

Национальный корпус русского языка

Зачем нужен поэтический корпус и как его использовать

Кирилл Михайлович Корчагин, Институт русского языка им. В. В. Виноградова РАН
(Россия, Москва), stivendedal@gmail.com

DOI: 10.31857/S013161170007630-2

АННОТАЦИЯ: Поэтический корпус в составе Национального корпуса русского языка — инструмент для исследователей русской поэзии и поэтического языка. Корпус содержит обширную коллекцию русской поэзии XVIII–XX веков, отражает все заметные поэтические направления и продолжает пополняться. В нем присутствуют два типа разметки — грамматическая и стиховедческая. Если первая совпадает с разметкой в основном корпусе НКРЯ, то вторая — специфична именно для поэтического корпуса. Эта разметка содержит ключевые параметры поэтического текста, среди которых метр, строфика, схема рифмовки и другие. Эти параметры позволяют производить поиск и структурировать массивы русских поэтических текстов с совпадающими формальными параметрами, автоматическим образом выполняя задачу, для решения которой раньше требовались специализированная справочная литература и/или трудоемкая обработка текстов. Использование стиховедческой разметки вместе с грамматической позволяет пользователю корпуса решать задачи, лежащие на границе лингвистики и литературоведения. Среди таковых есть и традиционные для русского стиховедения задачи — такие, как исследование истории поэтических форм, — и задачи, которые могут представлять интерес для историков языка, например исследование русского ударения. В статье показаны примеры решения обоих типов задач, а также кратко описан функционал поэтического корпуса.

КЛЮЧЕВЫЕ СЛОВА: поэтический корпус, Национальный корпус русского языка, стиховедение, digital humanities, история русского ударения, лингвистическая поэтика

для цитирования: Корчагин К. М. Зачем нужен поэтический корпус и как его использовать // Русская речь. 2019. № 6. С. 113–127. DOI: 10.31857/S013161170007630-2.

БЛАГОДАРНОСТИ: Статья написана при поддержке гранта Президента Российской Федерации МК-2219.2018.6 «Типология синтагматических членений в русском стихе XVIII–XX века».

Russian National Corpus

What Russian Poetic Corpus is Made for and How to Use It

Kirill M. Korchagin, Vinogradov Russian Language Institute (Russian Academy of Sciences)
(Russia, Moscow), stivendedal@gmail.com

ABSTRACT: A poetic sub-corpus of the Russian National Corpus is a useful tool for scholars of Russian poetry and language. The poetic corpus contains a spacious collection of the 18th–20th century’s Russian poetry, it represents all its major trends and movements, and its creators update and enrich it from time to time.

There are two layout types in the corpus: those of grammar and of versification. If the first is the same as in the “main” corpus, the second is specific only for the poetic one. The latter layout consists of key parameters of any text of poetry which includes poetic meter, stanza structure, rime pattern, and so on. These parameters allow to search for Russian poetic texts and automatically structure them. Previously to perform such tasks special reference literature or laborious text processing was required.

The user of the corpus can combine these two kinds of parameters to solve tasks lying on the border between linguistics and theory of literature. For instance, it can be an investigation into the history of Russian accentuation. The present paper briefly regards possible ways to resolve such tasks through the prism of the poetic corpus functionality.

KEYWORDS: poetic corpus, National Russian Corpus, theory of verse, digital humanities, history of Russian prosody, linguistic poetics

FOR CITATION: Korchagin K. M. What Russian Poetic Corpus is Made for and How to Use It. *Russian Speech = Russkaya Rech'*. 2019. No. 6. Pp. 113–127. DOI: 10.31857/S013161170007630-2.

ACKNOWLEDGEMENTS: The paper was supported by grant of president of Russian Federation, project МК-2219.2018.6. “The typology of Syntagmatic Segmentation in Russian poems of the 18th — 20th century”.

Среди всех дисциплин, пользующихся лингвистическими корпусами, теория стиха занимает особое место. С одной стороны, с первых шагов «научного» стиховедения в XIX веке, когда исследователи впервые стали отступать от догматических и нормативных способов описания стиха, они начали работать с массивами данных, которые мы сейчас назвали бы корпусами. Но с другой стороны, создание первых современных корпусов, то есть систем, состоящих не только из текстов, но и из инструментов их анализа, никак не было связано с исследованием литературных текстов, тем более поэтических. Это несовпадение двух, казалось бы, предельно близких, научных идеологий примечательно. Фактически только в последнее десятилетие получило распространение использование корпусов за пределами лингвистических задач, но даже сейчас, когда научная идеология *digital humanities* получила широкое распространение, поэтические корпуса часто не получают заслуженного внимания.

