

МАРКЕРЫ ПРОЗЫ XIX ВЕКА (СРАВНЕНИЕ ЧЕРЕЗ СТОЛЕТИЕ)

Наш частотный словарь прозы середины XIX века¹ с 15 миллионами словоупотреблений сравнивается с грандиозным словарем О.Н. Ляшевской и С.А. Шарова², а точнее – с жанром «Художественная литература», охватывающим период 1950–2007 гг. и включающим 35 миллионов словоупотреблений. Все данные в последнем словаре даются в виде относительных частот на миллион словоупотреблений (ipm – items per million), что и сохраняется в следующей табличке:

| | XIX в. | XX в. | | XIX в. | XX в. |
|---------|--------|-------|----------|--------|-------|
| большой | 801 | 811 | однако | 487 | 330 |
| вы | 7802 | 4216 | письмо | 510 | 304 |
| дорогой | 127 | 213 | проблема | 1 | 116 |
| ли | 2108 | 1288 | сволочь | 11 | 54 |
| машина | 32 | 602 | сказать | 3946 | 4115 |
| милый | 315 | 138 | телефон | – | 231 |
| мой | 3322 | 2137 | товарищ | 207 | 369 |
| мужчина | 173 | 333 | человек | 3154 | 2670 |
| но | 5880 | 6155 | черт | 229 | 218 |

Для каждого из двух корпусов можно статистическими средствами выделить лексические маркеры при помощи формулы: $S = (f - m - 1) / \sqrt{m}$, где f – абсолютная частота слова, m – математическое ожидание этой частоты в рамках данной нулевой гипотезы.

При сравнении двух корпусов обнаружено 1267 маркеров XIX века (при $S > 8$) и 714 маркеров XX века (при $S > 4$). Из 18 слов нашей таблички семь оказались маркерами XIX века: *вы* $S=42$, *ли* $S=56$, *милый* $S=28$, *мой* $S=64$, *однако* $S=22$, *письмо* $S=41$, *человек* $S=15$. Маркерами XX века являются: *дорогой* $S=11$, *машина* $S=48$, *мужчина* $S=16$, *но* $S=20$, *проблема* $S=21$, *сволочь* $S=11$, *сказать* $S=11$, *телефон* $S=30$, *товарищ* $S=15$. Слова *большой* и *черт* не показали значимых расхождений между двумя словарями.

В докладе будут рассмотрены наиболее интересные семантические группы маркеров XIX века.

¹ Шайкевич А.Я., Андриющенко В.М., Ребецкая Н.А. Дистрибутивно-статистический анализ языка русской прозы 1850–1870-х гг. Т. 1. М., 2013. С. 155–495.

² Ляшевская О.Н., Шаров С.А. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М., 2009.