

# Многозначные слова в контексте и вне контекста

Доклад на семинаре «Некоторые применения  
математических методов в языкознании»  
(МГУ, 16.11.2013)

*Борис Иомдин ([iomdin@ruslang.ru](mailto:iomdin@ruslang.ru))*

*Институт русского языка*

*им. В. В. Виноградова РАН*

*НИУ «Высшая школа экономики»*

*Школа анализа данных Яндекса*

# План доклада

- Разрешение многозначности как проблема АОТ
- Корпусный подход к выделению значений
- Разрешение многозначности вне контекста?
- Типы контекстов: текст, ситуация, говорящий
- Эксперименты с информантами: ранжирование лексем, словарные и несловарные значения
- Все ли лексемы существуют одновременно?
- Абсолютная vs. относительная многозначность
- Дальнейшие направления исследований и возможные применения

# Разрешение лексической многозначности (WSD)

- Использование словарных толкований
  - алгоритм Леска (совпадающие элементы в толкованиях слов в контексте). Ср. закон семантического согласования Апресяна (*Хороший кондитер не жарит хворост на газовой плите*)
- Использование корпусов
  - аннотированных
  - неаннотированных (кластеризация контекстов в корпусе)

# Ранжирование значений по частотности

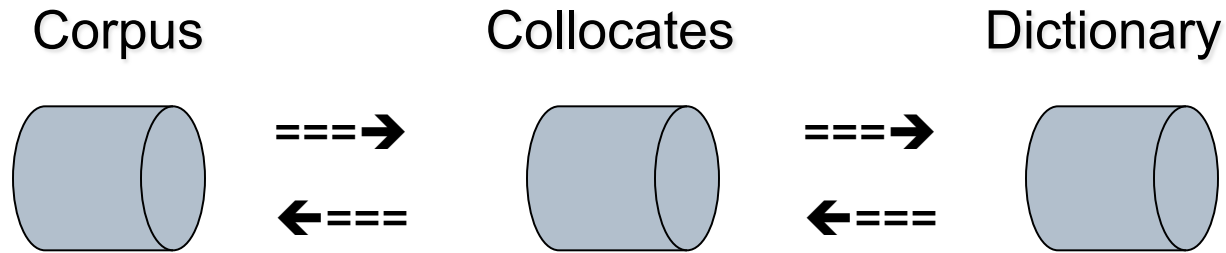
- Совместная встречаемость слова и его синонимов по Wordnet (Lapata and Keller 2007)
- Сравнение сочетаемости слова в разных значениях и его аналогов по тезаурусам (Mohammad and Hirst 2006)

# Проблемы

- Нехватка хороших словарей
- Словари: широкие значения (слишком мало), Wordnet: узкие значения (слишком много)
- Несистемность описаний
- Недискретность полисемии
- Несогласие экспертов (inter-judge variance)

Обзор WSD: Navigli 2009

# Корпусный подход



**Collocates, Constructions, Domains = CoCoDo**

(Kilgariff 2013)

# Sketch Engine (A. Kilgariff)

Sketch  Engine

search in **huge** text corpora



# Sketch Engine (A. Kilgariff)

- Корпусы 52 языков
  - Обзоры сочетаемости слова
  - Частотные списки
  - Создание и сравнение корпусов
  - и др.
- Самый большой корпус: RuTenTen (20 млрд. слововхождений)



# ВЫШКА

 ruTenTen11 freq = [91305](#) (4.5 per million)

subject_of	3575	0.0	a_modifier	13218	0.1	prec_prep	26981	0.1	и/или	7188	0.0	pp_obj_na	6781	0.3
стоять	<a href="#">274</a>	0.75	нефтяной	<a href="#">1344</a>	7.04	на	<a href="#">10107</a>	0.85	лес	<a href="#">327</a>	2.61	судья	<a href="#">454</a>	4.14
находиться	<a href="#">196</a>	0.02	буровой	<a href="#">673</a>	8.86	с	<a href="#">7124</a>	0.95	мачта	<a href="#">226</a>	6.63	часовой	<a href="#">205</a>	5.28
возвышаться	<a href="#">70</a>	4.5	смотровой	<a href="#">658</a>	8.5	под	<a href="#">537</a>	0.5	подъемник	<a href="#">220</a>	5.78	подниматься	<a href="#">181</a>	3.05
светить	<a href="#">41</a>	3.46	наблюдательный	<a href="#">639</a>	7.6	возле	<a href="#">134</a>	2.78	трамплин	<a href="#">184</a>	6.82	стоять	<a href="#">155</a>	1.41
торчать	<a href="#">41</a>	3.41	парашютный	<a href="#">608</a>	9.31	вокруг	<a href="#">100</a>	1.38	кран	<a href="#">180</a>	3.79	сидеть	<a href="#">155</a>	1.19
располагаться	<a href="#">36</a>	0.72	телевизионный	<a href="#">549</a>	6.2	около	<a href="#">97</a>	1.75	башня	<a href="#">158</a>	3.14	подняться	<a href="#">154</a>	3.07
<b>Чикатило</b>	<a href="#">33</a>	7.67	сторожевой	<a href="#">461</a>	8.3	про	<a href="#">80</a>	0.12	провода	<a href="#">139</a>	3.99	забираться	<a href="#">103</a>	5.31
виднеться	<a href="#">33</a>	3.72	нижегородский	<a href="#">305</a>	4.38	мимо	<a href="#">71</a>	2.62	труба	<a href="#">106</a>	1.46	охранник	<a href="#">85</a>	3.66
устанавливаться	<a href="#">29</a>	0.52	пожарный	<a href="#">305</a>	4.02	напротив	<a href="#">46</a>	2.56	лестница	<a href="#">88</a>	1.96	залезть	<a href="#">82</a>	5.14
строиться	<a href="#">21</a>	0.67	диспетчерский	<a href="#">203</a>	6.79	вместо	<a href="#">35</a>	0.16	мост	<a href="#">78</a>	1.25	забраться	<a href="#">70</a>	5.16
грозить	<a href="#">19</a>	1.08	судейский	<a href="#">194</a>	6.41	вблизи	<a href="#">29</a>	1.88	люлька	<a href="#">64</a>	5.45	залезать	<a href="#">70</a>	4.8
рухнуть	<a href="#">17</a>	1.79	деревянный	<a href="#">187</a>	3.24	вроде	<a href="#">14</a>	0.34	платформа	<a href="#">62</a>	1.11	пулемет	<a href="#">67</a>	4.05
показаться	<a href="#">16</a>	0.7	дозорный	<a href="#">179</a>	8.24	насчет	<a href="#">14</a>	0.29	забор	<a href="#">61</a>	2.35	дежурить	<a href="#">56</a>	4.72
раскачиваться	<a href="#">15</a>	4.07	спасательный	<a href="#">159</a>	6.23	внутри	<a href="#">13</a>	1.18	опора	<a href="#">58</a>	1.66	пржектор	<a href="#">55</a>	4.73
выситься	<a href="#">15</a>	4.03	метровый	<a href="#">157</a>	6.74	из-под	<a href="#">12</a>	0.04	барак	<a href="#">56</a>	3.97	наблюдатель	<a href="#">55</a>	2.65
взорваться	<a href="#">12</a>	2.42	телескопический	<a href="#">153</a>	6.93				эстакада	<a href="#">40</a>	4.61	взбираться	<a href="#">52</a>	4.83
кивнуть	<a href="#">11</a>	0.25	сотовый	<a href="#">147</a>	3.64				вышка	<a href="#">40</a>	4.01	снайпер	<a href="#">51</a>	4.53
сгореть	<a href="#">10</a>	1.76	обзорный	<a href="#">104</a>	5.27				антенна	<a href="#">39</a>	2.09	полезть	<a href="#">51</a>	3.92
ставиться	<a href="#">10</a>	0.24	осветительный	<a href="#">103</a>	6.0				колючка	<a href="#">37</a>	5.22	лезть	<a href="#">44</a>	2.78
загореться	<a href="#">9</a>	1.74	передвижной	<a href="#">101</a>	5.23				горка	<a href="#">36</a>	2.44	авария	<a href="#">44</a>	1.01
комплектоваться	<a href="#">9</a>	1.72	специальный	<a href="#">98</a>	0.31				лабаз	<a href="#">35</a>	6.62	взрыв	<a href="#">42</a>	0.92
качаться	<a href="#">8</a>	2.77	Буровая	<a href="#">95</a>	7.59				трубопровод	<a href="#">34</a>	1.27	арбитр	<a href="#">40</a>	3.64
закачаться	<a href="#">7</a>	4.25	строительный	<a href="#">93</a>	1.4				стремянка	<a href="#">32</a>	5.59	подъем	<a href="#">40</a>	0.62
сооружаться	<a href="#">7</a>	3.88	контрольный	<a href="#">92</a>	3.06				металлоконструкция	<a href="#">31</a>	3.17	взобраться	<a href="#">39</a>	5.42
затонуть	<a href="#">7</a>	2.97	пограничный	<a href="#">91</a>	4.08				крыша	<a href="#">31</a>	0.0	солдат	<a href="#">33</a>	0.57

