Компьютерные алгоритмы лингво-текстологической стратификации Повести временных лет

Д. А. Добровольский 1 , С. М. Михеев 2 Российский государственный гуманитарный университет, Москва 2 Институт славяноведения РАН, Москва

The paper proposes a new method of textual study of the Russian Primary Chronicle. The method includes analysis of the graphics, vocabulary, and syntax of the Laurentian copy of The Chronicle with the use several simple algorithms written in the Python 3 programming language. The idea is to generate a thesaurus and analyze the distribution patterns of the numerous words and spellings in order to identify the specific usages typical of each of the Chronicles' authors.

Еще в XIX в. исследователи пришли к заключению, что Повесть временных лет [далее — ПВЛ] не была первым местным сочинением по истории Руси. Однако восстановление предшествующих ПВЛ летописных сводов представлялось авторам позапрошлого века делом «скорее художественным, чем ученым» [Костомаров, 1861, с. 3]. Научный подход к проблеме, ставящий во главу угла требования проверяемости и воспроизводимости, был предложен А.А.Шахматовым, чьи «Разыскания о древнейших русских летописных сводах» [Шахматов, 2002 (1908)] и неоконченное издание ПВЛ [Шахматов, 2003 (1916); Шахматов, 1940] являются классикой современного летописного источниковедения. Вместе с тем, опыт А.А. Шахматова также оказался неоднозначным, о чем свидетельствует появление значительного числа альтернативных версий начальной истории русского летописания, оспаривающих не только частности, но и базовые тезисы предшественника, включая атрибуцию ПВЛ Нестору Печерскому, двукратное редактирование летописи во втором десятилетии XII в. и даже предположение о том, что свод-предшественник ПВЛ сохранился в составе Новгородской I летописи младшего извода [Истрин, 1923–1924; Бугославский, 2006 (1941); Алешковский, 1971; Кузьмин, 1977 и др.]. «Общепризнанных» идей в современном летописеведении практически нет.

Наблюдаемой диссоциации летописного источниковедения как научной дисциплины в немалой степени способствовал специфический состав сохранившихся источников. Текст Начальной летописи (т.е. Повести временных лет и предшествовавших ей сводов XI в.) представлен значительным количеством списков, среди которых есть и относительно ранние. не испытавшие существенного воздействия со стороны позднейших редакторов, деятельность которых стала особенно активной в XV-XVI вв. Однако все эти списки, за исключением Новгородской I летописи XIII-XIV вв., либо восходят к одному протографу, которым была ПВЛ, как она сложилась после 1115 г., либо испытали определяющее воздействие этого сочинения. В свою очередь, новгородская традиция сохранила более ранний текст лишь частично, причем число этих «частей» тоже стало предметом споров: одни говорят о двух фрагментах [Шахматов, 2002 (1908), с. 26], другие только об одном [Алешковский, 1971, с. 25-31, 71-83; Гиппиус 1997, с. 50-55]. Описанное положение ограничивает возможности сопоставления летописей. Какую-то пищу для рассуждений могут дать грамматические нестыковки, вроде остатков двойственного числа в известии о захвате Олегом Киева [ПСРЛ I], но число таких ситуаций ограничено. Наконец, нестыковки в хронологии могут иметь текстологическое значение лишь в том случае. если доказано, что летописное повествование развивалось строго линейно и целью книжников было дать точные датировки событий, а наблюдения, например, над записями 6571-6575 гг. о новгородском походе Всеслава говорят скорее об обратном [Добровольский, 2009]. Пробелы соблазнительно заполнить соображениями о «нелогичности», «непоследовательности» и «противоречиях» летописного текста. Очевидно, однако, что такого рода построения опираются на определенную интерпретацию изучаемого источника, которую можно принимать, а можно и не принимать. Логика средневекового нарратива столь тонка и сложно (на современный взгляд) организована, что получить

сколько-нибудь обязывающие результаты путем ее анализа практически невозможно. Необходимо искать новые методики текстологической стратификации ПВЛ.

