

Санкт-Петербургский государственный университет

На правах рукописи

Чепига Валентина Петровна

СРАВНИТЕЛЬНО-СТИЛИСТИЧЕСКИЙ АНАЛИЗ
ПРОИЗВЕДЕНИЙ РОМЕНА ГАРИ И ЭМИЛЯ АЖАРА

Специальность 10.02.05 – романские языки

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата филологических наук

Санкт-Петербург 2008

Диссертация выполнена на кафедре романской филологии факультета филологии и искусств Санкт-Петербургского государственного университета и в учебно-научном подразделении «Наука о языке» Университета Новая Сорбонна Париж-3

Научные руководители: доктор филологических наук,
профессор Марусенко
Михаил Александрович

доктор филологических наук,
профессор Феноглио Ирэн

Официальные оппоненты: Заслуженный деятель науки РФ, доктор
филологических наук, профессор
Пиотровский Раймонд Генрихович

Кандидат филологических наук,
Гребенников Александр Олегович

Ведущая организация: Московский государственный университет
им. М.В. Ломоносова

Защита состоится «__» _____ 2008 года в «___» часов на заседании диссертационного совета Д.212.232.48 по защите докторских и кандидатских диссертаций при Санкт-Петербургском государственном университете по адресу: 199034, г. Санкт-Петербург, Университетская наб., 11, ауд. ____.

С диссертацией можно ознакомиться в Научной библиотеке имени М. Горького Санкт-Петербургского государственного университета по адресу: 199034, г. Санкт-Петербург, Университетская наб., 7/9.

Автореферат разослан «___» _____ 2008 г.

Ученый секретарь

диссертационного совета

С.Т. Нефедов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Ромен Гари (1914–1980), один из крупных французских писателей XX в., привлекает внимание как широкого круга читателей, так и исследователей, прежде всего своей экспериментальной деятельностью в области художественной литературы. Интерес для исследователей представляет разносторонность творчества Р. Гари: он был прозаиком, работал журналистом, занимался написанием пьес, сценариев, вел политическую деятельность, представляя интересы Франции за рубежом.

Многогранностью творчества Р. Гари обуславливается разнообразие работ, посвященных этому писателю, а также широкая тематика научных исследований, касающихся его литературного наследия. Многие работы в западном литературоведении затрагивают проблемы стиля писателя (Lecarme-Tabone, Piton), изучены его дипломатическая (Larat, Sacotte) и общественная деятельность (Rosse).

Первым крупным событием для разрозненных исследователей творчества Р. Гари стал международный симпозиум, посвященный жизни и произведениям автора, организованный по инициативе ассоциации «Les Mille Gary» в 2002 г., в задачи которой входит повышение интереса к творчеству Р. Гари; из последних работ ассоциации можно назвать выпуск в 2007 г. неизданной главы «Le Socialisme» из произведения Э. Ажара «Gros-Câlin».

В отечественном литературоведении и лингвистике к творчеству Р. Гари начали обращаться с недавнего времени, и в российской современной критике наличествуют лишь работы ознакомительного характера, в основном краткие предисловия к русским переводам его книг, среди которых не представляется возможным выделить какую-либо глубокую работу аналитического порядка.

Наибольший интерес вызывает проблема авторства четырех произведений, написанных под псевдонимом Эмиль Ажар (1974–1980). Стили двух авторов настолько отличны друг от друга, что при жизни писателя у его

современников и критиков не возникало сомнений в том, что Р. Гари и Э. Ажар являлись двумя разными людьми. После выхода в свет романа П. Павловича «L'Homme que l'on croyait» (1981 г.), а также посмертной публикации Р. Гари «Vie et mort d'Émile Ajar» (1981 г.) литературный мир потряс тот факт, что Р. Гари и Э. Ажар могут быть одним и тем же человеком. Таким образом, писатель оказался единственным автором, дважды получившим Гонкуровскую премию. В западном литературоведении появился ряд работ по истории псевдонимов писателя (Obergoeker), исследованию общности мотивов, характерных для романов Р. Гари и Э. Ажара (Lalande, Spivey), и стилю произведений Э. Ажара (Bellos, Tournier, Abdeljaouad). Следует выделить 2 биографических исследования, посвященных Р. Гари (Anissimov, Bona) и три романа, героем которых является писатель (Blach, Huston, Pavlowitch). Тем не менее, работы по атрибуции псевдонимных романов Э. Ажара сводятся к двум публикациям:

