

На правах рукописи

УДК 81'32

РОДИОНОВА

Елена Сергеевна

ЛИНГВИСТИЧЕСКИЕ МЕТОДЫ АТРИБУЦИИ И ДАТИРОВКИ  
ЛИТЕРАТУРНЫХ ПРОИЗВЕДЕНИЙ  
(К ПРОБЛЕМЕ «КОРНЕЛЬ – МОЛЬЕР»)

Специальность 10.02.21 – прикладная и математическая лингвистика

А В Т О Р Е Ф Е Р А Т

диссертации на соискание ученой степени  
кандидата филологических наук

Санкт-Петербург

2008

Работа выполнена на кафедре математической лингвистики  
Федерального государственного образовательного учреждения  
высшего профессионального образования  
«Санкт-Петербургский государственный университет»

Научный руководитель: доктор филологических наук, профессор  
МАРУСЕНКО Михаил Александрович

Официальные оппоненты: доктор филологических наук, профессор  
БЕЛЯЕВА Лариса Николаевна

кандидат филологических наук,  
старший научный сотрудник  
ДМИТРИЕВА Нина Львовна

Ведущая организация: Балтийский государственный технический  
университет «Военмех» им. Д. Ф. Устинова

Защита состоится «\_\_\_» \_\_\_\_\_ 2008 г. в \_\_\_ ч. на заседании  
диссертационного совета Д 212.199.17 по защите диссертаций на соискание  
ученой степени доктора наук в Российском государственном педагогическом  
университете им. А. И. Герцена по адресу: 191186, г. Санкт-Петербург, наб.  
р. Мойки, д. 48, корп. 14, ауд. 314.

С диссертацией можно ознакомиться в фундаментальной библиотеке  
Российского государственного педагогического университета  
им. А. И. Герцена.

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2008 г.

Ученый секретарь  
диссертационного совета  
доктор филологических наук,  
профессор

Л. А. Пиотровская

### Общая характеристика работы

Реферируемая диссертация посвящена разработке языковедческой теории в области атрибуции и датировки литературных произведений на материале стихотворных пьес, приписываемых Мольеру.

Вопрос об авторстве комедий, опубликованных под псевдонимом Мольера, был поднят в начале XX века и на сегодняшний день широко обсуждается как во Франции, так и за ее пределами. Согласно существующим гипотезам, в число возможных авторов пьес, приписываемых Мольеру, входят такие известные французские драматурги, как П. Корнель и Ф. Кино. В разные годы исследователи творчества Мольера приводили различные литературные и биографические доводы, подтверждающие ту или иную точку зрения [P. Louys, H. Poulaille, H. Wouters, G. Forestier]. Наиболее жаркие дебаты разгорелись вокруг возможного написания П. Корнелем лучших комедий в стихах, приписываемых Мольеру. Гипотеза о принадлежности перу П. Корнеля произведений, приписываемых Мольеру, в трудах различных исследователей стала традиционно обозначаться как проблема «Корнель – Мольер». Одно из последних исследований этой проблемы было проведено Д. Лаббе, чья работа базируется на изучении лексического состава языка обоих писателей. Однако анализ одного лишь лексического уровня не может служить достаточно достоверным основанием для атрибуции текстов, поскольку именно лексический состав языка, связанный с темой повествования, лучше всего поддается имитации. Возникла ситуация, когда лингвистическая задача не может быть решена методами, базирующимися исключительно на филологическом анализе, или с использованием методов количественного описания лексического состава текстов.

Описание индивидуальных авторских стилей с целью перевода информации, содержащейся в неформализованном виде в тексте, на формализованный искусственный язык на основе математической статистики заключается в выявлении специфических языковых признаков текста и использовании строгих правил параметризации [Н. А. Морозов, В. В. Виноградов, П. Вашак, И. П. Севбо, Л. В. Милов, Ю. В. Сидоров и др.]. Параметризация стихотворных текстов XVII века на французском языке с использованием априорного словаря параметров синтаксического уровня представляет собой особую лингвистическую задачу.

Вопрос датировки пьес, приписываемых Мольеру, неизменно возникает при рассмотрении проблемы авторства. При установлении времени написания литературных произведений возможны различные приемы, связанные с графологическим и лексическим анализами [Д. С. Лихачев, К. Ф. Тарановский, М. И. Шапир]. На сегодняшний день отсутствуют универсальные методики датировки, основанные на математическом описании авторских стилей.

**Актуальность темы исследования** определяется проблематикой, связанной с математическим моделированием текстов при помощи многомерных классификаций и с автоматизацией лингвистических исследований. В диссертации методы математического моделирования и количественного описания индивидуально-авторских стилей на синтаксическом уровне нашли свое применение при решении проблемы «Корнель – Мольер». Особое значение работа приобретает в связи с возросшим за последнее время интересом к проблеме «Корнель – Мольер» и большой значимостью вопроса об истинном авторстве комедий, приписываемых Мольеру, как для французской литературы, так и для мировой культуры в целом.

Актуальность исследования связана также с разработкой нового математического метода датировки на основе стилеразличающих синтаксических параметров с учетом тренда параметров индивидуального авторского стиля, под которым в данной работе понимается основная тенденция изменения параметров индивидуального авторского стиля на протяжении литературного творчества писателя, очищенная от случайных влияний и индивидуальных особенностей отдельных периодов. С помощью нового метода датировки в работе осуществляется датировка стихотворных пьес, приписываемых Мольеру, и проверяется выдвинутая автором гипотеза о существовании тренда параметров индивидуального авторского стиля.

Таким образом, диссертационная работа, посвященная атрибуции и датировке стихотворных пьес, приписываемых Мольеру, соответствует современной научной проблематике и является актуальной.

**Теоретической основой** исследования являются следующие положения.

Индивидуальный авторский стиль представляет собой сложную иерархическую систему и должен описываться набором стилеразличающих параметров методом многомерной классификации. Проблема атрибуции анонимных и псевдонимных текстов представляет собой одну из задач распознавания образов [С. Ватанабе, М. А. Марусенко].

В настоящее время стиль рассматривается прежде всего как категория структурно-синтаксическая [А. И. Ефимов, И. П. Севбо, Г. М. Мартыненко]. Особенности индивидуального авторского стиля наиболее отчетливо проявляются в синтаксисе: в отборе конструкций, их расположении и объединении в сложное целое.

Индивидуальный авторский стиль может меняться как с годами, так и в зависимости от жанра [И. П. Севбо, М. А. Марусенко]. Проблема датировки может решаться как задача ранжирования многомерных математических моделей с учетом тренда параметров индивидуального авторского стиля.

**Объектом исследования** являются характеристики индивидуальных авторских стилей Мольера, П. Корнеля и Ф. Кино.

**Предметом исследования** в данной работе является установление авторства и датировка спорных произведений, приписываемых Мольеру.

В диссертации выдвигаются три **гипотезы**.

