

УДК 51-78: 821.111

КОНТРАСТИВНЫЙ СТИЛЕМЕТРИЧЕСКИЙ АНАЛИЗ ВИКТОРИАНСКИХ РОМАНОВ: ГЕНДЕРНЫЙ АСПЕКТ¹

Валерия Игоревна Мелкозерова

стажер-исследователь Лаборатории языковой конвергенции

НИУ «Высшая школа экономики» – Санкт-Петербург

190068, Санкт-Петербург, наб. канала Грибоедова 123. vimelkozerova@edu.hse.ru

Кристина Дмитриевна Росляк

студент департамента филологии

НИУ «Высшая школа экономики» – Санкт-Петербург

190068, Санкт-Петербург, наб. канала Грибоедова 123. kdroslyak@edu.hse.ru

Исследование посвящено изучению гендерных особенностей в произведениях английских писателей XIX века – сестер Эмили и Шарлотты Бронте, Чарльза Диккенса и Уильяма Теккерея. С помощью автоматического анализа текстов статистическим инструментом *stylo2* на языке R сравнивается отличительная лексика «женских» и «мужских» романов рассматриваемого периода. Результаты исследования показали наличие существенных различий в используемой лексике: в произведениях писательниц больше внимания уделяется описанию эмоций и чувств, природы и телесности, в то время как в текстах писателей преобладают слова, связанные с передвижениями, денежными отношениями и городом.

Ключевые слова: гендерные исследования; викторианский роман; стилеметрия; лексическая асимметрия; художественная картина мира.

Введение

С появлением и распространением феминистского движения стал применяться гендерный подход к осмыслению речи и литературы. Гендерный метод в лингвистике подразумевает изучение взаимосвязи языка и биологической сексуальности его носителей [Пушкарева 1999]. По словам Н.В. Шаминой, такой подход свидетельствует о том, что «сфера чувств, эмоциональные реакции, поведенческие стереотипы, специфика взаимоотношений с миром, формы освоения и присвоения культурных ценностей существуют в мужском и женском вариантах» [Шамина 2005].

Немаловажное влияние на развитие гендерных исследований также оказала когнитивная лингвистика. Когнитивисты исходят из того, что устройство языка может объяснить, как работают познание, интерпретация действительности, категоризация и другие мыслительные процессы [Скребцова 2018]. Соответственно, представляется, что гендерный подход в рамках когнитивной лингвистики позволяет понять, есть ли различия между «женской» и «мужской» картинами мира на основе лексической асимметрии и предпочтении определенных слов.

В настоящей работе проводится анализ гендерных признаков в текстах авторов Англии XIX в. – сестер Эмили и Шарлотты Бронте, а также Чарльза Диккенса и Уильяма Теккерея.

Выбор материала обусловлен тем фактом, что викторианский роман был отправной точкой для развития английской реалистической прозы. Одновременно с наследованием предыдущих литературных традиций – литературе просвещения и жанру готического романа – разрабатывались новые темы, образы, стиль и приемы.

Важную роль в этом процессе сыграли женщины-писательницы, чьи произведения стали фундаментом для становления викторианского романа. Именно в английской литературе Англии XIX в. женская тема является одной из доминирующих, писательницы выходят на первый план, наравне с мужчинами вносят вклад в развитие литературы и приобретают особый социальный вес, а также имеют возможность самостоятельно зарабатывать благодаря творчеству [Шамина 2005]. До этого периода в истории литературы женщины-писательницы редко становились столь известными, как сестры Бронте. Более того, писательницы предыдущей эпохи часто подражали мужскому стилю при написании произведений. В эпоху викторианства в Англии, однако, убежденность женщин-писательниц в необходимости изменения идеологических установок, социокультурного уклада и привычных понятий о женском образе сильно повлияла на их творчество и на литературный процесс. Н.В. Шамина в диссертации «Женская проблематика в виктори-

анском романе» анализирует эволюцию «женского вопроса» через романы Дж. Остен, Ш. Бронте, Э. Бронте и Дж. Элиот, а также присущие этим романам художественно-эстетические особенности. Исследователь отмечает, что произведения сестер Бронте свидетельствуют о том, что писательницы «ощущали необходимость показать процесс становления новой женщины, эволюцию ее жизненных установок». Таким образом, в романах викторианских писательниц становится актуальной женская проблематика на фоне социокультурных, исторических и идеологических изменений эпохи, голос женщины становится громче, приобретает индивидуальные характеристики.

Исследования художественной прозы с применением гендерного подхода в основном сосредоточены на стилистических особенностях текстов женских и мужских авторов, а также на способе выражения эмоций. Так, например, в статье Е.В. Сурковой «Гендерные особенности авторской репрезентации эмоций художественных персонажей» были изучены гендерные различия в описании эмоций персонажей и связанные с этим гендерные стереотипы. Автор статьи делает вывод, что авторы-мужчины и авторы-женщины по-разному описывают эмоциональные реакции своих персонажей. Например, мужчинам свойственна экспликация гнева, тогда как в связи с табуированностью проявлений гнева у женщин «вектор эмоции гнева у персонажей-женщин часто направлен вовнутрь» [Суркова 2005].