1) диссертационному труду 1998 г., проведенному польским исследователем А. Павловски, который использовал модель временных рядов. Результатом работы А. Павловски стала атрибуция четырех произведений Э. Ажара Ромену Гари. При этом исследователь не учитывал жанровое несоответствие романов Э. Ажара и привлек в качестве потенциальных авторов ряд современников Р. Гари, таких как Р. Кено, Л. Арагон и М. Турнье, основываясь только на ряде газетных заметок, появившихся до установления того факта, что роль Э. Ажара исполнял П. Павлович, двоюродный племянник Р. Гари;

2) труду 2007 г. французского лингвиста Д. Лаббе, который использовал метод «межтекстового расстояния» применительно к пьесам П. Корнеля и Мольера. В приложении исследователь продемонстрировал новый метод на четырех произведениях Э. Ажара, сравнивая их с романами Р. Гари. При этом неоднозначность полученных результатов привела к тому, что некоторые романы самого Э. Ажара оказались близки к выделенному Д. Лаббе разряду, в

который попадают произведения, имеющие статистически значимые различия, на основе чего делается вывод о том, что либо у текстов разные авторы, либо один автор написал тексты, относящиеся к разным жанрам.

Настоящая диссертация представляет собой первое отечественное исследование, посвященное сравнительному анализу стилей двух писателей, а также установлению авторства псевдонимных произведений Э. Ажара методом распознавания образов, что обуславливает **актуальность** работы. Диссертация является совместным проектом (Санкт-Петербургский государственный университет, университет Новая Сорбонна – Париж-3, Французский Национальный центр научных исследований CNRS), что подчеркивает ее актуальность. Впервые данный метод применяется на материале современного французского языка, что обуславливает **научную новизну** работы. Исследование автографов Р. Гари осуществлялось в сотрудничестве с IMEC (Institut Mémoires d'édition contemporaine, г. Кан), в котором хранится рукописное наследие писателя.

Цель настоящей работы заключается в комплексном исследовании индивидуальных авторских стилей, текстологическом анализе рукописей писателей и установлении авторства псевдонимных произведений Э. Ажара. Достижение поставленной цели предполагает решение следующих **задач**:

- изучение генезиса текстов произведений Р. Гари и Э. Ажара;
- привлечение архивных документов – рукописных черновых и беловых редакций произведений Р. Гари и Э. Ажара;
- исследование методов атрибуции литературных произведений;
- выделение языковых параметров, характерных для французского языка;
- проведение атрибуции псевдонимных произведений Э. Ажара методом распознавания образов.

Из всего комплекса проблем, связанных с литературным творчеством Р. Гари, **объектом исследования** явилось изучение характеристик индивидуальных авторских стилей Р. Гари и Э. Ажара. **Предметом**

исследования стала атрибуция псевдонимных произведений и генезис рукописных текстов.

Материалом для исследования послужили три произведения Р. Гари: «Au-delà de cette limite votre ticket n'est plus valable», 1975 г., «Clair de femme», 1977 г., «Les Cerfs-Volants», 1980 г.; три произведения Э. Ажара: «Gros-Câlin», 1974 г., «La Vie devant soi», 1975 г., «L'Angoisse du roi Salomon», 1979 г.; два произведения П. Павловича: «Un autre monde», 2004 г., «Tom», 2005 г.

Для решения поставленной задачи атрибуции оказалось необходимым сочетать **принципы** филологического анализа, а именно: документально-рукописный (археологический), исторический (биографии, сведения современников), идейно-стилистический и методы теории распознавания образов. При решении конкретной задачи атрибуции псевдонимных произведений **методологической базой** диссертации послужила теория распознавания образов.

Основные положения, выносимые на защиту. В результате проведенного исследования установлено, что индивидуальные авторские стили возможных авторов произведений, написанных под псевдонимом Э. Ажар, существенно различаются по определенным синтаксическим характеристикам. Изучение генезиса рукописных текстов, формирующих корпусы исследуемых произведений, а также биографические и документальные источники позволили сформулировать атрибуционную гипотезу о том, что авторами трех романов, написанных под псевдонимом Э. Ажар («Gros-Câlin», «La Vie devant soi», «L'Angoisse du roi Salomon»), могли быть Р. Гари и/или П. Павлович. Проверка атрибуционной гипотезы позволила исключить авторство П. Павловича и установить, что вероятным автором романов является Р. Гари.

Теоретическая значимость исследования. Впервые на материале современного французского языка была разработана методика параметризации текста.

Применение существующей методики атрибуции анонимных и псевдонимных текстов к французскому материалу позволяет рассматривать ее как универсальную, не зависящую от языковых и типологических характеристик изучаемых языков.

Результаты исследования показывают, что латентные характеристики индивидуального авторского стиля обладают значительной статистической устойчивостью и, даже при целенаправленных усилиях писателя скрыть свое авторство, позволяют с достаточной вероятностью установить истинного автора псевдонимного или анонимного текста.