<u>gen_modifier</u>	<u>6272</u>	<u>0.0</u>	<u>pp_obj_c</u>	<u>4971</u>	<u>0.3</u>	<u>object4_of</u>	<u>2492</u>	<u>0.1</u>	<u>pp_v</u>	<u>2075</u>	<u>0.0</u>	<u>object2_of</u>	<u>1710</u>	<u>0.1</u>
связь	<u>1863</u>	2.93	прыжок	<u>819</u>	6.52	построить	<u>106</u>	1.3	бассейн	<u>110</u>	1.77	стоять	<u>153</u>	1.4
тур	<u>471</u>	3.62	прыгать	<u>308</u>	5.57	устанавливать	<u>101</u>	0.31	море	<u>89</u>	0.26	ставить	<u>59</u>	0.28
оператор	<u>437</u>	3.22	прыгнуть	<u>152</u>	5.74	поставить	<u>83</u>	1.33	залив	<u>84</u>	2.43	поставить	<u>48</u>	0.54
станция	<u>91</u>	0.54	охота	<u>151</u>	2.84	установить	<u>65</u>	1.75	парок	<u>36</u>	0.1	страивать	<u>34</u>	1.05
спасатель	<u>86</u>	4.24	бассейн	<u>108</u>	1.74	поставлять	<u>51</u>	0.14	норвеги	<u>17</u>	7.87	строить	<u>27</u>	0.8
башня	<u>62</u>	1.79	кабан	<u>94</u>	5.21	строить	<u>42</u>	1.43	бинокль	<u>13</u>	3.05	изолировать	<u>24</u>	3.31
электропередача	<u>53</u>	3.99	нырять	<u>64</u>	5.26	соорудить	<u>34</u>	5.16	техас	<u>11</u>	3.21	торчать	<u>18</u>	2.24
аэропорт	<u>45</u>	0.51	спускаться	<u>56</u>	2.87	сносить	<u>24</u>	3.26	норвегия	<u>10</u>	0.31	гореть	<u>11</u>	0.13
ретранслятор	<u>42</u>	5.98	часовой	<u>54</u>	3.38	уничтожать	<u>23</u>	0.81				сносить	<u>10</u>	2.01
маяк	<u>29</u>	3.45	спуститься	<u>50</u>	3.28	сооружать	<u>20</u>	3.08	<u>pp_obj_v</u>	<u>1834</u>	<u>0.0</u>	виднеться	<u>9</u>	1.88
буровая	<u>25</u>	4.1	прыгун	<u>45</u>	6.28	закончить	<u>20</u>	0.37	поступать	<u>99</u>	1.4	пушить	<u>8</u>	6.44
телецентр	<u>24</u>	5.3	забор	<u>42</u>	1.82	страивать	<u>20</u>	0.28	поступление	<u>72</u>	1.21	взрывать	<u>8</u>	1.87
мегафон	<u>24</u>	3.54	провода	<u>36</u>	2.05	взорвать	<u>17</u>	3.86	поступить	<u>39</u>	0.7	взорвать	<u>7</u>	2.61
аэродром	<u>24</u>	2.0	парашют	<u>31</u>	3.4	захватывать	<u>15</u>	0.48	врезаться	<u>27</u>	3.05	возвышаться	<u>7</u>	1.2
шахта	<u>18</u>	1.1	стрельба	<u>31</u>	1.32	снести	<u>14</u>	3.24	путевка	<u>12</u>	0.2			
антенна	<u>18</u>	0.98	попрыгать	<u>24</u>	5.32	демонтировать	<u>13</u>	3.32						
-туры	<u>16</u>	6.35	наблюдатель	<u>24</u>	1.47	уничтожить	<u>12</u>	1.64						
пограничник	<u>16</u>	2.82	прожектор	<u>22</u>	3.45	перемещать	<u>12</u>	0.75						
терем	<u>15</u>	4.27	пулемет	<u>21</u>	2.4	возвести	<u>10</u>	2.74						
охранник	<u>15</u>	1.16	снайпер	<u>20</u>	3.21	передвигать	<u>10</u>	0.99						
-тура	<u>14</u>	6.16	стрелять	<u>20</u>	1.3	возводить	<u>10</u>	0.88						
радиостанция	<u>13</u>	0.72	слезать	<u>19</u>	4.56	окончить	<u>10</u>	0.29						
стадион	<u>13</u>	0.06	слезть	<u>19</u>	4.29	обстрелять	<u>9</u>	3.94						
снайпер	<u>12</u>	2.45	выстрел	<u>18</u>	0.89	теснить	<u>9</u>	3.61						
обозрение	<u>11</u>	1.7	спрыгнуть	<u>16</u>	3.55	подорвать	<u>8</u>	3.0						

# Пушить вышки

заберет еще 1 вышку, дракона, задавит бот с джанглером и это еще не полный список чем можно заняться

Paintik, 24 янв 2013

---

Пушат слабаки, тру папки фармятся до фулл закупа ЖВ

Riazlar, 24 янв 2013

#17

InSwift24 нравится это.