Статья В.А. Гришина и Е.А. Ермолаевой также посвящена анализу дифференциации в выражении эмоций относительно половой принадлежности, но касается эмоциональных проявлений персонажей разного гендера. Оказалось, что в художественной литературе Великобритании женские персонажи в большем объеме и с большей вариативностью проявляют эмоции, а также в невербальном аспекте выражают эмоции более открыто (при помощи слез и физического контакта), «в то время как мужчинам свойственно выражение экспрессии негативных эмоций посредством резких движений» [Гришина, Ермолаева 2020].

В работе А.А. Гвоздевой гендерные особенности были изучены с точки зрения языковых картин мира авторов. Исследование подтверждает, что на основе выбора определенных единиц лексикона, предпочтении тех или иных частей речи, синтаксиса и стилистики, можно сделать вывод о дифференциации языковых картин мира мужчин и женщин: женщинам-авторам свойственно более красочное описание речи женских персонажей и их эмоций, «склонность к оценке деятельности окружающих»; мужчины-авторы чаще вы-

бирают глаголы активного залога, указывающие активную позицию, а также сниженную лексику, демонстрирующую грубость и агрессию, и лексику, связанную с армейской, политической, военной и научной тематикой [Гвоздева 2004]. Результаты описанных исследований последовательно демонстрируют, что разница в лексиконе писательниц и писателей действительно существует.

Таким образом, предметом исследования являются различия в изображении художественного мира в текстах викторианских писательниц и писателей на лексическом уровне.

Метод исследования

В настоящей работе будет проведен автоматический стилеметрический анализ произведений. Попытки стилеметрии осуществлялись с XIX в., но развитие и распространение метод получил с появлением компьютеров. С конца 1970-х гг. Дж. Берроуз исследовал художественные тексты, используя частотности отдельных слов и метод главных компонент (PCA) [Burrows 1987]. Он показал, что неполнозначные служебные слова, такие как союзы, предлоги и артикли, могут служить маркерами авторского стиля. В 2001 г. исследователь представил инструмент Delta, который позволяет определять наиболее вероятного автора анонимного текста с помощью сопоставления частотностей служебных слов в этом тексте и в корпусе в целом [Burrows 2002]. Впоследствии был придуман и модифицирован метод Zeta, который позволяет сравнивать два корпуса текстов [Burrows 2006; Craig, Kinney 2009].

Для выявления гендерных сигналов в текстах в настоящем исследовании был выбран метод стилеметрического контрастного анализа текстов с помощью пакета stylo2 [Eder, Rybicki, Kestemont 2016]. Стилеметрия изучает стилистические особенности текста с помощью количественных (статистических) характеристик. Как правило, этот метод используется для атрибуции, т. е. определения автора текста, однако пакет stylo2 позволяет, например, сравнивать два корпуса текстов или кластеризовать несколько текстов по сходству [там же 2016]. В ходе исследования реализованы следующие задачи: апробация инструмента автоматического стилеметрического инструмента stylo2 [там же]; формирование выборки для эксперимента; визуализация данных; формирование таблиц с тематическими группами наиболее частотной отличительной лексики; анализ полученных результатов.

Для стилеметрического анализа использовалась функция `oppose()`² [там же: 118], которая выполняет сравнительный анализ между двумя

заданными наборами текстов и генерирует два списка слов (наиболее предпочтительные и избегаемые слова для первого автора или первой группы авторов). Для подсчетов привлекалась метрика Craig's Zeta (см. [Craig, Kinney 2009]).

Получившиеся списки предпочитаемых слов были проанализированы и упорядочены по тематическим блокам. Подобный подход наследует принципы составления функционального тезауруса, который, по мнению М.Л. Гаспарова, и является «художественным миром» [Гаспаров 1988].

Материал исследования

Анализ осуществляется на материале избранных текстов на языке оригинала – английском. «Женская» часть корпуса состоит из двух викторианских романов XIX в.: «Грозовой перевал» Эмили Бронте (1847 г.) и «Джейн Эйр» Шарлотты Бронте (1847 г.), в то время как представителями «мужской» стороны стали романы «Приключения Оливера Твиста» Чарльза Диккенса (1838 г.) и «Ярмарка Тщеславия» Уильяма Теккерея (1848 г.).

Важно отметить, что выбор текстов, написанных в период с 1838 по 1848 гг., был осуществлен таким образом, чтобы исключить хронологи-

ческие факторы влияния на лексическом уровне. Кроме того, «Оливер Твист» Ч. Диккенса и «Джейн Эйр» Ш. Бронте – это романы воспитания, которые повествуют о жизни юноши и девушки с их раннего возраста, описывают их взросление и становление как личностей, а потому романы похожи еще и в жанровом аспекте.

Так как объем словоупотреблений в оригинальном тексте У. Теккерея почти в два раза превышал объемы других текстов, было принято решение сократить «Ярмарку тщеславия» и взять отрывки из начала, середины и конца произведения, чтобы максимально сохранить возможные данные и сбалансировать корпус для получения точных результатов.