Одну из возможных методик предложил А.А.Гиппиус Гиппиус, 1996; Гиппиус, 2006 и др.]. Изучая язык Синодального списка Новгородской I летописи, ученый заметил, что некоторые примечательные графико-орфографические и лексикосинтаксические явления распределены в этом источнике неравномерно. В плане истории выявленные «швы» обнаружили высокую степень корреляции со сменой новгородских архиепископов. Очевидно, заключил ученый, ведение летописей в Новгороде было занятием официальным и осуществлялось под контролем церковных властей, а языковая неоднородность связана с тем, что новый владыка, как правило, передавал дело историописания новому писцу. Аналогичная связь между вероятной сменой руки и появлением / исчезновением определенного круга слов и оборотов была выявлена при изучении способов введения прямой речи в Начальной летописи [Гиппиус, 2001]. Границы зон распространения тех или иных языковых явлений в большинстве случаев предельно наглядны, что обеспечивает воспроизводимость полученных выводов. «Лингвистическая карта» может дать ответы на многие вопросы из истории текста ПВЛ.

Однако такой карты до сих пор не существует: наблюдения самого А.А.Гиппиуса касаются одного пласта лексики (verba dicendi), а недавняя работа А.Тимберлейка [Timberlake, 2005] ограничена не только по числу изучаемых показателей (это оформление заголовков известий), но и по охвату текста (рассмотрены только статьи 1050-1110 гг.). А.А.Гиппиус пишет, что соотношение тематических и атематических форм аориста от глагола речи (разъ и раша vs. рекохъ и рекоша) «является, по-видимому, наиболее ярким» среди «текстологически релевантных морфологических признаков» [Гиппиус, 2001, с. 165]. Однако далее ученый анализирует еще несколько признаков, каждый из которых «обладает, так сказать, собственным текстологическим "рисунком"» [Там же, с. 170]. Очевидно, число маркеров индивидуального стиля летописцев может увеличиваться и дальше.

Работа по выявлению таких маркеров весьма трудоемка. Существует несколько словоуказателей к ПВЛ, дающих более или менее точное представление о лексике памятника и позволяющих выявить употребления того или иного конкретного слова [Творогов, 1997 (1984); Gröber, Müller, 1977—1986]. Однако необходимо не только найти все вхождения данной лексемы в текст, но и оценить равномерность их распределения, выявить максимумы и минимумы, а также найти слова, распределение которых по тексту окажется сходным (чтобы исключить случайные результаты и разного рода статистические «артефакты»). Решить все эти задачи без использования компьютера не представляется возможным.

Предлагаемые алгоритмы реализованы на языке Python3 [http://pvthon.org/download/releases/3.1.2/], выбор которого обусловлен тремя соображениями. Во-первых, данный язык способен без дополнительной настройки и преобразований совершать операции над строками в кодировке Unicode, что дает возможность учитывать большинство орфографических особенностей изучаемого текста. Во-вторых, Python позволяет наглядно обрабатывать так называемые словари, или коллекции пар «ключ» — «значение» (в качестве ключа выступают слова, как они представлены в источнике, в качестве значения ссылки на те столбцы 1-го тома Полного собрания русских летописей, где эти слова, а точнее написания, встречаются). Втретьих, Python является интерпретируемым языком; написанные на нем программы хранятся в виде исходного кода, что позволяет оперативно корректировать используемые алгоритмы. В-четвертых, наконец, Python распространяется под очень либеральной лицензией [http://python.org/psf/license/], что дает возможность свободно скачивать и устанавливать необходимые интерпретаторы и среду разработки, а также распространять написанные программы. Объектом обработки стал текст ПВЛ по Лаврентьевской летописи, заимствованный из проекта Д. Бирнбаума "The e-PVL" [http://clover.slavic.pitt.edu/pvl/] и преобразованный в файл CSV, каждая строка которого содержит три поля: ссылку на столбец издания, ссылку на строку в столбие и собственно текст источника.

