

## ОТЗЫВ

**официального оппонента о диссертации Зиянгирова Руслана Олеговича «Конкорданс произведений С. Т. Аксакова: актуальные контексты изучения» (Екатеринбург — 2016), представленной на соискание ученой степени кандидата филологических наук по специальности 10.01.01 — русская литература в диссертационный совет Д 212.285.15 при ФГАОУ ВПО «Уральский федеральный университет имени первого Президента России Б. Н. Ельцина»**

Диссертация Руслана Олеговича Зиянгирова построена на соединении двух самостоятельных областей — традиционного литературоведения и компьютерных подходов к анализу текстовой информации. На языке современной науки это сопряжение получило наименование *digital humanities*, не нашедшее пока своего русского аналога. То, что представление диссертационной работы к защите приходится на зазор между появлением англоязычного термина и оформлением его отечественного соответствия, однозначно свидетельствует в пользу *актуальности* рецензируемого исследования.

Действительно, именно современная ситуация успехов в компьютерной сфере спровоцировала применение опробованных на других данных методов к культурно значимым текстам. К этому нужно прибавить и лавинообразный рост объемов оцифрованных текстов, то есть собственно предмета изучения, который не всегда возможно охватить аналитическим взглядом одного исследователя. Человеческие возможности конечны, и современные филологи начинают осознавать, что с помощью компьютера они способны свои возможности расширить. Эта идея дала толчок еще одной непосредственно связанной с *digital humanities* областью современной гуманитарной науки, *distant reading* («дальнее чтение» или «отвлеченное чтение»), то есть комплексу подходов, которые позволяли бы делать выводы о тенденциях, организующих такое количество литературных текстов, которые один человек просто не способен прочесть и проанализировать традиционными филологическими методами. В диссертации оба эти термина не упомянуты, но несомненно то, что и термины, и стоящие за ними понятия имеют до неё прямое касательство.

Не может не импонировать, что все эти современные и отчасти новаторские подходы звучат в одном контексте с наследием С. Т. Аксакова. Во-первых, потому что до сих пор Аксаков не был в фокусе внимания ни тех, кто развивает компьютерные подходы к литературе, ни лексикографов, и проделанная соискателем работа представляет несомненную *научную новизну*. Во-вторых, компьютерная сфера в гуманитарной науке пока еще имеет статус новости, а Аксаков благодаря своим взглядам и поэтике, скорее, ассоциируется с традиционностью и консерватизмом (сам диссертант пишет о «патриархальности»). Такой

неявный контраст материала и метода создает дополнительную интригу, сопровождающую представленную на рассмотрение диссертацию.

Еще один контекст работы составляет заслуженная и уважаемая область филологического труда под общим названием авторской лексикографии, занимающаяся составлением словарей на основе текстов отдельных писателей. В этой сфере и именно сейчас тоже делается очень много: защищена докторская диссертация Л. Л. Шестаковой (2012), в Институте русского языка им. В. В. Виноградова действует семинар «Авторская лексикография и история слов», публикуются тематические сборники. На таком устойчивом *теоретическом базисе* и основывается диссертация Р. О. Зиянгирова.

При этом важно отметить, что увлечение инновациями не приводит к ослеплению исследовательской оптики, автор работы остается способен критически отнестись к новым методам, оговариваясь: «создание объективной фактологической базы данных для литературоведческого анализа не может являться самоцелью, нуждаясь в дальнейшей исследовательской интерпретации» (с. 11). Такой взвешенный подход обеспечивает *теоретически значимые результаты* диссертации, в которой соискатель не останавливается на освоении модных методов, а двигается дальше, соединяя эти методы с проверенными временем аналитическими и герменевтическими практиками, получает новое знание о предмете. Резюмируя, можно сказать, что Р. О. Зиянгиров счастливо избежал сопутствующих искушений и не забыл о главной цели своей работы, изучении словарной составляющей художественной системы С. Т. Аксакова.

Работа состоит из трех глав, прозрачно реализующих общий композиционный замысел. Первая глава «Конкорданс как информационно-исследовательский ресурс: теория и практика разработки» являет собой жанрово необходимый обзорный раздел, наименее важный с точки зрения новизны и собственно вклада в достижение нового знания. И к этому, с позиций ценностной шкалы исследователя наименее значимому разделу, как раз имеется наибольшее число вопросов. Внимательное чтение показывает, что большинство решений, о которых упоминает диссертант, были актуальны во второй половине 2000-х годов, а на текущий момент должны быть признаны устаревшими, они практически не востребованы ни исследователями, ни компьютерной индустрией. Так, обзоревая реализацию технологии автоматического морфологического анализа соискатель останавливается на библиотеке «RCO Morphology Professional SD SDK» и системе «АОТ» (с. 14—15). Если в свое время «АОТ» действительно дала мощный импульс развитию компьютерной лингвистики, то сейчас лидерами в своей области всё же являются программные модули Mystem, PyMorphu, TreeTagger, а упоминание в этом контексте библиотеки от Microsoft, поставляемой в виде dll следует считать не более чем недоразумением.