Тексты не проходили специальную предобработку (токенизацию; стемматизацию или лемматизацию), поскольку это не требуется запуска работы stylo2 [Eder, Rybicki, Kestemont 2016].

В ходе работы были сопоставлены материалы общих «женских» и «мужских» корпусов и отдельно романы воспитания Ш. Бронте и Ч. Диккенса. В Таблице 1 представлена характеристика наборов текстов.

Таблица 1

Характеристика выборки

Текст	Объем, словоупотребления	Объем общего корпуса
«Джейн Эйр» Ш. Бронте	~188 000	~305 500
«Грозовой перевал» Э. Бронте	~117 500	
«Ярмарка Тщеславия» (отрывки) У. Теккерея	~157 500	~315 500
«Приключения Оливера Твиста» Ч. Диккенс	~158 000	

Результаты

Результатом работы stylo2 являются списки предпочитаемых и избегаемых слов для первого корпуса текстов относительно второго корпуса. По желанию программа генерирует визуализацию, на которой отображены наиболее отличительные вхождения для обоих списков (см. Рис. 1, 2). На визуализацию попадают имена собственные, особенности орфографии (например, в корпусе «женских» текстов чаще встречаются сокращения составных времен).

Поскольку первым был загружен корпус с «женскими» романами, далее речь пойдет именно об особенностях женской прозы относительно мужской. При первом сравнении, которое проходило на материале объединенных «женских» и «мужских» корпусов текстов, на материале списков слов было выделено семь тематических групп: прилагательные и наречия, цвета; чувства, эмоции и качества; денежные категории; пространство, помещения; люди; образы, явления, природа, понятия; глаголы, действия. Далее рассмотрим эти группы более детально.

1. Описания: прилагательные и наречия, цвета. В Таблицу 2 по указанным выше категориям приведена лексика, которую предпочитают и избегают употреблять женщины-писательницы в своих романах, на фоне лексики романов писателей мужчин.

В категории «Описания: прилагательные, наречия, цвета» «женские» тексты отличают слова *strange* ‘странный’ (здесь и далее перевод наш – В.М. и К.Р.), *wild* ‘дикий’, *stranger* ‘чужой’, ‘незнакомец’ (см. Табл. 2). Возможно, это говорит о замкнутости художественного мира. Известно, что действие в этих произведениях происходит в почти безлюдных местах – усадьбах и поместьях, при этом появление нового персонажа, что может расцениваться как появление чего-то чужого и незнакомого, становится сюжетной завязкой. Слово *wild* могло употребляться как по отношению к женщине и ее «дикому нраву» (героиня «Грозового перевала» отличалась особой темпераментностью), так и по отношению к окружающему миру.