1. Сложная атрибуционная гипотеза:

– тексты стихотворных пьес, приписываемых Мольеру, полностью принадлежат Мольеру (нулевая гипотеза);

– тексты стихотворных пьес, приписываемых Мольеру, полностью принадлежат П. Корнелю (первый вариант альтернативной гипотезы);

– тексты стихотворных пьес, приписываемых Мольеру, являются произведениями П. Корнеля, Ф. Кино и одного или нескольких неизвестных авторов (второй вариант альтернативной гипотезы).

2. Существует тренд параметров индивидуального авторского стиля, выражающийся в том, что на протяжении литературного творчества писателя значения параметров устойчиво изменяются в одном направлении.

3. Стихотворные пьесы, входящие в театр Мольера, были написаны ранее, чем предполагалось.

Основная **цель** работы состоит в применении математического метода атрибуции анонимных и псевдонимных произведений для решения проблемы «Корнель – Мольер». Другой целью работы является разработка математического метода датировки и установление действительной последовательности написания спорных стихотворных пьес.

Достижение поставленных целей предполагает решение следующих **конкретных задач**:

1) провести обзор состояния проблемы «Корнель – Мольер», выбрать объект атрибуции и сформировать атрибуционную гипотезу;

2) сравнить и систематизировать существующие лингвистические методы атрибуции с целью обоснования применения математического метода атрибуции анонимных и псевдонимных произведений, основанного на теории распознавания образов;

3) сформировать априорные классы и априорный словарь параметров с перечислением правил параметризации стихотворных текстов XVII века на французском языке с примерами для каждого параметра;

4) с помощью специально разработанного программного обеспечения определить информативный набор параметров, обеспечивающих разбиение объектов на классы, и создать математические модели анализируемых текстов путем описания объектов на языке информативных параметров;

5) осуществить проверку атрибуционной гипотезы;

6) изучить существующие лингвистические методы датировки и разработать универсальный алгоритм датировки литературных произведений;

7) установить последовательность написания стихотворных пьес, приписываемых Мольеру, с помощью нового метода датировки литературных произведений.

**Материалом** для данного исследования стали следующие тексты:

1) 13 комедий в стихах, приписываемых Мольеру (“L'Étourdi”, “Le dépit amoureux”, “Sganarelle”, “Dom Garcie de Navarre”, “L'École des maris”, “Les

Fâcheux”, “L'École des femmes”, “La Princesse d'Élide”, “Tartuffe”, “Le Misanthrope”, “Mélicerte”, “Pastorale comique”, “Les Femmes savantes”);

2) 11 комедий в стихах П. Корнеля (“Mélite”, “La Veuve”, “La Galerie du Palais”, “La Suivante”, “La Place royale”, “Comédie des Tuileries”, “L'Illusion comique”, “Le menteur”, “La Suite de Menteur”, “Don Sanche d'Arago”, “Tite et Bérénice”);

3) 3 комедии в стихах Ф. Кино (“Les Rivaies”, “L'Amant indiscret”, “La Mère Coquette”).

**Методы исследования.** В работе изучаются возможности применения математических и прикладных методов к исследованию языка и используется семиотическое и математическое моделирование естественного языка на основе теории вероятностей и математической статистики. Кроме того, в работе применяются методы обработки данных, стилеметрии, лингвостатистики, системно-описательный метод, а также методы филологического анализа, в частности методы анализа состава и структуры предложения.

#### **Основные положения, выносимые на защиту.**

1. Стихотворные пьесы, приписываемые Мольеру, являются произведениями П. Корнеля (“Le dépit amoureux”, “Sganarelle”, “L'École des maris”, “Les Fâcheux”, “L'École des femmes”, “Tartuffe”, “Le Misanthrope”, “Mélicerte”, “Pastorale comique”, “Les Femmes savantes”), Ф. Кино (“L'Étourdi”) и неизвестного автора (“Dom Garcie de Navarre”, “La Princesse d'Élide”).

2. Предложенная методика определения значений параметров на материале текстов XVII века на французском языке позволяет решать задачу определения авторства анонимных и псевдонимных текстов данного периода.

3. Существует тренд параметров индивидуального авторского стиля, выражающийся в том, что на протяжении литературного творчества писателя значения параметров устойчиво изменяются в одном направлении.

4. Метод датировки на основе стилеразличающих синтаксических параметров, предложенный в работе, позволяет решать задачи по установлению последовательности написания литературных текстов и может использоваться в дальнейших исследованиях.

5. В результате датировки спорных пьес, атрибутированных П. Корнелю и Ф. Кино, подтверждена гипотеза о более раннем написании стихотворных пьес, входящих в театр Мольера.

#### **Научная новизна** работы заключается в следующем:

- осуществлена атрибуция стихотворных пьес, приписываемых Мольеру, методами математического анализа текстов с применением теории распознавания образов;

- разработаны правила параметризации текста на материале текстов XVII века на французском языке;

- выявлены новые возможности автоматизации лингвистического исследования в области установления авторства;

- разработан метод датировки литературных произведений;

– новый метод датировки применен при установлении последовательности написания пьес, приписываемых Мольеру.

**Теоретическая значимость** проведенного исследования заключается в том, что в нем решена проблема авторской принадлежности текстов классиков французской литературы. Результаты исследования способствуют развитию математических и прикладных методов изучения языка и вносят вклад в развитие теории и практики атрибуции литературных произведений методами распознавания образов, а также теории и практики датировки литературных произведений математическими методами.

**Практическая значимость** исследования состоит в возможности использования полученных данных о составе априорного словаря параметров в качестве примера при составлении правил параметризации текстов на французском материале. Полученные данные о существовании тренда параметров индивидуального авторского стиля могут использоваться в дальнейших работах по определению последовательности написания литературных произведений с целью их датировки.

**Рекомендации по использованию результатов** исследования. Полученные в ходе работы результаты можно использовать при чтении специальных курсов «Анализ данных» и «Лингвистические методы атрибуции анонимных и псевдонимных произведений», а также курсов по прикладной лингвистике, квантитативной стилистике и истории французской литературы.

**Апробация работы.** Основные положения диссертационного исследования обсуждались на заседаниях кафедры математической лингвистики филологического факультета Санкт-Петербургского государственного университета (2005–2007 гг.) и прошли апробацию в ряде докладов, прочитанных на XXXV и XXXVI Международных филологических конференциях (СПбГУ, 2006-2007), на Всероссийской научной конференции «Квантитативная лингвистика: исследования и модели» (Новосибирск, 2005), Всероссийской научной конференции с международным участием «Актуальные проблемы теоретической и прикладной лингвистики и оптимизация преподавания иностранных языков» (Тольятти, 2005), на Международной научной конференции «MegaLing'2007. Горизонты прикладной лингвистики и лингвистических технологий». (Украина, Партеит, 2007), а также получили отражение в 7 публикациях общим объемом 2 п. л.

**Объем и структура работы.** Диссертационное исследование содержит 183 страницы машинописного текста, оно состоит из введения, четырех глав с выводами в конце каждой, заключения, библиографии, включающей 128 наименований, из них 46 на иностранных языках, и списка источников. Основные результаты исследования обобщены в 41 таблице и 1 рисунке. В качестве приложения приводятся таблицы, описывающие состав театров Мольера, П. Корнеля и Ф. Кино.

