

Шабалина Анастасия Вадимовна

### **ОПРЕДЕЛЕНИЕ АВТОРСТВА "ЧЕТВЕРТОГО ПРОДОЛЖЕНИЯ" ЦИКЛА О ГРААЛЕ**

В статье рассматривается вопрос об определении авторства "Четвертого продолжения" ("Продолжения Жербера") цикла о Граале путем проверки атрибуционных гипотез, полученных на основании рассмотрения существующих исторических данных о возможных авторах данного произведения. Приводятся результаты применения двух подходов к параметризации авторских текстов: на основании параметров рифм и параметров описания старофранцузского языка с применением к параметризованным текстам математических методов машинного обучения, кластерного анализа и теории распознавания образов.

Адрес статьи: [www.gramota.net/materials/2/2016/10-3/43.html](http://www.gramota.net/materials/2/2016/10-3/43.html)

Источник

### **Филологические науки. Вопросы теории и практики**

Тамбов: Грамота, 2016. № 10(64): в 3-х ч. Ч. 3. С. 156-160. ISSN 1997-2911.

Адрес журнала: [www.gramota.net/editions/2.html](http://www.gramota.net/editions/2.html)

Содержание данного номера журнала: [www.gramota.net/materials/2/2016/10-3/](http://www.gramota.net/materials/2/2016/10-3/)

### **© Издательство "Грамота"**

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: [www.gramota.net](http://www.gramota.net)

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: [phil@gramota.net](mailto:phil@gramota.net)

## SPECIFICS OF REPRESENTING THE CONCEPT "INTELLECTUALS" IN A. P. CHEKHOV'S ORIGINAL WORLDVIEW

**Suleimanova Al'mira Kamilovna**, Doctor in Philology, Associate Professor  
*Ufa State Petroleum Technological University*  
sgaziz@ufanet.ru

**Ziyatdinova Evgeniya Vyacheslavovna**, Ph. D. in Pedagogy  
*Engineering Lyceum N 83 named after M. S. Pinsky of the Ufa State Petroleum Technological University*  
evg-zi21@yandex.ru

The article examines the structure of the concept "Intellectuals", analyzes some means to verbalize this concept in the Russian language at the near periphery level, discovers the specifics of its representation in A. P. Chekhov's original worldview, identifies the synonymous and antonymous lexemes predetermining the usage context. The paper also considers author's intonation and the possible transformation of the basic lexeme meaning in the context of the whole literary text.

*Key words and phrases:* concept; pre-nuclear zone; representant; author's original worldview; lexeme; associative field.

УДК 81'32

*В статье рассматривается вопрос об определении авторства «Четвертого продолжения» («Продолжения Жербера») цикла о Граале путем проверки атрибуционных гипотез, полученных на основании рассмотрения существующих исторических данных о возможных авторах данного произведения. Приводятся результаты применения двух походов к параметризации авторских текстов: на основании параметров рифм и параметров описания старофранцузского языка с применением к параметризованным текстам математических методов машинного обучения, кластерного анализа и теории распознавания образов.*

*Ключевые слова и фразы:* цикл о Граале; «Четвертое продолжение»; атрибуция; старофранцузский язык; рифма; теория распознавания образов; кластерный анализ; машинное обучение.

**Шабалина Анастасия Вадимовна**

*Санкт-Петербургский государственный университет*  
avshabalina@gmail.com

### ОПРЕДЕЛЕНИЕ АВТОРСТВА «ЧЕТВЕРТОГО ПРОДОЛЖЕНИЯ» ЦИКЛА О ГРААЛЕ

«Четвертое продолжение», или «Продолжение Жербера», является последним из сохранившихся до нашего времени продолжений знаменитого цикла о Граале. Цикл о Граале включает в себя оригинальный незавершенный роман Кретьена де Труа и три «продолжения»: «Первое продолжение», известное также как «Продолжение Гавейна (Говена)», или «псевдо-Воше», «Второе продолжение», существующее в короткой и длинной версиях и посвященное дальнейшим приключениям Персеваля, «Третье продолжение», или «Продолжение Манессье», завершающее повествование, и, собственно, «Четвертое продолжение», или «Продолжение Жербера», существующее только в двух из дошедших до нас девятнадцати манускриптов и вставленное переписчиками как интерполяция между «Вторым» и «Третьим» продолжениями.