Практическая значимость исследования. Научно-практическое значение диссертации определяется возможностью использовать ее результаты в ходе дальнейшего изучения творчества Р. Гари, а также в общих и специальных лекционных курсах по текстологии для студентов по применению математических методов в лингвистике.

Работа по систематизации и классификации архивов позволила дать доступ к новым автографам для исследователей рукописного наследия Р. Гари.

Апробация работы. Основные положения диссертации и ее отдельные аспекты обсуждались на научных конференциях и семинарах:

- XXXIII международная филологическая конференция (Санкт-Петербург, 2004);
- международная междисциплинарная конференция (Пуатье, 2005);
- XXXV международная филологическая конференция (Санкт-Петербург, 2006);
- XXXVI международная филологическая конференция (Санкт-Петербург, 2007);
- аспирантский семинар Института современных текстов и рукописей ITEM (Institut des Textes et Manuscrits modernes) «Manuscrits et linguistique» (Париж, 2007);

- международный симпозиум «Мультилингвизм и генезис текста: лингвокультурный аспект» (Москва, ИМЛИ, 2007).

Принципиальные положения диссертации изложены в 9 публикациях.

Цели и задачи, поставленные в диссертации, определили **структуру** работы. Диссертация состоит из введения, четырех глав, заключения, списка цитируемой литературы (161 наименование), списка используемых материалов и одного приложения «Параметризация текста на французском языке». Основное содержание работы изложено на 138 страницах. Общий объем исследования 201 страница.

ОСНОВНЫЕ ПОЛОЖЕНИЯ ДИССЕРТАЦИИ

Во **введении** аргументируется правомерность постановки проблемы, ее актуальность и научная новизна, раскрываются цели и задачи исследования, излагаются общие и частные методологические принципы работы с исследуемым материалом, определяется структура работы.

Первая глава «История вопроса. Основные проблемы» посвящена анализу биографических источников, способствовавших поискам литературных псевдонимов Р. Гари. Параллельно представлен краткий обзор зарубежных трудов, посвященных творчеству писателя (см. выше), выполненный в свете изучаемой проблематики.

Во **второй главе** «Сравнительный анализ черновых и беловых редакций романов Р. Гари и Э. Ажара» проводится сопоставительный анализ различных приемов работы над рукописями двух авторов (1974–1980). Рукописное наследие писателя находится в литературных архивах ИМЕС (Institut Mémoires d'édition contemporaine) в г. Кан (Франция).

Используются все известные источники анализируемых текстов (печатные и рукописные), расшифрованы и прочтены черновики, приведены все редакции, проанализированы и систематизированы их варианты.

Возможность рассмотрения всего рукописного корпуса одного произведения писателя как «единого текста» была предложена Д. С. Лихачевым, поэтому становится возможным наблюдать творческую манеру работы писателя в ее движении, не разбивая корпус произведения на отдельные тексты. Особое внимание уделено носителям, что позволило реализовать четкую классификацию разных типов документов, выявить различные этапы редакций и осветить процесс работы над рукописью.

Результатом систематизации архивных документов, составляющих корпусы произведений Р. Гари и Э. Ажара, стал сравнительный анализ приемов работы двух «авторов» и сделаны следующие замечания:

1. Все романы Э. Ажара написаны в форме сказа, этим они отличаются от романов Р. Гари. В силу этого образ рассказчика колеблется, иногда расширяясь до пределов «образа автора». Динамика форм данного соотношения лежит в основе постоянного изменения функций основных словесных сфер сказа, делает их колеблющимися, семантически многоплановыми. Рассказчик и автор перекрывают и сменяют один другого, вступая в разные отношения с образами персонажей. Так, анализ автографов Э. Ажара показывает тенденцию писателя к колебаниям между собственно авторской речью и речью «рассказчика». От редакции к редакции писатель вносит корректировку в речь «рассказчика», приближая ее к речи выбранного им на эту роль героя. В общем, надо отметить, что главные персонажи романов Э. Ажара принадлежат к социальным низам: таковы Момо в «*La Vie devant soi*», Жанно в «*L'Angoisse du roi Salomon*». Это приводит к тому, что писатель зачастую пользуется элементами разговорной речи.

2. Для обоих авторов характерно «прогрессирующее письмо» (в отличие от «программированного письма» (Науе), которое характеризуется наличием определенных этапов – набросков, разбивок на эпизоды и т.п.), то есть для них не характерна работа по сбору подготовительных документов к написанию романа, как-то: планов, путевых заметок, рисунков и т.д. В каждой новой

версии романов авторы опираются только на предыдущие отредактированные версии.