Построение работы определяется целями, задачами и материалом исследования.

## Основное содержание работы

Во **введении** обосновывается актуальность исследования, дается обзор последних работ по его проблематике, определяются объект, предмет, гипотезы, цель задачи и материал исследования, перечисляются методы исследования, раскрывается его научная новизна, теоретическая и практическая значимость, предлагаются рекомендации об использовании его результатов, формулируются положения, выносимые на защиту.

В **первой главе** «К истории вопроса» произведен обзор состояния проблемы «Корнель – Мольер», выбран объект атрибуции и сформирована атрибуционная гипотеза.

Проблема «Корнель – Мольер» была впервые затронута в 1919 году известным французским поэтом П. Луи, большим знатоком поэзии XVII века. Досконально изучив творчество великого французского драматурга П. Корнеля (1606-1684), автора 34 стихотворных пьес, он неожиданно для себя обнаружил большое сходство его стихов со стихами другого знаменитого драматурга Ж.-Б. Поклена (1622-1673), работавшего под псевдонимом Мольер. Театр Мольера составляют 33 произведения, написанных как в стихотворной, так и в прозаической форме. Исходя из стилистических характеристик работ двух драматургов, он приписал П. Корнелю работу над такими шедеврами Мольера, как “Le Misanthrope”, “L'École des femmes”, “Tartuffe”, “Dom Juan”. П. Луи высказал предположение, что одной из причин, по которой П. Корнель мог отказаться от своих произведений в пользу актера Мольера, было его желание рассказать в комедиях о своей любви, жизни, чувствах, но сделать это анонимно, не раскрывая истинного своего имени.

Основные тезисы, намеченные в статьях П. Луи, были доработаны и обоснованы в 1957 году в книге романиста А. Пулайля «Корнель под маской Мольера». А. Пулайль пишет альтернативную биографию Мольера, начиная с его раннего детства. Исследователь сопоставляет факты биографий Мольера и П. Корнеля и доказывает возможность и несомненность этого творческого союза. В качестве основной причины, побудившей величайшего поэта своего времени П. Корнеля воспользоваться услугами Мольера, А. Пулайль называет желание П. Корнеля свести счеты со своими многочисленными врагами, изобразив их в злободневных сатирических пьесах.

Эта работа не нашла в свое время отклика ни у простых почитателей таланта Мольера и П. Корнеля, ни у профессиональных мольеристов и корнелистов, и до 1990 года, когда проблема авторства театра Мольера была вновь поднята двумя адвокатами из Брюсселя, И. Вутером и К. де Вилль де Гуйе, лишь немногие исследователи французской литературы XVII века слышали о гипотезе П. Луи.

В своей работе И. Вутер и К. де Вилль де Гуйе подробно останавливаются на идейно-стилистических характеристиках таких пьес, как “Les Précieuses ridicules”, “Les Fâcheux”, “L'École des femmes”, “Tartuffe”,



“Dom Juan”, “Amphitryon”, и, принимая во внимание также историю написания этих пьес, доказывают невозможность авторства Мольера. Создание длинного списка работ, подписанных Мольером, потребовало бы от него усиленной работы с утра до вечера, но Мольер никак не располагал таким количеством времени, поскольку постоянно работал в качестве директора театра, режиссера-постановщика и актера. Исследователи указывают также на затрудненное финансовое положение П. Корнеля, связанное с зависимостью от королевских денежных выплат. Именно финансовая зависимость от двора, по их мнению, заставила П. Корнеля высмеивать своих врагов под чужим псевдонимом. Помимо П. Корнеля, в числе возможных авторов, работавших под именем Мольера, они называют драматурга Ф. Кино, чей стиль близок комедиям-балетам, приписываемым Мольеру.

В приложении к этой книге находится работа выдающегося этнолога Ф. Вернауда. Проведя лексикографическое и стилистическое исследование пьес, он обнаружил большое сходство текстов Корнеля и Мольера: сходство словаря, стихосложения и стиля. Например, в пьесах Мольера им были обнаружены нормандские термины и большое число терминов и выражений, связанных с юриспруденцией, которые мог использовать только П. Корнель, житель Руана и блестящий юрист. Кроме того, в них содержатся цитаты из мало известных произведений на латыни и ссылки на религиозную литературу, которая отсутствует в библиотеке Мольера и которую, несомненно, хорошо изучил набожный П. Корнель. Между текстами пьес Мольера и П. Корнеля Ф. Вернауд находит такие совпадения, как ссылки на Аристотеля и Горация, упоминания Армении или армян, использование редкого имени Никандр и другие.

Современные литературоведы и театральные деятели отзываются об этих исследованиях очень критично и приводят свои контраргументы. Наиболее полный критический обзор основных положений сторонников сотрудничества Мольера и П. Корнеля был проведен заведующим кафедрой исследований театра XVII века в Сорбонне Ж. Форестье. По его мнению, театр Мольера является достаточно однородным по своему составу, несмотря на обилие жанровых подвидов: в любой большой комедии Мольера можно найти влияние фарсов. Способность Мольера писать большое количество пьес в сжатые сроки Ж. Форестье объясняет тем, что многие из его работ представляют собой короткие дивертисменты и написаны в прозе.

Однако у сторонников атрибуции П. Корнелю пьес Мольера, помимо идейно-стилистических и биографических аргументов, появилось первое математическое доказательство.

В 2001 году профессор Института политических исследований в Гренобле, специалист по анализу речи Д. Лаббе предложил математический метод атрибуции, заключающийся в вычислении «межтекстового расстояния». В 2003 году он применил новый метод при исследовании лексического состава театров Мольера и П. Корнеля. Вычисление

«межтекстового расстояния» позволило ему атрибутировать П. Корнелю около 18 комедий, подписанных Мольером.

Исследование Д. Лаббе получило широкую огласку во Франции и за ее пределами и вызвало огромный интерес со стороны специалистов различных областей знаний – как математиков, так и литературоведов. Кроме того, работа Д. Лаббе послужила толчком к новым исследованиям проблемы авторства театра Мольера.

В 2003 году гипотеза тайного сотрудничества Мольера и П. Корнеля была подтверждена результатами расследования Ф. Видаля, проведенного на основе анализа биографических данных двух драматургов.

Еще через год вышла в свет книга Д. Буасье, представляющая собой наиболее полное и структурированное досье по делу Мольера. Вслед за предыдущими исследователями творчества Мольера и П. Корнеля, Д. Буасье утверждает, что пьесы, составляющие его театр, были получены им тремя путями: часть пьес была составлена его сотрудниками из французских, испанских и итальянских комедий; другая часть была куплена у различных поэтов, нуждавшихся в деньгах, или у их вдов; и, наконец, какая-то часть пьес была заказана у П. Корнеля.