Хронологически оригинальный роман был написан в 1181-1191 гг. [4], по другим источникам – 1168-1190 гг. [10], «Первое продолжение» датируется концом XII века, «Второе продолжение» было создано между 1205 и 1210 гг., датировка «Третьего» является менее точной – от 1214 до 1230 гг. [7, s. 188-122; 10], «Четвертое», наиболее вероятно, было написано в 1226-1230 гг.

#### Формирование атрибуционных гипотез

На основании детального анализа существующих предположений об авторстве, анализа исторических данных о времени и месте написания «Четвертого продолжения», особенностей авторского стиля предполагаемых авторов были выдвинуты три основные атрибуционные гипотезы:

1. Нулевая гипотеза  $H_0$ : автором «Четвертого продолжения» является Жербер де Монтрёй.

Гипотеза  $H_0$  являлась приоритетной перед началом процесса атрибуции, поскольку подкреплялась множеством как субъективных предположений, так и объективных данных, выдвинутых исследователями. При этом данная гипотеза не может однозначно считаться доказанной, так как некоторые доводы ее сторонников опровергаются более поздними исследованиями, часть из которых не выдерживает критики с точки зрения здравого смысла, другие представляют собой доказательную базу под заранее сформулированный результат, а некоторые, при этом достаточно убедительные, не могут считаться определяющими [6; 9; 13, p. 180].

2. Альтернативная гипотеза, первый вариант  $H_1$ : автором является Манессье.

Гипотеза, предположительно, менее вероятная, чем первая, однако тесное переплетение текстов двух продолжений в манускрипте T, географические, диалектические и временные рамки их написания, а также тот факт, что их сюжеты дополняют друг друга, стали поводом для некоторых исследователей [12] утверждать,

что «Манессье» и «Жербер» хорошо знали о труде друг друга. Достаточно логичным представляется предположение, что у обоих продолжений один автор и завершение в «Четвертом продолжении» отсутствует потому, что его автор решил развить и углубить некоторые сюжетные ходы в отдельном произведении. Окончательное объединение двух продолжений произошло, видимо, уже после смерти автора стараниями переписчиков.

3. Альтернативная гипотеза, второй вариант  $H_2$ : автором «Четвертого продолжения» является некий неизвестный автор.

Исходя из существующих данных о значительном искажении авторских текстов в процессе их последовательного многократного копирования [5], можно предположить, что таким автором возможно является коллектив переписчиков манускриптов, объединявших разрозненные куски различных предположений по своему усмотрению.

Исходя из данных атрибуционных гипотез, был сформирован корпус текстов, необходимый для составления априорного алфавита классов, состоящий из двух классов:  $\Omega_1$  – Жербер де Монтрёй и  $\Omega_2$  – Манессье. К сожалению, существует только два текста, которые с уверенностью можно отнести к данным классам: «Роман о Фиалке» (объем класса – 1901 предложение) и «Продолжение Манессье» (объем класса – 2400 предложений). В соответствии с этим для удобства дальнейшего описания названия априорных классов преобразуются в  $\Omega_1$  – «Роман о Фиалке» и  $\Omega_2$  – «Продолжение Манессье». Атрибутируемый класс «Неизвестный автор» будет обозначаться как  $\Omega_3$  – «Четвертое продолжение» (объем класса – 3565 предложений). Необходимо принять во внимание, что во всех классах учитываются только предложения авторского текста, не содержащие прямую речь.

#### Формирование параметрического пространства

Для математического описания данных классов используется два различных типа параметров, основанных на описании авторского стиля и особенностей используемых им рифм и на основании параметров описания его языка. Оба набора параметров позволяют решить задачу определения сходства между различными текстами с учетом особенностей, присущих средневековым текстам, таких как, например, достаточно значительный вклад переписчиков в тексты переписываемых манускриптов. Список первичных параметров, предложенный в [3], модифицируется на основании имеющихся знаний о старофранцузском языке путем исключения следующих параметров: **X38** – «число причастных оборотов», **X39** – «число членов причастных оборотов», **X40** – «число распространенных причастных определений» и **X41** – «число членов распространенных причастных определений», – так как они отражают грамматические понятия, не относящиеся к группе романских языков, соответственно, их невозможно использовать для французского и старофранцузского языков. Кроме того, исключаются параметры **X46** – «число обособленных членов» и **X45** – «число членов в группах обособленных членов», так как они могут быть использованы только для современного французского языка или французского языка XVII-XVIII веков, когда уже применялись правила пунктуации. В ходе проведения мероприятий по снижению размерности параметрического пространства (см. далее) были выделены четыре информативных параметра, см. Таблицу 1.