Оба автора, не имея ни плана, ни предварительных записей, пишут, не останавливаясь (об этом мы можем судить по наличию периодической пагинации), целые страницы. Такая же манера работы была характерна, например, для Л.-Ф. Селина. В первых черновых редакциях романов Р. Гари и Э. Ажара четко прослеживается тенденция к сокращению слов, к оставлению недописанных слов, как будто писатель торопится поставить точку, закончить начатый абзац. Например, слово *monsieur* у обоих авторов превращается в **msie*, в инфинитиве «проглатываются» окончания: **parti*, **mange*, восстановленные путем контекстного анализа.

Есть черновые наброски, но они уместаются преимущественно на нескольких первых (и иногда последних) страницах тетрадей, далее начинается связный текст.

3. Корпус текстов состоит из следующих элементов: черновики – черных бухгалтерских тетрадях (биографы писателя упоминают о том, что такая манера работы характерна только для Э. Ажара); нескольких рукописных версий; машинописной версии с исправлениями, сделанными от руки.

Одинаково оформлены тексты черновой тетрадной версии – нет деления на главы, письмо быстрое, исправления последовательные, вносимые одновременно с написанием непосредственно текста произведения. Нет отступа для полей, верхняя строка остается незаполненной. Нередко использование тетрадях для написания самого первого чернового текста – такая манера работы характерна, например, для М. А. Булгакова, М. И. Цветаевой.

Отсутствует четкая пагинация первых черновых редакций текстов, что хорошо разграничивает два типа связности текста – «локальную» и «глобальную» (Ковтунова). Так, периодическая пагинация отражает локальную связность текста, то есть связность последовательностей в тексте, в первую очередь предложений и далее – более крупных единиц. Периодическая

пагинация отражает «рабочий сеанс» писателя, т.е. количество написанных листов за определенный промежуток времени (Михайлов). Для Р. Гари и Э. Ажара количественное выражение рабочего сеанса находится в интервале [10–20]. Последовательная пагинация характерна на заключительных этапах работы над текстом – в последних черновых и беловых редакциях текстов, что отражает принцип глобальной связности текста, то есть обеспечения единства текста как целого.

4. Авторы используют в основном перьевую ручку черного цвета. Для исправлений используются фломастеры черного и/или красного цветов, ручка ярко-синего цвета, простой карандаш.

5. Авторы используют фотокопии одной версии (как рукописной, так и машинописной) в последующей версии, благодаря чему становится возможной параллельная работа над несколькими версиями.

6. Авторы используют фотокопии машинописных версий для множественных исправлений одного и того же фрагмента романа.

Эта видимая общность работы интересна с точки зрения дальнейшего сравнительного анализа черновых и беловых рукописей писателей. По-видимому, своего рода «гомогенность» (Pickering) носителя для письма, а также определенной марки «Registres Le Dauphin» характерна для писателей на подсознательном уровне.

На примерах двух романов – «Clair de femme» и «L'Angoisse du roi Salomon» – нами была продемонстрирована работа Р. Гари и Э. Ажара над словом. Мы использовали компьютерную программу EDITE, позволяющую провести сравнительный анализ разных редакций того или иного произведения.

Неоднократная правка текста у Р. Гари материально выражается в виде различных графических наложений. Так, первый слой правки представляет собой правку от руки, повторяющую расположение печатных строк, соблюдается равномерный пробел. Следующий слой правки – это записи на полях, а также надстрочные и междустрочные вставки, пометы поперек

страницы, сопровождаемые цифрами и астерисками отсылок, нарушающие ее «геометрию» (Бюстарре). Далее идут пометы фломастером черного цвета, что еще более затрудняет восприятие текста: помимо единовременных исправлений – перечеркиваний слова, вводятся перечеркивания слева направо всей страницы.

В черновиках Р. Гари мы можем условно выделить три временных слоя исправлений: первый сделан, видимо, во время написания текста – это добавления смыслового и стилистического характера, а также единовременные написания нескольких вариантов первой и последней главы; второй и третий слои – это авторская правка уже после написания рукописи, в основном эти два слоя характеризуются зачеркиваниями и сокращениями.

Первая черновая редакция романов, представленная тетрадами, является очень характерной: так, в ней мы находим только «сиюминутную правку», то есть зачеркивание и написание нового варианта сверху исправленного слова, либо на полях.

Выбор носителя тесно связан с техникой письма. Так, Р. Гари во время доработки текста наклеивал полоски бумаги и листы формата А4 в местах правки.