На сегодняшний день кроме русского поэтического корпуса существует несколько других корпусов такого рода, но количество их очень мало, а те, которые есть, как правило, представляют собой коллекции текстов без специфической стиховедческой разметки¹. Доступны корпуса русской

¹ Среди таких корпусов можно упомянуть корпус английской поэзии *Gutenberg English Poetry Corpus*, объем которого — 2 миллиона словоупотреблений (для сравнения: в поэтическом подкорпусе НКРЯ — около 12 миллионов). Корпус доступен по адресу: <https://github.com/aparrish/gutenberg-poetry-corpus>. Среди других — корпус среднеанглийской поэзии (*Parsed Corpus of Old English Poetry*; <https://pcmer.net/>) объемом около 200 тысяч словоупотреблений, а также корпус староанглийской поэзии, разработанный специалистами из Йорка и Хельсинки (*The York-Helsinki Parsed Corpus of Old English Poetry*; <http://www-users.york.ac.uk/~lang18/pcorpus.html>) и содержащий 17 тысяч словоупотреблений. Также интерес представляет параллельный корпус русских и французских поэтических текстов первой трети XIX века, подготовленный Б. В. Ореховым (<http://nevmenandr.net/fr/>).

и чешской поэзии (<http://www.versologie.cz/en/kcv.html>), а также башкирский поэтический корпус (<http://web-corpora.net/bashcorpus/search/>), во многом экспериментальный. В какой-то мере поэтическим корпусом можно считать электронную библиотеку «Персей» (<http://www.perseus.tufts.edu/hopper/>), где собраны тексты европейской античности. Русский поэтический корпус на фоне этих проектов выделяется и количественно, и качественно: представленный в нем массив текстов отражает практически все течения в русской поэзии XVIII–XX веков и продолжает пополняться, а специализированная стиховедческая разметка отличается большой подробностью.

Поэтический корпус, насколько возможно, стремится к полноте. Однако это не значит, что в него попадают все стихотворения, написанные на русском языке. Напротив, авторы проходят достаточно строгий отбор. Критериями для попадания в корпус оказывается появление академических собраний (прежде всего, в серии «Библиотека поэта», но не только), присутствие в антологиях русской поэзии, наличие исследовательской литературы, а для поэтов последних пятидесяти лет — литературы критической [Корчагин 2015]. Все вместе это позволяет создать относительно «взвешенную» картину развития русской поэзии. Можно сказать, что авторы, представленные в поэтическом корпусе, делятся на три группы. В первую входят те, которые представлены полными или наиболее полными из доступных собраниями сочинений, чаще всего это наиболее известные русские поэты — от Ломоносова и Пушкина до Марины Цветаевой и Бориса Слуцкого. Во вторую группу входят поэты, которые играли достаточно заметную роль в литературном процессе эпохи, но никогда не занимали в нем центральных мест. Наконец, третья группа — поэты, которые присутствуют в корпусе в виде отдельных стихотворений. Чаще всего объем сочинений таких авторов ограничен объективными причинами — например, плохой сохранностью (как в случае стихов ленинградского поэта Алика Ривина) или мимолетным появлением поэта на литературной сцене.

Если грамматическая разметка в поэтическом корпусе совпадает с разметкой в основном, то стиховедческая устроена иначе. Именно она превращает поэтический корпус в особый, не имеющий аналогов инструмент. Чтобы получить доступ к этой разметке, пользователь должен зайти на страницу поэтического подкорпуса (<http://ruscorpora.ru/search-poetic.html>) и выбрать ссылку «Задать подкорпус». На той же странице расположена подробная инструкция к корпусу (<http://ruscorpora.ru/new/instruction-main.pdf>), к которой лучше всего обратиться для более подробного ознакомления с функционалом корпуса (рис. 1–2).

национальный корпус
РУССКОГО
ЯЗЫКА

главная

Поэтический корпус: выбор подкорпуса

Вы можете задать подмножество корпуса, по которому в дальнейшем будет вестись поиск. Подробнее о параметрах текста см. в разделе «[Параметры текста](#)». О специальных параметрах и терминологии, используемых в корпусе поэтических текстов, см. [терминологический указатель](#). См. также [список авторов](#), произведений которых размещены в Корпусе

основной синтаксический

газетный

параллельный

обучающий

диалектный

поэтический

устный

акцентологический

мультимедийный

мультиязык

исторический

НООС

использование корпуса

Основные параметры текста ?

