

УДК 821.512.141.08
ББК 83.3(2Рос=Баш)
О 654

*Утверждено к печати
Федеральным государственным бюджетным учреждением науки
Институтом языкознания Российской академии наук*

Рецензенты:

член-корреспондент РАН *А. В. Дыбо*
кандидат филологических наук *К. М. Корчагин*

Орехов Б. В.

О654 Башкирский стих XX века. Корпусное исследование / Б. В. Орехов. –
СПб.: Алетейя, 2019. – 344 с.: ил.

ISBN 978-5-907189-29-4

Представлены результаты всестороннего количественного исследования башкирской системы версификации в XX веке. С использованием современных статистических инструментов подвергаются анализу все уровни организации стихотворного текста от фоники до лексики и грамматики, с особым вниманием к метру и ритму. Количественные данные получены на корпусе текстов 103 башкирских поэтов общим объемом в 1.77 млн словоупотреблений. Анализ предварен подробным обзором науки о тюркском стихе, начиная с 1950-х годов. Утверждается, что основную роль в башкирском стихосложении XX века играют силлабические формы фольклорного происхождения узун-күй и кыска-күй, первая из которых специфична для поволжско-кыпчакского слогосчитающего стиха. Приводится подробное сопоставление башкирского стиха с киргизским.

Книга завершается примерами поэтических текстов на башкирском языке, сгенерированных с использованием искусственных нейронных сетей.

УДК 821.512.141.08
ББК 83.3(2Рос=Баш)

ISBN 978-5-907189-29-4



© Б. В. Орехов, 2019
© Издательство «Алетейя» (СПб.), 2019

Оглавление

Список глосс	9
0. Введение	11
1. Историко-культурный контекст башкирской поэзии XX века	17
2. Проблемы и термины тюркского стиховедения.	22
2.1. Тюркское стихосложение в науке XX века	22
2.1.1. Первая половина XX века.	22
2.1.2. 1950-е	23
2.1.3. 1960-е	25
2.1.4. 1970-е	31
2.1.5. 1980-е	32
2.1.6. 1990-е	37
2.2. Стиховедческие исследования поэзии на поволжско-кыпчакских языках.	43
2.2.1. У истоков татарского стиховедения	44
2.2.2. Основополагающий труд о башкирском стихосложении.	45
2.2.3. Татарское стиховедение 1960–1970-х.	47
2.2.4. Татарский стих в историческом контексте	49
2.2.5. Обобщающий труд о башкирском стихе	51
2.2.6. Татарское стиховедение 2000-х.	53
2.2.7. Обобщающая работа об истории тюркского стиха	55
2.2.8. Статистический подход к истории башкирского стиха	56
2.2.9. Отдельные замечания о башкирском стихе	57
3. Башкирская поэзия XX века. Характеристика корпуса	61
3.1. Поэтические корпуса и башкирский поэтический корпус	61
3.2. Репрезентативность и сбалансированность корпуса	63
3.2.1. Оценка сбалансированности	63
3.2.2. Оценка репрезентативности	66
4. Квантитативный стих в истории башкирской поэзии	70
4.1. Система аруза на башкирской почве.	70
4.2. Открытые и закрытые слоги в башкирском языке	74
4.3. Башкирский аруз и его формы	76

5. Метр башкирской силлабики	87
5.1. Вопрос об изосиллабизме	87
5.2. Метрический и строфический контекст	91
5.3. Метрический репертуар башкирской поэзии	93
5.4. Динамика употребления размеров во времени	99
5.5. Метрический профиль автора	103
6. Ритм башкирской силлабики	110
6.1. Метр и ритм	110
6.2. Длина строки и длина слова	111
6.3. Ритм размеров с четным числом слогов	116
6.3.1. 10-сложник	116
6.3.2. 8-сложник	123
6.3.3. Прочие размеры четной длины	127
6.4. Ритм размеров с нечетным числом слогов	128
6.4.1. 9-сложник	128
6.4.2. Правило Болотова	132
6.4.3. Прочие размеры нечетной длины	134
6.5. Гипотеза паритивного счета	136
7. Ритм и грамматика	140
7.1. Ритм и язык	140
7.2. Морфологический анализ башкирского текста	141
7.2.1. Автоматический разбор башкирской словоформы	141
7.2.2. Снятие морфологической омонимии	146
7.3. Ритмический словарь частей речи	148
7.4. Морфологический профиль строки	152
7.5. Глагольное время в стихе	158
8. Функционирование силлабической метрики в тексте	163
8.1. Метрический контекст отрезков длиннее строки	163
8.2. Изосиллабическая поэзия	166
8.3. Узун-кюй	172
8.4. Кыска-кюй	175
8.5. Лексическая специфика метра	177
8.6. Прочие парные чередования	181
8.7. Совместная встречаемость размеров в стихотворении	183
8.8. Вольный силлабический стих	185
8.9. Русские переводы кыска-кюй	189
9. Цезура	193
9.1. Цезура в силлабическом стихе	193
9.2. Словоразделы	194
9.3. Цезура в изосиллабических текстах	198
9.4. Цезура в паралогических текстах и вольном стихе	202