При проведенном анализе становится очевидным, что Р. Гари много работал над версиями романа, что не было явным при поверхностном исследовании версий и заставляло предполагать, что писатель небрежно обращается с текстом. Для Р. Гари практически не характерны большие текстовые замены, но характерны текстовые перестановки.

Переработка текстов Э. Ажара сводится к стилистической отделке – к словесной яркости и насыщенности, которые были так высоко оценены современниками писателя. Автор отбрасывает вялые обороты первой редакции, добиваясь эффекта живой разговорной речи. Напротив, для Р. Гари не характерны просторечные выражения.

Структура сложного синтаксического целого у Э. Ажара расчленена на ряды простых предложений. Для Р. Гари свойственен стиль, вписывающийся в норму литературного языка и характеризующийся наличием книжных оборотов (например, герундием).

В общем, работа над текстом Э. Ажара является не очень насыщенной исправлениями – основные замены автор вносит уже на этапе переписывания первой версии, в дальнейшем работа ведется над стилистикой, оформлением и подчисткой текста.

В результате проведенного анализа корпуса рукописей двух авторов, сделан вывод о существовании видимых отличий работы над текстом Р. Гари и Э. Ажара, при общности инструментов литературной мастерской – носителей и выбранного цвета письма, отсутствия предварительной работы по сбору материала (было обнаружено не более десятка газетных вырезок и заметок – в том числе, для романа «Au-delà de cette limite votre ticket n'est plus valable» Р. Гари).

На основе изученных материалов сформированы гипотезы по атрибуции псевдонимных произведений. При этом нулевая гипотеза сформулирована следующим образом: (H_0) – все тексты романов Э. Ажара полностью принадлежат Р. Гари. Альтернативные гипотезы сформулированы так: (H_a^1) – тексты романов Э. Ажара являются произведением П. Павловича и (H_a^2) – тексты романов Э. Ажара являются совместным произведением Р. Гари и П. Павловича с определенной долей участия каждого из них. Вопрос об атрибуции того или иного романа Э. Ажара решается в каждом случае индивидуально.

Для практической реализации процедуры атрибуции сформированы априорные классы: Ω (Gary) и Ω (Pavlowitch).

В **третьей главе** «Методы атрибуции» делается обзор различных методов атрибуции анонимных и псевдонимных произведений.

Подробно рассматриваются методы лексического и синтаксического анализа при исследовании стилистики текстов, метод автоматической обработки данных, метод «межтекстового расстояния», предложенный французским исследователем Д. Лаббе, и метод распознавания образов.

Общий вывод заключается в том, что применение математических моделей при лингвостилистическом анализе делает метод атрибуции анонимных и псевдонимных произведений на основе теории распознавания образов наиболее приемлемым из объективных методов атрибуции.

Индивидуальный стиль писателя наиболее явно проявляется в синтаксисе: в отборе, расположении и объединении в сложное целое синтаксических конструкций. Так, за широким диапазоном речевых перевоплощений творческой индивидуальности всегда можно увидеть структурообразующий стержень творчества писателя (Гиндин). Таким образом, в данной работе, исходя из требований объективности и полноты описания авторского стиля, в основу исследования по определению авторства романов Э. Ажара была положена методика атрибуции анонимных и псевдонимных произведений, разработанная М. А. Марусенко, которая предусматривает применение многомерной классификации, основанной на теории распознавания образов, и описания индивидуального авторского стиля в синтаксическом аспекте.

В **четвертой главе** «Атрибуция романов Э. Ажара» осуществляется проверка атрибуционной гипотезы (см. выше) с использованием математического метода атрибуции анонимных и псевдонимных произведений. Состав априорного алфавита классов и формирование алфавита атрибутируемых объектов определен в соответствии с требованием жанровой, стилевой и хронологической (1974–1980) однородности исследуемых текстов. Из работ Р. Гари, П. Павловича и Э. Ажара были отобраны произведения, представляющие собой романы, написанные от первого лица. Проведена процедура разбиения речевого материала на авторскую и прямую речь, при

которой именно авторская речь явилась дальнейшим предметом изучения. Параметрическое пространство составили 54 параметра из априорного словаря параметров, и следующий этап исследования заключался в отборе из полученного априорного словаря небольшого числа информативных параметров. 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 в 200 предложений для каждого априорного класса. Результаты эксперимента были представлены в виде объектно-признаковых матриц данных размерностью $n \times N = 200 \times 54$, где n – число параметров, а N – число объектов. Общее число элементов матриц данных составило по 10800.