Кроме того, не совсем понятно, зачем вообще эти инструменты становятся предметом перечисления, если ни один из них не использовался в создании стоящего в центре диссертации конкорданса.

Не всегда можно согласиться с теми оценками, которые дает диссертант обозреваемой области: «если проблема морфологического анализа уже частично решена <...>, то задачи синтаксического и особенно семантического анализа решены далеко не в полной мере» (с. 13). В действительности можно говорить о том, что на сегодняшний день задачи семантического и синтаксического анализа решены в той же степени, что и задачи морфологического анализа. Возможно, не лучше, но и не хуже. Чтобы убедиться в этом, достаточно взглянуть на результаты работы системы АBBYU Compeno, ресурсы вроде «RusVectōrēs» и соревнование синтаксических анализаторов в рамках конференции по компьютерной лингвистике «Диалог» 2012 года.

Реконструируя логику диссертанта, можно предположить, что он, по всей видимости, опирается на перечни авторитетных в своё время разработок, которые в настоящий момент морально устарели и не могут быть ориентиром в создании инновационных инструментов филологического анализа. Проблема и в том, что соискатель учитывает исключительно отечественные достижения, игнорируя то, что сделано в области digital humanities в мировой практике. Только этим можно объяснить то, что в обзоре упомянуты инструменты, касающиеся низкого уровня организации текста: морфологический, синтаксический анализ, но отсутствуют указания на сервисы нового поколения, построенные на нейронных сетях, дистрибутивной семантике и векторных моделях, реализующих разного рода стратегии визуализации содержания (вроде тепловой карты текста, графа персонажей и т. д.).

Вот наиболее показательный пример из первой главы: «мы стали свидетелями возникновения принципиально новых, самостоятельных научных практик», — совершенно справедливо провозглашает диссертант (с. 19). Но при этом приводит почему-то в крайне малой степени новаторский пример: разработку электронных библиотек. А ведь наряду с этим, безусловно, особенным по сравнению с традиционной библиотечной деятельностью, видом филологической работы, возникли гораздо более яркие аналитические практики, касающиеся именно анализа и интерпретации текста на основе автоматической обработки больших текстовых массивов: в частности, исследования, объединяемые под общей маркой distant reading, проводимые в передовой в своей области Stanford Literary Lab. Здоровый консерватизм именно в филологической работе порой совершенно необходим, но, кажется, там, где одним из движителей должны быть современные тенденции и стремление соответствовать духу времени, такой консерватизм должен нести известные ограничения.

Не вполне справедливыми представляются и оценки, которыми сопровождается обзор

«Фундаментальной электронной библиотеки "Русская литература и фольклор"» (ФЭБ). ФЭБ — также морально устаревший ресурс. Необходимый для филолога, крайне полезный, но лишённый хороших (или даже просто достаточных для исследователя) поисковых механизмов, которые низводят эту библиотеку до ранга «холодного» (в терминах computer science) хранилища информации. Там можно искать слова, о чём и пишет Р. О. Зиянгиров. Но, скажем, нельзя учитывать при поиске богатую метаинформацию, входящую в библиографическое описание документа. Проще говоря, слова и словосочетания найти можно, а вот слова, опубликованные именно в 1857 году — нет. Если в середине 2000-х это было простительно, то для современного ресурса такой недостаток представляется существенным.

С чем нельзя решительно согласиться, так это с тем, что «разработчики рассматривают ФЭБ не только как действующий продукт, но и как полигон для дальнейших экспериментов и исследований, направленных на решение многочисленных и разнообразных задач, связанных с проблемами внедрения современных информационных технологий в сферу гуманитарных наук» (с. 22). На самом деле, разработчики ФЭБ крайне консервативны, и не приемлют никаких новых решений и надстроек над имеющимся сервисом. Как показывает опыт, все попытки предложить им новые подходы и решения обречены на неудачу.

В том же ключе прочитывается и такой тезис диссертации: «Особое место в ФЭБ занимает <...> Каталог сетевых ресурсов» (с. 21). Соискатель воздает ему хвалу, однако следует признать, что в следствие бурного развития сети и усиления влияния поисковых систем на климат Интернета, каталоги любого рода превратились в вымирающий жанр: когда-то (в 1990-е и начале 2000-х) бывшие популярными ресурсы Яндекс.Каталог и Dmoz, сейчас не востребованы ни массовым, ни профессиональным пользователем.