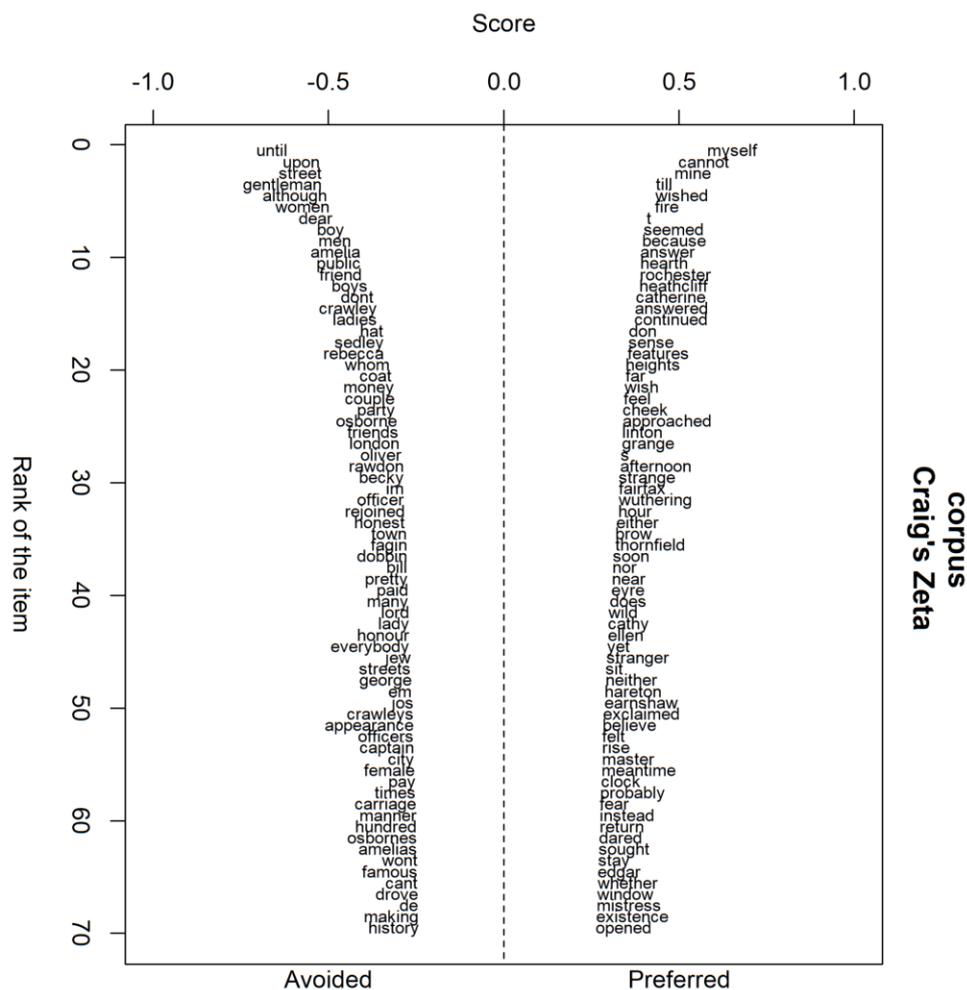


Рисунок 1. Предпочитаемая и избегаемая лексика в «женских» романах относительно лексики «мужских» романов

Кроме того, слово *pure* ‘чистый’, вероятно, могло характеризовать женских персонажей, в то время как, по сравнению с «мужской» выборкой, слова *gentle* ‘нежный’, *tender* ‘ласковый’, ‘мягкий’ не используются. Предварительно можно сделать вывод, что женщины скорее будут описывать героинь как чистых (что часто соотносится со значением добродетели), а мужчины как кротких и ласковых/чутких.

2. **Чувства, эмоции и качества.** Можно заметить, что в женских романах чаще встречается лексика, связанная с желанием, чувствованием, памятью: *wish* ‘желание’, ‘желать’, *wished* ‘желал’, и *desire* ‘сильное желание’, *grin* ‘страсть’ (последнее слово имеет более эмоциональные коннотации. Исследователи отмечают, что в романе «Грозовой перевал» Э. Бронте описана крайне эмоциональная, стремящаяся к власти девушка, что принципиально отличает ее от «традиционных» женских персонажей предыдущей литературной эпохи [Иоскевич 2011: 16–30]. В книге «Джейн Эйр» акцент сделан на внутрен-

ний мир героини: она духовно утонченная, с чувством собственного достоинства, независимая и благородная женщина.

Таким образом, ориентированность на описание духовно-психологических аспектов выражается в лексике: героини желают, чувствуют, верят, боятся, помнят (воспоминания также могут быть эмоциональными), сомневаются. Примечательно также слово *strength* ‘сила’, так как писательницам, по всей видимости, было важно указать на внутреннюю силу своих персонажей.

Писательницы не используют лексику, которая присуща мужскому понятию о чести и добродетели: *honest* ‘честный’, *honour* ‘честь’. По сравнению с мужскими романами, в женских гораздо больше лексики, связанной с эмоциями и выражением чувств, в то время как в мужских романах разнообразие слов, выражающих чувственную составляющую, более скудное.

3. **Денежные категории.** В предпочтительной для писательниц лексике нет понятий, связанных с денежными отношениями (такие слова как

money ‘деньги’, *bill* ‘счет’, *pay* ‘плата’, ‘платить’, *deal* ‘сделка’ и т. д.). Это можно связать с социальным положением женщин в Англии XIX в.: они были экономически зависимы от мужчин, которые управляли всеми финансами в семье. Кроме того, акцент у женщин-писательниц смещен на внутренние достоинства и благодетели, а потому они не употребляют лексику, связанную с внешними признаками успеха и ассоциирующуюся с расчетом.

4. **Пространство, помещения.** При рассмотрении наиболее предпочитаемых слов выстраивается интересный образный ряд: *grange* ‘усадьба’, *grounds* ‘угодья’, *kitchen* ‘кухня’, *library* ‘библиотека’, *chamber* ‘кабинет’, *moors* ‘болота’. Это вновь напоминает о том, что действие в женских романах происходит на фоне сельской местности, во владениях, находящихся за городом, а также наводит на мысль о замкнутости пространства художественного мира. В «женских» романах нет упоминаний большого количества людей и построек, городских развлечений, улиц, архитектурных описаний.

5. **Люди.** В женской прозе самое часто употребляемое предпочитаемое слово – *myself* (которое в русском языке может быть аналогом слов ‘сама’, ‘себе’, ‘себя’ и ‘собой’). Это говорит о сфокусированности главных героинь на себе. При этом, напротив, нет большого количества названий групп людей: *boys* ‘мальчики’, *women* ‘женщины’, *public* ‘общество’, *lady* ‘леди’, *gentlemen* ‘джентльмены’, *crowd* ‘толпа’, *fellows* ‘односельчане/людишки’, *fathers* ‘отцы’, *mothers* ‘матери’, *husbands* ‘мужья’, *friends* ‘друзья’ – вероятно, в «мужских» романах сильнее выражен социальный контекст.

6. **Образы, явления, природа, понятия.** Заметен акцент на телесности в «женских» романах: *brow* ‘бровь’, *knee* ‘колено’, *hair* ‘волосы’, *eye* ‘глаз’, *lips* ‘губы’, *curls* ‘кудри’, *tongue* ‘язык»; также выделяются образы природы и слова, указывающие на погоду: *fire* ‘огонь’, *nature* ‘природа’, *winter* ‘зима’, *spring* ‘весна’, *earth* ‘земля’, *cloud* ‘облако’, *rain* ‘дождь’, *wind* ‘ветер’.