Название

Цикл

Книга стихов

Автор текста без приписываемого

Пол: любой мужской женский

Год рождения: с по

Автор оригинала (для переводных текстов)

Год создания: с по

Место написания — авторская запись

Место написания — унифицированное название (например, Ленинград, СПб -> Петербург)

Число словоформ: с по

Рис. 1. Поэтический корпус: выбор подкорпуса и доступные параметры поиска (начало)

Fig. 1. Poetic corpus: choosing subcorpora and available search options (the beginning of the procedure)

Место написания — авторская запись

Место написания — унифицированное название (например, Ленинград, СПб -> Петербург)

Число словоформ: с по

Число строк: с по

Жанр текста [выбрать](#)

Язык оригинала (для переводных текстов) [выбрать](#)

Метр [выбрать](#) (введите другое)

Длина строки (стопы, икты в тонике, слоги в силлабике) [выбрать](#) (введите другое)

Клаузула [выбрать](#) (введите другое)

Строфика [выбрать](#) (введите другое)

Графическая строфика [выбрать](#) (введите другое)

Рифма [выбрать](#) (введите другое)

Формула

Дополнительные параметры [выбрать](#)

Рис. 2. Поэтический корпус: выбор подкорпуса и доступные параметры поиска (продолжение)

Fig. 2. Poetic corpus: choosing subcorpora and available search options (proceeding with it)

Специфические характеристики поэтических текстов, с которыми может работать пользователь, делятся на две группы: те, что касаются обстоятельств создания текстов, и те, что описывают структуру стиха. К первому типу относятся название текста, цикл и/или книга стихов, в которые он входит, имя автора, а в случае если текст переводной — имя переводчика и язык оригинала. Здесь также можно выбрать тексты, написанные только женщинами или только мужчинами, задать дату рождения автора. Далее идут параметры, связанные с обстоятельствами создания текстов — год и место написания, а также объем текста в строках или словоформах. Затем указывается жанр текста: пользователь корпуса может выбрать из большого репертуара жанров — от басни и баллады до песни и кантаты. Разметка по жанрам применяется в основном к поэзии XVIII–XIX веков, но иногда она появляется и в более поздних поэтических текстах.

Эта предварительная, самая «грубая» разметка поэтического корпуса, однако даже ее использование может быть информативно. Например, можно получить сведения о распространении жанра баллады в русской поэзии: всего баллад в корпусе 308 по версии 2019 года. В корпусе этих баллад можно производить поиск грамматическими инструментами основного корпуса, а список баллад можно дополнительно упорядочить, например, по году создания (рис. 3). При этом стиховедческие характеристики каждой из баллад также будут доступны.

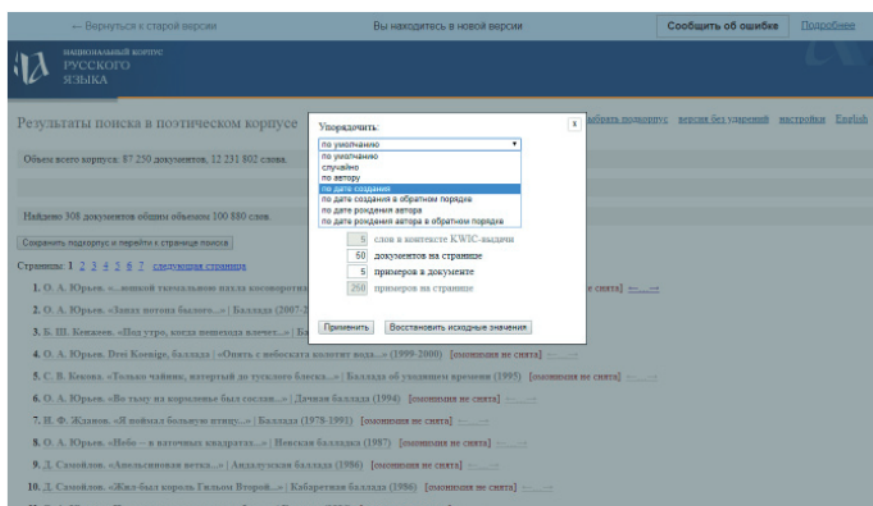


Рис. 3. Сортировка подкорпуса баллад | Fig. 3. The sorting of ballade subcorpus

Самая ранняя (1791) баллада в корпусе принадлежит Николаю Карамзину и написана четырехстопным ямбом с перекрестной рифмовкой (рис. 4).

