

Секция «Иностранные языки и регионоведение»

**Многомерная классификация стихотворных текстов (на материале лирики
Р. Фроста, Э. Робинсона, Э. Лоуэлл, В. Линдсей)**

Николаева Елена Александровна

Аспирант

*Смоленский государственный университет, Физико-математический факультет,
Смоленск, Россия*

E-mail: leno4kator@mail.ru

Одной из задач при исследовании индивидуального стиля является сопоставление стилей письма различных авторов. Широкое распространение получили различные квантитативные методы изучения идиостиля, которые направлены на выделение ряда параметров, характеризующих текст, и выявление степени их представленности в произведениях авторов.

В качестве материала исследования нами привлечены ямбические стихотворные тексты широко известных американских поэтов конца XIX – начала XX века Роберта Фроста, Эдвина Робинсона, Вэчела Линдсей и Эми Лоуэлл. Как известно, стихотворный текст значительно отличается от художественной прозы большей упорядоченностью, единообразно заданным метром, более сложным взаимодействием различных уровней.

Для исследования был привлечен ряд параметров, отражающих ритмические, синтаксические, морфологические, фонетические признаки, а также признаки стихотворного синтаксиса.

Синтаксический уровень в нашей работе представлен следующими параметрами: синтаксическая функция слова, занимающего первую и последнюю сильную позицию в стихотворной строке, количество придаточных предложений в составе сложноподчиненного и простых предложений в составе сложносочиненного предложения. Морфологические признаки в данном исследовании определяются по частеречной принадлежности слов в первой и последней сильных позициях в строке. Выделение параметров «поэтического синтаксиса» основано на понятиях инверсии, переноса, разрыва стиха синтаксической паузой и стилистической маркированности конца строки. В качестве параметров фонетического уровня нами привлекается количество слогов в словах в первой сильной позиции. Ритмический уровень представлен такими признаками, как наличие пропусков ударения на первом, втором и последнем сильных местах, количество слогов в анакрусе, наличие отягчения анакрusy и размер, которым написана строка.

В качестве основного метода исследования нами был использован метод многомерного статистического анализа данных – дискриминантный анализ. В результате обработки исследуемого материала нами были получены данные, свидетельствующие о том, что исследуемые признаки позволяют с достаточно высокой точностью расклассифицировать тексты.

Степень успешности дифференциации стилей четырех авторов была проверена путем проведения апостериорной классификации текстов поэтов. Автоматическое разбиение текстов на четыре класса в 87% случаев правильно отнесло тексты к «своим» авторам. Вероятность случайного совпадения автоматической и естественной группировки всех текстов по четырем классам составляет всего 25%, в связи с чем полученный

результат, более чем в три раза превышающий случайную возможность правильного решения, может рассматриваться как весьма хороший.

Было установлено, что наиболее значимыми для дискриминации текстов оказались синтаксические и морфологические характеристики.

В исследовании также проводился учет распределения некоторых признаков в начале и конце стихотворной строки. Из данных дискриминантной модели следует, что признаки начала строки вносят больший вклад в дифференциацию классов, чем признаки конца строки.

Среди параметров дискриминантной модели, для которых был невозможен учет распределения в стихотворной строке, наибольший вклад вносит признак поэтического синтаксиса «частичная инверсия».