Гипотеза о тайном сотрудничестве Мольера и П. Корнеля стала столь популярна, что официально вошла в историю литературы: в 2005 году в учебное пособие по истории французской литературы был включен соответствующий параграф под заголовком «Написал ли Корнель пьесы Мольера?».

Проведенный обзор истории вопроса атрибуции театра Мольера свидетельствует о все возрастающем интересе к данной проблеме и о существовании разных подходов к ее решению.

В реферируемом исследовании, исходя из требования соблюдения жанрово-стилевой однородности текстов, в класс атрибутируемых объектов были включены 13 комедий в стихах, приписываемых Мольеру (см. с.5).

При формировании атрибуционной гипотезы в число возможных авторов могут быть включены драматурги Мольер, П. Корнель, Ф. Кино и неопределенное число неизвестных авторов. Литературно-критическая атрибуционная гипотеза может быть сформулирована следующим образом.

Нулевая гипотеза ( $H_0$ ): тексты стихотворных пьес, приписываемых Мольеру, полностью принадлежат Мольеру.

Альтернативная гипотеза имеет следующий вид:

( $H_a^1$ ): тексты стихотворных пьес, приписываемых Мольеру, полностью принадлежат П. Корнелю;

( $H_a^2$ ): тексты стихотворных пьес, приписываемых Мольеру, являются произведениями П. Корнеля, Ф. Кино и одного или нескольких неизвестных авторов. В этом случае необходимо определить число авторов и возможную долю участия каждого из них.

**Вторая глава** носит название «Методы атрибуции». В ней произведен анализ субъективных и объективных методов атрибуции, рассмотрены методы атрибуции, основанные на анализе лексики и синтаксических структур, изучены некоторые аспекты применения автоматической обработки данных и обосновано применение математического метода атрибуции анонимных и псевдонимных работ.

Долгое время в практике атрибуции доминировали историко-документальные и филологические методы исследования. Для выявления авторских особенностей применялась субъективная методика атрибуции, в соответствии с которой субъективно отбирались внешние детали авторского стиля, такие как любимые слова, термины, выражения.

Первым отечественным ученым, использовавшим математический аппарат для решения задачи атрибуции, считается Н. А. Морозов, опубликовавший в 1915 г. статью «Лингвистические спектры». В отличие от предшествующих исследователей, филологов-классиков, опиравшихся при атрибуции на частоту употребления знаменательных слов, Н. А. Морозов полагал, что для индивидуального стиля писателя показательными являются именно служебные слова, поскольку они никак не связаны с темой и содержанием книги. Метод, предложенный Н. А. Морозовым, лег в основу многих исследований по лексическому составу языка писателей, однако не может служить основой для полного достоверного стилистического исследования, поскольку не выходит за рамки лексического анализа и состава предложения. Критические обзоры метода и результатов исследования Н. А. Морозова были сделаны А. А. Марковым и В. Э. Сеземаном.

В классификации методов атрибуции, предложенной академиком В. В. Виноградовым, были противопоставлены субъективные и объективные принципы, применяемые в целях атрибуции. В этой работе группа объективных принципов выделялась в рамках традиционной лингвистики, а объективным лингвостатистическим методам атрибуции отводилась лишь вспомогательная роль, поскольку к началу 60-х годов XX века отсутствовали эффективные методы атрибуции, основанные на формальном математическом аппарате.

С 60–70-х годов XX века при описании индивидуального стиля лингвоматематические методы стали применяться все шире, благодаря чему накапливались данные о свойствах единиц языка и формировался специальный научный аппарат атрибуции текстов. Работы А. Л. Гришунина, А. Якубайтис, А. Н. Скляревича, А. П. Василевича посвящены применению методов статистики к лексике и грамматике.

Использование специально разработанных индексов для оценки лексической структуры текста было обусловлено стремлением разработать новый универсальный аппарат для объективного анализа лексики. Многие ученые разрабатывают новые оценки лексического состава, отличающиеся теми или иными недостатками, не рассматривая уже готовые, существующие решения в смежных областях знания, например в математике, где для

описания объектов различной природы давно и успешно применяется теория распознавания образов.

Опыт количественно-лингвистических исследований был обобщен в монографии Ю. Тулдавы в 1987 году. В этой работе Ю. Тулдава сформулировал два основных принципа изучения лексики в количественном аспекте: принцип системности и вероятностно-статистический характер организации лексики, а также высказал идею о связи признаков, которая явилась предпосылкой для разработки математического аппарата оценки связей между параметрами.

Одно из последних исследований по атрибуции текстов, основанных на лексическом анализе, было проведено Д. Лаббе, предложившим в 2001 году формулу вычисления «межтекстового расстояния», которое подразумевает анализ лексического состава двух текстов и определение меры их близости или удаленности друг от друга по следующей формуле:

$$D_{V_a,b(u)} = \sum_{V \in (A,B)} |F_{ia} - E_{ia(u)}|, \quad (1)$$

где  $F_{ia}$  – частота вокабулы  $i$  из множества  $A$ ,  $E_{ia(u)}$  – математическое ожидание вокабулы  $i$ .

В работах Д. Лаббе лексический анализ текстов происходит с помощью автоматической процедуры морфологического анализа, в соответствии с которой каждое слово представляется в виде записи, состоящей из трех компонентов: словоформы, вокабулы и соответствующей части речи. По полученным данным вычисляется «межтекстовое расстояние», и результаты отображаются в виде древовидной классификации. Исследования Д. Лаббе обнаруживают серьезные недостатки методологического и статистико-вероятностного характера, связанные в первую очередь с недостоверностью результатов атрибуции, осуществленной на основе анализа одного лишь лексического уровня. При стилистическом анализе в целях атрибуции изучение лексического состава текста должно быть дополнено данными и о других языковых уровнях, и в первую очередь – о синтаксической структуре анализируемого текста.

В 70-е – 80-е годы XX века в отечественной лингвистике был проведен ряд исследований, посвященных количественно-структурному изучению текстов на синтаксическом уровне. Методы стилистической диагностики, основанные на анализе графов синтаксических связей, представлены в исследованиях И. П. Севбо и Г. Я. Мартыненко. Диагностические параметры, предложенные И. П. Севбо, и меры сложности, анализируемые в работах Г. Я. Мартыненко, связаны с характеристиками предложения, а не текста, но именно анализ текста должен лежать в основе эффективной методики фиксации авторского стиля.

Период с конца 70-х годов XX века до настоящего времени отмечен бурным развитием вычислительной техники и программного обеспечения, в связи с чем все больше исследователей проявляют интерес к применению

компьютерной обработки данных при анализе текстов, как в синтаксическом, так и в грамматическом, лексическом аспектах. Одна из первых методик установления авторства, основанная на анализе текста с автоматизированным получением частотных словарей и статистических данных, была предложена в работе норвежского филолога Г. Хетсо в 1978 году. В исследованиях древних текстов, проводимых под руководством Л. В. Милова, обработка текстового материала заключается в построении графов «сильных связей» по матрице частот парной встречаемости грамматических классов слов и происходит с помощью специально разработанной компьютерной программы. Обязательное применение автоматической обработки данных в целях определения авторства лежит в основе работ Ю. В. Сидорова, И. О. Тарнопольской, Д. В. Хмелева.