1. Н. М. Карамзин. «Во тьме ночной врывается буря...» Раиса (1791) [оношения не сета] [оношения не сета]	Автор	Н. М. Карамзин
2. А. Ф. Мерзляков. «Сур»	Пол автора	муж
3. С. С. Бобрин. «Там, где»	Дата рождения автора	1766
4. Н. И. Дмитриев. «Как»	Название	«Во тьме ночной врывается буря...» Раиса
5. В. А. Жуковский. «Ран»	Дата создания	1791
6. В. А. Жуковский. «От»	Сфера функционирования	художественная
7. В. А. Жуковский. «Всё»	Жанр художественной литературы	баллада стихотворение
8. Н. Ф. Грамматин. «Ря»	Предлоžený	45
9. В. А. Жуковский. «От»	Словофори	405
10. В. А. Жуковский. «Ран»	Матр	Я
11. В. А. Жуковский. «Всё»	Длина строки	4
12. В. А. Жуковский. «На	Клаузула	ни регулярная
13. В. А. Жуковский. «Ден»	Метрическая формула	Яйям
14. В. А. Жуковский. «От»	Рифма	0
15. В. А. Жуковский. «Нак»	Строфия	2
16. В. А. Жуковский. «На	Число строк	88
17. В. А. Жуковский. «Зач»	Возраст аудитории	≠возраст
18. В. А. Жуковский. «Вла»	Уровень образования аудитории	≠уровень
19. В. А. Жуковский. «Всё»	Размер аудитории	большая
20. А. С. Пушкин. «После»	Тип носителя	выгта
	Источники	Н. М. Карамзин. Полное собрание стихотворений. Библиотека поэта. Большая серия. – Л.: Советский писатель. 1969.
	Подкорпус	поэтический

Рис. 4. Баллада «Раиса»: стиховедческие параметры | Fig. 4. The ballade “Raisa” and its poetic parameters

Те параметры, которые описывают формальную структуру стиха, могут быть более интересны для стиховеда и предоставляют возможность для более тонкой настройки поисковых запросов. Каждое стихотворение в поэтическом корпусе снабжается такой же разметкой, как баллада на рис. 4. Эта разметка состоит из двух уровней, или слоев: первый — это обобщающие характеристики текста, доступные пользователю корпуса; второй — построчная разметка, в результате которой в каждом тексте отмечаются позиции ударений, рифмующихся окончаний строк, а в текстах, написанных неклассическим стихом, также ритмические формы. К этой разметке пользователь имеет лишь частичный доступ: он может, например, увидеть распределение ударений в любом из текстов заданного им подкорпуса (рис. 5)².

Набор параметров стиховедческой разметки в целом общеупотребителен во всех исследованиях стиха. Это **метр**, **длина строки**, **клаузула**

² Речь, разумеется, идет не о реальных ударениях, а о тех, которые допускаются схемой размера — так называемых иктах. В корпусе они обозначаются значком грависа (`), а не акута (´), как языковые ударения.

[перейти на страницу поиска](#) [выбрать подкорпус](#) [версия без ударений](#) [настройки](#)

Н. М. Карамзин. Раиса (1791)

Раиса

Древняя баллада

Во тьме ночной ярилась буря;
Сверкал на небе грозный луч;
Гревели грома в чёрных тучах,
И сильный дождь в лесу шумел.

Нигде не видно было жизни;
Сокрылось всё под вёрный кров.
Раиса, бедная Раиса,
Скиталась в тёмноте одна.

Нося отчаяние в сердце,
Она не чувствует грозы,
И бури страшный вой не может
Её стенаний заглушить.

Она бледна, как лист увядший,
Как мёртвый цвет, уста её;
Глаза покрыты тёмным мраком,
Но сильно бьётся сердце в ней.

С её открытой белой груди,
Являемой ветвями дерёв,
Текут ручьи кипящей крови
На зёлень влажных землёв.

Рис. 5. Баллада «Раиса»: версия с ударениями | Fig. 5. The ballade "Raïsa" with accents

(то есть окончание строки), **строфика** и ее графическое представление, **способ рифмовки** и разного рода дополнительные особенности структуры, учет которых может оказаться важным.

