состав ПВЛ в монографиях Ф. П. Филина и А. С. Львова в XX веке. Целью авторов являлся анализ лексики памятника, который послужил основой для решения вопроса о характере и происхождении древнерусского литературного языка [Филин 1949; Львов 1975].

4. Для точного выбора интересующих нас контекстов необходимо правильно, учитывая будущую форму результата, сделать запрос, после выполнения которого электронное издание ПВЛ предоставляет все необходимые для изучения контексты (фрагменты текста) (URL: http://manuscripts.ru/mns/portal.main?p1=12&p_lid=1). Использование запросной формы оказывается более удобным и быстрым способом поиска материала, чем выборка его вручную по печатному изданию, где не сохранены графико-орфографические особенности текста.

Для нашей работы необходимо было сделать выборку всех контекстов, где употребляются слова с корнем *добр-*. Для этого мы должны были:

- 1) выбрать с помощью запросных форм исследуемый текст;
- 2) задать в запросной строке искомую форму, в данном случае – основу *добр*;

- 3) поставить в запросной строке после заданной искомой формы знак %, для того чтобы система выдала все возможные формы с данной основой, в том числе и двусловные образования;
- 4) выбрать форму результата – текст или указатель;
- 5) указать диапазон листов;
- 6) задать поиск контекстов с искомым элементом.

В результате был получен перечень контекстов в виде фрагментов текста с выделением искомого элемента и в виде ссылки на фрагмент, где употреблена искомая единица. Ссылка связана с одним из указателей. Так, в количественном указателе дано количество словоформ в тексте (в нашем случае – с корнем *добр-*); в прямом и обратном указателях система выдает список форм, расположенных в алфавитном порядке (в обратном указателе это инверсированный порядок) со ссылкой на контекст через адрес словоформы.

При использовании многотекстовой запросной формы система дает возможность выборки сразу по нескольким рукописям, что позволяет выявить специфику употребления интересующих форм в разных текстах. В этом режиме предоставляется возможность использовать в качестве формы результата не только перечисленные выше перечни, но еще и списочный и табличный указатели. В последнем материал представлен в виде таблицы, в столбцах которой помещены: список словоформ и их адреса для каждого текста отдельно. Списочный указатель содержит искомые формы, их количество в анализируемых текстах, аббревиатуры рукописей, адреса контекстов.

Следует отметить, что в зависимости от требований пользователя маска поиска может быть разной – *основа*, *основа + суффикс*, *приставка + основа*, *вся форма* и др. Важно правильно сформировать запрос, используя маски: % (процент) – для любого количества любых символов и _ (нижнее подчеркивание) – для любого одного символа.

5. При работе с текстом ПВЛ мы поставили цель выявить специфику функционирования слов с корнем *добр-* как одного из способов репрезентации представления о добре.

При выявлении значений для исследуемых единиц мы опирались, главным образом, на концепцию В. В. Колесова, в которой раскрывается специфика представления о добре в средневековом сознании [Колесов 2001].

В ходе исследования мы выяснили, что слова с общим корнем *добр-* в исследуемом памятнике имеют семантику, которая обусловлена, во-первых, этическим аспектом ('красивый':), во-вторых, прагматической функцией ('полезный, пригодный':). Рассмотренные примеры свидетельствуют о сложной семантической структуре слов с общим корнем *добр-*. Репрезентация того или иного значения зависит прежде всего от условий контекста.

В результате проведенного анализа контекстов нами выявлены основные значения понятия *добро*, такие, как 'польза', 'красота', 'истина', при этом установлено, что представление о добре изменялось. Иначе говоря, первично употребление слова *добро* в значении 'польза' (

29 л.). Позднее *добро* как 'польза' переосмысливается в *добро* как 'красоту' (

17 об.). Одновременно развивается и другое значение, вытекающее из признаков Блага – 'истинный, праведный' (_____ 18 л.).

Архаическое (дохристианское, языческое) сознание отражает представление о добре как о земном воплощении всего положительного, отсюда внешние проявления этого добра - то, что можно увидеть (например, внешняя красота), использовать для себя, сделать пригодным (отсюда польза). Христианство предлагает новое представление о добре как о благе: земное добро – это не только путь к небесному благу. И именно эта идея находит свое отражение в тексте «Повести временных лет».

В заключение следует сказать, что использование таких систем обработки древних текстов позволяет не просто познакомиться с графико-орфографическими особенностями текста, но и осуществить определенные операции при работе с ним для последующего лингвистического анализа материала. Такие системы позволяют также не только работать с текстом, исследовать его, но и получать всю необходимую информацию о нем. Современные технологии должны обеспечить доступность текстов современным исследователям в области гуманитарных наук, в том числе в сети Интернет.

Благодарности

Создание Интернет-коллекции «Повести временных лет» в трех списках летописи: Лаврентьевском (1377 г.), Ипатьевском (первая четверть XV в.) и Радзивилловском (XV в) выполнено в рамках проекта, поддержанного Российским фондом гуманитарных исследований (РФГНФ), грант № 07-04-12140в, а также в рамках тематического плана Ижевского государственного технического университета (исследование текста и языка).