На этом фоне подкупает, что диссертант вполне точно воспроизводит в одном из частных примеров на с. 27 соотношение частотностей слов в языке, ориентируясь на современный источник, «Частотный словарь современного русского языка» О. Н. Ляшевской и С. А. Шарова, это показывает зрелый уровень Р. О. Зиянгирова как исследователя. Правда, присутствующая на стр. 52 диссертации ссылка на тот же словарь заставляет искать в нём слово «мать» на стр. 39, в то время как искомого слова там нет и быть не может: на стр. 39 размещена часть алфавитного списка лемм от «владелица» до «внесудебный», куда лемма «мать», разумеется не входит. «Мать» в упомянутом словаре находится на совсем другой странице, а именно на стр. 160. Ради справедливости отметим, что показатель IPM для слова «мать» в диссертации приведён корректно, правда, соискатель не учитывает, что IPM рассчитывался по общезыковому корпусу текстов, в то время как значимым было бы сравнение показателей частотности у Аксакова и его современников, ведь за прошедшие два

века язык серьезно менялся, особенно в такой подвижной сфере, как распределение частотностей слов.

В первой главе соискатель приводит и сконструированный им перечень основных требований, «предъявляемых к информационным системам, рассчитанным на филологическую работу с текстами» (с. 17). В целом он не вызывает возражений, однако смущает неполная чёткость следующей формулировки: «система должна подчиняться ясно сформулированной концепции, в которой отражаются цели и задачи системы» (там же). В целом соглашаясь с диссертантом в том, что ресурс нужно строить на понятных принципах, важно при таком подходе не забыть и важное свойство электронных систем: эмерджентность, то есть способность порождать такие связи, которые не планировались при их создании, и благодаря которым могут возникнуть возможности постановки новых задач, которые не только раньше не ставились, но и не приходили никому в голову.

Переходя к собственно теме конкорданса, Р. О. Зиянгиров почему-то представляет его как электронный инструмент, хотя конкордансы возникли задолго до появления компьютеров и в своем бумажном виде имеют длинную историю. Далее это будет упомянуто в диссертации, но первичная импликация прямой связи конкордансов с компьютерными средствами вызывает недоумение. Может быть, это связано с неудачно выбранной формулировкой, а, может быть, с недостаточным знакомством со специальной литературой по авторской лексикографии. Так, в списке литературы отсутствует фундаментальная монография Л. Л. Шестаковой «Русская авторская лексикография: Теория, история, современность» (2011), хотя частная работа того же автора, посвященная именно аксаковскому словарю, в библиографическом перечне упомянута.

В главе имеется и ряд других, обращающих на себя внимание недочётов. Например, вопреки тексту диссертации далеко не все затруднения при построении конкорданса, перечисленные на стр. 27, связаны с законом Ципфа, а только последняя. Степенное распределение никак не влияет, скажем, на то, что «частота слова в одном тексте может повлиять на его позицию в словаре», это уже другая проблема, являющаяся отражением закономерности функционирования слов в разных документах корпуса.

Далее, почему-то постструктурализм отнесён к числу точных методов науки, что по факту совершенно не так. Не вполне удачно и то, что упоминаемый несколько раз В. М. Живов сопровождается предикатами в настоящем времени, что создает ненужное впечатление будто бы исследователь ещё активен в то время как Виктор Маркович ушёл из жизни в апреле 2013 года.

Однако, повторимся, сама глава, хотя и представляет собой легкую мишень для критики, занимает в общей конструкции диссертации второстепенное место, и должна

рассматриваться как вспомогательная по отношению к собственно исследовательским главам второй и третьей.

Так, вторая глава «Контекст охоты в произведениях С. Т. Аксакова» демонстрирует детальный филологический анализ словоупотребления соответствующей лексемы и дериватов этого слова, основанный, по сути, на корпусной методике. Это зрелое исследование, сочетающее в лучшем смысле филологическую внимательность, щепетильность и способность работать с данными различной структуры и организации. После обзора критических работ об Аксакове автор переходит к собственно аналитическому этапу, развертывая перед читателем картину, на которой оказываются качественно прорисованы основные особенности аксаковского текста. Единственная проблема, о которой здесь можно упомянуть, это отсутствие в добротном в целом параграфе «Цифровой анализ и литературоведческая интерпретация охотничьих текстов С. Т. Аксакова» собственно цифровой составляющей. Автор действительно грамотно обращается с частотностями слов, опирается на них в своих рассуждениях и выводах, со знанием дела говорит о структуре текста, но того, что можно было бы ожидать от цифрового исследования, здесь практически нет (скажем, математическая статистика или визуализация данных). Наоборот, возможности, которые предоставляет цифровая среда, часто игнорируются. Например, на стр. 67—69 приводится перечень дериватов слова «охота», которые служат основой для дальнейших рассуждений автора. Однако от современной работы мы вполне могли бы ожидать не только списка словоформ, но и списка лемм, который тоже был бы по-своему информативен. Кроме того, перечень оказывается неполным, так как соискатель, очевидно, учитывал только формы, стоящие рядом в алфавитном списке. Меж тем, если с помощью компьютера поискать в составленном Р. О. Зиянгирове конкордансе, находится и неучтённая им словоформа «поохотничьи» (221:1; 308:1).

