

ОТЗЫВ

официального оппонента Орехова Бориса Валерьевича,
кандидата филологических наук, доцента факультета филологии
ФГАОУ ВПО Национальный исследовательский университет

«Высшая школа экономики»

на диссертацию Ковригиной Любови Юрьевны

«Негауссовое моделирование лексико-статистической структуры вариативности
текста (на примере «Сказания о Мамаевом побоище»)», представленную на
соискание ученой степени кандидата филологических наук по специальности
10.02.21 — Прикладная и математическая лингвистика

Диссертация Л. Ю. Ковригиной представляет собой фундаментальное научное исследование, выполненное на интересном материале, который в сочетании с предложенными для исследования методами становится особенно любопытной научно-практической площадкой для формирования будущих исследовательских практик, дополняющих традиционные филологические подходы.

Представленная на соискание ученой степени кандидата филологических наук работа обладает несомненной *актуальностью* в свете все большего проникновения точных методов в традиционные гуманитарные дисциплины, даже такие, которые до самого последнего времени считались исключительно традиционными в своем обращении с материалом. Среди таких дисциплин следует назвать и текстологию, проблемы которой наравне с собственно математическим статистическим инструментарием входят в сферу интересов диссертанта.

Работа посвящена крайне сложному объекту исследования: вариативному тексту, отношения между вариантами которого могут быть описаны строго математически, но в реальной научной практике такие попытки не предпринимались. В этом состоит очевидная *научная новизна* представленной работы. Осмысленные в диссертации трудности, возникающие на пути применения математического моделирования в этой области филологии, представляют основной *теоретический вклад* соискателя в соответствующих

дисциплинах.

Положения, выносимые диссертантом на защиту, сформулированы убедительно и находят подтверждение в тексте работы. Они в самом деле отражают основной сюжет исследования, формулируя наиболее значимые результаты моделирования лексических распределений вариативного текста.

Отдельного упоминания заслуживает весьма внушительный *теоретический базис* исследования, включающий не только локальные методические достижения отдельных научных сфер, с которыми тематически соприкасается диссертация, но и генеральный общеметодологический подход, имеющий в виду рассмотрение поликомпонентных объектов в науке в целом. Это дает работе значительные преимущества, сообщает ей большую фундированность и позволяет получить *теоретически значимые результаты*, которые могут быть использованы в гораздо более широких сферах, чем те, которые напрямую имеют отношение к тематике исследования.

Диссертация состоит из четырех глав.

Первая глава ««Сказание о Мамаевом побоище» как вариативный текст и объект исследования» достаточно скрупулезно разбирает существующую ситуацию с вариантами «Сказания», а также теоретически обосновывает понятие вариативного текста. В этой части работы автор получает возможность продемонстрировать глубокие познания в вопросах текстологии древних памятников и разнообразных нюансов этой предметной области (таких, как «психология ошибок» переписчика, с. 48), что особенно импонирует читателю, если учитывать, что основной инструментарий применяемых методов лежит в совсем другой области, ориентированной на математическую статистику. Однако гуманитарная сторона исследуемого вопроса не представляется сколько-нибудь ущербно. Напротив, все необходимые к учету аспекты темы учтены и проанализированы. Так, показывая осведомленность в вопросе вариативности написания разнообразных графем средневековых рукописей диссертант специально (и, на наш взгляд, совершенно справедливо) говорит о

необходимости их унификации для последующих подсчетов и приводит алгоритм этой унификации (с. 57).

Вторая глава «Моделирование структуры текста как поликомпонентного объекта» переходит непосредственно к методам исследования. Для создания лингвостатистической модели используются статистические методики, основанные на негауссовом распределении единиц в выборке. Глава представляет собой весьма добросовестный обзор такого рода подходов, их историю в науке, особенности применения. Надо сказать, что традиционная лингвостатистика привыкла оперировать закономерностями, выявленными на нормальном распределении, так что активное введение в научный оборот количественной лингвистики именно негауссовых закономерностей представляется крайне плодотворным, особенно учитывая такую важную особенность распределения именно лексических единиц, как закон Ципфа. Разбору возможностей гауссовой и негауссовой статистики диссертант посвящает специальное место на стр. 105-109. Сопоставительное рассмотрение статистических методов в табличной форме на стр. 153 имеет, на наш взгляд, самостоятельную методическую ценность.