Таблица 2

Предпочитаемая и избегаемая лексика в романах писательниц относительно романов писателей: «Описания: прилагательные и наречия, цвета»

Предпочитаемая лексика	Избегаемая лексика
«Описания: прилагательные и наречия, цвета»	
strange, wild, stranger, dared, pure, tired, hardly, soft, cold, gathered, fastened, feared, handsome, bitter, impossible, wet, continually, smooth, clear	dear, pretty, appearance, famous, young, poor, fat, worthy, dearest, smiling, old, gentle, tender, cocked, tight, instant, british, dirty, great, big, previous, military, best, timid
«Чувства, эмоции, качества»	
wished, seemed sense, wish, feel, believe felt, fear, opened, desire, strength, remember, judgment, feeling, trouble, grace, liberty, wishes, solitude, recall, charm, oh	honest, honour, manner, sex, grin
«Денежные категории»	
	money, bill, paid, pay, hundred, pound, deal, pounds, score
«Пространство, помещения»	
grange, library, kitchen, moors, moor, chamber, hall, threshold, grounds	street, party, officer, town, streets, officers, city, regiment, army, country, places, houses
«Люди»	
myself, mine, miss, human, cousin, fiend	gentleman, women, boy, men, public, friend, boys, ladies, couple, friends, lord, lady, everybody, jew, captain, female, gentlemen, persons, himself, fellows, crowd, fathers, mothers, husbands, man, names, person, sir
«Образы, явления, природа, понятия»	
fire, hearth, spirit, present, rest, brow, nature, winter, frame, spring, cloud, wind, death, trees, knee, hills, reed, hair, rain, eye, aspect, path, nay, approach, lips, curls, earth, tongue, existence, evening	hat, coat, history, fact, waistcoat, legs, beer, dinner, bottle, box, story, effect, boots, conversation, board, length, ball, ill, pace, shop, events, alarm, bumble, pocket, creature, wine
«Глаголы, действия»	
answer, answered, continued, approached, rise, return, sought, stay, seek, commenced, endure, shut, seat, suffer, stayed, tell, sat, observed, ask, stood, explain, saw, turn, resolved, aid, enter, reflected, comprehend, talk, hurt, struck, stir, interrupted	drove, making, presenter, stop, arrived, nodded, carried, carry, occurred, bow, caused, drinking

В избегаемой лексике присутствуют слова, связанные с описанием одежды: *hat* ‘шляпа’, *coat* ‘пальто’, *waistcoat* ‘жилет’, *pocket* ‘карман’, *boots* ‘сапоги’; и быта: *beer* ‘пиво’, *dinner* ‘обед’, *box* ‘коробка’, *conversation* ‘разговор’, *shop* ‘магазин’, *wine* ‘вино’. Это может быть связано как со спецификой пространства, в котором разворачиваются сюжеты книг (усадьба и город), так и с наследованием в женских текстах традиции готического романа связывать чувства и настроение с погодой. Более того, в представлениях XIX в. все еще была распространена идея о естественности женщины и ее близости с природой.

7. **Глаголы, действия.** Интересно, что в «женских» романах предпочтительны глаголы,

связанные с разговором и диалогом: *talk* ‘говорить’, *ask* ‘спрашивать’, *answer* ‘отвечать’, *explain* ‘объяснять»; с поиском, постижением: *seek* ‘искать’, *observed* ‘наблюдать’, *reflected* ‘размышлять’, *comprehend* ‘постигать’; эмоциональным восприятием, *stir* ‘будоражить’, *strike* ‘поразить’. В то же самое время «женским» текстам не свойственны глаголы движения – мотив жажды приключений характерен для главных героев мужчин.

Кроме того, было осуществлено второе сопоставление романов-воспитания Ч. Диккенса и Ш. Бронте, в которых авторы описывают существенную часть жизни центральных героев от их детства до взросления (см. Рис. 2).

