

МЕТОД ФОРМАЛЬНОГО ВЫДЕЛЕНИЯ ТЕМАТИЧЕСКИ НЕЙТРАЛЬНОЙ ЛЕКСИКИ (НА ПРИМЕРЕ СТАРОСЛАВЯНСКИХ ТЕКСТОВ)

А.А. Кретов

В статье введены понятия тематически маркированной и тематически нейтральной лексики; предложен метод системного взвешивания слов по двум функциональным параметрам: прямому (частотному - Q-параметр) и косвенному (длина слова - F-параметр). Первый параметр характеризует функционирование слова в данном тексте, второй - его функционирование на продолжительном отрезке времени - настолько продолжительном, чтобы функционирование успело повлиять на длину слова; введен Индекс тематической маркированности слова (**ИнТеМ**), вычисляемый по формуле $\text{ИнТеМ} = \text{Q-вес} - \text{F-вес}$, где Q-вес - вес слова по частоте, а F-вес - вес слова по длине. Установлено, что в словаре-источнике более 98% слов с отрицательным значением **ИнТеМа** относятся к тематически нейтральной лексике.