Стремление ученых к применению автоматической стилистической диагностики и автоматизированного поиска индивидуальных характеристик авторского стиля приводит к тому, что предпочтение в стилистических исследованиях отдается анализу любых других языковых уровней, кроме синтаксического. Зависимость стилистического анализа от компьютерной обработки данных и от методов, для нее предназначенных, приводит к упрощению методологической основы исследований, что, в конечном итоге, делает методы атрибуции текста менее эффективными.

В диссертационной работе, исходя из требований объективности и полноты описания авторского стиля, в основу исследования по определению авторства пьес Мольера была положена методика атрибуции анонимных и псевдонимных произведений, разработанная М. А. Марусенко, которая предусматривает применение многомерной классификации, основанной на теории распознавания образов, и описание индивидуального авторского стиля в синтаксическом аспекте. Данная методика позволяет определять характеристики текста, а не отдельного предложения на разных уровнях языковой системы. На этапе описания атрибутируемых объектов на языке параметров из априорного словаря параметров предусмотрена ручная обработка данных. Статистическая обработка полученных данных и реализация алгоритма проверки атрибуционной гипотезы производится с помощью компьютерных программ.

**Третья глава** диссертационной работы «Атрибуция стихотворных пьес Мольера» посвящена проверке атрибуционной гипотезы и интерпретации ее результатов.

Проверка атрибуционной гипотезы предусматривает отбор информативных параметров из априорного словаря параметров и процедуру распознавания.

Инвентарь параметров, составляющих исходное описание, был составлен из 51 параметра, релевантного для описания текстов XVII века на французском языке. Язык параметров представляет собой систему определения и фиксации ряда наиболее существенных структурных особенностей организации предложений. Значения параметров, полученные

в результате описания априорных классов, являются признаками, характеризующими соответствующие классы.

В работе сформулированы общие правила анализа текста на французском материале и введены правила параметризации текста по каждому параметру.

Состав априорного алфавита классов определяет математические модели априорных классов, относительно которых производится атрибуция спорных произведений. В данной работе сформированы два априорных класса: П. Корнеля –  $\Omega_1$ (Corneille) и Ф. Кино –  $\Omega_2$ (Quinault), мощностью 11 и 3 текста соответственно.

Для определения информативного набора параметров был проведен эксперимент по описанию априорных классов на языке параметров из априорного словаря параметров, для чего были сделаны «прикидочные» случайные выборки объемом по 100 предложений. Результаты эксперимента были представлены в виде двух объектно-признаковых матриц данных размерностью  $n \times N = 100 \times 51$ , где  $n$  – число параметров, а  $N$  – число объектов.

При формировании набора информативных параметров была применена схема М. М. Бонгарда, предусматривающая двухступенчатое свертывание параметрического пространства. На первом этапе происходит разбиение априорного набора информативных параметров на два подмножества параметров, релевантных и нерелевантных для различения априорных классов. Релевантность параметров для различения двух априорных классов определяется по  $t$ -критерию Стьюдента (формула 2), пороговое значение которого при уровне значимости  $\alpha = 0,05$  равно 1,96.

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\left( \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)^{1/2}}, \quad (2)$$

где  $\bar{x}_1, \bar{x}_2$  – средние арифметические,  $\sigma_1, \sigma_2$  – стандартные отклонения, а  $n_1, n_2$  – размеры выборок.

Значения  $t$ -критерия для пяти параметров оказались больше критического, что позволило определить параметры X02 (число элементарных предложений), X04 (число сочиненных предложений), X21 (число спрягаемых форм глагола), X31 (число подлежащих), X32 (число местоимений-подлежащих) как информативные.

Второй этап схемы М. М. Бонгарда предусматривает процедуру свертывания параметрического пространства на подмножестве информативных параметров. Результаты вычисления с помощью программного обеспечения критерия эффективности каждого параметра показали, что дальнейшего уменьшения числа диагностирующих параметров не произошло, и информативный набор составили пять параметров, полученных на первом этапе.

Этап перехода от реального объекта к его математической модели заключается в описании этого объекта на языке параметров рабочего словаря системы распознавания. Следующим этапом исследования стало формирование матриц данных атрибутируемых объектов и априорных классов, чей объем был заранее определен. Полученные координаты эталонов классов приведены в табл. 1.

Таблица 1. Координаты эталонов классов на 0-й итерации

Параметр	Класс			
	$\Omega_1^0 (Corneille)$		$\Omega_2^0 (Quinault)$	
	$\bar{x}_i$	$\sigma_i$	$\bar{x}_i$	$\sigma_i$
X02	1,91	1,08	2,02	1,38
X04	0,56	0,98	0,92	1,31
X21	1,90	1,25	2,01	1,49
X31	1,69	1,13	1,83	1,44
X32	1,25	1,01	1,40	1,16
	N=1122		N=465	

Алгоритм распознавания включает в себя детерминированную и вероятностную атрибуцию. Детерминированный алгоритм распознавания определяет принадлежность атрибутируемых объектов к тому или иному классу текстов. В качестве классифицирующей функции детерминированного алгоритма распознавания использовался  $t$ -критерий Стьюдента (формула 2 на с. 14). Положительное решение об отнесении объекта к тому или иному классу принималось только в том случае, если наблюдаемые значения  $t$ -критерия меньше критического во всех пяти измерениях пространства классификации. Классификация объектов при помощи детерминированного алгоритма распознавания является итерационной процедурой, в которой после каждой итерации происходит изменение состава и мощности классов, а также перерасчет координат классов. Остановка детерминированного алгоритма произошла на 3-ей итерации. В результате работы детерминированного алгоритма была определена авторская принадлежность шести из тринадцати анализируемых объектов: пьесы “Le dépit amoureux”, “L'École des maris”, “Les Fâcheux”, “L'École des femmes”, “Tartuffe”, “Les Femmes savantes” были атрибутированы П. Корнелю с вероятностью, находящейся в интервале между  $0,95$  и  $0,95+0,05^5$ , причем реальное значение вероятности ближе к правой границе интервала.

После остановки детерминированного алгоритма остались не отнесенными ни к одному классу семь атрибутируемых объектов. Реализация вероятностного алгоритма распознавания предусматривает преобразование исходной матрицы данных в матрицу евклидовых расстояний между априорными классами и неатрибутированными объектами.

Исходная матрица данных строится по средним значениям  $\bar{X}_i$  каждого из пяти диагностирующих параметров.

Поскольку значения используемых диагностирующих параметров могут заметно отличаться, для приведения всех параметров к единой шкале необходимо привести матрицу данных к стандартному виду, при котором средние значения всех параметров равны нулю, а дисперсии равны единице. Стандартизованная матрица (см. табл. 2) обозначается как  $X = \{x_{ij}\}$ .