---

Всегда пушить и сносить вышки ,при этом нужно переносить "линию вардов" глубже в джангл противника полностью отрезая от фарма . Враг остается без мап контроля.

maddamage, 24 янв 2013

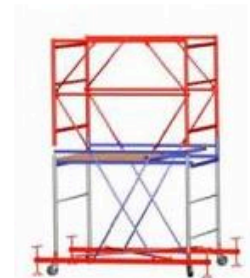
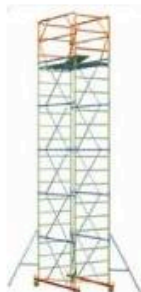
---

разожрался - пуш, бомжуешь - фармись

# Вышка-тура

Вышка-тура (передвижная вышка) — это металлическая сборно-разборная конструкция башенного типа. Ее основные элементы — плоские рамы-площадки и оснащенные лестничными ступенями. Из этих рам при помощи стяжек с втулками и специальных штырей собирают секции. Основанием конструкции служат четыре колеса и четыре винтовые опоры, установленные на деревянные подкладки. Для размещения рабочих и необходимых материалов на нужной высоте выполнения работ вышка-тура оснащается комплектом двух типов настилов: с решетчатой и сплошной поверхностями.

Высота передвижной вышки варьируется от 2,8 до 20,8 м. Чтобы обеспечить конструкции максимальную устойчивость, каждая вышка-тура высотой более 6,4 м снабжается стабилизаторами, прикрепленными к основной части с помощью хомутов.



<u>pp_c</u>	<u>1621</u>	<u>0.1</u>
пулемет	<u>137</u>	5.14
часовой	<u>108</u>	4.4
прожектор	<u>103</u>	5.75
автоматчик	<u>89</u>	7.48
антенна	<u>36</u>	2.0
охранник	<u>29</u>	2.14
колокол	<u>21</u>	2.64
пулеметчик	<u>17</u>	5.32
автомат	<u>17</u>	0.04
видеокамера	<u>14</u>	1.28
вертухаями	<u>12</u>	7.77
телекамера	<u>12</u>	3.58
шатриком	<u>11</u>	7.73
ретранслятор	<u>11</u>	4.39
парашют	<u>11</u>	1.95
талевой	<u>10</u>	7.38
трамплин	<u>10</u>	2.78
передатчик	<u>10</u>	1.65

<u>pp_dля</u>	<u>1600</u>	<u>0.2</u>
прыжок	<u>552</u>	5.97
наблюдение	<u>112</u>	1.69
ныряние	<u>81</u>	7.76
спасатель	<u>22</u>	2.32
бурение	<u>21</u>	2.08
радиомаяк	<u>15</u>	6.28
отстрел	<u>12</u>	3.52
скалолазания	<u>11</u>	4.54
антенна	<u>11</u>	0.29
стрелок	<u>9</u>	1.7
караул	<u>8</u>	2.12

<u>pp_на</u>	<u>1305</u>	<u>0.1</u>
крыша	<u>38</u>	0.31
вершина	<u>34</u>	0.67
кабан	<u>16</u>	2.72
холм	<u>15</u>	0.6
стадион	<u>14</u>	0.18
миде	<u>11</u>	7.66
горизонт	<u>11</u>	0.28
горе	<u>10</u>	0.52
вышка	<u>7</u>	1.6
шельф	<u>7</u>	1.31
топ	<u>7</u>	0.92
аэродром	<u>7</u>	0.26

<u>pp_obj_k</u>	<u>721</u>	<u>0.1</u>
направиться	<u>25</u>	1.61
вышка	<u>15</u>	2.71
приговаривать	<u>12</u>	2.59
приговорить	<u>11</u>	2.49
подъехать	<u>8</u>	1.8
зашагать	<u>7</u>	2.73
подъезжать	<u>7</u>	1.28
побежать	<u>7</u>	0.67

<u>pp_obj_ot</u>	<u>531</u>	<u>0.1</u>
метр	<u>42</u>	0.15
удаленность	<u>14</u>	2.87
помеха	<u>8</u>	0.53

<u>pp_po</u>	<u>504</u>	<u>0.0</u>
угол	<u>116</u>	1.07
периметр	<u>83</u>	3.83

<u>pp_obj_po</u>	<u>421</u>	<u>0.0</u>
экзамен	<u>30</u>	0.36
лекция	<u>25</u>	0.46
триангуляция	<u>21</u>	7.0
местоположение	<u>13</u>	0.68
препод	<u>12</u>	5.59

<u>pp_obj_pod</u>	<u>315</u>	<u>0.2</u>
лезть	<u>8</u>	0.36

<u>pp_obj_из</u>	<u>303</u>	<u>0.1</u>
вылететь	<u>21</u>	2.25
вылет	<u>18</u>	1.23
вылетать	<u>13</u>	1.7

<u>passive</u>	<u>263</u>	<u>0.1</u>
снабжать	<u>17</u>	1.73
оборудовать	<u>14</u>	0.66
оснащать	<u>12</u>	0.35

<u>pp_obj_do</u>	<u>250</u>	<u>0.1</u>
добраться	<u>25</u>	1.82
добираться	<u>23</u>	0.75
доходить	<u>20</u>	0.61

<u>pp_obj_co</u>	<u>199</u>	<u>0.1</u>
забор	<u>13</u>	0.15
шатер	<u>12</u>	2.4

<u>pp_co</u>	<u>152</u>	<u>0.1</u>
снайпер	<u>21</u>	3.39
спасатель	<u>17</u>	1.97

<u>pp_obj_o</u>	<u>135</u>	<u>0.0</u>
диплом	<u>36</u>	0.74
корочка	<u>7</u>	1.79