Метр — один из основных параметров стиховедческой разметки. Все метры в русской поэзии относятся к одной из трех систем стихосложения — силлабо-тонической, тонической или силлабической. Силлабических текстов в русской поэзии меньше всего: это стих «доклассической» русской поэзии, XVII и первой четверти XVIII века, потом поэты возвращались к нему лишь изредка — в подражаниях старой поэзии или в экспериментальных переводах с других языков. Силлабо-тонических текстов, напротив, больше всего: у этой системы практически не было конкурентов начиная со второй половины XVIII века и заканчивая первой четвертью XX века. Это лучше всего изученный тип русского стиха. Наконец, под именем тонической системы объединяется ряд метров, образующих своего рода континуум — от дольников, наиболее урегулированных метров, до акцентных и свободных стихов, не подчиняющихся никаким заранее заданным схемам распределения ударения. Этот тип стиха начинает использоваться в русском модернизме и затем широко распространяется. Несколько в стороне от этих трех систем стихосложения стоит так назы-

ваемый гетерометрический стих, в котором размеры, принадлежащие к разным системам, могут смешиваться друг с другом.

От метра зависит распределение ударений в стихотворении, поэтому этот параметр можно назвать ключевым для стиховедческой разметки. Тип разметки зависит от того, к какой системе стихосложения относится текст. В силлабических текстах в общем случае однозначно определяется только позиция конечного ударения строки, так называемой ударной константы. В силлабо-тонических с высокой точностью можно предсказать почти каждое ударение. Тонические тексты с этой точки зрения устроены неравномерно: в дольнике ударение можно определить почти с такой точностью, как в классической силлаботонике, с меньшей уверенностью это можно сделать в тактовике и с еще меньшей — в акцентном стихе.

Чтобы показать, как это может работать, возьмем относительно редкую в русской поэзии форму творительного падежа *толпами*, в которой потенциально возможно два ударения — на первом и на втором слоге. Всего эта форма употребляется 172 раза в 160 текстах, то есть очень редко по сравнению с общим объемом корпуса (рис. 6). Но уже по этой выборке видно, как одна акцентная модель сменяется другой. У поэтов XVIII века (Ломоносова и других) обнаруживается исключительно форма *толпа́ми*: *Толпа́ми на́глыё на вёрх взбега́ли к на́м* (Ломоносов); *Вы лёгкими́ меня́ толпа́ми о́кружите́* (Муравьев); *И к вóротáм егò толпа́ми приходя́щих* (Хемницер) и т. д. То же и у поэтов начала XIX века: *Нарòды ли́ковáть стека́лись толпа́ми...* (Жуковский); *Смех о́бщий ùм отвéт; над мра́чными́ толпа́ми* (Пушкин); *Какими́ стра́шными́ толпа́ми* (Языков). Причем в это время форма чаще возникает в позиции рифмы: действительно, *толпа́ми* можно срифмовать, например, со *струями* или с *врагами*. Поэты XVIII века избегали таких рифм, предпочитая более богатые созвучия вроде *толпа́ми ~ глаза́ми* или *толпа́ми ~ слова́ми*, так как во всех трех словах присутствует согласный *л* справа от ударения, придающий рифме глубину. Но уже в поколении Жуковского такие строгие правила рифмовки выходят из употребления.

Во второй половине века форма употребляется довольно часто, но, видимо, постепенно начинает ощущаться как поэтическая вольность, возникая либо в начале строки в ямбических размерах, где поэту могло понадобиться слово с ударением на втором слоге (*Толпа́ми о́бразòв и тёплых, ù лучи́стых* (Случевский)), либо снова в позиции рифмы, где оно с легкостью находило рифменную пару (*Когда́ же за́поётè вы, толпа́ми* (Апухтин), где *толпа́ми* рифмуется с *вами*). По всей видимости, в середине строки оно бы ощущалось как намеренный и тяжеловесный архаизм, и этим пользовались поэты рубежа XIX–XX веков, не чуравшиеся

ческий контекст, то поэт, особенно если он не чужд стилизации, может употребить и старую форму *толпа́ми*. Если Сергей Соловьев, друг юности Белого, пишет *В Лакёдемѡн толпа́ми шёл наро́д*, демонстративно употребляя архаичную форму в середине строки, то описывающий городскую суету Саша Черный использует более новую форму: *Киша́т конце́рты толпа́ми прохо́жих* (1922).

