

ПРОБЛЕМА «КОРНЕЛЬ – МОЛЬЕР»: К ВОПРОСУ ОБ АВТОРСТВЕ И ДАТИРОВКЕ СТИХОТВОРНЫХ ПЬЕС, ПРИПИСЫВАЕМЫХ МОЛЬЕРУ

*Работа представлена кафедрой математической лингвистики
Санкт-Петербургского государственного университета.
Научный руководитель — доктор филологических наук, профессор М. А. Марусенко*

Статья посвящена атрибуции и датировке стихотворных пьес, приписываемых Мольеру. В основу исследования была положена методика атрибуции анонимных и псевдонимных произведений, которая предусматривает применение многомерной классификации, базирующейся на теории распознавания образов. В работе выдвигается гипотеза о существовании тренда параметров индивидуального авторского стиля и описывается новый математический метод датировки литературных произведений.

The article is devoted to the problem of attribution and dating of plays ascribed to Molière. The research is based on mathematical methods of attribution of anonymous and pseudonymous works with multidimensional classification and aspects of the pattern recognition theory. The author of the article assumes that there is a trend of the writer's individual style parameters. She also describes a new mathematical method of literary works' dating.

Вопрос об авторстве комедий, опубликованных под псевдонимом Мольер, был поднят в начале XX в. 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 зрения¹. 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».

Вопрос датировки пьес, приписываемых Мольеру, неизменно возникает при рассмотрении проблемы авторства. Известно, что придворные драматурги старались пи-

сать 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. П. Луи датирует пьесы, составляющие театр Мольера, годами молчания П. Корнеля в 1630-е и 1650-е гг.² Согласно его теории, шедевры, приписываемые Мольеру, были созданы за 20–30 лет до их первых постановок в театре. Наряду с различными пьесами, постановки которых происходили вскоре после их написания, П. Корнель писал сатирические «острые» комедии, которые не могли быть представлены на суд публики от его имени. И только после того, как он вывел на сцену Мольера, все его бессмертные шедевры были наконец поставлены в придворном театре.

Итак, существуют различные гипотезы о принадлежности и датировке пьес, составляющих театр Мольера.

1. Атрибуция стихотворных пьес, приписываемых Мольеру. В основу исследования по определению авторства пьес Мольера была положена методика атрибуции анонимных и псевдонимных произведений, разработанная М. А. Марусенко³, которая предусматривает применение многомерной классификации, основанной на теории распознавания образов, и описания 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.

При формировании атрибуционной гипотезы в число возможных авторов могут быть включены драматурги Мольер, П. Корнель, Ф. Кино и неопределенное число неизвестных авторов. В настоящей работе мы имеем дело с проблемой сосуществования противоречащих мнений, и литературно-критическая атрибуционная гипотеза может быть сформулирована следующим образом: нулевая гипотеза (H_0): тексты пьес, составляющие театр Мольера, полностью принадлежат Мольеру.

Альтернативная гипотеза имеет следующий вид:

- (H_a^1): тексты пьес, приписываемые Мольеру, полностью принадлежат П. Корнелю;

- (H_a^2): тексты пьес, приписываемые 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.

Материалом данного исследования стали 13 комедий в стихах, приписываемых Мольеру, а также 11 комедий в стихах П. Корнеля и 3 комедии в стихах Ф. Кино.

Проверка атрибуционной гипотезы выполняется средствами теории распознавания образов и предусматривает реализацию двух независимых этапов: отбор информативных параметров и процедуру распознавания.

В нашем исследовании информативные параметры должны разделять два априорных класса: Корнеля – Ω (Pierre Corneille) и Кино – Ω (Philippe Quinault). Априорный словарь параметров составил 51 параметр, релевантный для описания текстов XVII в. на французском языке.

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. Бонгарда⁴, предусматривающая двухступенчатое свертывание параметрического пространства. На первом этапе происходит разбиение априорного набора информативных параметров на два подмножества параметров, релевантных и не релевантных для различения априорных классов. Релевантность параметров для различения двух априорных классов определяется по t -критерию Стьюдента (формула 1), пороговое значение которого при уровне значимости $\alpha = 0,05$ равно 1,96.

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

где \bar{x}_2, \bar{x}_1 – средние арифметические, σ_1, σ_2 – стандартные отклонения, а n_1, n_2 – размеры выборок.

Значения t -критерия для пяти параметров оказались больше критического, что позволило определить параметры X02 (чис-

ло элементарных предложений), X04 (число сочиненных предложений), X21 (число спрягаемых форм глагола), X31 (число подлежащих), X32 (число местоимений-подлежащих) как информативные. На втором этапе схемы М. М. Бонгарда, предусматривающем процедуру свертывания параметрического пространства на подмножестве информативных параметров при помощи критерия эффективности, дальнейшего уменьшения числа информативных параметров не произошло.

