

Корпуса:

Тематический корпус – диалектный корпус НКРЯ
Объем – 285281 токен

Контрастный корпус – основной корпус НКРЯ
Объем – 283431966 токенов

w_i	Тип	$Count_{SpecC}$ (частота в Specific Corpora)	$Count_{RefC}$ (частота в контрастном корпусе (Reference Corpora))	LogLikelihood	Ранг	Вес 2 Wierdness	Ранг
разговор	Тематич.	36515	1	5,2	2	3,65	2
квартира	Тематич.	8000	4	1,01	1	2,01	1
язык	Контраст.	30547	11	1,48	3	2,6	3
хата	Контраст.	691	18	-4,69	4	0,38	4

Посмотрев на таблицу, можно увидеть, что тематическое слово для первого и второго корпуса получило самую высокую оценку, а второе контрастное слово – самую низкую. Важно отметить, что слова такие слова все слова в контрастном корпусе были в малом количестве.