

## АНАЛИЗ СОДЕРЖАНИЯ НАУЧНОГО ЖУРНАЛА С ПОМОЩЬЮ ЛЕКСИКО-СТАТИСТИЧЕСКОГО МЕТОДА

Клементьева А.А.

Московский государственный университет им. М.В. Ломоносова, Филологический ф-т,  
каф. Русского языка, Россия, 119415, г. Москва, ул. Проспект Вернадского, 37, Тел.:  
+7(927)191-27-42, E-mail: aklementyeva@mail.ru

Автоматизированный анализ текстов исследовательских статей и научных журналов – мировой тренд. С помощью такого анализа делаются попытки выявить фокусы интереса и структурировать онтологии различных дисциплин.

В докладе метод лексико-статистического анализа, т.е. подсчета частотности употребления в тексте отдельных слов, словоформ и словосочетаний будет применен для исследования содержания научного журнала «Социальная и клиническая психиатрия», архив которого за 2014 – 2019 г. был предоставлен редакцией.

Текстовый корпус, на котором проводилось исследование, включает 394 статьи. Объем корпуса составил 948 413 словоупотреблений. Компьютерная лемматизация и ручная «чистка» позволили получить словник из 19 684 уникальных слов, а также данные по частотности их грамматических форм и списки биграмм, триграмм и тетраграмм, встречающихся в корпусе. Слова в словнике были объединены в лексические семьи.

Изучение и интерпретация полученных данных совместно лингвистами и психиатрами позволили установить:

1) если из 50 самых частотных слов в частотном словаре русского языка Ляшевской – Шарова [1] (далее – ЛШ) только три полнозначительных (год, человек, сказать), то в исследуемом корпусе их 17;

2) высокая частотность слов *уровень*, *показатель*, *результат* говорит о том, что журнал публикует исследовательские статьи, а присутствие слов *терапия* и *лечение* в пятом десятке свидетельствует о том, что материалы журнала носят практико-ориентированный характер;

3) частотность лексических семей, входящих в название журнала (*социальный*, *клинический*, *психиатрия*) говорит о соответствии материалов, публикуемых в журнале, его заявленной проблематике;

4) из всех психических заболеваний наибольшее внимание получают расстройства шизофренического спектра и депрессии: лексическая семья «Шизофрения» включает 20 членов с частотностью от 2827 до 1, в лексической семье «Депрессия» 22 члена с частотностью от 1823 до 1.

Авторы исследования предполагают, что использованный метод дает возможность установить (1) соответствие материалов, публикуемых журналом, его тематике, заявленной в названии; (2) научно-исследовательскую направленность журнала.