# варежка/рукавица

ruTenTen11 freqs = 25

варежка 6.0 4.0 2.0 0 -2.0 -4.0 -6.0 рукавица

и/или	7613	6776	0.2	0.2
крага	0	<u>100</u>	0.0	8.4
респиратор	0	<u>112</u>	0.0	7.8
спецодежда	0	<u>168</u>	0.0	6.3
нарукавник	0	<u>21</u>	0.0	6.3
унт	0	<u>23</u>	0.0	6.2
спецобувь	0	<u>17</u>	0.0	6.1
наплечник	0	<u>18</u>	0.0	6.1
полушубок	0	<u>38</u>	0.0	6.1
наколенник	0	<u>18</u>	0.0	5.7
фартук	<u>6</u>	<u>126</u>	2.1	6.5
варежка	<u>14</u>	<u>66</u>	4.1	6.4
телогрейка	<u>7</u>	<u>21</u>	4.0	5.7
прихватка	<u>15</u>	<u>30</u>	5.3	6.4
валенок	<u>105</u>	<u>131</u>	6.5	6.9
рукавица	<u>66</u>	<u>61</u>	6.3	6.2
перчатка	<u>1334</u>	<u>1237</u>	8.1	8.0
шапка	<u>851</u>	<u>376</u>	7.1	5.9
кисет	<u>19</u>	<u>7</u>	5.7	4.4
шапочка	<u>206</u>	<u>66</u>	5.9	4.3
носок	<u>1257</u>	<u>229</u>	8.1	5.6
шарф	<u>855</u>	<u>81</u>	8.6	5.2
шарфик	<u>44</u>	0	5.9	0.0
пинетками	<u>21</u>	0	6.5	0.0
носочков	<u>40</u>	0	7.1	0.0
носочек	<u>290</u>	0	9.0	0.0

object2_of	58	159	0.0	0.0
снимать	0	<u>72</u>	0.0	0.3

object4_of	2828	2713	0.4	0.3
стянуть	0	<u>28</u>	0.0	5.2
напрялить	0	<u>6</u>	0.0	4.9
сдернуть	0	<u>8</u>	0.0	4.5
стаскивать	0	<u>10</u>	0.0	4.0
надеваться	0	<u>6</u>	0.0	3.0
натянуть	<u>10</u>	<u>59</u>	3.0	5.5
натягивать	<u>20</u>	<u>51</u>	2.9	4.2
скинуть	<u>9</u>	<u>22</u>	2.4	3.7
шить	<u>110</u>	<u>212</u>	5.3	6.3
снять	<u>113</u>	<u>218</u>	2.9	3.8
одеть	<u>6</u>	<u>10</u>	2.4	3.1
надевать	<u>211</u>	<u>334</u>	4.6	5.2
сшивать	<u>11</u>	<u>16</u>	2.6	3.1
стащить	<u>9</u>	<u>8</u>	3.9	3.7
сшить	<u>7</u>	<u>6</u>	3.8	3.6
одевать	<u>66</u>	<u>51</u>	3.2	2.9
надеть	<u>15</u>	<u>9</u>	3.3	2.6
вязать	<u>359</u>	<u>46</u>	6.8	3.8
связать	<u>26</u>	0	3.5	0.0
захлопывать	<u>7</u>	0	3.8	0.0
разинуть	<u>16</u>	0	5.7	0.0
бует	<u>7</u>	0	5.7	0.0
раззявить	<u>10</u>	0	6.7	0.0
раззявливать	<u>12</u>	0	7.0	0.0
разевать	<u>75</u>	0	7.5	0.0

pp_из	477	491	0.4	0.4
двунитки	0	<u>9</u>	0.0	9.0
спилка	0	<u>6</u>	0.0	7.8
брезент	0	<u>31</u>	0.0	5.8
мешковина	0	<u>6</u>	0.0	4.0
шкурка	0	<u>6</u>	0.0	1.7
шкура	0	<u>17</u>	0.0	1.6
люфы	<u>9</u>	<u>12</u>	8.0	8.4
сизалья	<u>10</u>	<u>11</u>	7.7	7.8
шелк	<u>6</u>	0	0.3	0.0
овчина	<u>16</u>	<u>12</u>	5.1	4.7
шерсть	<u>80</u>	<u>49</u>	3.0	2.3
пряжа	<u>7</u>	0	1.5	0.0
мех	<u>81</u>	<u>26</u>	3.6	2.0
пух	<u>14</u>	0	2.2	0.0
замша	<u>6</u>	0	2.7	0.0
норка	<u>12</u>	0	2.8	0.0
флиса	<u>7</u>	0	5.5	0.0
мохер	<u>9</u>	0	5.8	0.0
микроволокна	<u>10</u>	0	6.6	0.0
микрофибры	<u>16</u>	0	6.9	0.0

# Значение определяется КОНТЕКСТОМ

You shall know a word by the company it keeps

(Firth 1957)

# МНОГОЗНАЧНОСТЬ ВНЕ КОНТЕКСТА?

Context is the only means to identify the meaning of a polysemous word. Therefore, all work on sense disambiguation relies on the context of the target word to provide information to be used for its disambiguation.

(Ide and Véronis 1998)



# Существуют ли значения слов?

Do meanings also exist outside the transactional contexts in which they are used? I would argue that they do not.

(Hanks 2013: 73)

# Нет смысла вне контекста?

Without context, words mean nothing.

Adam Kilgariff, Lexicom-2013

# Предметные области (domains)

- MUDDY WATERS
- MIXERS AND DIRECTIONERS

Little Mix



One Direction



# Прайминг

Вот список слов: *джаз, рок, поп, блюз*.

Сколько слов из него есть в «Толковом словаре живого великорусского языка»

В. И. Даля, изданном в 60-е годы XIX века?

(И. А. Рубанов)

# Прайминг

- Which flights **serve** breakfast?
- Does Lufthansa **serve** Philadelphia?

# The “zeugma” test

- Which flights **serve** breakfast?
- Does Lufthansa **serve** Philadelphia?
- ?Does Lufthansa serve breakfast and San Jose?
- Since this conjunction sounds weird,
  - we say that these are **two different senses of “serve”**

(Jurafsky 2013)

# В контексте

- *Чтобы не пришлось любимой плакать, крепче за баранку держись, шофер  
(via М. Ахметова)*
- *Переберем ребра графа Делоне  
(via С. Стенин)*
- *– Я пока без голоса, но лечу.  
– Выздоровливай! А куда летишь?*
- *Я кое-что вытянул из посудомойки*
- *Надо распечатать письмо*
- *Пора сдавать номер*

# Ситуационный контекст



**Сергей Волков**

2 ч. назад 

О многозначности слов.