Ввиду большого объема выборки было использовано обратное построение матриц данных, при котором набор значений параметров представляется соответствующей строкой матрицы, а значения всех параметров на объектах (предложениях) – соответствующим столбцом. Затем были вычислены статистические характеристики анализируемых объектов: среднее арифметическое (\bar{x}_i) и стандартное отклонение (σ_i) для каждого класса. На первом этапе произошло разбиение априорного набора информативных параметров на два подмножества параметров, релевантных и не релевантных для различения априорных классов. Для этого была произведена автоматическая классификация параметров в одномерном пространстве. Релевантность параметров для различения априорных классов была определена по t -критерию Стьюдента, пороговое значение которого при уровне значимости $\alpha = 0,05$ равно 1,96. Процесс классификации заключался в сравнении значений критерия с пороговым значением. Если наблюдаемое значение критерия больше порогового, то параметр относился к числу информативных, в противном случае он исключался из дальнейшего рассмотрения.

Статистически значимой оказалась разность средних для параметров X02, X03, X06, X07, X08, X11, X19, X20, X21, X23, X24, X25, X28, X29, X30, X31 (наблюдаемые значения t -критерия больше критического). Все остальные параметры оказались нерелевантными для различения априорных классов.

В дальнейшем подмножество диагностических параметров формировалось на основе группы из шестнадцати параметров, релевантных для разделения известных классов.

На основе объектно-признаковой матрицы была сформирована корреляционная матрица связей параметров, элементами которой являются выборочные коэффициенты корреляции. Полученная матрица данных имеет размерность 54x54. Содержательный критерий информативности набора параметров заключается в слабой корреляции информативных параметров между собой и сильной их корреляции с остальными параметрами, не вошедшими в эту группу. На основе корреляционной матрицы были определены средняя внутригрупповая корреляция и средняя внегрупповая корреляция каждого параметра. Затем были вычислены критерии эффективности каждого параметра.

Подмножество из шестнадцати параметров разбилось на подмножество из трех параметров (X23, X24, X28), значения для эффективности которых лежат в интервале [0,509–0,951], и на подмножество из тринадцати параметров, значения для эффективности которых лежат в интервале [0,318–0,467].

Таким образом, из подмножества шестнадцати параметров удалось выделить подмножество из трех параметров, удовлетворяющих сформулированным в диссертации требованиям сильной корреляции с другими параметрами совокупности и способности к разделению объектов на классы. В дальнейшем в эксперименте по классификации текстов каждый объект (роман) характеризовался набором из трех диагностических (информативных) параметров, а классификация объектов производилась в 3-мерном пространстве, осями которого явились данные параметры. При определении

координат распознаваемых объектов и эталонов априорных классов были применены методы выборочного обследования текста.

Необходимые объемы выборок для определения координат атрибутируемых объектов и эталонов классов определялись по наибольшим значениям параметров. В настоящем эксперименте они соответствуют значениям параметра X24 (кроме романа «Gros-Câlin», в котором наибольшее значение имеет параметр X23) – число союзов. Были сформированы матрицы данных атрибутируемых объектов и априорных классов.

Применяемый алгоритм распознавания включил в себя детерминированную и вероятностную атрибуции. Детерминированный алгоритм выявил существенные статистические различия между атрибутируемыми объектами и обоими априорными классами, в результате чего три романа оказались не отнесенными ни к одному априорному классу.

Реализация вероятностного алгоритма распознавания предусматривает преобразование исходной матрицы данных в матрицу евклидовых расстояний между объектами, подлежащими распознаванию, и эталонами априорных классов.

При наличии двух априорных классов решающее правило сформулировано следующим образом: $X_i \in \Omega_2, P(X_i \in \Omega_2) > \sum_{n=1}^2 P(X_i \in \Omega_n)$. Таким образом, $X_i \in \Omega_j, P(X_i \in \Omega_j) > 0,5$, то есть при данном решающем правиле пороговое значение вероятности принадлежности объекта к одному из априорных классов должно превышать 0,5. По результатам вероятностной атрибуции классу Ω (Gary) смогли быть атрибутированы все три объекта: «Gros-Câlin», «La Vie devant soi» и «L'Angoisse du roi Salomon».

Заключительным этапом процедуры распознавания стала оценка качества полученной классификации, могущая иногда повлечь за собой корректировку полученных классов. Оценка качества классификации основана на выявлении закономерностей для каждого класса, которые могут быть определены как

подмножества из заданного множества объектов, значения параметров которых одинаковы для большинства объектов анализируемого класса и отличаются от значений параметров других классов. Правильность и правомочность полученных результатов после проведения вероятностной атрибуции были проверены, исходя из оценки однородности априорного класса до и после добавления к нему атрибутируемых объектов. Критерием однородности послужило вычисление среднего квадрата расстояния между объектами $\bar{d}^2(\Omega_N)$. Была построена матрица данных, объединившая исходные данные трех объектов, составляющих класс Ω (Gary), и всех трех атрибутируемых этому классу объектов. В результате эксперимента, поскольку среднеквадратичный разброс $\bar{d}^2(\Omega_{Gary}) > \bar{d}^2(\Omega_{Gary, Ajar})$, сделан вывод о том, что полученная классификация объектов улучшилась и приблизилась к естественной, что подтвердило результаты работы вероятностного алгоритма и принадлежность априорному классу трех атрибутируемых объектов.