10. Длина произведения	206
10.1. Длина текста и жанр.	206
10.2. Длина текста как элемент авторского стиля.	210
11. Вопрос о тонизации башкирского стиха.	214
11.1. Башкирское стихосложение в условиях русской культурной гегемонии.	214
11.2. Автоматическое распознавание силлабо-тонического размера	218
11.3. Силлабо-тоническая интерпретация силлабического стиха.	221
12. Свойства башкирской рифмы	229
12.1. Рифма и аллитерация.	229
12.2. Вокалические и консонантные рифмы	232
12.3. Грамматическая рифма	238
12.4. Рифма в тексте	239
13. Гармония гласных как прием	248
14. Контрастивные данные: современный киргизский стих	254
14.1. Контрастивное стиховедение и киргизский стих	254
14.2. Метр и ритм киргизского стиха в сопоставлении с башкирским.	257
15. Стиховой профиль и сетевой анализ истории башкирской поэзии	262
15.1. Стиховой профиль: статистические выбросы	262
15.2. Сеть башкирского стиха	265
15.3. Стих и гендер	268
16. Лексика стиха и семантический ореол метра	271
16.1. Дистрибутивно-семантическая модель башкирской поэзии	271
16.2. Тематическое моделирование и семантический ореол метра	275
16.3. Словарь как индикатор стиля	283
17. Генерация башкирского стиха с помощью искусственных нейронных сетей.	285
17.1. Искусственные нейронные сети.	285
17.2. Поэтический тест Тьюринга.	287
18. Заключение	293
Приложение 1. Авторы, включенные в башкирский поэтический корпус	295
Приложение 2. Наиболее частотные рифменные пары в башкирской поэзии	301
Приложение 3. Материалы к частотному словарю башкирской поэзии.	304
Приложение 4. Башкирские стихотворения, порожденные искусственной нейронной сетью	316
Приложение 5. Анкета для эксперимента с искусственно порожденными стихотворениями	319
Литература	325
Указатель.	342

0. Введение

В 1826 году видный филолог-классик И. Г. Я. Герман объявил, что филологии, которая занималась бы именно башкирской культурой, просто не может существовать: «Ибо именно творения духа в первую очередь придают тем предметам их ценность. Если бы предметы сами по себе имели такую ценность, то нельзя было бы не признать, почему бы, напр., башкиры, древний народ, называвшийся у греков Βέχαιρες, не имели полного права требовать для себя башкирских предметных филологов. Ведь если бы предметное знание исчерпывало собственную сущность филолога, то кто-то мог бы стать изрядным филологом и не понимая греческого и латыни»¹ (перевод С. М. Шаулова).

Даже если принять во внимание множественность трактовок предмета и назначения филологии (см. [Орехов 2010А]), несомненно, что время уже опровергло утверждение Германа (см., к примеру, сборник статей «Вопросы башкирской филологии» [Вопросы 1959]). Поэтому наша цель состоит не столько в том, чтобы показать, что башкирская предметная