В настоящее время реализованы два модуля. Первый разбирает исходный файл Lav.csv, создавая словарь, где в качестве

ключей выступают словоформы, а в качестве значений ссылки на столбцы, в которых эти формы выявлены. Результатом работы модуля является еще один файл CSV, представляющий собой словоуказатель к ПВЛ по Лаврентьевскому списку. Второй модуль выявляет и группирует слова, распределение которых оказывается тождественным или сходным (с определенным допуском). Это позволяет оперативно и наглядно проверять любые догадки относительно истории летописного текста, и — более того — сравнивать догадки между собой по числу соответствующих им явлений. В дальнейшем планируется разработать модуль грамматического анализа, использование которого, впрочем, представляется факультативным, поскольку текстологическое значение может иметь не столько употребление слова как такового, сколько употребление его в конкретном времени / лице / числе / падеже, свидетельствующее о склонности / антипатии книжника к конструкциям того или иного типа.

По мере накопления опыта работы с материалом отдельные наблюдения составят полную «лингвистическую карту» ПВЛ.

Список литературы

- Алешковский, 1971 Алешковский М. Х. Повесть временных лет: судьба литературного произведения в Древней Руси. М., 1971.
- Бугославский, 2006 (1941) Бугославский С. А. «Повесть временных лет»: (Списки, редакции, первоначальный текст) // Текстология Древней Руси. М., 2006. Т. 1. С. 281–312.
- Гиппиус, 1996 Гиппиус А.А. Лингво-текстологическое исследование Синодального списка Новгородской Первой летописи: автореф. дис. ... канд. филол. наук. М., 1996.
- Гиппиус, 1997 Гиппиус А. А. К истории сложения текста Новгородской Первой летописи // Новгород. историч. сб-к. Вып. 6 (16). СПб., 1997. С. 3–72.
- Гиппиус, 2001 Гиппиус А. А. Рекоша дружина Игореви...: к лингвотекстологической стратификации Начальной летописи // Russian Linguistics. Vol. 25. 2001. P. 147–181.
- Гиппиус, 2006 Гиппиус А. А. История и структура оригинального древнерусского текста (XI-XIV вв.): комплексный анализ и реконструкция: автореф. дис. ... д-ра филол. наук. М., 2006.
- Добровольский, 2009 Добровольский Д. А. Когда Всеслав ходил на Новгород: (опыт интерпретации летописной статьи 6575 г.) // Вспомогательные исторические дисциплины в пространстве гуманитарного знания: материалы XXI междунар. науч. конф. М., 2009. С. 146–149.

- Истрин, 1923—1924 Истрин В. М. Замечания о начале русского летописания: (по поводу исследований А. А. Шахматова в области древнерусской летописи) // Изв. Отд. рус. яз. и слов. РАН. Т. 26. С. 45–102. Т. 27. Пг.; Л., 1923–1924. С. 207–251.
- Костомаров, 1861 Костомаров Н. И. Лекции по русской истории. Ч. 1: Источники русской истории. СПб., 1861.
- Кузьмин, 1977 Кузьмин А. Г. Начальные этапы древнерусского летописания. М., 1977.
- ПСРЛ I Полное собрание русских летописей [Репринт. изд.]. Т. 1: Лаврентьевская летопись. М., 1997.
- Творогов, 1997 (1984) Творогов О.В. Лексический состав «Повести временных лет»: (Словоуказатели и частотный словник): (Кроме «Поучения» Владимира Мономаха) // ПСРЛ. М., 1997. Т. 1. С. 581–733.
- Шахматов, 1940 Шахматов А. А. «Повесть временных лет» и ее источники // Тр. Ин-та рус. лит. АН СССР. Отд. др.-рус. лит. Л., 1940. Т. 4. С. 9–150.
- Шахматов, 2002 (1908) Шахматов А. А. Разыскания о древнейших русских летописных сводах // История русского летописания. СПб., 2002. Т. 1. Кн. 1. С. 20–483.
- Шахматов, 2003 (1916) Шахматов А.А. Повесть временных лет // История русского летописания. СПб., 2003. Т. 1. Кн. 2. С. 527–977.
- Gröber, Müller, 1977–1986 Gröber B., Müller L. Vollständiges Wörterverzeichnis zur Nestorchronik / Handbuch zur Nestorchronik / Hrsg. von Ludolf Müller. Bd. III. München, 1977–1986. Lief. 1–4.
- Timberlake, 2005 Timberlake A. Intervals of The Kiev Chronicle (1050–1110) // Zeitschrift für slavische Philologie. Vol. 64. 2005. S. 51–70.