В **приложении 1** «Параметризация текста на французском языке» предложены правила определения значений следующих 54 параметров, образующих в совокупности исходное описание изображения объекта: X01 – число слов в простом самостоятельном предложении, X02 – число элементарных предложений в цельном предложении, X03 – число главных предложений, X04 – число сочиненных предложений, X05 – число сочиненных предложений без спрягаемой формы глагола, X06 – число подчиненных предложений, X07 – число подчиненных предложений 1-й степени, X08 – число подчиненных предложений 2-й степени, X09 – число подчиненных предложений 3-й степени, X10 – число подчиненных предложений 4-й и высших степеней, X11 – число элементарных предложений без номинативного подлежащего, X12 – число подчиненных предложений без спрягаемой формы глагола, X13 – число вставных предложений, X14 – число охватывающих предложений, X15 – число слов 1-й группы (знаменательных), X16 – число слов 2-й группы (служебных), X17 – число существительных, X18 – число

прилагательных, X19 – число местоимений, X20 – число спрягаемых форм глагола, X21 – число именных форм глагола, X22 – число наречий, X23 – число предлогов, X24 – число союзов, X25 – число подчинительных союзов, X26 – число сочинительных союзов, X27 – число предикативов, X28 – число прямых дополнений, X29 – число косвенных дополнений, X30 – число подлежащих, X31 – число местоимений-подлежащих, X32 – число групп однородных членов, X33 – число членов однородных групп, X34 – число однородных сказуемых, X35 – число однородных групп дополнений, X36 – число герундиальных оборотов, X37 – число членов герундиальных оборотов, X38 – число распространенных причастных определений, X39 – число членов распространенных причастных определений, X40 – число согласованных определений, X41 – число причастий - согласованных определений, X42 – число несогласованных определений, X43 – число существительных - несогласованных определений, X44 – число обособленных членов, X45 – число членов в группах обособленных членов, X46 – число абсолютных оборотов, X47 – число членов в группах абсолютных оборотов, X48 – число инфинитивных оборотов, X49 – число членов инфинитивных оборотов, X50 – число существительных без группы, X51 – число групп существительных, X52 – число членов групп существительных, X53 – число знаменательных слов в группах существительных, X54 – число служебных слов в группах существительных.

Список параметров, входящих в априорный набор, получен в результате обобщения и унификации параметров, предложенных разными исследователями в процессе изучения текстов методами статистико-комбинаторной лингвистики (Марусенко). Выделенные параметры имеют универсальный характер, но, применительно к каждому конкретному языку, параметризация (определение значений параметров) содержит свои особенности; так, не все параметры могут оказаться релевантными для какого-либо языка (например, для русского языка параметры 46 и 47 (число

абсолютных оборотов и число членов в группах абсолютных оборотов)) являются нерелевантными.

Для формализации процедур определения значений параметров – с целью минимизации случайной ошибки – принимались единообразные решения в отношении параметров, относящихся к спорным вопросам грамматики французского языка.

Заключение.

В диссертации проведено исследование характеристик индивидуальных авторских стилей с целью атрибуции трех псевдонимных романов Э. Ажара методами теории распознавания образов, которая подразумевает отнесение псевдонимного текста, характеризующегося набором числовых параметров, к определенному классу совокупности произведений одного автора.

В диссертации показано, что наряду с объективными методами атрибуции, такими как документально-рукописные, исторические и идейно-стилистические, необходимо установление наличия или отсутствия статистически значимых различий между текстами атрибутируемых произведений и корпусами текстов предполагаемых авторов.

Произведен обзор творческого пути Р. Гари, в результате которого показана склонность писателя к использованию различных псевдонимов. Подробно рассмотрены псевдонимные произведения Э. Ажара. Учтены точки зрения современников Р. Гари на возможное авторство произведений Э. Ажара. На основе объективных исторических данных (биографии, сведения современников и т.п.) показана несостоятельность выдвижения гипотезы о следующих возможных авторах спорных текстов: Л. Арагона, Ж. Ланцмана, Р. Кено, М. Курно и М. Турнье. В работе по атрибуции использованы тексты П. Павловича, двоюродного племянника Р. Гари, поскольку только он мог участвовать в написании романов под псевдонимом «Ажар».

