

Тематический корпус – газетный корпус НКРЯ

<http://www.ruscorpora.ru/search-paper.html>

Объем – 228521421 токен

Контрастный корпус – основной корпус НКРЯ

<http://ruscorpora.ru/search-main.html>

Объем – 283431966 токенов

w_i	Тип	$\text{Count}_{\text{SpecC}}$	$\text{Count}_{\text{RefC}}$	LogLikelihood	Ранг	Weirdness	Ранг
передавать	темат.	30527	16166	8122.12	2	2.342	2
корреспондент	темат.	65340	16393	42864.81	1	4.944	1
поиск	общеупотр.	23946	20579	1495.11	4	1.443	3
уходить	общеупотр.	30036	55372	- 3162.47	3	0.673	4

В данном случае Weirdness лучше отражает «тематичность» слов, тогда как LogLikelihood показывает зависимость от частотности слова – у «уходить» показатель больше, хотя перекося его частоты в сторону общеупотребительного корпуса больше, чем у «поиск».