Таблица 2. Стандартизованная матрица данных  $X = \{x_{ij}\}$

		Параметр				
		x2	x4	x21	x31	x32
Объект	1М	0,39	1,18	0,12	0,30	0,73
	3М	-1,13	-0,77	-1,22	-1,17	-1,08
	4М	1,65	0,01	1,57	1,49	1,29
	8М	1,17	0,29	1,39	1,20	0,45
	10М	-0,05	-0,55	-0,16	0,16	0,64
	11М	-0,80	-0,94	-0,97	-0,77	-0,80
	12М	-1,32	-1,05	-1,01	-1,45	-1,73
Класс	$\Omega_1$	-0,17	-0,10	-0,05	-0,13	-0,10
	$\Omega_2$	0,24	1,91	0,33	0,37	0,59

С помощью функции евклидова расстояния (формула 3) на основе полученной стандартизованной матрицы данных была построена матрица евклидовых расстояний между априорными классами и неатрибутированными объектами (см. табл. 3).

$$d(a,b) = \left[ \sum_{i=1}^n (x_{aj} - x_{bj})^2 \right]^{0,5}, \quad (3)$$

где  $n$  – размерность евклидова пространства,  $a$  и  $b$  – две точки в пространстве  $E_n$  с координатами  $a(x_{a1}, x_{a2}, \dots, x_{an})$ ,  $b(x_{b1}, x_{b2}, \dots, x_{bn})$ .

Таблица 3. Матрица евклидовых расстояний между классами и неатрибутированными объектами  $d(X_i, \Omega_i)$

		$d(X_i, \Omega_i)$						
		Объект						
Класс		1М	3М	4М	8М	10М	11М	12М
$\Omega_1$		1,70	2,18	3,24	2,47	0,93	1,69	2,75
$\Omega_2$		0,79	4,08	2,97	2,30	2,53	3,76	4,66



Вероятность принадлежности  $j$ -го объекта к  $i$ -тому классу (см. табл. 4) вычисляется на основе матрицы евклидовых расстояний между объектами и априорными классами по следующей формуле:

$$P_{ji} = \frac{1}{d_{ji}} \left( \sum_k \frac{1}{d_{jk}} \right)^{-1}, \quad (4)$$

где  $d_{ji}$  – расстояние между  $j$ -тым объектом и  $i$ -тым классом, а  $d_{jk}$  – расстояние между  $j$ -тым объектом и остальными классами классификации.

Таблица 4. Матрица вероятностей принадлежности объектов априорным классам  $P(X_i \in \Omega_j)$

		$P(X_i \in \Omega_j)$						
		Объект						
Класс		1М	3М	4М	8М	10М	11М	12М
$\Omega_1$		0,32	0,65	0,48	0,48	0,73	0,69	0,63
$\Omega_2$		0,68	0,35	0,52	0,52	0,27	0,31	0,37

При наличии двух априорных классов решающее правило может быть сформулировано следующим образом:

$$X_i \in \Omega_j, P(X_i \in \Omega_j) \geq 0,5. \quad (5)$$

При данном решающем правиле пороговое значение вероятности принадлежности объекта к одному из априорных классов должно равняться или превышать 0,5. В результате работы вероятностного алгоритма распознавания Ф. Кино были атрибутированы пьесы “L'Étourdi”, “Dom Garcie de Navarre” и “La Princesse d'Élide”, а П. Корнелю были приписаны пьесы “Sganarelle”, “Le Misanthrope”, “Mélécerte”, “Pastorale comique”.

Заключительным этапом процедуры распознавания стала оценка качества полученной классификации, основанная на выявлении закономерностей для каждого класса.

Для признания законченности процедуры распознавания необходимо добиться выполнения условия соответствия полученного разбиения на классы содержательному понятию кластера и условия, при котором полученный кластер может называться сгущением. Кластер и сгущение – две качественные градации, которые позволяют определить, насколько тесно расположены объекты внутри класса по сравнению с расположением объектов внутри всей исходной совокупности. Образованные в процессе действия распознающего автомата классы будут отвечать содержательному понятию кластера в том случае, если средний квадрат расстояния между объектами класса будет меньше среднего квадрата расстояния между объектами исходной совокупности, т. е.  $\bar{d}^2(\Omega_n) < \bar{d}^2(\Omega_N)$ . Сгущение в таком

случае реализуется, когда максимальный квадрат расстояния между объектами данного класса меньше среднего квадрата расстояния между объектами исходной совокупности, то есть  $d_{\max}^2(\Omega_n) < \bar{d}^2(\Omega_N)$ . На этом этапе были подтверждены выводы, сделанные по результатам вероятностного алгоритма распознавания.

Следующим этапом оценки качества классификации стало сравнение однородности априорных и апостериорных классов. В качестве критерия однородности было использовано среднее расстояние между объектами  $\bar{d}(\Omega_N)$ . Оценка однородности состава априорных классов полностью подтвердила результаты, полученные в ходе детерминированного и вероятностного алгоритма распознавания для класса  $\Omega_1$ (Corneille), которому с разной степенью вероятности были атрибутированы 10 объектов. Что касается  $\Omega_2$ (Quinault), то по результатам проведенного анализа качества классификации из трех атрибутированных классу объектов лишь один объект, пьеса “L'Étourdi” может быть атрибутирована ему с вероятностью 0,68. Два других объекта формируют отдельный апостериорный класс  $\Omega_3$ , как не принадлежащие ни одному из априорных классов, но составляющие один класс, удовлетворяющий условиям кластера и сгущения.

Итак, в результате проведенного исследования получила свое подтверждение гипотеза ( $H_a^2$ ): тексты пьес Мольера являются произведениями П. Корнеля, Ф. Кино и неизвестного автора.

Представим полученные данные в виде матрицы решений, где 0 – отказ от распознавания, 1 – положительное решение по детерминированному алгоритму атрибуции, 1\* – положительное решение по вероятностному алгоритму атрибуции (см. табл. 5).

Таблица 5. Матрица решений

Код	Наименование	$\Omega_1$ (Corneille)	$\Omega_2$ (Quinault)	$\Omega_3$
1M	L'Étourdi	0	1*	0
2M	Le dépit amoureux	1	0	0
3M	Sganarelle	1*	0	0
4M	Dom Garcie de Navarre	0	0	1*
5M	L'École des maris	1	0	0
6M	Les Fâcheux	1	0	0
7M	L'École des femmes	1	0	0
8M	La Princesse d'Élide	0	0	1*
9M	Tartuffe	1	0	0
10M	Le Misanthrope	1*	0	0
11M	Mélicerte	1*	0	0
12M	Pastorale comique	1*	0	0
13M	Les Femmes savantes	1	0	0

В четвертой главе «Датировка стихотворных пьес Мольера» проведен обзор методов датировки литературных произведений, предложен математический метод датировки на основе стилеразличающих параметров и рассмотрен вопрос датировки пьес, приписываемых Мольеру.