К исследованию привлечены черновые рукописи Р. Гари и Э. Ажара, в результате исследования которых показаны различия в работе двух авторов.

По результатам биографических данных и сравнительного исследования рукописей Р. Гари и Э. Ажара сформированы атрибуционные гипотезы. Для определения координат эталонов априорных классов сформирован корпус текстов, отвечающих требованиям синхронии и жанрово-стилевой однородности.

Выделены 54 параметра, образующие в совокупности исходное описание изображения объекта. Проведена формализация определения значений параметров. Каждый из выделенных 54 параметров проиллюстрирован примерами из исследуемых текстов. Рабочий алфавит параметров распознающей системы по атрибуции романов «Gros-Câlin», «La Vie devant soi» и «L'Angoisse du roi Salomon» включил 3 диагностирующих параметра: X23 – число предлогов, X24 – число союзов, X28 – число прямых дополнений.

Атрибуция текстов распознающим автоматом была произведена в два этапа: при помощи детерминированного алгоритма, затем, после остановки последнего, при помощи вероятностного алгоритма распознавания. Детерминированный алгоритм выявил существенные статистические различия между атрибутируемыми объектами и обоими априорными классами, в результате чего три романа оказались не отнесенными ни к одному априорному классу. Работа вероятностного алгоритма позволила отнести к априорному классу Ω (Gary) все три атрибутируемых объекта.

В результате проведения оценки качества классификации и выделения апостериорного класса, сделан вывод о том, что полученная классификация объектов улучшилась и близка к естественной, что подтвердило результаты работы вероятностного алгоритма и принадлежность априорному классу трех атрибутируемых объектов.

Тот факт, что все результаты получены путем вероятностной атрибуции, может быть интерпретирован следующим образом: Р. Гари, скрывавший свое

авторство, намеренно старался изменить стиль псевдонимных произведений. Именно этим может объясняться отсутствие результатов детерминированной атрибуции. Но, поскольку писатель не может значительно изменять характеристики латентных синтаксических структур при создании больших объемов текста, вероятностная атрибуция смогла однозначно указать на истинного автора.

Результаты исследования служат еще одним показателем того, что характеристики индивидуального авторского стиля носят объективный характер и не зависят от воли пишущего.

По теме диссертации опубликованы следующие работы:

1. Чепига В. П. Особенности деления текста на абзацы в произведении Э. Ажара «Страхи царя Соломона» (От авторской рукописи к опубликованному тексту) // Материалы XXXIII международной филологической конференции. Вып. 12. СПб, 2004. С. 73–78;
2. Chepiga Valentina. Romain Gary/Émile Ajar: deux auteurs, un seul écrivain? // Document pédagogique. École thématique «Analyses Pluridisciplinaires de l'Écriture. Convergences et Débats». Poitiers, 2005. S. n.;
3. Чепига В. П. Ромен Гари и Эмиль Ажар. Сравнительный анализ рукописей авторов с помощью программы EDITE // Материалы XXXV международной филологической конференции. Вып. 21. СПб, 2006. С. 32–36;
4. Чепига В. П. Сравнительный анализ произведений Ромена Гари и Эмиля Ажара (1974–1980 гг.): лексическое богатство словаря авторов // Материалы XXXVI международной филологической конференции. Вып. 10: Прикладная и математическая лингвистика. СПб, 2007. С. 84-88;
5. Chepiga Valentina. Les archives de Romain Gary // Genesis. № 29. P., 2007. P. 141-143;
6. Чепига В. П. Сравнительный анализ черновых и беловых текстов Ромена Гари и Эмиля Ажара // Древняя и Новая Романия. Вып. 7. СПб, 2007. С. 226–236;

7. Чепига В. П. Проблема «Ромен Гари – Эмиль Ажар»: атрибуция романов, опубликованных под псевдонимом Эмиль Ажар // Известия РГПУ им. А. И. Герцена. № 19(45). СПб, 2007. С. 284–290;
8. Чепига В. П. Иноязычные заимствования в черновых и беловых редакциях произведений Ромена Гари // Материалы международного симпозиума 3–5 октября 2007. М.: ИМЛИ РАН, 2007. С. 97–99;
9. Чепига В. П. К проблеме атрибуции романов, написанных под псевдонимом Эмиль Ажар: лексическое богатство словаря, метод «межтекстового расстояния», метод распознавания образов // Известия РГПУ им. А. И. Герцена. № 27 (53). СПб, 2008. С. 262-277.