Другие параметры стиховедческой разметки также содержат достаточно богатый репертуар формальных характеристик стиха. Здесь можно выбрать разные типы строф (рис. 7), среди которых встречаются также особые строфы, получившие известность в европейском стихосложении — например, октава или сапфическая строфа. Здесь же перечислены и наиболее распространенные твердые формы, в основном восходящие к романскому (сонет, триолет, рондо) или восточному стихосложению (газелла, рубаи). Из других специфических параметров важное место занимает способ рифмовки (рис. 8), где указаны самые распространенные типы рифмовок в русской поэзии. Среди них лидирующая с большим перевесом «перекрестная» рифмовка (схема *абаб*), популярная в XVIII веке «парная» (*аа*), более изысканные «охватная» (*абба*) и «затянутая» (*абаба*, *аббаб* и другие варианты), а также все остальные, не имеющие общеупотребительных названий.

Одна из возможных здесь задач — изучение того, как распространяются и закрепляются в русском стихосложении некоторые формы иностранного происхождения. Например, можно извлечь из корпуса сведения об

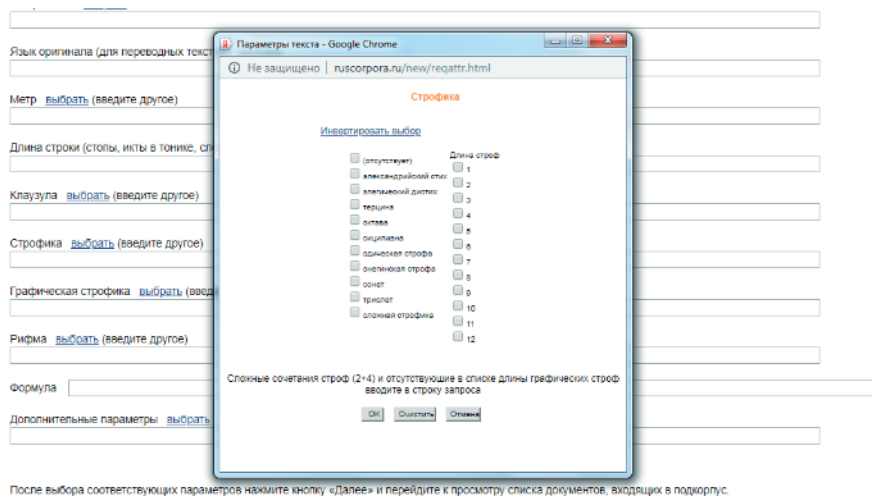


Рис. 7. Репертуар строф | Fig. 7. Repertoire of stanzas

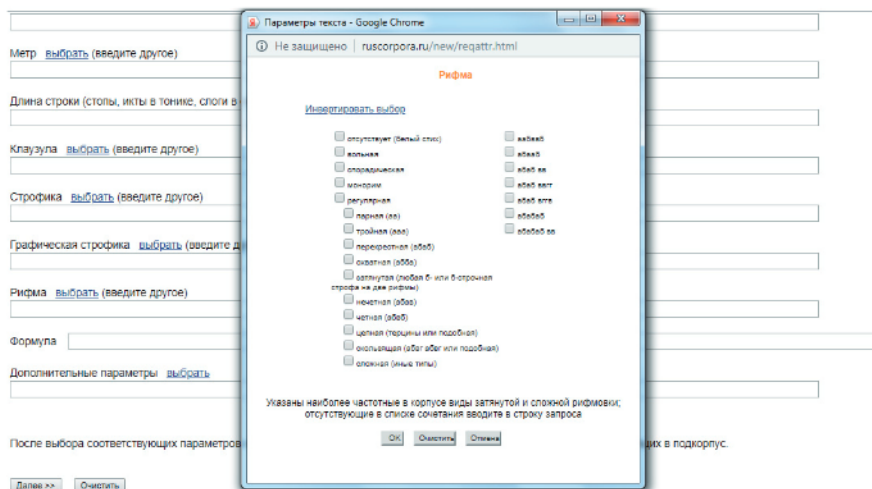


Рис. 8. Репертуар схем рифмовки | Fig. 8. Repertoire of rhyme schemes

истории русской октавы. Этой строфе посвящены отдельные работы: фундаментальная статья М. И. Шапира [Шапир 2009] и кандидатская диссертация его ученицы А. С. Белоусовой [Белоусова 2013]. Данные, представленные в корпусе, возможно, менее подробны, но они могут составить базу для подобного исследования и служить первым приближением к более глубокому труду. Всего в корпусе около 130 октав: самые ранние принадлежат Феофану Прокоповичу (К сочинителю сатир, 1730) и Ипполиту Богдановичу (Понеже, 1761). После этого следует разрыв: снова октава начинает употребляться уже в пушкинское время и далее периодически используется поэтами вплоть до второй половины XX века. Среди поздних октав — три стихотворения Игоря Холина, где бурлескные коннотации этой формы вновь оказываются уместны, стихотворения Светланы Кековой и Тимура Кибирова. Разобьем все это пространство по десятилетиям, оставив в стороне неточно датированные тексты, и получим в черновом варианте историю распространения русской октавы (рис. 9).