"Ну что, сдаем номер?" -- спрашивает меня младший сын. И у мгновенно все холодеет внутри, потому что статьи в номер я еще не прочел, корректору не отправил, ничего не сверстано и не вычитано, картинки не найдены и так далее... А речь-то всего о том, что надо выезжать из гостиницы. Уф, перевожу дыхание.

У главных редакторов ежемесячных журналов свои тараканы. Интересно, у кого что еще такого же сорта?





1980



2010

Парень клеит модель в клубе

# Актуальность значений

двушка

перевод

энциклопедии

русский язык

[проверка орфографии](#)

Быть может, вы искали «девушка»?

## Толковые словари

[ДÉВУШКА](#), девушки, жен. 1. Лицо женского пола, достигшее половой зрелости, но еще не вступившее в брак. 2. Служанка, работница (разг. устар.).

[dic.academic.ru](#) › Толковый словарь Ушакова

[Показать все результаты по запросу «девушка» \(60 результатов\)](#)

## Толковые словари

[ДВУ́ШКА](#) -и; ж. Монета номиналом в две копейки.

[dic.academic.ru](#) › Большой толковый словарь русского языка. - 1-е изд-е: СПб.: Норинт

[Дву́шка](#). Ж. разг. Монета достоинством в две копейки.

[dic.academic.ru](#) › Толковый словарь Ефремовой

[ДВУШКА](#). ДВУ́ШКА, -и, жен. (прост.). Двухкопеечная монета.

[dic.academic.ru](#) › Толковый словарь Ожегова

(ср. Кронгауз 2013)

# Психолингвистические эксперименты

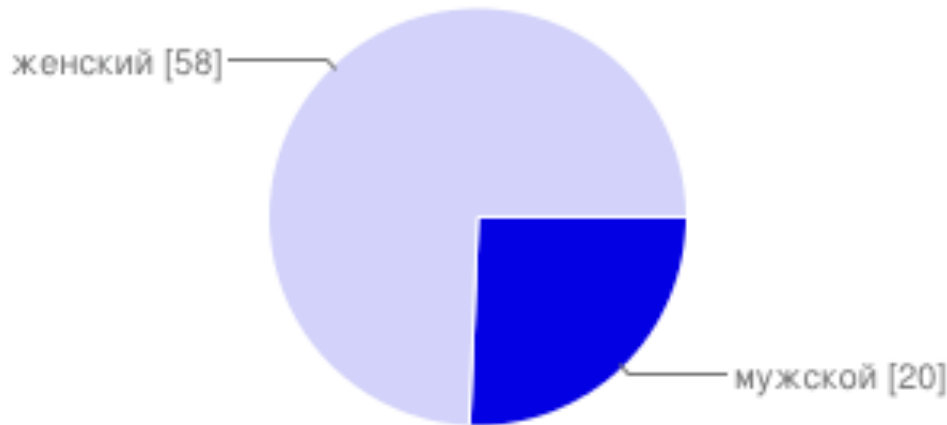
act	Freq	answer	Freq
pretense/performance	37	response	81
to perform	30	solution	18
to take action	16		
division	12		
a deed	3		

Эксперимент на свободные ассоциации  
(Durkin and Manning 1989)

# Эксперимент на ЛЛШ-2013

82 участника, средний возраст: 21 год

**пол**



мужской	<b>20</b>	26%
женский	<b>58</b>	74%

Благодарность ассистентам А.А.Лопухиной и П.К.Васильеву!

# Суть эксперимента

Даны слова:

*автомат, банан, бой, бумага, бюро, вид, визитка, действие, деталь, дипломат, картонка, книжка, крем, майка, мешок, пакет, пилот, таблетка, трубка, ящик.*

Любым способом укажите для каждого слова его значения, нумеруя их в том порядке, в котором они вам приходят в голову.

# *автомат: БТС*

1. Самодействующее устройство (аппарат, машина, прибор), производящее работу по заданной программе без непосредственного участия человека. // Разг. О телефоне-автомате.
2. Ручное автоматическое скорострельное оружие.
3. О человеке, действующем по выработанному шаблону, безучастно и механически совершающем что-л. *Это не человек, а автомат.\**

\*Здесь и далее **красным цветом** выделены словарные значения, не встретившиеся в эксперименте, или наоборот.

# *автомат*: эксперимент

оружие 93%

зачет 30%

коробка передач 24%

телефон 5%

...

# *банан*: Словарь Ефремовой

**I м.**

- 1.** Высокое тропическое и субтропическое травянистое растение с большими листьями...
- 2.** Плод такого растения.

**II м. разг.**

Юбка из разноцветных клиньев, по форме напоминающая банан I 2.



# банан: эксперимент

плод	100%
плавсредство	41%
двойка	12%
растение	10%
наушники	7%
единица	5%
...	



# банан 'двойка'

квартирант "солдат, демобилизованный по приказу, но еще находящийся в воинской части", козырь "солдат, служащий при штабе", курок "тайник", курковать "прятать что-либо", базис "обмен", базисовать "обмениваться с кем-либо чем-нибудь", чифанить "есть" и т.п.; в студенческом жаргоне: стёпа, степон, стипеша "стипендия". преп "преподаватель". банан "двойка". забананить

Лексика современных говоров (Лукьянова 1983)

# банан 'единица'

Nouns may be formed from numerals to denote quantity in specific meanings. Grades in Russian number from 5 (excellent) to 1 (fail). Grades are listed below with their official and slang names.

<b>Grade</b>	<b>Official names</b>	<b>Student slang</b>
5	пятёрка (отлично - <i>excellent</i> , очень хорошо - <i>very good</i> )	отл, пятák
4	четвёрка (хорошо - <i>good</i> )	хор
3	трóйка (удовлетворительно - <i>satisfactory</i> )	уд, трóяк
2	двóйка (неудовлетворительно - <i>unsatisfactory</i> )	пáра, неуд
1	<b>единица</b> ( <i>fail</i> )	<b>кол, банáн</b>

Russian Reference Grammar (Watzke and Sweigert 1997)

Поиск

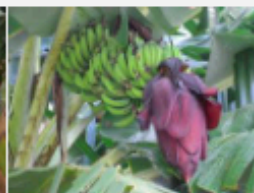
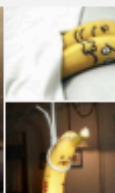
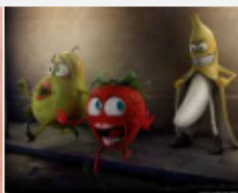
**Картинки**