Таблица 1.

#### Информативные параметры

Обозначение параметра	Название параметра
X18	Число имен прилагательных
X29	Число слов в косвенном падеже
X35	Число однородных групп дополнений глагола
X52	Число членов групп существительных

Идея использования набора параметров, основанных на особенностях используемых автором рифм, базируется на том факте, что к рифмам переписчики относились более осторожно и, в отличие от прозаического текста, предпочитали их не изменять, так как это влекло за собой необходимость изменения всей рифмованной пары слов. При параметризации исследуемого корпуса текстов использовались параметры, предложенные в [5] и включающие в себя лексическую, фонетическую и силлабическую структуру языка поэта. Перечень используемых параметров приведен в Таблице 2.

Таблица 2.

#### Список параметров для атрибуции при помощи рифмованных строк

Параметр	Описание параметра
P1	Женская рифма
P2	Мужская рифма
P3	Единственное число
P4	Множественное число
P5	Идентичная рифма
P6	Богатая рифма
P7	Достаточная рифма
P8	Бедная рифма

Параметр	Описание параметра
P9	Ассонанс
P10	Простой леонинский стих
P11	Завершенный леонинский стих
P12	Рифмованные слова
P13	Рифмованные звуки
P14	Рифмованные слоги
P15	Часть речи рифмованного слова
P16	Часть предложения рифмованного слова

Для удобства дальнейшего изложения эти параметры были разделены на две условные группы: «качественные параметры» и «количественные параметры». Под «количественными параметрами» будут подразумеваться параметры, для определения которых необходимо подсчитать количество включений какой-либо сущности в выборке из элементов параметризируемого класса, под «качественными» – параметры, характеризующие особенности рифм.

### Методы атрибуции

Для решения задачи атрибуции были применены четыре различных метода:

1. Атрибуция с использованием теории распознавания образов.

В соответствии с определением автора метода М. А. Марусенко атрибуция «в общем случае представляет собой задачу преобразования входной информации, представляемой некоторыми параметрами распознаваемых объектов, в выходную информацию, представляющую собой заключение о том, к какому классу принадлежит рассматриваемый объект» [3, с. 106]. Алгоритм применения данного метода подразумевает выполнение нескольких последовательных этапов:

1) снижение размерности параметрического пространства за счет определения информативных параметров путем применения схемы Бонгарда с двухступенчатым свертыванием параметрического пространства [1]. Данная схема включает в себя два этапа: этап определения релевантных и нерелевантных параметров для различения априорных классов и этап свертывания параметрического пространства на подмножестве информативных параметров. Этот этап не применяется при использовании параметризации рифмованных пар;

2) определение координат рассматриваемых объектов;

3) выполнение атрибуции с использованием детерминированного, а в случае необходимости и вероятностного алгоритмов.

В случае использования параметрического пространства, основанного на параметрах старофранцузского языка, сработал детерминированный алгоритм, причем подтвержденной оказалась гипотеза  $H_1$  об авторстве Манессье.

В случае использования параметров рифм для атрибуции при помощи теории распознавания образов нерационально использовать группу «количественные параметры», так как в таком случае размерность вектора, соответствующего каждому классу, будет несоразмерно велика относительно входящих в него ненулевых элементов, что объясняется принципом формирования тестовой выборки. Параметрическое пространство, используемое для атрибуции при помощи теории распознавания образов, включает в себя параметры с P1 по P11. При атрибуции детерминированный алгоритм авторства не выявил, а вероятностный подтвердил гипотезу  $H_0$  об авторстве Жербера де Монтрёй с вероятностью 0,6.