Полная датировка произведения заключается в установлении времени написания и публикации произведения. В случае с театром Мольера даты написания пьес точно установить невозможно, поскольку не сохранилось ни одной рукописи Мольера. Время написания комедий, составляющих его театр, определяется исходя из сведений о первых их постановках. Однако у многих исследователей его творчества возникали сомнения по поводу правильности принятого способа датировки. Так, Вольтер считал, что комедия “*Les Précieuses ridicules*” впервые была поставлена еще в провинции, а не 15 ноября 1659 года. И. Вутер указывает на то, что пьеса “*Le dépit amoureux*”, постановка которой произошла годом позже, сильно уступает комедии “*L'Étourdi*” по стилю, что свидетельствует о более позднем написании “*L'Étourdi*”. Известно также, что придворные драматурги старались подготовить часть пьес заранее, чтобы быстро реагировать на желания короля ставить новые пьесы на придворные празднования. Возможно, что и комедии, приписываемые Мольеру, руководившему «труппой короля», были написаны несколько в иной последовательности, чем полагается в настоящее время. Наиболее спорная гипотеза была выдвинута П. Луи, который приписывает П. Корнелю авторство лучших пьес Мольера, и датирует их годами молчания П. Корнеля в 30-е и 50-е годы XVII века.

Итак, существуют различные мнения о датировке пьес, составляющих театр Мольера, анализ которых позволяет выдвинуть гипотезу о том, что стихотворные пьесы, приписываемые Мольеру, были созданы ранее, чем предполагалось.

В настоящее время практически отсутствуют работы, посвященные математическому анализу авторского стиля с целью определения его развития во времени. В разные годы исследования по датировке литературных произведений были проведены В. М. Жирмунским, К. Ф. Тарановским, М. И. Шапиром. Ввиду отсутствия универсальных методик датировки в рамках настоящей работы был разработан новый математический метод датировки на основе стилеразличающих синтаксических параметров, в основе которого лежит гипотеза о существовании тренда параметров индивидуального авторского стиля на протяжении литературного творчества автора. Проверка этой гипотезы предусматривает выполнение следующих процедур: отбор объектов, последовательность создания которых надлежит установить (ниже – датируемый объект), отбор информативных параметров, формирование математических моделей датируемых объектов, ранжирование объектов на основании стандартизированной матрицы данных и сопоставление результатов датировки с начальными данными.

Непосредственно процедура определения датировки литературных произведений осуществляется путем построения стандартизированной матрицы данных  $X = \{x_{ij}\}$ , на основе которой производится ранжирование значений  $j$ -го параметра по каждому объекту. Затем значения рангов суммируются, и по полученной сумме рангов происходит ранжирование датируемых объектов, которое и определяет последовательность их написания. Датировка происходит при сопоставлении полученной последовательности работ автора с первоначальными сведениями о времени их написания и публикации. Результатом может стать новая гипотеза о написании литературных произведений раньше или позже предполагаемого срока.

С помощью метода датировки на основе стилеразличающих синтаксических параметров проведен анализ двух совокупностей пьес, полученных по итогам работы алгоритма распознавания в третьей главе. Одна совокупность сформирована из пьес, принадлежащих П. Корнелю, и из атрибутированных ему произведений из театра Мольера. Другая совокупность состоит из работ Ф. Кино и атрибутированной ему пьесы Мольера "L'Étourdi". Под датировкой в данном случае понимается установление последовательности написания комедий и сопоставление полученных результатов с общепринятой датировкой по годам их первых постановок.

При анализе работ Ф. Кино и атрибутированной ему пьесы из театра Мольера "L'Étourdi" на основе стандартизированной матрицы данных была построена таблица ранжированных значений пяти диагностических параметров по каждому объекту (см. табл. 6). Затем было произведено ранжирование датируемых объектов по ранговым суммам  $\sum$  (см. табл. 7).

Таблица 6. Ранжирование объектов класса  $\Omega_2$ (Quinault) и атрибутируемого объекта 1М (L'Étourdi) по значениям параметров  $x_2, x_4, x_{21}, x_{31}, x_{32}$

Объект	$x_2$	Ранг $x_2$	$x_4$	Ранг $x_4$	$x_{21}$	Ранг $x_{21}$	$x_{31}$	Ранг $x_{31}$	$x_{32}$	Ранг $x_{32}$
1Q	1,41	1	1,50	1	1,47	1	1,43	1	1,35	1
2Q	-0,54	3	-0,60	4	-0,59	3	-0,45	3	-0,37	3
3Q	-0,85	4	-0,47	3	-0,69	4	-0,86	4	-1,02	4
1М	-0,01	2	-0,42	2	-0,19	2	-0,13	2	0,05	2

Таблица 7. Ранжирование объектов класса  $\Omega_2$ (Quinault) и атрибутируемого объекта 1М (L'Étourdi) по значениям ранговых сумм  $\sum$

Объект	Автор	Наименование	Год постановки	$\sum$	Ранг
1Q	Кино	Les Rivaies	1653	5	1
1М	«Мольер»	L'Étourdi	1658	10	2
2Q	Кино	L'Amant Indiscret	1654	16	3
3Q	Кино	La Mère Coquette	1665	19	4

Результаты применения метода датировки при анализе комедий Ф. Кино свидетельствуют, что временное направление совпадает с увеличением ранга. Таким образом, был установлен повышательный тренд параметров, под которым в данной работе понимается изменение тренда параметров индивидуального авторского стиля в сторону увеличения значений параметров. Полученные данные позволяют говорить о возможном написании пьесы “L'Étourdi” драматургом Ф. Кино до 1654 года.

Совокупность пьес, принадлежащих П. Корнелю, и атрибутированных ему произведений из театра Мольера оказалась разделенной на три группы: первую группу составили ранние пьесы и героические комедии П. Корнеля, вторую группу – пьесы, принадлежащие П. Корнелю и приписываемые Мольеру, постановки которых происходили в 1631–1643 годы и 1658–1672 годы соответственно, и в третью группу вошли наиболее слабые пьесы, приписываемые Мольеру.

С одной стороны, подобное смешение дат может быть обусловлено большим разнообразием жанровых подвидов пьес, составляющих данную совокупность, и сильной зависимостью результатов эксперимента от однородности и качества исследуемого материала. С другой стороны, близость пьес П. Корнеля, написанных в 30-е годы, с атрибутированными ему пьесами из театра Мольера может указывать на то, что лучшие стихотворные пьесы, приписываемые Мольеру, были созданы П. Корнелем задолго до их постановки на сцене, а именно в период с 1630 до 1650-х годов. С некоторой долей уверенности можно говорить о подтверждении в ходе исследования гипотезы, выдвинутой П. Луи, о более раннем, чем принято считать, создании ряда комедий, приписываемых Мольеру.

## Заключение

В заключении обобщаются результаты проведенного исследования, намечаются перспективы дальнейших исследований в данной области и делаются следующие **выводы**.