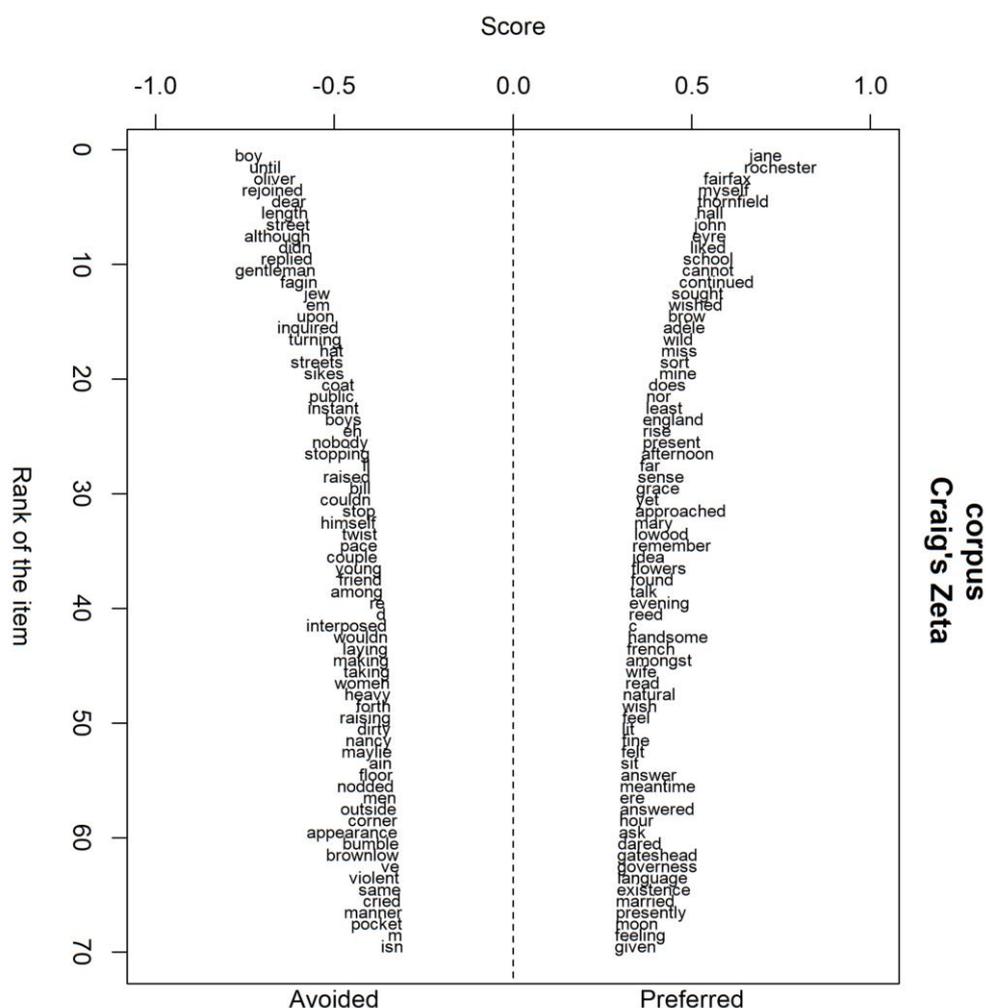


Рисунок 2. Предпочитаемая и избегаемая лексика в романе Ш. Бронте «Джейн Эйр» относительно лексики романа Ч. Диккенса «Приключения Оливера Твиста»

Проанализируем некоторые тенденции по результатам второго сопоставления (полный функциональный тезаурус размещен в Приложении). Обратим внимание, что, как и в предыдущем случае, в тексте женщины-писательницы присутствует больше лексики, связанной с эмоциями,

чувствами и желаниями, а также чаще употребляются положительные качества и описания, в то время как в «Оливере Твисте» присутствует лексика с более отрицательными коннотациями: *cruelty* ‘жестокость’, *cunning* ‘хитрец’, *sneer* ‘усмешка’, *dirty* ‘грязный’, *hasty* ‘поспешный’,

poor 'бедный'). Кроме того, в тексте Бронте, по сравнению с текстом Диккенса, присутствует описание цветов – черный, серый, голубой.

Отдельно нужно представить некоторые слова, которые были выведены программой друг за другом, для более наглядного понимания художественного мира двух авторов.

«Джейн Эйр»: *sort* 'упорядочивать', *mine* 'мой', *sense* 'чувство', *grace* изящество, *remember* 'помнить', *idea* 'идея', *flowers* 'цветы', *found* 'найти', *talk* 'говорить', *evening* 'вечер', *handsome* 'привлекательный', *wife* 'жена', *read* 'читать', *natural* 'естественный', *wish* 'желать', 'желание', *feel* 'чувствовать', *lit* 'освещенный', *fine* 'хорошо'. При прочтении этого словесного списка может выстроиться ассоциативный ряд: привлекательный мужчина проявляет благосклонность к женщине или жене; разговор при свете огня, пространство окрашено светом сумерек, кто-то из персонажей читает книгу, а героиня обязательно что-то чувствует. Кроме того, нельзя не заметить, что эту группу слов объединяет аллитерация (s, f, w, l-m-n), что создает эффект восприятия, близкий к стихотворному. В категорию «люди» можно вынести класс лексики, связанной с семьей: *sisters*, *daughter*, *brother*, *wife*, *family*. Можно сделать вывод, что понятие семьи имело особое значение для писательницы.

«Приключения Оливера Твиста»: *boy* 'мальчик', *rejoined* 'воссоединенный', *street* 'улица', *gentleman* 'джентельмаен', *jew* 'еврей', *inquired* 'осведомлялся', *turning* 'поворачивался', *hat* 'шляпа', *streets* 'улицы', *coat* 'плащ', *public* 'общество', *instant* 'моментальный', *boys* 'мальчики', *stop* 'остановиться', *pace* 'темп', *couple* 'пара', *friend* 'друг', *laying* 'лежащий', *making* 'делающий', *taking* 'берущий', *women* 'женщины', *forth* 'вперед', *raising* 'поднимающий', *dirty* 'грязный', *floor* 'пол', *nodded* от 'кивать', *men* 'мужчины', *outside* 'снаружи', *corner* 'угол', *appearance* 'внешность', *violent* 'насильственный', *cried* причастие или форма прошедшего времени от 'плакать'. Здесь также наблюдается большое количество глаголов и лексики, связанной с городом. Возникает образ человека в пальто и шляпе, активно передвигающегося вдоль улиц, поднимающегося по этажам, наблюдающего грязь и кивающего прохожим.