При этом выводы главы убедительны, сама исследовательская стратегия выглядит фундировано и позволяет достичь интересных результатов.

Глава 3 «Гендерный контекст «Семейной хроники» С. Т. Аксакова», несмотря на некоторую провокативность названия, как и вся остальная работа, не просто дань научной моде. Соискателю и здесь удалось обойти опасную область постмодернистских переинтерпретаций, замешанных на понятии гендера. Р. О. Зиянгиров строго следует в хорошем смысле традиционной филологической устновке реконструкции художественного мира Аксакова, подкрепляя каждый выдвигаемый тезис примерами из текста. Скажем, убедительна интерпретация «Семейной хроники» как пространства, в котором напряжение создается между патриархальной средой и вторжением в неё новой хозяйки.

Что вызывает решительное несогласие, это «Приложение». Ещё на стр. 36 диссертации

говорится о том, что программная реализация конкорданса «выполнена в рамках <идеологии? — Б. О.> открытого исходного кода», затем, уже в приложении упоминается Ричард Столлман и разработанные им принципы свободного и открытого программного обеспечения (с. 152), однако тут же, семью страницами ниже, выясняется, что программный код автором не открыт. Он не только не выложен в свободном доступе на одном из распространённых в программистской среде хостингов кода вроде [github.com](https://github.com) или [bitbucket.org](https://bitbucket.org), как это принято при создании свободного программного обеспечения, но и «изменен до нерабочего состояния в целях защиты прав интеллектуальной собственности», что радикально противоречит идеологии свободного и открытого программного обеспечения, таким образом, положение со стр. 36 диссертации следует признать не соответствующим действительности. Кроме того, хочется успокоить автора, и ответственно заявить, что сама ситуация кражи его кода, написанного на практически вышедшем из употребления языке программирования, решающего не самую сложную задачу, которая по силам любому первокурснику даже не профильной специальности, представляется практически невероятной.

Скажем несколько слов и о самом конкордансе, работа с которым стала ядром диссертационного исследования. В целом он выполняет возложенные на него функции. Большой ложкой дёгтя является наличие назойливой рекламы на сайте ЭНИ «С. Т. Аксаков», которая затрудняет использование ресурса (в частности, модульное окно, которое появляется при заходе на сайт, невозможно закрыть в течение первых 10 секунд его показа). При этом в диссертации (с. 9) говорится о том, что конкорданс может быть использован в корпусной лингвистике. Это положение не очевидно и требует подтверждения или хотя бы иллюстрации. Сама попытка распространить эти слова была бы полезна ещё и потому, что значительно его прояснила: корпус может представить любой текст в виде конкорданса, и дополнительный конкорданс, по сути, корпусу не нужен, так как не обладает теми возможностями, которые есть у корпуса (поиск по лемме, поиск слов на определённом расстоянии друг от друга, поиск слов в сочетании с определёнными грамматическими категориями).

Как видно, основные возражения рецензента вызвали технические или принудительно обусловленные жанром части работы. С содержательной исследовательской стороны диссертация представляет собой добротное научное произведение, в котором филологически корректно ставятся и разумно решаются осмысленные задачи, выбранные пути этих решений не вызывают возражений, а строгость и основательность выводов не могут не вызывать эмпатии профессионального читателя.

Подводя итог еще раз подчеркнем, что кандидатская диссертация Р. О. Зиянгирова «Конкорданс произведений С. Т. Аксакова: актуальные контексты изучения» - законченное самостоятельное исследование, содержащее оригинальный подход к исследуемой области. Диссертация представляет собой научно-квалификационную работу, которая соответствует критериям, установленным п. 9 Положения о присуждении ученых степеней, утвержденным Постановлением Правительства Российской Федерации о порядке присуждения ученых степеней от 24 сентября 2013 г. № 842, а ее автор заслуживает присуждения ученой степени кандидата филологических наук по специальности 10.01.01 — русская литература.

Кандидат филологических наук,  
доцент Школы лингвистики  
ФГАОУ ВПО Национальный исследовательский университет  
«Высшая школа экономики»  
«8» сентября 2016 г.

Б. В. Орехов

Орехов Борис Валерьевич  
Адрес: 101000, г. Москва, ул. Мясницкая, д. 20  
ФГАОУ ВПО Национальный исследовательский университет  
«Высшая школа экономики»  
Тел.: +7 (495) 772-95-90  
e-mail: borekhov@hse.ru