1. Изучение документально-исторических фактов и данных филологического анализа позволило сформировать сложную атрибуционную гипотезу, описывающую возможность написания спорных пьес, приписываемых Мольеру, П. Корнелем, Ф. Кино и неизвестными авторами. Класс атрибутируемых объектов составили 13 комедий в стихах, приписываемых Мольеру.

2. Обзор истории развития научной мысли в области параметризации авторского стиля позволяет выделить следующие основные тенденции: переход от одномерных классификаций к описанию объектов в многомерном признаковом пространстве, все более широкое использование компьютерной обработки данных, а также возникший в последние десятилетия интерес исследователей к применению синтаксического анализа при описании авторского стиля. Эффективный метод стилистического анализа в целях определения авторства подразумевает применение многомерных

классификаций, определение характеристик текста, а не отдельного предложения, и описание текста на разных уровнях языковой системы с учетом как лексического состава текста, так и его структуры. Всем этим требованиям отвечает математический метод атрибуции анонимных и псевдонимных произведений, основанный на теории распознавания образов.

3. В работе представлен опыт применения математических методов атрибуции на основе теории распознавания образов при атрибуции текстов на французском языке. Предложенная методика определения значений параметров на материале текстов XVII века на французском языке позволила решить задачу определения авторства спорных пьес, приписываемых Мольеру. Априорный словарь параметров составили 51 параметр.

Два априорных класса –  $\Omega_1$ (Corneille) и  $\Omega_2$ (Quinault) – составили 11 комедий в стихах П. Корнеля и 3 комедии в стихах Ф. Кино соответственно.

4. Набор информативных параметров составили 5 параметров: X02 (число элементарных предложений), X04 (число сочиненных предложений), X21 (число спрягаемых форм глагола), X31 (число подлежащих), X32 (число местоимений-подлежащих). Проведенная атрибуция подтвердила высокую информативность и стиледифференцирующие возможности параметров синтаксического уровня. На этапе отбора информативных параметров в работе показана возможность автоматизации лингвистического исследования в области установления авторства.

5. В результате работы детерминированного алгоритма была определена авторская принадлежность шести из тринадцати анализируемых объектов: пьесы “Le dépit amoureux”, “L'École des maris”, “Les Fâcheux”, “L'École des femmes”, “Tartuffe”, “Les Femmes savantes” были атрибутированы П. Корнелю с вероятностью больше 0,95. В результате работы вероятностного алгоритма распознавания Ф. Кино была атрибутирована пьеса “L'Étourdi”, а П. Корнелю были приписаны пьесы “Sganarelle”, “Le Misanthrope”, “Mélicerte”, “Pastorale comique” с разной степенью вероятности (от 0,63 до 0,73). Результаты вероятностного алгоритма были откорректированы в процессе процедуры оценки качества классификации. Две пьесы, “Dom Garcie de Navarre” и “La Princesse d'Élide”, составили апостериорный класс. Итак, получил свое подтверждение вариант альтернативной гипотезы ( $H_a^2$ ): тексты пьес, приписываемые Мольеру, являются произведениями П. Корнеля, Ф. Кино и одного неизвестного автора.

6. Произведенный обзор существующих методов датировки текстов показал отсутствие универсальных математических методов установления последовательности написания текстов. Применение предложенного метода датировки на основе стилеразличающих синтаксических параметров показало целесообразность такого подхода, а также позволило установить возможность математической оценки тренда параметров индивидуального авторского стиля на протяжении литературного творчества писателя.

7. В результате датировки пьес, атрибутированных П. Корнелю и Ф. Кино, была подтверждена гипотеза о более раннем написании стихотворных пьес, входящих в театр Мольера.

Перспективы дальнейшего исследования состоят в разработке проблем описания индивидуальных авторских стилей и датировки литературных произведений с использованием предложенного в данной работе математического метода датировки.

**Основные положения диссертационной работы нашли отражение в следующих публикациях.**

1. Марусенко М. А., Мельникова Е. Е., Родионова Е. С. Атрибуция анонимных и псевдонимных статей, опубликованных в журналах «Время» и «Эпоха» в 1861 – 1865 годах // Квантитативная лингвистика: исследования и модели (КЛИМ-2005) : материалы Всероссийской научной конференции (6 – 10 июня 2005 г.) / редкол.: А. А. Поликарпов, Г. Г. Сильницкий, В. В. Поддубный. – Новосибирск : Новосибирский гос. пед. ун-т, 2005. – С. 283–294 (0,45 п. л., авт. текст – 0,2 п. л.).

2. Марусенко М. А., Мельникова Е. Е., Родионова Е. С. Об авторстве статей, приписываемых Ф. М. Достоевскому // Актуальные проблемы теоретической и прикладной лингвистики и оптимизация преподавания иностранных языков : материалы Всероссийской научной конференции с международным участием (11 – 13 октября 2005 г.) / сост.: Ю. И. Горбунов, Г. В. Артамонова, Т. Н. Андреюшкина. – Тольятти : Тольяттинский гос. ун-т, 2005. – С. 23–29 (0,3 п. л., авт. текст – 0,15 п. л.).

3. Родионова Е.С. Проблема авторства драматургических произведений Мольера // Материалы XXXV Международной филологической конференции (13 – 18 марта 2006 г.). – СПб : Филол. фак. С.-Петерб. гос. ун-та, 2006. – Вып. 21 : Секция математической лингвистики / отв. ред. А. С. Герд, Т. Г. Скребцова : в 2 ч. – Ч. 2. – С. 11–16 (0,25 п. л.).

**4. Родионова Е.С. Параметризация стилей: отбор информативных параметров при атрибуции пьес Мольера // Вестник Санкт-Петербургского университета. Сер. 9. Филология. Востоковедение. Журналистика. – Вып. 2. – Ч. 2. – 2007. – С. 61–67 (0,5 п. л.).**

5. Родионова Е. С. Отбор информативных параметров при атрибуции стихотворных пьес Мольера // Материалы XXXVI Международной филологической конференции (12 – 17 марта 2007 г.). – СПб : Филол. фак. С.-Петерб. гос. ун-та, 2007. – Вып. 10 : Прикладная и математическая лингвистика / под ред. Т. Г. Скребцовой. С. 67–74 (0,3 п. л.).

6. Родионова Е.С. Атрибуция стихотворных пьес Мольера методами распознавания образов // MegaLing'2007. Горизонты прикладной лингвистики и лингвистических технологий : доклады международной научной конференции, 24 – 28 сентября 2007 г., Украина, Крым, Партенит = Horizons of Applied Linguistics And Linguistic Technologies : proceedings of international conference, 24 – 28 September 2007, Ukraine, Crimea, Parthenit /

ред. В. А. Широков, С. С. Дикарева. – Симферопль : ДиАйПи, 2007. – С. 313–314 (0,15 п. л.).

**7. Родионова Е. С. Проблема «Корнель – Мольер»: к вопросу об авторстве и датировке стихотворных пьес, приписываемых Мольеру // Известия Российского государственного педагогического университета им. А. И. Герцена. – 2008. – № 27 (61) [ апрель ]. – С. 216 – 220 (0,45 п. л.).**