Обсуждение результатов

Таким образом, сопоставления текстов демонстрируют разницу в лексическом составе произведений, написанных женщинами и мужчинами в викторианскую эпоху. Выбранный метод отчасти позволил решить поставленные задачи.

Было выявлено, что в текстах женщин писательниц чаще встречаются описания

чувств, эмоций, желаний, присутствует больше описаний природы и телесности; также по сравнению с текстами авторов-мужчин, в текстах, написанных женщинами, более явно прослеживается образ собственного «я» (*myself*), фокус обращен на внутренние благородные качества, а денежные категории не так важны. Кроме того, «женские» тексты, судя по выборке глаголов, более статичны – в основном употребляются глаголы, связанные с разговором и поиском. Стоит также отметить, что писательницы избегают обобщений и указания категорий людей. При этом в их текстах специфическое место занимают героини – члены семьи.

Для косвенного подтверждения результатов о гендерном сигнале далее можно было бы продолжить работу с пакетом *stylo*, а именно обратиться к функции *classify()* и автоматически классифицировать несколько фрагментов из романов, написанных женщинами и мужчинами [Eder, Rybicki, Kestemont 2016: 115–117]. Если программа сможет успешно разделить отрывки романов, значит гендерный сигнал действительно можно считать статистически.

Однако следует помнить и об ограничениях количественного метода. Во-первых, интерпретации могут быть неточны, поскольку методы «дальнего чтения» позволяют охватить большие объемы, но не предполагают внимания к деталям. Не всегда точно можно сказать, с чем связан выбор той или иной лексической единицы – для получения более достоверных результатов необходимо анализировать, в каком контексте и в каком значении каждое слово употребляется, чтобы избежать ложных обобщений. Инструмент *stylo2* не учитывает контекстные значения.

Во-вторых, на выбор слов, помимо исторических и гендерных факторов, могут влиять иные культурные характеристики и особенности самих текстов, которые трудно предугадать и исключить. Например, в текущей выборке основное действие в «мужских» романах происходит в городе, а в «женских» – за городом, в усадьбах и поместьях, соответственно, еще предстоит выяснить, случайность это или закономерность. Чтобы нивелировать влияние различных факторов, необходимо повторить исследование на более представительных, сбалансированных выборках большего объема.

Примечания

¹ Исследование выполнено в ходе реализации проекта «Литература и общество: цифровая платформа СОЦИОЛИТ», поддержанного программой «Научный фонд» НИУ ВШЭ в 2024 г.

² Примером использования функции *oppose()* может быть исследование Т.Ю. Литвиновой, посвященное изучению гендерного аспекта речи

участников экстремистского форума [Литвинова 2019]. Исследователю удалось подтвердить ранее полученные данные: женщины используют преимущественно местоимения (я), интенсификатор (очень), отдельные слова группы «чувства» (сердце, люблю), «литература» (стихи, публикации), а мужчины – слова тематической группы «война», «правоохранительные органы» (милиция, правоохранительных). Это исследование подтверждает уместность выбранного метода.

Список литературы

Гаспаров М.Л. Художественный мир писателя: тезаурус формальный и тезаурус функциональный // Проблемы структурной лингвистики. 1988. Т. 1984. С. 125–137.

Гвоздева А.А. Языковая картина мира: лингвокультурологические и гендерные особенности (на материале художественных произведений русскоязычных и англоязычных авторов): дис. ... канд. филол. наук. Краснодар, 2004. 151 с.

Гришина В.А., Ермолаева Е.А. Гендерные особенности выражения экспрессии эмоций (на материале художественной литературы Великобритании) // Иностранные языки: проблемы преподавания и риски коммуникации. Саратов: Саратовский источник, 2020. С. 61–67.

Иоскевич М.М. Интерпретация романа Э. Бронте «Грозовой перевал»: аспекты, итоги, перспективы / Гродненск. гос. ун-т им. Я. Купалы. Гродно, 2011. 173 с.

Литвинова Т.А. Стилеметрическое исследование текстов участников экстремистского форума: гендерный аспект // Известия Воронежского государственного педагогического университета. 2019. №. 4. С. 227–236.

Пушкарева Н.Л. Гендерные исследования: рождение, становление, методы и перспективы в системе исторических наук // Женщина. Гендер. Культура / отв. ред. З.А. Хоткина, Н.Л. Пушкарева, Е.И. Трофимова; Рос. лет. шк. по гендер. исслед., Моск. центр гендер. исслед. М., 1999. С. 15–33.