Глава 3 «Лексико-статистическая структура списков «Сказания о Мамаевом побоище»» возвращает диссертанта к основному предмету исследования — спискам древнерусского текста. Рассмотренные в предыдущей главе методики последовательно применяются к выделенному материалу. Методика исправления эффекта рангового искажения по В.П. Маслову, аппроксимация эмпирических кривых H -распределением Б.И. Кудрина, скользящий коэффициент Г.Я. Мартыненко, метод структурно-типологической динамики В.В. Фуфаева и др. каждый по-своему позволяют получать новое знание о помещенных в модель текстах. Хотя в целом этот раздел следует считать своего рода «полигоном», апробацией рассматриваемых методик, результаты которых мы можем оценить, сравнив их с результатами обычного качественного анализа данных. Так, для выделения высокочастотной служебной лексики не

обязательно использовать точку h Хирша (см. 195), так как такого рода лексика уже заранее известна и может быть задана списком до обращения к тексту.

В четвертой главе «Многомодельное представление лексико-статистической структуры «Сказания о Мамаевом побоище» как вариативного текста» осуществляется статистическое сравнение списков памятника. Отметим ряд содержательных выводов, сделанных из этого сравнения. Среди них наиболее интересные касаются особенностей сокращения текста во времени за счет редко встречающихся в тексте слов. Здесь же нужно упомянуть и устойчивость относительных объемов частотных классов слов (с. 232).

Специально обратим внимание на аккуратность формулировок и обстоятельность изложения в тексте диссертации, что также следует считать ее несомненным достоинством.

Список литературы насчитывает 229 позиций и содержательно выглядит вполне соответствующим рассматриваемому предмету. Диссертация также включает объемный иллюстративный материал в виде текстов, таблиц и графиков, к которым следуют отсылки из основного текста работы.

По итогам рассмотрения содержания работы мы заключаем, что *цель исследования достигнута, поставленные задачи решены и достоверность представленных в работе результатов не вызывает сомнений. Практическая значимость исследования заключается в теоретической платформе для создания технологии быстрой применимости к другим аналогичным произведениям письменности.*

Назовем несколько замечаний, имеющих, скорее, дискуссионный характер. Их появление мы склонны были бы связывать не с виной диссертанта, а со сложностью структуры работы, в которой неизбежно должны были появиться дискуссионные аспекты.

Во-первых, вызывает вопросы следующий тезис автора работы: «В любую эпоху при работе над текстом появляются его варианты, но в силу культурных особенностей, полноценные вариативные тексты *значительного объема*

обнаруживаются только в средневековой литературе» (с. 26). Диссертант ссылается на распространенное представление о единичности современного художественного текста, что не отменяет вариативности, например, редакций и вариантов больших романов, как в черновых рукописях (см. соответствующие тома полного собрания сочинений Л.Н. Толстого), так и в разных изданиях, в том числе прижизненных для автора (см. эдиционную историю «Евгения Онегина»). Вариативностью также обладают множественные переводы на один язык текстов другого (ср., например, переводы, включенные в Параллельный корпус в составе Национального корпуса русского языка).

Во-вторых, нам кажется более правильным подходом абстрагирование от конкретных инструментов исследования. Строго говоря, не слишком важно, в уме сложил исследователь два числа, посчитал на калькуляторе или написал специальную программу на одном из языков программирования. Подробное описание очевидных при обработке текста операций в тексте диссертации можно было бы и опустить. В частности, лишними нам кажутся стр. 50-56 (описание работы в системе «СКАТ») и стр. 61-64 (описание работы с лингвопроцессором NooJ).

В-третьих, хотелось бы более артикулированного ответа на вопрос о том, могут ли примененные методики анализа в их совокупности или по отдельности решать какие-то насущные задачи текстологии?

В качестве несущественного для общей концепции работы отметим оформительский недостаток, проявленный в некоторой непоследовательности именования программных продуктов. Так, морфологический анализатор «mystem» в тексте каждый раз упоминается как выпущенный компанией «Яндекс» в то время, как для остальных программ выпустившие их компании не называются. Причина такой разницы не совсем понятна.

Перечисленные замечания не снижают достоинств диссертационного исследования, которое следует охарактеризовать как выполненное на высоком научно-теоретическом уровне.

Автореферат и опубликованные в рецензируемых изданиях работы отражают содержание диссертации.

Таким образом, диссертационная работа Любови Юрьевны Ковригиной «Негауссовое моделирование лексико-статистической структуры вариативности текста (на примере «Сказания о Мамаевом побоище»)» полностью соответствует всем требованиям п. 9-14 Постановления Правительства Российской Федерации «О порядке присуждения ученых степеней» от 24 сентября 2013 г. №842 в части требований, предъявляемым к кандидатским диссертациям, а ее автор заслуживает присуждения искомой степени кандидата филологических наук по специальности 10.02.21 — Прикладная и математическая лингвистика.

Официальный оппонент,
Орехов Борис Валерьевич,
кандидат филологических наук,
доцент факультета филологии
ФГАОУ ВПО Национальный исследовательский университет
«Высшая школа экономики»

«22» января 2015 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