Список источников

ЛЛ – «Повесть временных лет» по Лаврентьевской летописи : Полное собрание русских летописей. Т. 1. : Лаврентьевская летопись. – Репринт. изд. 1926–1928 гг. – М. : Языки русской культуры, 1997. – 733 с.

л. – лицевой лист.

об. – оборотный лист.

ПВЛ – «Повесть временных лет».

Список литературы

Баранов 2006 – *Баранов, В. А.* От оцифрованных коллекций средневековых рукописей к электронным многофункциональным Интернет-библиотекам / В. А. Баранов // Современные информационные технологии и письменное наследие: от древних рукописей к электронным текстам. – Ижевск, 2006. – С. 3-9.

- Душечкина 1973 – *Душечкина, Е. В.* Художественная функция чужой речи в русском летописании / Е. В. Душечкина // Учен. зап. Тарт. ун-та. – 1973. – Вып. 306 : Тр. по рус. и славян. филологии. Т. 21. – С. 65–104.
- Еремин 1966 – *Еремин, И. П.* «Повесть временных лет» как памятник литературы / И. П. Еремин // Еремин И. П. Литература Древней Руси : Этюды и характеристики / И. П. Еремин. – М.; Л. : Изд-во АН СССР, 1966. – С. 43–47.
- Ковалев 1980 – *Ковалев, Г. Ф.* Собираемые этнонимы в «Повести временных лет» / Г. Ф. Ковалев // Сравнительно-исторические исследования русского языка. – Воронеж : Изд-во Ворон. ун-та, 1980. – С. 69–74.
- Колесов 2001 – *Колесов, В. В.* Древняя Русь: наследие в слове : в 5 кн. Кн. 2. Добро и зло. / В. В. Колесов. – СПб. : Филол. ф-т СПбГУ, 2001. – 304 с.
- Колосов 1872 – *Колосов, М. А.* Очерк истории звуков и форм русского языка с XI по XIV столетия / М. А. Колосов. – Варшава, 1872. – 192 с.
- Лавровский 1852 – *Лавровский, П. А.* О языке северных русских летописей / П. А. Лавровский. – СПб., 1852. – 163 с.
- Лихачев 1979 – *Лихачев, Д. С.* Поэтика древнерусской литературы / Д. С. Лихачев. – М. : Наука, 1979. – 359 с.
- Львов 1975 – *Львов, А. С.* Лексика «Повести временных лет» / А. С. Львов. – М. : Наука, 1975. – 367 с.
- Некрасов 1896 – *Некрасов, Н. П.* Заметки о языке «Повести временных лет» по Лаврентьевскому списку / Н. П. Некрасов // Изв. Отд. рус. яз. и словесности. – 1896. – Т. 1. – С. 832–927; 1897. – Т. 2, кн. 1. – С. 104–174.
- Творогов 1981 – *Творогов, О. В.* Литература Древней Руси / О. В. Творогов. – М. : Просвещение, 1981. – 128 с.
- Творогов 2001 – *Творогов, О. В.* Лексический состав «Повести временных лет» (словоуказатели и частотный словник) / О. В. Творогов // Лаврентьевская летопись. (Полное собрание русских летописей.

Том первый.) – 2-е изд. – М. : Языки славянской культуры, 2001. – 496 с.

Филин 1949 – *Филин, Ф. П.* Лексика русского литературного языка древнекиевской эпохи (по материалам летописей) / Ф. П. Филин // Учен. зап. Ленингр. пед. ин-та. – Т. 80. – 1949. – 288 с.

Хабургаев 1979 – *Хабургаев Г. А.* Этнонимия «Повести временных лет» в связи с задачами реконструкции восточнославянского глоттогенеза / Г. А. Хабургаев – М. : Изд-во Моск. ун-та, 1979. – 231 с.

Шахматов 1940 – *Шахматов, А. А.* «Повесть временных лет» и ее источники / А. А. Шахматов // Тр. Отд. древнерус. лит-ры. – 1940. – Т. 4. – С. 9–140.

Манускрипт – *Повесть временных лет* [Электронный ресурс] // «Манускрипт: письменное наследие» / Лаборатория по автоматизации филологических работ УдГУ ; кафедра лингвистики ИжГТУ. – 2004–2008. – Режим доступа : http://manuscripts.ru/mns/portal.main?p1=12&p_lid=1, свободный. – Загл. с титул. страницы.

Searching and presentation of linguistic data in a digital edition as the basis for studying the lexicon and semantics of Old Russian (using material from the Primary Chronicle)

Regina A. Anikina

Izhevsk State Technical University, Izhevsk, Russia

This paper deals with the electronic edition of the Russian Primary Chronicle (Laurentian codex) published in the Manuscript informational-analytical system. Use of the Manuscript informational-analytical system is explained for a sample query, with a technical explanation of how to derive a list of results of one word form in various contexts.