Скребицова Т.Г. Когнитивная лингвистика: классические теории, новые подходы // М.: Издательский Дом ЯСК, 2018. С. 43–100.

Суркова Е.В. Гендерные особенности авторской репрезентации эмоций художественных персонажей (на материале современной массовой литературы России и Германии): дис. ... канд. филол. наук. Волгоград, 2005. 191 с.

Шамина Н.В. Женская проблематика в викторианском романе 1840–1870-х годов (Джейн Остен, Шарлотта и Эмили Бронте, Джордж Элиот): дис. ... канд. филол. наук. Казань, 2005. 235 с.

Burrows J.F. Computation into criticism: A study of Jane Austen's novels and an experiment in method. Oxford: Clarendon Press, 1987. 270 p.

Burrows J.F. 'Delta': a measure of stylistic difference and a guide to likely authorship // Literary and linguistic computing. 2002. Т. 17, № 3. С. 267–287.

Burrows J.F. All the way through: testing for authorship in different frequency strata // Literary and Linguistic Computing. 2006. Т. 22, № 1. С. 27–47.

Craig H., Kinney A.F. (eds.). Shakespeare, computers, and the mystery of authorship. Cambridge: Cambridge University Press, 2009. 256 p.

Eder M., Rybicki J., Kestemont M. Stylometry with R: a package for computational text analysis // The R Journal. Vol. 8(1). 2016. P. 107–121.

CONTRASTIVE STYLOMETRIC ANALYSIS OF VICTORIAN NOVELS FROM A GENDER PERSPECTIVE

Valeria I. Melkozerova

Research Assistant, Linguistic Convergence Laboratory
Higher School of Economics – Saint Petersburg

Kristina D. Roslyak

Student, Department of Philology
Higher School of Economics – Saint Petersburg

The research is devoted to the study of gender peculiarities in the works of 19th century English authors – sisters Emily and Charlotte Brontë, as well as Charles Dickens and William Thackeray. Using automatic text-matching, the distinctive vocabularies of “female” and “male” novels are compared via the statistical tool Stylo in the R language. The study demonstrates that there are significant differences between vocabularies used by women and men: in the works of female writers more attention is paid to the description of emotions and feelings, nature and corporeality, while in the texts of male writers words related to movement, money relations and the city prevail.

Keywords: gender studies; Victorian novel; stylometry; lexical asymmetry; artistic worldview.

**Приложение. Функциональный тезаурус лексики, характерной и нехарактерной для романа
«Джейн Эйр», относительно романа «Приключения Оливера Твиста»**

Категория	Предпочитаемая лексика	Избегаемая лексика
Описания: прилагательные и наречия, цвета	strange, wild, handsome, natural, fine, dared, strange, charm, clear, pure, reader, quietly, soft, almost, grey, black, continually, smooth, blue, beauty, daily, probably, broad, swept, glorious, bitter, bent	dear, young, interposed, dirty, appearance, particular, hurried, hastily, impatiently, astonishment, old, gradually, latter, inside, short, catching, anxious, violently, cocked, pulling, poor, quickly, softly, lighted, pretty, determined, slight, recent, breathless, accompanied, directly
Чувства, эмоции и качества	wished, seemed, sense, wish, feel, felt, existence, feeling, believe, love, desire, recall, grace, doubt, affection, delight, pleasure, genial, strength	violent, manner, intelligence, dodger, grin, trembling, shaking, growled, sternly, hearted, satisfaction
Денежные категории	wealth	bill, pound
Пространство, помещения	grange, hall, school, england, evening, governess, schoolroom, gates, moor, library, grounds, chamber	street, streets, floor, outside, corner, noise, workhouse, stairs, ground, gate, haven
Люди	myself, mine, miss, wife, self, sisters, daughter, brother, servant, human, mistress, family, madam, servants	boy, gentleman, jew, public, boys, nobody, himself, couple, friend, women, men, doctor, monks, lady, companion, itself, themselves, robber, person, thieves, heads, shop, anybody, crowd
Образы, явления, природа, понятия	fire, hearth, present, idea, flowers, lit, brow, hour, moon, liberty, nature, spirit, marriage, marble, movement, sky, rest, fire, flesh, habits, character, features, task, movements, curls, fortune, cloud, trees, moonlight, choice, wood, presence, world	hat, coat, pace, heavy, forth, pocket, purpose, waistcoat, alarm, consideration, devil, throat, paces, pockets, knees, heels, body, stick, opportunity, history, beer, expedition, shop, birth, pot, effect, thief, collar, fact, behavior, mouth, remark
Глаголы, действия	continued, sought, rise, approached, remember, found, talk, read, sit, answer, answered, ask, married, given, stay, aid, return, leave, aware, live, examined, pass, saw, marry, learn, forget, endure, gave, used, become, find, shall, meet, noticed, marked, lifted, asked, comprehend, allowed, pronounced, enter	rejoined, replied, inquired, turning, stopping, raised, laying, making, taking, raising, nodded, bumble, cried, retorted, pray, remained, walked, carried, hold, won, muttered, returned, darted, starting, drinking, walking, lying, reached