Уже этот график позволяет сделать предварительные выводы. После почти случайных опытов XVIII века интерес к октаве возникает во второй половине 1810-х годов — примерно в то же время, когда Байрон пишет октавой поэмы «Беппо» (1818) и «Дон Жуан» (1819), хотя, возможно, ранние октавы Дельвига (1817) и Кюхельбекера (1819) еще не связаны с этим контекстом. Однако в 1820-е байроническая поэма получает широкое распространение и закрепляется в русской поэзии после «Домика в Коломне» Пушкина, написанного в самом конце десятилетия (1830). При этом

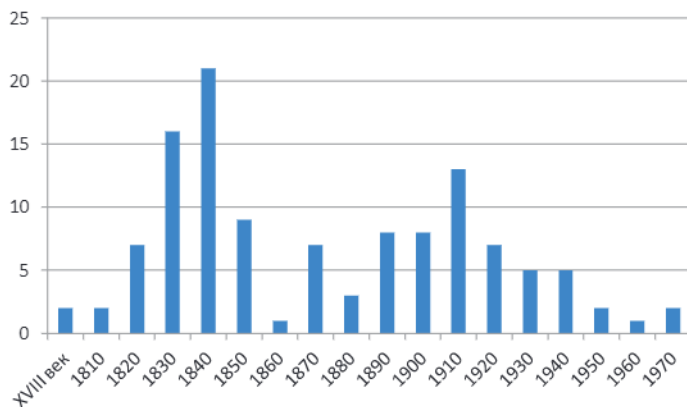


Рис. 9. Русская октава | Fig. 9. Russian ottava rima

мода на нее возникает с запозданием, достигая пика уже в 1840-е годы и сохраняя уверенные позиции в следующем десятилетии.

Затем в 1860-е годы следует катастрофический спад, вызванный, по всей видимости, усталостью от октавы и ее сатирических (или, как выражался М. И. Шапир, «ирои-комических») коннотаций, и далее на протяжении 1870–1900-х годов она употребляется изредка как знак принадлежности к «пушкинской традиции». Причем даже открытие памятников Пушкину в Москве в 1880 году и в Петербурге в 1884-м, хотя и широко освещалось в тогдашних медиа, не оказало существенного влияния на интерес к октаве. То же можно сказать и о столетнем юбилее поэта: по данным П. Н. Беркова, сложная общественная ситуация на фоне голода в некоторых российских губерниях не позволяла интеллигенции полностью солидаризироваться с торжественным поводом [Берков 1937: 402–403]. Употребление строфы заметно вырастает в 1910-е годы, когда в целом в литературных кругах ощущается необходимость в реактуализации русской поэзии пушкинского времени. Более низкие показатели отмечаются в следующие два десятилетия, 1920–1930-е годы, когда приближается столетие со дня смерти Пушкина, отмечавшееся в сталинском государстве с большой помпой и, видимо, заставившее поэтов, среди прочего, вспомнить и об октаве. Этот же интерес, впрочем не очень ярко выраженный, сохраняется и в сороковые, но снижается во второй половине века, когда для полноценного возрождения октавы не находится существенных причин вплоть до новых бурлескных опытов Тимура Кибирова.

Стиховедческая разметка может быть полезна не только стиховеду, но и историку языка. Так, в корпусе стихотворений Антиоха Кантемира,