2. Атрибуция с использованием критериев близости.

Подразумевает проверку атрибуционной гипотезы с использованием критериев сходства. В данной работе применяется так называемый критерий Сёрнсена-Чекановского [11]. Для определения значения критерия также выполняются первые два этапа предыдущего метода, после чего производится подсчет и сравнение значений критерия для априорных классов и атрибутируемого объекта. Критерий Сёрнсена-Чекановского определяется по формуле (1) [ibidem] и принимает значения от 0 до 1, причем он равен 1, когда два объекта полностью совпадают, и 0, когда два объекта полностью не совпадают.

$$K_S = \frac{2 \sum_{i=1}^n \min(a_i b_i)}{\sum_{i=1}^n a_i + \sum_{i=1}^n b_i} \quad (1)$$

В случае использования параметрического пространства, основанного на параметрах старофранцузского языка, подтверждается атрибуционная гипотеза  $H_1$  со значением критерия 0,536.

В случае использования параметров рифм атрибуция ведется по каждому параметру отдельно, причем по параметрам P1-P11, P15, P16 и P12 подтверждается гипотеза  $H_0$ , по P13 и P14 – гипотеза  $H_1$ .

3. Атрибуция с использованием вероятностных классификаторов.

В качестве классификатора используется наивный байесовский классификатор [8], как один из самых простых и хорошо зарекомендовавших себя при обработке натуральных языков. Обучение и тестирование классификатора проводится на корпусе определенных значений информативных параметров, после чего производится проверка атрибуционной гипотезы.

В случае использования параметрического пространства, основанного на параметрах старофранцузского языка, «Четвертое продолжение» атрибутируется Манессье и подтверждается гипотеза  $H_1$ .

В случае использования параметров рифм, в процессе проведения исследования на подготовительной стадии был исключен параметр P15 из-за возникновения ошибок классификации тестовых выборок. Устойчивый результат классификации получается только при использовании параметров P13, P14 и P16, при этом для первых двух подтверждается гипотеза  $H_0$ , а для последнего –  $H_1$ .

#### 4. Атрибуция с использованием иерархической аггломеративной кластеризации.

Для атрибуции строятся дендрограммы из попарно объединяемых кластеров, сформированных из объекта атрибуции и априорных классов, и фиксируются характер и направление их объединения. Алгоритм иерархической кластеризации не подразумевает наличия обучения, таким образом, отсутствует необходимость делать априорные предположения о составе классов. Для его применения все априорные классы и атрибутируемый объект разбиваются на кластеры равного размера и сравниваются между собой, в результате чего наиболее близкие вектора попарно объединяются и сравнение повторяется. В качестве метрики близости используется Евклидово расстояние.

В случае использования параметрического пространства, основанного на параметрах старофранцузского языка, кластеры, сформированные из класса  $\Omega_1$ , объединяются практически только с кластерами того же класса, тогда как кластеры, полученные из классов  $\Omega_2$  и  $\Omega_3$ , смешиваются между собой случайным образом, что является подтверждением атрибуционной гипотезы  $H_1$ .

В случае использования параметрического пространства, основанного на свойствах рифм, результаты, полученные при помощи алгоритма иерархической кластеризации, не дают в данном случае такой же четкой картины, как при использовании параметрического пространства, основанного на параметрах старофранцузского языка, однако из полученных дендрограмм можно сделать вывод о том, что по параметрам P1-P11, P15, P16 подтверждается гипотеза  $H_0$  об авторстве Жербера де Монтрёй.

Исследование проводилось в два последовательных этапа. На первом этапе была проведена параметризация текстов с использованием первой группы параметров, после чего, на втором этапе, к ним применялись последовательно все перечисленные методы атрибуции.

Несмотря на то, что в целом различие между значениями классификаторов или критериев для случаев сравнения классов «Продолжение Манессье» и «Роман о Фиалке» с атрибутируемым объектом в некоторых случаях было несущественно, тем не менее все четыре метода однозначно подтвердили гипотезу об авторстве Манессье. Таким образом, полученные на первом этапе исследований результаты по всем четырем методам дают возможность утверждать, что верна гипотеза  $H_1$ : автором «Четвертого продолжения» является Манессье.