Карты

Ещё ▾

Инструменты поиска

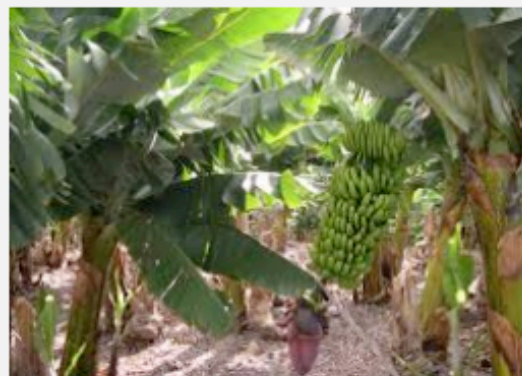
Безопасный поиск ▾



Банан Прикол

Банан Дерево

Банан Растение

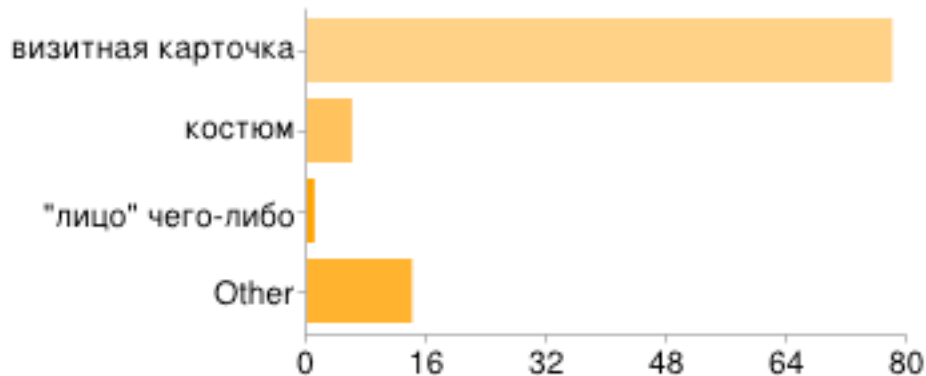


# *визитка*: словари

- Род полуфрака, кафтанчика особого покроя, придуманного щеголями, для утренних разъездов. (Даль)
- 1. Однобортный короткий сюртук...  
2. разг. = Визитная карточка (БТС)
- I Старинная мужская одежда...  
**II Небольшая кожаная - обычно мужская - сумочка на ремешке для документов, денег, сигарет, ключей и т.п.; барсетка.**
- III *разг.* Визитная карточка. (Ефремова)

# *визитка*: эксперимент

## **ВИЗИТКА**



**визитная карточка 78**

**костюм 6**

**"лицо" чего-либо 1**

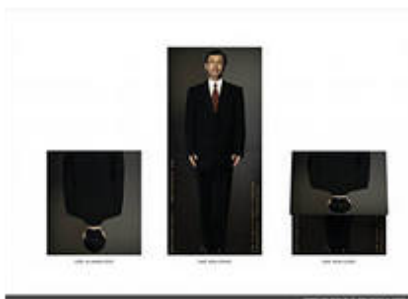
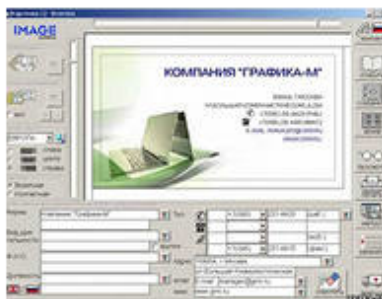
**Other 14**

Поищите также: [визитка образец](#), [визитка шаблон](#), [макет визитки](#)



КОСИМ  ЗАБИВАЕМ

Куряефф, наркоделец  
Угол Невского и Садовой



# *визитка* 'представление'

Яндекс

визитка 1



Найти

визитка 1 класса в стихах

визитка 10 класса сценарий

визитка 11 класса

визитка 1 класса

визитка 10 класса

визитка 11 класса сценарий

визитка 10 класса в стихах

визитка 1с

визитка 1+1

визитка 1 класса в стихах и песнях



# *КНИЖКА: БТС*

1. =Книга (1-2 зн.); небольшая книга. Телефонная к. К. стихов. Записная к.
2. Отдельный номер толстого (ежемесячного и т.п.) журнала.
3. В составе названий различных документов...  
Трудовая к. Зачётная к. Орденская к. Чековая к.
4. Разг. = Сберегательная книжка
5. Анат. Один из четырёх отделов желудка жвачных животных...

# КНИЖКА: эксперимент

книга	82%
документ	32%
электронная	11%
диван	5%
стол	1%
...	

# *май*ка: словари

- МАЙКА. см. *май* и *маять* (Даль)

- МАЙКА

1. Летняя трикотажная рубашка без рукавов или с короткими рукавами.

2. (зоол.). Бескрылый жук с длинным телом, питающийся растениями. (Ушаков)

- МАЙКА

1. Трикотажная, обычно нижняя, рубашка без рукавов и воротника [первонач. спортивная].

2. То же, что футболка (Ожегов)

# майка: эксперимент

одежда	94%
белье	35%
футболка	24%
оба	8%

пакет 30%

...

# *пакет*: БТС

- 1.** Упакованный в бумажную или иную обёртку какой-л. предмет (предметы); свёрток. // Бумажный кулёк для упаковки каких-л. предметов, продуктов и т.п. Продажа овощей в пакетах. Молоко в пакетах.
- 2.** Конверт с письмом официально-делового содержания. *Контрольный пакет* (акций)
- 3.** Комплект документов, официальных бумаг.
- 4.** Спец. Стопка ящиков или одинаковых деталей, строительных материалов и т.п., уложенных на специальный поддон для погрузки, перевозки и т.п.

# *пакет*: эксперимент

пластиковый 93%

услуг 15%

бумажный 8%

**данных 5%**

...

# *таблетка*: словарь Ефремовой

**I устар.**

**1.** Тонкая пластинка - обычно из слоновой кости - для записи чего-либо на память.

**2.** Дощечка, пластинка - обычно из обожженной глины - для письма.

**II** Твёрдая дозированная лекарственная форма, получаемая прессованием медикаментов.

**III** Небольшая, обычно двухсторонняя, икона, написанная на холсте, покрытом левкасом...

**IV разг.** Женский головной убор в виде небольшого плоского берета; шляпка-таблетка.

# *таблетка: эксперимент*

лекарство 100%

ключ 10%

«кряк» 7%

шляпа 7%

планшет 6%

...



# Эксперимент на факультативе для школьников при РГГУ и ВШЭ

Даны слова:

*батарея, ведро, кирпич, туника*

Любым способом укажите для каждого слова его значения, нумеруя их в том порядке, в котором они вам приходят в голову.