¹ «Denn die Geisteswerke sind es ja erst, die jenen Sachen ihren Werth geben. Wären es die Sachen an und für sich selbst, die einen solchen Werth hätten, so wäre nicht abzusehen, warum z. B. die Baschkiren, ein uraltes Volk, Βέχαιρες bey den Griechen genannt, nicht eben das Recht hätten, für sich Baschkirische Sachphilologen zu verlangen. Ja wenn die Sachkenntniß das eigentliche Wesen des Philologen ausmachte, so könnte jemand ein ziemlicher Philolog werden auch ohne Griechisch und Lateinisch zu verstehen». (Hermann 1826: 9). Подробнее о контексте этого высказывания: [Тротман-Валлер 2009] По всей видимости, после заграничного похода русской армии, в котором участвовали башкирские полки (проходя через Веймар, кто-то из башкир подарил Гёте свой боевой лук, о чём Гёте потом увлечённо рассказывал Эккерману), в Европе сформировалась своя «башкирская мифология», которая и привела к этой знаменательной ошибке Германа: он отождествил башкир и бехиров, народ, согласно «Аргонавтике» Аполлония Родосского (Apollon. 2,394), проживавший на черноморском побережье.

филология возможна, сколько в том, чтобы продемонстрировать, что она может быть построена на современных основаниях, к которым, в частности, следует отнести исследовательскую стратегию *distant reading*, оформленную в работах Ф. Моретти [Moretti 2013] и его последователей. *Distant reading* ('дистанцированное, отвлеченное чтение', переводчики русского издания остановились на варианте «дальнее чтение» [Моретти 2016]) противопоставлено *close reading* (в русской традиции — 'медленное чтение'), то есть традиционному для литературоведения (технологическому [Венедиктова 2007] или позитивно-критическому [Назиров 2015: 94]) способу взаимодействия читателя и текста, при котором исследователь, опираясь на свой профессиональный опыт, на накопленные аналитические наблюдения предшественников и, самое главное, на конкретное словесное представление, входит в герменевтический круг, строит интерпретации и описывает структуры. «Отвлеченное чтение» предполагает, что литературовед не сноситя с текстом напрямую, а пытается уловить значимые для литературы тенденции опосредованно, через модели, в основу которых положена извлеченная из исходного текста и систематизированная информация. Классический пример — работа Моретти «Корпорация стиля: размышления о 7 тысячах заглавий (британские романы 1740–1850)»², в которой автор с помощью количественных мер описывает эволюцию английского романа, опираясь при этом на изменение длины названий во времени и распределение в них определенных слов.

Такой подход не означает неуважения к писательскому труду или литературе вообще. Во-первых, это просто честное признание ограниченности человеческих возможностей. Исследователь способен прочесть за отведенное ему время конечное число художественных произведений, а посвятить полному объему сложно организованных текстов достойное количество времени и сил для выявления и осмысления всех нюансов — задача нереализуемая. Меж тем литературная традиция в целом как система, как комплексный объект, — предмет, вызывающий к изучению в не меньшей степени, чем отдельный текст. И отвлеченное чтение дает возможность обозреть традицию (или хотя бы масштабный набор текстов) целиком. Во-вторых, *distant reading* — это результат традиционного для гуманитария поиска новых методов исследования его материала, при этом привлекаемые методы вполне привычны для естественных наук, в которых ценится репрезентативность объема исходных данных: чем больше данных, тем точнее выявленные в них закономерности и надежнее выводы. И если

² Русский перевод работы «Style Inc.: Reflections on 7,000 titles (British Novels 1740–1850)».

астрофизики в своих исследованиях используют наблюдения за тысячами и даже сотнями тысяч звезд, это вовсе не свидетельствует, что им неинтересны отдельные звезды как предмет. Даже наоборот, как считается: «Чтобы делать что-то в области отвлеченного чтения, нужно хорошо освоить медленное чтение» [Spivak 2006: 102].

Основное отличие исповедуемого в этой книге подхода от отвлеченного чтения в том, что на дальнейших страницах производится поиск неочевидных закономерностей текста (и шире — всей текстовой традиции). Многие из системных отношений, описанных методами Ф. Моретти, представляют собой знание, которое можно было бы получить с помощью традиционных филологических подходов (например, интроспекции³). Нас как исследователей скорее интригует «встреча» значимой для литературоведа информации и привычных для точных и естественных наук методов, поэтому специально уточним: результаты исследований, изложенные в этой работе, не могли быть получены иначе, нежели с помощью подсчетов, тренировки компьютерных моделей и визуализации.