На втором этапе аналогичные действия были предприняты по отношению к объектам, параметризованным с использованием параметров из второй группы. Так как параметры, основанные на использовании рифм, предложены в данной работе и применяются в задачах атрибуции впервые, то в процессе проведения исследования было рассмотрено несколько вариантов совместного и отдельного использования различных параметров и групп параметров в комбинации с используемыми алгоритмами поиска близости между классами с целью определения оптимальных параметров для решения задачи поиска авторства.

Для того чтобы прояснить результаты атрибуции по рифмам и однозначно решить вопрос об авторстве произведения, необходимо отметить следующие факты, приведенные в истории вопроса:

1. Все три произведения имеются в наличии исключительно в виде копий, составленных в разных скрипториях.

2. «Четвертое продолжение», или «Продолжение Жербера», имеется только в двух манускриптах, принадлежащих одному скрипторию, географически расположенному в районе Пикардии.

3. «Четвертое продолжение» в обоих случаях подверглось существенному редактированию, прежде всего, для того чтобы оно могло быть помещено в текст перед «Третьим продолжением».

4. Оба продолжения написаны примерно в одно и то же время, предположительно, «Третье продолжение» было написано немногим ранее «Четвертого».

5. «Четвертое продолжение» в значительной степени посвящено прояснению спорных вопросов и неточностей, в том числе допущенных в «Продолжении» Манессье, и объясняет, почему именно Персеваль достоин узнать тайну Грааля.

Исходя из вышеперечисленного, можно предположить, что оба произведения могли быть написаны одним и тем же автором, условно называемым Манессье. Фактические причины, по которым один автор решил написать два продолжения, невозможно определить точно, но можно привести некоторые из них:

1. В период времени, примерно соответствующий написанию «Четвертого продолжения», патронесса Манессье, Иоанна Фландрская, была отстранена от власти в Бретани и не могла обеспечивать трубадуру средства к существованию. В связи с этим он был вынужден обратиться за покровительством к другому сеньору. Такое покровительство автору знаменитого «Третьего продолжения» могло быть обеспечено в обмен на обещание написать еще одно продолжение для нового патрона.

2. Со временем у автора могло возникнуть желание вернуться к переосмыслению личности героя и сюжета легенды, в том числе в связи с описанными выше неточностями и «недочетами».

Вследствие этого сохранились оба продолжения, а в манускриптах *T* и *V* еще и сформировался единый сюжет, а большая распространенность «Третьего продолжения» вызвана тем, что его делали по заказу более влиятельного сеньора (в 1233 г.) Иоанна Фландрская вернула себе власть), а «Четвертое продолжение» не приобрело известности за пределами Пикардии.

Когда впоследствии сохранившийся в Пикардии текст был вставлен местными переписчиками в полный текст «Повести о Граале», более известный текст с уже общепризнанной концовкой было решено оставить в конце манускрипта без дополнительного редактирования, а менее известный был отредактирован и вставлен перед ним, таким образом, чтобы получилось непротиворечивое по смыслу произведение.

До появления в XV веке школы Великих риториков и лабораторий рифм во Франции средневековые рифмы представляли собой широкий, но ограниченный набор клише, призванный сохранять литературную традицию, столь важную для средневекового общества. Мастерство поэта оценивалось по его способности искусно воспроизводить известные образные и композиционные структуры, лексические топоры того времени [2]. Эволюция средневековой литературы в существенной степени зависела от смены конфигурации готических моделей, при этом рифмы продолжительное время оставались статичным инструментарием.

Исходя из этого и с учетом того, что средневековый переписчик в некотором смысле мог считаться полноценным соавтором текста [5] и в связи с этим вносить существенные корректировки в текст копируемых произведений, и, принимая во внимание тот факт, что Жербер де Монтрёй на момент составления копии был крупнейшим, а возможно, и единственным известным поэтом в данном регионе, вполне разумно предположить, что переписчик попытался имитировать его поэтический стиль в рамках редактирования «Четвертого продолжения» с целью улучшить его в меру своего понимания, что объясняет и некоторые лексические параллели, и результаты применения методов атрибуции на основе параметризации при помощи рифм.