# *батарея*: БТС

1. Основное артиллерийское или миномётное подразделение из нескольких орудий.

*Зенитная батарея*

2. Укреплённая позиция для нескольких таких орудий.

*Был ли ты на батарее?*

3. Соединение нескольких одноптипных приборов, аппаратов, устройств, элементов в единую систему для совместного действия. *Аккумуляторная батарея, солнечная батарея*

- Об устройстве, служащем для обогрева помещения.

*Батарея парового отопления*

4. Шутл. Ряд каких-л. (обычно однородных) предметов.

*Батарея бутылок, подарков*

# *ведро: БТС*

**1.** Сужающийся книзу или цилиндрический сосуд с двумя ушками и дугообразной ручкой для ношения и хранения жидкостей, сыпучих веществ.

*Цинковое, эмалированное ведро. Ведро для мусора*

- О количестве, объёме (от 8 до 10 литров).

*Набрать ведро клюквы. Наловить ведро окуней*

**2.** Старинная русская мера объёма жидкостей, равная  $1/40$  бочки или 12 литрам (применялась до введения метрической системы)

# *кирпич: БТС*

**1.** Прямоугольный брусок из обожжённой глины, используемый для строительных работ.

*Груда кирпичей. Кладка в два кирпича.*

- собир. Необожжённый кирпич. Воз кирпича.

**2.** Предмет, имеющий форму такого бруска.

*Кирпич торфа. Кирпич чая. Кирпич ржаного хлеба.*

**3.** Разг. Дорожный знак с изображением прямоугольника, запрещающий въезд, проезд транспорта.

*Здесь не проедем: кирпич висит.*

# *туника*: БТС

1. **В Древнем Риме:** род белой шерстяной или льняной рубахи с короткими рукавами, носившейся под тогой.
2. =Тюник (тюника).
3. Зоол. Наружный слой хордовых животных.
4. Ботан. Наружный слой конусовидной верхушки растения.

**ТЮНИК**, -а; м.; **ТЮНИКА**, -и; ж. [франц. tunique] Театр.

Традиционный костюм классической балетной танцовщицы; удлинённая пышная лёгкая многослойная юбка и облегающий лиф.



## Как вы назовете этот предмет одежды? Открытое голосование

блузка	9	1.5%
кофточка	32	5.2%
платье	403	66.1%
сарафан	10	1.6%
туника	138	22.6%
другое (напишу комментарий)	18	3%

Проголосовало **610** человек. [Получить код](#)

# Игра в «шляпу»

- Чем мажутся – крем
- Жидкое мыло для душа – гель
- Плоский диван у врача на приёме – кушетка
- Что надеваешь, когда идешь в ванную – халат
- Некрасивая обувь, носят в плохую погоду – боты
- Такая расческа более старинная – гребенка
- Одежда, которую женщины и девочки носят на ногах, а на мальчиках не видела – туфли (Женя, 5 лет)
- Шейное украшение – бусы ожерелье кольцо
- Сыр на палочке с огурчиком – канапе
- Фотографическая... фото... что получается, когда снимаешь... что я держу в руках... – ??? (карточка)

(Иомдин 2014)

# Терминология vs. бытовое словоупотребление

«**Курок** – часть пистолета, на которую надо нажать, чтобы произвести выстрел.

**Мимоза** – растение, цветущее весной мелкими желтыми соцветиями, в некоторой степени символ 8 марта.

**Паранойя** – психическое заболевание, мания преследования.

**Плечо** – часть руки, присоединяющая руку к туловищу.

**Предплечье** – часть руки между локтем и плечевым суставом.

**Целлофан** – материал из пластмассы, используемый для изготовления пакетов.

**Шизофрения** – психическое заболевание, «раздвоение личности».

(Иомдин 2012)



# Различия идиолектов

43 произведения А. Марининой:  
*джерпер 46, пуловер 3*

131 произведение Д. Донцовоу:  
*джерпер 6, пуловер 232*

(Иомдин 2010)

# *Своеобычный*: Стругацкие

- ...прошел Выбегалло, я испытал к нему даже нечто вроде симпатии, потому что у него была своеобразная пшенная каша в бороде...
- ...но тут пришел доктор Р. Квадрига и заказал своеобразную бутылку рома...
- Лицо у Комова было непривычно озабоченное, ... без этой своеобразной, оскомину набившей ледяной сосредоточенности
- ...он без всяких предисловий и своеобразных в таких случаях ... шуточек доложил Тойво, что...

# Вышка: Солженицын (НКРЯ)

35 вхождений, из них

- 33 'сторожевая вышка'
- 1 'вышка церкви'
- 1 'высшая мера', с комментарием:

*Очень часто не помилование получали они, а вышку (так называют арестанты «высшую меру», они не терпят высоких слов и все называют как-нибудь погрубей и покороче)*

# *Вышка: Навальный (ЖЖ)*

17 вхождений, из них

- 13 'ВШЭ'
- 4 'смотровая вышка' (в одном тексте)
- 1 'телебашня'

# Альтернативные модели управления

- *Ольга предлагает отвезти тете Мане самовар (Г. Щербакова)*
- *Через два дня позвонили со студии и предложили приехать и заключить договор (Л. Вертинская)*
- *Он даже предложил мне давать Юре уроки французского языка и платить за урок тарелкой супа (В. Гроссман)*

(Iomdin and Iomdin 2012)

# Альтернативные модели управления

- *На пирожные он выменивал хлеб, муку, масло, пшено, табак – весь состав своего пайка, за исключением сахара: сахар он оставлял себе (В. Ходасевич, Диск).*
- *Но, как видно, руссы сильно желали выменивать на свои товары арабские монеты, диргемы, которые везде и во всяком значении имели большую ценность (С. Соловьёв, История России с древнейших времён)*

# Альтернативные модели управления

- *Он два дня не ел хлеба, затем выменял на хлеб большой фибровый чемодан*  
(В. Шаламов, Тифозный карантин).
- *Выменять ножик на удочку*  
(Словарь Ожегова и Шведовой).

# Многозначность относительна

- Родственные языки
- Диалекты
- Региональные варианты
- Социальные варианты  
(возраст, профессия, социальный статус, ...)
- Идиолекты



# Отношения лексем многозначного слова

- Общеязыковая многозначность
- Свободное варьирование
  - активное vs. пассивное владение
  - разное ранжирование в идиолектах
- Дополнительная дистрибуция:
  - банан* ‘двойка’ – ‘единица’
  - фуфайка* ‘поддевка’ – ‘свитер’ – ‘ватник’ – ‘футболка’
  - мимоза* ‘*Mimosa pudica*’ – ‘*Acacia dealbata*’

# Методы исследования

- Эксперименты
- Корпусные исследования
  - кластеринг в подкорпусах
  - нужна метаразметка!