В принципе, упреки в очевидности выводов уже стали привычной проблемой для новых областей гуманитарных наук, пытающихся использовать статистику для конструирования гипотез в не совсем привычной для этого сфере. Так, кроме отвлеченного чтения, можно упомянуть «культуромику» [Michele et al. 2010], дисциплину, представители которой делают выводы о динамике общественных явлений по истории слов, отраженных в базе данных отсканированных в рамках проекта Google Books книг. Однако следует учесть, что очевидные выводы для молодой науки скорее не недостаток, а необходимый этап зарабатывания доверия. Так, гораздо труднее было бы положиться на не согласующиеся с нашими интуитивными или с признанными традиционной наукой представлениями выводы, особенно если сделаны они были в рамках экспериментального применения непривычных методов. В то же время поиск «неочевидных закономерностей (в чем, собственно, и состоит научное познание)» остается основной задачей филолога [Новиков 2005], и эту задачу в меру сил автора пытается решать эта книга. Таким образом, один из основных ее сюжетов — это проблема применения точных методов в традиционно гуманитарной сфере изучения поэзии.

В то же время автор рассчитывает, что и проблематика книги выходит за пределы локальной области башкирской (или даже тюркской)

³ Так, многое из того, о чем говорят нам графы в работе «Теория сетей, анализ сюжета», в принципе, постижимо и без построения сети персонажей.

поэтической традиции⁴. Главная сфера этой работы — стиховедение, прежде всего в той своей части, которая называется «русским методом». Речь идет о лингвостатистическом подходе к описанию стихотворных закономерностей: «Метр реализуется в языке и предстает в разнообразии многих ритмических вариаций; исследователь их описывает и подсчитывает, предъявляя результаты в объективных статистических показателях, и только на этой основе строит в дальнейшем теоретическое описание механизмов ритмообразования» [Бейли 2004: 12]. В последние годы развитие «русского метода» органично соединилось с успехами корпусной лингвистики и привело к появлению его особого извода: корпусного анализа метрики [см. Корпусный анализ 2013, 2014] и шире — корпусной поэтики, то есть исследовательского инструментария получения лингвостатистических данных с помощью поэтического корпуса. Поэтический корпус — это репрезентативная электронная коллекция поэтических текстов, снабжённых поисковой системой, способной производить поиск с учётом широкого набора лингвистически и стиховедчески релевантных параметров.

В области лингвистики появление корпусов уже отразилось и на постановке исследовательских задач, и на значимости получаемых результатов [Плунгян 2008А]. В ближайшее время следует ожидать таких же методологических сдвигов и в стиховедении, и тогда книга очерка башкирского стиха впишется в ряд соположенных исследований стиха на корпусной основе.

Башкирский поэтический корпус создан коллективом исследователей под руководством автора этой книги в 2012–2013 годах и доступен в Интернете по адресу <http://web-corpora.net/bashcorpus/>. На материале этого корпуса и базируется настоящее исследование. Мы надеемся, что эта книга способна показать незаслуженно забытые и просто не освоенные пути в стиховедении, которыми можно будет пойти при изучении других поэтических традиций, в том числе и русской.

Кроме того, обращение именно к башкирской поэзии имеет свой особенный смысл в контексте стиховедческих исследований. К настоящему

⁴ В этом смысле мы хотели бы солидаризироваться с А. Б. Лордом с поправкой на этническую идентификацию: «Несмотря на то что лишь две ветви индоевропейских народов, греки и славяне (точнее, носители сербскохорватского и болгарского языков), рассматриваются здесь сколько-нибудь подробно, я надеюсь, что книга не получилась узко специальной. Узость — будь то узость этническая, географическая, религиозная, социальная или даже научная — всегда недостаток, тем более в наш космический век» [Лорд 1994: 10].

моменту сложилась ситуация, в которой имеющийся стиховедческий аппарат оказывается адаптирован прежде всего к материалу русской силлабо-тоники, для нее описаны репертуар и динамика исторических изменений метров, ритмов, рифм, твердых форм. В то же время силлабическая система стихосложения имеет свою специфику, и механический перенос на силлабику исследовательских практик, изначально разработанных для иноструктурного материала, приведет к тому, что утрачено будет само понимание этой специфики. Таким образом, башкирские данные, представленные в этой книге, должны служить своего рода опытным образцом для отработки методов исследования специально силлабического стиха.

