

КВАНТИТАТИВНЫЙ АНАЛИЗ ПРОИЗВЕДЕНИЙ ПОЭТОВ XVIII–XIX ВЕКОВ

В данной работе в качестве корпуса текстов для исследования взяты произведения 34-х поэтов Золотого века и Серебряного века, в отношении подлинности и авторства которых сомнений нет.

Золотой век русской поэзии – первая половина XIX века, Серебряный век русской поэзии – период с 1890-х годов по начало 1920-х годов. Невозможно рассматривать одно в отрыве от другого, Серебряный век без наследия Золотого, а Золотой – без продолжения его в Серебряном.

Квантитативный анализ текста предполагает расчет ряда некоторых количественных характеристик корпуса текстов, например, энтропии. Энтропия (от греч. *entropía* – поворот, превращение) – понятие, впервые введенное в термодинамике (введено Р. Клаузиусом в 1865 г.) для определения меры необратимого рассеяния энергии. Энтропия широко применяется и в других областях науки. В информатике под энтропией понимается мера неопределенности или непредсказуемости информации, неопределенность появления какого-либо символа первичного алфавита. В лингвистике, например, под первичным алфавитом можно понимать традиционный алфавит, расширенный знаками препинания и специальными разделителями (пробел, разметка, непечатные невидимые на экране символы), или полный набор ЛЕ (лексических единиц – словоформ, текстоформ, лемм и т. п.). При отсутствии информационных потерь энтропия равна количеству информации на символ передаваемого сообщения. К. Шеннон предположил, что прирост информации равен утраченной неопределенности, и задал требования к ее измерению: 1) мера должна быть непрерывной, т. е. изменение значения величины вероятности на малую величину должно вызывать малое результирующее изменение функции; 2) в случае, когда все варианты (в нашем случае – ЛЕ) равновероятны, увеличение количества вариантов ЛЕ должно всегда увеличивать значение функции; 3) должна быть возможность сделать выбор ЛЕ в два шага, в которых значение функции конечного результата должно являться суммой функций промежуточных результатов. Таким образом, показатель энтропии количественно характеризует уровень информационной упорядоченности текста как системы, т. е. чем больше энтропия, тем менее упорядочен текст, при этом энтропия – это функция состояния: любому состоянию системы можно придать вполне определенное значение энтропии [Shannon 1948; Колмогоров 1965; Пиотровский и др. 1977; Бектаев 1978; Арапов 1988]. Величина энтропии текста – это свойство этого текста: степень его упорядоченности, степень его отклонения от состояния полного хаоса, при котором все ЛЕ имели бы равную вероятность, а текст превратился бы в бессмысленный набор ЛЕ. Упорядоченность текста будет тем больше, чем

больше различие вероятностей и чем больше вероятность последующего события будет зависеть от вероятностей предыдущих событий.

В процессе количественного анализа значений накопленной энтропии, индексов итерации, исключительности, плотности, дистрибуции и предсказуемости была произведена оценка вклада поэтов в русский язык и литературу, в культуру России. В итоге получилось, что А.А. Блок, А.С. Пушкин, Н.А. Некрасов, М.Ю. Лермонтов, В.А. Жуковский внесли больший вклад, чем остальные поэты.

Трактовка индексов дистрибуции, итерации, исключительности, предсказуемости, плотности требует дополнительных исследований, и не только на выборке поэтов Золотого века и Серебряного века. Например, целесообразно сопоставить творчество поэтов с творчеством прозаиков, и не только с классиками русской художественной литературы, но и с другими представителями отечественной литературы.

Литература

Арапов М.В. Количественная лингвистика. М.: Наука, 1988.

Бектаев К.Б. Статистико-информационная типология тюркского текста. Алма-Ата: Наука КазССР, 1978.

Бектаев К.Б., Пиотровский Р.Г. Математические методы в языкознании. Ч. 2. Математическая статистика и моделирование текста. Алма-Ата, 1974.

Колмогоров А.Н. Три подхода к определению понятия «количество информации» // Проблемы передачи информации. Т. I, вып. 1. 1965. С. 3–11.

Колмогоров А.Н., Гельфанд И.М., Яглом А.М. К общему определению количества информации. ДАН СССР, 1956. Т. 111. С. 745.

Пиотровская А.А., Пиотровский Р.Г., Разживин К.А. Энтропия русского языка // Вопросы языкознания. 1962. № 6. С. 115–130.

Пиотровский Р.Г., Бектаев К.Б., Пиотровская А.А. Математическая лингвистика. М.: Высшая школа, 1977.

Шеннон К. Работы по теории информации и кибернетике. М.: Изд. иностр. лит., 2002.

Shannon C.E. A mathematical theory of communication. Bell System Technical Journal, vol. 27, pp. 379–423 and 623–656, July and October, 1948.