значительного поэта первой половины XVIII века, видно заметное расхождение с современными нормами акцентуации. Кантемир использовал исключительно силлабический стих, а следовательно, согласно принципам, принятым в поэтическом корпусе, среди всех ударений в его стихах более или менее однозначно могут быть установлены только конечные ударения в строке. «Классическая» русская силлабика вслед за польской требовала, чтобы последнее слово в строке содержало ударение на втором слоге от конца. Но, если исходить из известных правил акцентуации, складывавшихся после нормализации литературного языка, при чтении стихотворений Кантемира часто может показаться, что это правило нарушается (хотя на его обязательность сам поэт указывает в небольшом трактате о русском стихе «Письмо Харитона Макентина», опубликованном в 1744 году). Однако при более детальном исследовании выясняется, что идиолект Кантемира содержит специфические архаичные и региональные черты. Так, в корпусе сатир встречается рифменная пара *дру́же* (звательный падеж) ~ *у́же* (наречие), второй член которой отражает исконное древнерусское сочетание ортотонической формы *у* и энклитики *же*. Согласно А. А. Зализняку, правило, позволяющее получить акцентуированную древнерусскую словоформу, предполагает, что ударение падает *не* на энклитику. Именно так происходит у Кантемира, который в этом отношении полностью согласуется с нормами раннедревнерусской акцентной системы [Зализняк 1985: 35, 123].

Другой пример — рифменная пара *во́йдет* ~ *подо́йдет*. Казалось бы, проще всего предполагать здесь нарушение силлабического размера и читать оба слова с ударением на последнем слоге, как в современном литературном языке. Однако можно интерпретировать эти формы иначе. Дело в том, что Кантемир, как и почти все русские силлабисты, прошел украинскую выучку, а в украинском языке наблюдается перемещение ударения в приставочных глаголах на приставку в формах настоящего времени, если гласная в корне выпадает: *во́-јдет*, *подо́-јдет* [Зализняк 1985: 363]. Но это совпадает и с требованием силлабического размера сохранять постоянное ударение на предпоследнем слоге, объясняя широкое распространение форм третьего лица единственного числа в финальной позиции, где они, казалось бы, невозможны.

Таким образом, поэтический корпус может использоваться для самых разнообразных задач на стыке литературоведения и лингвистики, а специализированная разметка формальных параметров стиха делает его незаменимым инструментом для современного стиховеда. В этой статье было показано несколько способов работы с поэтическим корпусом, хотя ими далеко не исчерпываются те возможности, которые он предоставляет.

Литература

- Берков П. Н. Из материалов пушкинского юбилея 1899 года // Пушкин: Временник Пушкинской комиссии. М.–Л.: Издательство АН СССР, 1937. Вып. 3. С. 401–414.
- Белоусова А. С. Генезис и эволюция русской октавы. Дисс. ... канд. филол. наук / Московский государственный университет им. М. В. Ломоносова. М.: 2013. 198 с.
- Воронцова В. Л. Русское литературное ударение XVIII–XX веков. М.: Наука, 1979. 328 с.
- Зализняк А. А. От праславянской акцентуации к русской. М.: Наука, 1985. 432 с.
- Корчагин К. М. Поэзия XX века в поэтическом подкорпусе Национального корпуса русского языка: проблема репрезентативности // Труды Института русского языка им. В. В. Виноградова. 2015. Вып. 6. С. 235–256.
- Шапир М. И. Семантические лейтмотивы ирои-комической октавы: (Байрон – Пушкин – Тимур Кибиров) // Шапир М. И. Статьи о Пушкине. М.: Языки славянских культур, 2009. С. 124–190.
-

References

- Berkov P. N. [From Pushkin's jubilee materials]. *Pushkin: Vremennik Pushkinskoi komissii* [Pushkin: Pushkin commission's chronicles]. Moscow – Leningrad, AN SSSR Publ., 1937, vol. 3, pp. 401–414. (In Russ.)
- Belousova A. S. *Genezis i evolyutsiya russkoi oktavy*. Diss. kand. filol. nauk [Russian ottava rima: genesis and evolution]. Moscow, Nauka Publ., 2013. 198 p.
- Korchagin K. M. [The 20th century's poetry in the poetic subcorpus of Russian National Corpus: Issue of representativeness]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova RAN*. 2015, vol. 6, pp. 235–256. (In Russ.)
- Shapir M. I. [Sematic leitmotifs in the burlesque ottava rima: (Byron, Pushkin, and Timur Kibirov)] in Shapir M. I. *Stat'i o Pushkine* [Articles about Pushkin]. Moscow, Yazyki Slavyanskikh Kultur Publ., 2009, pp. 124–190. (In Russ.)
- Vorontsova V. L. *Russkoe literaturnoe udarenie XVIII–XX vekov* [Russian literary accentuation in the 18th–19th centuries]. Moscow, Nauka Publ., 1979. 328 p.
- Zaliznyak A. A. *Ot praslavyanskoy aktsentuatsii k russkoy* [From pre-Slavonic accentuation to Russian one]. Moscow, Nauka Publ., 1985. 432 p.