

ОБЩАЯ, НЕОРГАНИЧЕСКАЯ И ОРГАНИЧЕСКАЯ ХИМИЯ ЧЕРЕЗ ПРИЗМУ КОМПЬЮТЕРНОГО АНАЛИЗА ТЕКСТА

Мязин В.Д., Бадикова А.Г., Шматко А.С., Певцов Д.Н.

Московский государственный университет

В последние годы растет число иностранных студентов в РФ, в том числе химиков. Для изучения специальных предметов им необходимо читать учебную и научную литературу. Для полного понимания текста, по мировым данным, необходимо знать 95% слов, его составляющих. Если взять топ 20000 частотных слов русского языка, то покрытие текста среднего университетского учебника по химии будет меньше 95%, при этом выучить такое количество слов иностранцу практически нереально. Следовательно, актуальной является задача отбора наиболее часто используемых слов в учебниках. В нашем докладе представлены частотные словники, составленные на основе рекомендуемой литературы для курсов общей, неорганической и органической химии, а именно: 1) Третьяков Ю.Д. «Неорганическая химия», в 3-х томах; 2) Некрасов Б.В. «Основы общей химии»; 3) Ахметов Н.С. «Общая и неорганическая химия»; 4) Коренев Ю.М., Овчаренко В.П. «Общая и неорганическая химия. Курс лекций» 5) Реутов О.А., Курц А.Л., Бутин К.П. «Органическая химия» в 4-х томах; 6) Клайден Дж. «Органическая химия» в 3-х частях. Общий объем корпуса общей и неорганической химии – более 700 000 словоупотреблений, в органической – более 800 000, то есть суммарно более 1 500 000 словоупотреблений.

В докладе будут представлены результаты компьютерного анализа данных текстов, позволяющего выявить понятийные сходства и различия между этими дисциплинами. Сопоставление словников этих дисциплин также дает возможность оценить словарную нагрузку на студента, как русскоязычного, так и иностранного.