# Возможные применения

- Поисквые запросы (ср. на Яндексе персонализированный поиск, таргетинг рекламы Крипта)
- Учебники и обучающие программы
- Динамические словари на основе корпуса
- Частотные словари значений

# Персонализированный поиск Яндекса: история запросов, но не данные о социолектахх и идиолектах



# Применение в лексикографии

Проспект Словаря бытовой терминологии: самые частотные слова, называющие бытовые предметы ("Топ-50")

(Иомдин и др. 2013)

# Источники словаря

- Разные подкорпусы НКРЯ
- Корпус RuTenTen
- База данных юридических текстов «Консультант Плюс»
- Блоги
- Логи запросов пользователей Яндекса
- Собственные данные анкетирования информантов

# «Топ 20» бытовых предметов

*карта*

*телефон*

*права*

*часы*

*книга*

*банка*

*регистрация*

*сеть*

*тетрадь*

*план*

*форма*

*компьютер*

*журнал*

*файл*

*ключ*

*стол*

*письмо*

*билет*

*дневник*

*марка*

# Многозначность в «топе»

? карта

? телефон

– права

? часы

? книга

– банка

– регистрация

– сеть

+ тетрадь

– план

– форма

+ компьютер

? журнал

– файл

? ключ

+ стол

? письмо

+ билет

? дневник

– марка



# Проспект СБТ. Топ-50

*белье, брюки, бумага, бутылка, ведро,  
газета, джинсы, диван, документ, зеркало,  
карта, карточка, ключ, книга, книжка,  
компьютер, конверт, коробка, кошелек,  
крем, кресло, кровать, куртка, лекарство,  
мешок, нож, ноутбук, обувь, одежда, пакет,  
пальто, паспорт, платок, платье, плеер,  
ручка, сапоги, стакан, стол, стул, сумка,  
телевизор, телефон, туфли, учебник,  
холодильник, чемодан, штаны, экран, ящик*

# Литература

1. Durkin, K., and J. Manning. Polysemy and the subjective lexicon: Semantic relatedness and the salience of intraword senses. *Journal of Psycholinguistic Research* 18.6 (1989): 577-612.
2. Firth, J. R. 1957. *A Synopsis of Linguistic Theory 1930-1955*.
3. Hanks, P. 2013. *Lexical Analysis: Norms and Exploitations*.
4. Iomdin, L. and B. Iomdin.  
[Valency ambiguity interpretation: what can and what cannot be done](#). In: Proceedings of the 5th International Conference on Meaning-Text Theory. Barcelona, September 8–9, 2011. Ed. by Igor Boguslavsky and Leo Wanner. Barcelona: University Pompeu-Fabra, 2011.
5. Jurafsky, D.  
[Word Meaning and Similarity. Word Senses and Word Relations](#).
6. Kilgariff, A. 2013. [Word senses: the computational response](#).

# Литература

7. Lapata, M. and F. Keller. 2007.  
[An information retrieval approach to sense ranking](#). In Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL, Rochester, NY).
8. Lesk, M. 1986.  
[Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone](#). In Proc. of SIGDOC-86: 5th International Conference on Systems Documentation, Toronto, Canada.
9. Mohammad, S and G. Hirst. 2006.  
[Determining word sense dominance using a thesaurus](#). In Proceedings of the 11th Conference on European chapter of the Association for Computational Linguistics (EACL, Trento, Italy).
10. Navigli, R. [Word Sense Disambiguation: a Survey](#). ACM Computing Surveys, 41(2), ACM Press, 2009, pp. 1-69.

# Литература

11. Watzke J. and J. Sweigert, Jr. 1997. The Russian Reference Grammar: Core Grammar in Functional Context.
12. Ide, N. and J. Véronis. Introduction to the special issue on word sense disambiguation: the state of the art // *Computational linguistics*. 1998. Vol. 24. No. 1. 2-40.
13. Апресян Ю.Д. Лексическая семантика. Синонимические средства языка. М., 1974 (2-е изд.: М., 1995).
14. Иомдин Б. Л.  
[Материалы к словарю-тезаурусу бытовой терминологии. СВИТЕР: образец словарной статьи](#) // Слово и язык. Сборник статей к восьмидесятилетию академика Ю. Д. Апресяна. Отв. ред. И. М. Богуславский, Л. Л. Иомдин, Л. П. Крысин. М., 2011, с. 392–406.

# Литература

15. Иомдин Б. Л.  
[О "неправильном" использовании терминов: может ли язык ошибаться?](#) // Смыслы, тексты и другие захватывающие сюжеты: Сб. ст. в честь 80-летия И. А. Мельчука. / Под ред. Ю. Д. Апресяна, И. М. Богуславского, Л. Ваннера, Л. Л. Иомдина, Я. Миличевич, М.-К. Л'Омм, А. Польгера. М., 2012. С. 233–251.
16. Иомдин Б. Л. Наивные представления о значениях слов в русском языке // [Антропологический форум](#), 2014, №19 (в печати).
17. Иомдин Б. Л. Многозначные слова в контексте и вне контекста (готовится к публикации).
18. Иомдин Б. Л., Лопухина А. А., Панина М. Ф. и др.  
[Маг вел мот: изменения в языке на материале бытовой терминологии](#) // *Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2013» (Бекасово, 29 мая – 2 июня 2013 г.)*. Вып. 12 (19): в 2 т. Т. 1. М., 2013, с. 311–324.

# Литература

19. Кронгауз М. А. Актуальный счет, или Двухка, трёшка, пятак... // Russian Linguistics. 2013, Vol. 37, Issue 2, pp. 157-173
20. Лукьянова Н.А. Лексика современных говоров как объект изучения. Новосибирск, 1983.

## Словари

БТС – Большой толковый словарь русского языка / Сост., гл. ред. С.А. Кузнецов. СПб., 1998.

Даль – Даль В. И. Толковый словарь живого великорусского языка. Тт. I–IV. М., 1956.

Ефремова – Ефремова Т.Ф. Большой современный толковый словарь русского языка. В 3-х т. М., 2006.

Ожегов и Шведова – Ожегов С.И. и Н.Ю. Шведова. Толковый словарь русского языка. М., 1992.

Ушаков – Толковый словарь русского языка / Под ред. Д. Н. Ушакова. М., 1934–1940.

# Ресурсы

[Sketch Engine](#)

[WordNet](#)

[Технология Крипта](#)

[Персонализированный поиск Яндекса](#)

**Спасибо за внимание!**