Несмотря на то, что стиховедением интересовался А. Н. Колмогоров, внесший едва ли не решающий вклад в статистическую науку XX века, собственно математический аппарат стиховедения в основном остался на уровне 1960-х годов: специалисты ограничиваются процентами и линейными диаграммами, в редких случаях используется критерий χ -квадрат. Нам хотелось бы (пусть и не на «престижном» русском, а на имеющем ограниченную аудиторию башкирском материале) создать прецедент применения в стиховедении оценки регрессионных моделей, сетевого анализа, тематического моделирования, иерархической кластеризации и других давно вошедших в практику работы с данными методов.

Встреча количественных методов анализа данных, корпусной лингвистики и стиховедения неизбежно приводит к некоторому усложнению текста для гуманитарно ориентированного читателя. Вместо традиционно расплывчатых рассуждений о поэзии каждое утверждение здесь будет опираться на количественные методы, то есть на цифры, полученные из иногда простых и понятных, а иногда сложных статистических операций. Мы стремились по возможности объяснять используемые термины и строить визуализацию материала для облегчения восприятия, однако сделать чтение комфортным для того, кто привык к иной стилистике разговора о поэзии, удалось далеко не везде.

Главы с 1-й по 3-ю можно считать вводными, они подготавливают детальный разговор о башкирской поэзии. 4-я и 5-я главы посвящены метрике, главы с 6-й по 10-ю — ритму в силлабике, рассмотренному на разных уровнях организации произведения — от грамматики до целого текста. В главах с 11-й по 13-ю речь идет о понятой максимально широко фонике (к ней мы в данном случае относим и рифму). В 14-й главе предпринимается попытка широкого обобщения полученных данных о метрике в общетюркском контексте с помощью контрастивного киргизского материала, в 15-й главе, тоже обобщающей, полученные ранее данные становятся

основой для установления закономерностей на уровне литературного процесса. 16-я глава представляет некоторые результаты в направлении лингвистики и семантики стиха на башкирском материале. Наконец, завершающая книгу 17-я глава демонстрирует не столько пример анализа, сколько синтеза, использованного, однако, для получения дополнительных сведений о социокультурном статусе башкирской поэзии в наши дни.

Было бы несправедливо не поблагодарить за помощь в работе над этой книгой Азамата Абдрахмановича Галлямова, который сделал очень много для того, чтобы она появилась на свет: именно Азамату Абдрахмановичу мы обязаны оцифровкой большей части башкирских поэтических текстов и квалифицированными консультациями при создании анализатора башкирской морфологии. Также хотелось бы выразить признательность коллегам, которые консультировали нас по многим ключевым для книги вопросам и сделали замечания, позволившие существенно улучшить как содержание, так и способ его изложения. Это М. А. Алонцев, А. М. Буранчин, Г. Б. Вильданова, В. А. Гречачин, Е. Е. Земскова, М. А. Иосифян, С. А. Искандарова, А. Ф. Калинина, С. Л. Козлов, О. А. Маркелова, З. Н. Нигматьянова, Н. М. Перлина, В. А. Плунгян, И. Р. Саидбатталов, Э. А. Салихова, В. В. Файер, И. А. Шакиров, С. М. Шаулов, Е. В. Шаульский, Г. Юсупова. Наша специальная благодарность — О. Н. Ляшевской за неизменно критические замечания как стимул для профессионального роста (в каком качестве критика, надо сказать, получается мало у кого).

Трудно выразить нашу признательность уважаемым рецензентам К. М. Корчагину и А. В. Дыбо, которые взяли на себя труд прочесть книгу в рукописи и указали на недочеты, которые мы постарались исправить до того, как книга была отпечатана.

Большую техническую помощь при подготовке макета книги нам оказали Д. Ф. Муслимов, В. П. Фесенко и С. С. Белоусов.

Кроме специально оговоренных случаев, переводы башкирских стихотворных цитат принадлежат З. Н. Нигматьяновой.

В то же время все ошибки и недочеты этого труда целиком остаются на нашей совести.