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 1). 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. В результате работы детерминированного алгоритма была определена авторская принадлежность шести из тринадцати анализируемых объектов: пьесы «Le dйpit amoureux», «L'Йcole des maris», «Les Fвcheux», «L'Йcole des femmes», «Tartuffe», «Les Femmes savantes» были атрибутированы П. Корнелю с вероятностью, находящейся в интервале между 0,95 и 0,95 + 0,05⁵, 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 с помощью функции евклидова расстояния:

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

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

Вероятность принадлежности *j*-го объекта к *i*-му классу (табл. 1) вычисляется на основе матрицы евклидовых расстояний между объектами и априорными классами по следующей формуле:

$$P_{ji} = \frac{1}{d_{ji}} \left(\sum_k \frac{1}{d_{jk}} \right)^{-1}, \quad (3)$$

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

Таблица 1
Матрица вероятностей принадлежности объектов априорным классам $P(X_i \in \Omega_j)$

Класс	$P(X_i \in \Omega_j)$						
	Объект						
	1M	3M	4M	8M	10M	11M	12M
$\Omega(\text{Corneille})$	0,32	0,65	0,48	0,48	0,73	0,69	0,63
$\Omega(\text{Quinault})$	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:

$$X_i \in \Omega_j, \quad P(X_i \in \Omega_j) \geq 0,5. \quad (4)$$

В результате работы вероятностного алгоритма распознавания Ф. Кино были атрибутированы такие пьесы, как «L'Йtourdi», «Dom Garcie de Navarre» и «La Princesse d'Йlide», а П. Корнелю были приписаны пьесы «Sganarelle», «Le Misanthrope», «Mйlicerte», «Pastorale comique» с разной степенью вероятности. Результаты вероятностного алгоритма были откорректированы в процессе процедуры оценки качества классификации. Две пьесы, «Dom Garcie de Navarre» и «La Princesse d'Йlide», составили апостериорный класс.

В результате проведенного исследования получил свое подтверждение вариант альтернативной гипотезы (H_a^2): тексты пьес, приписываемые Мольеру, являются произведениями П. Корнеля, Ф. Кино и одного неизвестного автора.

2. Датировка стихотворных пьес, приписываемых Мольеру.

В основе предложенного нами метода датировки на основе стилеразличающих параметров 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 Корнелю, и из атрибутированных ему произведений из театра Мольера. Другая совокупность состоит из работ Ф. Кино и атрибутированной ему пьесы Мольера «L'Йtourdi». Под датировкой в данном случае понимается установление последовательности написания комедий и сопоставление полученных результатов с общепринятой датировкой по годам их первых постановок.

В результате применения метода датировки при анализе комедий Ф. Кино был установлен тренд параметров индивидуального авторского стиля в сторону увеличения значений параметров, а также нами была выдвинута гипотеза о написании пьесы «L'Йtourdi» драматургом Ф. Кино до 1655 г.

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

Смешение дат может быть обусловлено несколькими причинами, и основная из них – это сильная зависимость результатов эксперимента от однородности и качества исследуемого материала. Комедии, принадлежащие перу П. Корнеля, так же, как и атрибутированные ему комедии из театра Мольера, отличаются большим разнообразием жанровых подвидов. Близость пьес П. Корнеля, написанных в 1630-е гг., с атрибутированными ему шедеврами Мольера может указывать на то, что лучшие пьесы Мольера были созданы П. Корнелем задолго до их постановки на сцене, а именно в период с 1630 до 1650-х гг. Таким образом, 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.

В результате проведенного исследования получила свое подтверждение гипотеза о том, что тексты пьес Мольера являются произведениями П. Корнеля, Ф. Кино и еще одного неизвестного автора. Результаты проведенного эксперимента по датиров-

ке пьес, атрибутированных П. Корнелю и Ф. Кино, позволяют говорить о подтверждении выдвинутой автором гипотезы о существовании тренда параметров индивидуального авторского стиля и о подтверждении гипотезы о более раннем написании стихотворных пьес, входящих в театр Мольера. Применение метода датировки на основе стилеразличающих синтаксических параметров показало возможность и целесообразность такого подхода.

ПРИМЕЧАНИЯ

¹ *Poulaille H.* Corneille sous le masque de Molière. Paris: Grasset, 1957; *Wouters H., de Ville de Goyet C.* Molière ou l'auteur imaginaire? Bruxelles: Comptix, 1990; *Labbü D.* Corneille dans l'ombre de Molière. Histoire d'une recherche. Paris – Bruxelles: Les 'Impression nouvelles, 2003.

² *Loiyas P.* Broutilles, recueillies par Фрїдїриг Lachivre. P., 1938.

³ *Марусенко М. А.* Атрибуция анонимных и псевдонимных литературных произведений методами распознавания образов. Л.: Изд-во ЛГУ, 1990.

⁴ *Бонгард М. М.* Проблемы узнавания. М.: Наука, 1967.