Окончательный вывод об авторстве произведения звучит следующим образом: «Четвертое» и «Третье» продолжения имеют одного и того же автора, предположительно, некоего Манессье, но при этом «Четвертое продолжение» подверглось серьезному редактированию, вероятно, в скриптории города Арраса, при этом переписчик стремился подражать или отметить знаменитого поэта своего региона Жербера де Монтрёя.

Исходя из такого результата атрибуции, можно оценить подход к параметризации при помощи рифмованных строк как частично успешный, хотя и менее эффективный непосредственно для решения задач определения авторства анонимных произведений. Использование группы параметров, описывающих стихотворный стиль автора, является целесообразным в том случае, когда отсутствуют подозрения, что авторский стиль может подвергнуться имитации, либо же (в сочетании с атрибуцией на основе структуры предложения, когда результат атрибуции совпадает) в качестве дополнительного подтверждения правильности определения автора. Кроме того, целесообразно использовать такие параметры, предварительно дополнив их, в том случае, когда поэтический стиль исследуемых произведений более развит и совершенен, чем поэтический стиль средневековых трубадуров.

#### *Список литературы*

1. **Бонгард М. М.** Проблема узнавания. М.: Наука, 1967. 320 с.
2. **Косиков Г. К.** Средние века // Андреев Л. Г., Козлова Н. П., Косиков Г. К. История французской литературы: учебник. М.: Высшая школа, 1987. С. 9-140.
3. **Марусенко М. А.** Атрибуция анонимных и псевдонимных литературных произведений методами теории распознавания образов. Л.: ЛГУ, 1990. 164 с.
4. **Михайлов А. Д.** Французский рыцарский роман и вопросы типологии жанра в средневековой литературе / отв. ред. Н. И. Балашов. Изд-е 3-е. М.: Книжный дом «ЛИБРОКОМ», 2012. 349 с.
5. **Шабалина А. В.** Использование параметров рифм для атрибуции средневековых стихотворных текстов // Древняя и Новая Романия. 2015. Т. 15. № 1. С. 327-336.
6. **Birch-Hirschfeld A.** Die Sage vom Gral. Ihre Entwicklung und dichterische Ausbildung in Frankreich und Deutschland in 12 und 13 Jahrhundert. Leipzig: Verlag von F.C.W. Vogel, 1877. 291 S.
7. **Corley C. F. V.** Manessier's Continuation of Perceval and the Prose Lancelot Cycle // The Modern Language Review. 1986. Vol. 81. Issue 3. P. 574-591.
8. **Data classification: algorithms and applications** / ed. by C. C. Aggarwal. CRC Press, 2014. 680 p.
9. **Kraus F.** Über Gerbert de Montreuil und seine Werke. Erlangen: Fr. Junge (Junge and Sohn), 1897. 88 S.
10. **Marx J.** La légende arthurienne et le Graal. Paris: Presses Universitaires de France, 1952. 410 p.
11. **Mroz D., Reed D.** Resource assessment in forested landscapes. John Wiley and Sons, Inc., 1997. 387 p.
12. **Thether L. R.** Suspended Endings: The Mechanics of Medieval Continuation in the Perceval Continuations. Doctoral thesis. Durham University, 2009. 281 p.
13. **Wilmotte M.** Gerbert de Montreuil et les écrits qui lui sont attribués // Bulletins de l'Académie royale de Belgique: bulletins de la classe des lettres et des sciences morales et politiques de la classe des beaux-arts. 1900. P. 166-189.

#### **IDENTIFYING THE “FOURTH GRAIL CONTINUATION” AUTHORSHIP**

**Shabalina Anastasiya Vadimovna**  
Saint Petersburg University  
avshabalina@gmail.com

The article focuses on the problem of “Fourth (Gerbert’s) Grail Continuation” authorship. The researcher checks the attribution hypotheses developed on the basis of the available historical data concerning the alleged authors of this literary work. The paper describes two approaches to author’s texts parameterization: on the basis of the Old French rhythm parameters and description parameters, applying the machine learning mathematical methods, cluster analysis and image identification theory to parametrized texts.

*Key words and phrases:* Grail Cycle; “Fourth Continuation”; attribution; Old French; rhyme; image identification theory; cluster analysis; machine learning.