

СЛОВАРЬ «ЯЗЫКИ РУССКИХ ГОРОДОВ»: ПОДБОР ПРИМЕРОВ И ИНТЕРНЕТ

THE EXAMPLES FOR THE DICTIONARY OF THE VARIETIES OF URBAN RUSSIAN AND THE WWW

В. И. Беликов (VBelikoph@Yandex.ru)

Институт русского языка им. В.В. Виноградова РАН, Москва

В докладе рассматриваются функции иллюстраций и их подбор в традиционной лексикографии и роль Интернета при отборе иллюстративного материала для словаря «Языки русских городов».

Мы делаем здесь совместными усилиями самый настоящий словарь <...> Как правило, слова, которые мы обсуждаем, не попадают в толковые словари русского языка, но активно употребляются в разных регионах России.

*Преамбула форума «Городские диалекты»,
<http://forum.lingvo.ru/actualtopics.aspx?bid=26>*

Прежде чем говорить, как подбираются иллюстрации для словаря «Языки русских городов», необходимо остановиться на, казалось бы, ясных вопросах: какие иллюстрации вообще используются в толковых словарях, как они подбираются и каково их назначение.

В аннотациях к самым разным изданиям принято говорить «рассчитано как на специалистов, так и на самые широкие круги читателей». В применении к словарю *самые широкие круги читателей* следует понимать буквально. Специалистов же, чьи интересы призван удовлетворять словарь, я бы разделил на две категории: «(традиционные) лексикографы» и, условно говоря, «грамматисты». Не имея возможности вдаваться в детали, изложу их лексикографические интересы с очень сильным огрублением. «Лексикограф» заинтересован в наиболее точной фиксации лексического значения слова, во всем его многообразии, и в этом отношении он близок к рядовому пользователю словаря. «Грамматист» же всегда интересовался отражением в словаре парадигматики конкретной лексической единицы, а позднее и ее синтагматических свойств. Разумеется, реальные составители толковых словарей, будучи лингвистами, сделали большой шаг в направлении грамматикализации словаря. Достаточно сравнить первые статьи у Даля и в современных словарях. Словарь Даля начинается со статьи, посвященной букве А, а следующая, более короткая, «комплексно» описывает союз, частицу и междометие. В современных же толковых словарях служебным *a* уделено в десятки раз больше места, чем букве, при этом иногда (как в словаре Ожегова) буква *a* вообще обойдена вниманием. Не все слова интересуют «грамматиста» в равной степени, поэтому появились специализированные лексикографические издания, предназначенные исключительно для профессионалов. Специалист по семантике, вроде бы, должен иметь выраженные лексикографические интересы, однако в значении его интересуют скорее инварианты, чем идиосинкратические детали. В результате разница между *козлом* и *бараном* существенна для всех: для «грамматиста» они различаются акцентной парадигмой, для специалиста по семантике важны в первую очередь их переносные значения, сходные, но не идентичные. А что будет написано в словарных статьях *килька* и *тюлька*, теоретически должно волновать только «лексикографа»⁵.

Широкие круги читателей, напротив, обращаются к толковому словарю в основном с целью узнать *как правильно* (интересуясь буквой и значения, а также наличием ограничительных помет) или для уточнения значения малознакомого слова (в этом смысле различие между *килькой* и *тюлькой* ему интереснее, чем между *козлом* и *бараном*); из информации, близкой «грамматисту», наибольший интерес для значительной части непрофессионалов представляет проверка слуха, что «разрешили говорить *моё кофе*», а также акцентная парадигма «трудных» слов (не всегда, кстати, полностью выводимая из статей в толковом словаре); к статьям, толкующим «простые и ясные» слова типа *большой*, *если* или *стул* рядовой читатель обращается очень редко. Любопытство к иллюстративному материалу (если он не обладает самоценностью, как у Даля) целиком подчинено интересу к семантике.

Как известно, толковые словари снабжаются двумя типами примеров: словосочетаниями и предложениями; соотношение тех и других зависит от объема словаря. Ортогонально этому примеры делятся на модельные,

⁵ Волнения, впрочем, сугубо теоретические; так, в БТС эти слова толкуются, соответственно, как ‘мелкая промысловая рыба сем. сельдевых’ и ‘мелкая морская промысловая рыбка сем. сельдевых’.

порожденные самим лексикографом, и взятые из источников. Последние «вызывают, прежде всего, у самого лексикографа, большее доверие, чем придуманные» Но «ситуации, изображенные в литературных примерах, бывают подчас далеки от прототипических для описываемого слова» [Семенова 2003: 594]. Кроме того, вырванные из контекста цитаты могут быть чрезмерно длинны и, тем не менее, не вполне понятны. В результате в практику создания словарей «общего назначения» входит явно обозначаемое сокращение цитируемого примера и введение разъясняющих дополнений в квадратных скобках. Во все шире распространяющихся словарях, предназначенных для узкопрофессионального использования, допускается и «редактирование» цитат, точнее, создание модельных примеров, в которых все еще просматривается план содержания, принадлежащий, скажем, Пушкину (ср. приводимый С. Ю. Семеновой пример из семантического словаря РУСЛАН: *Цыгане шумной толпой по Бессарабии кочуют*⁶).

Разумеется, для решения сугубо лингвистических задач могут использоваться и общие словари, при этом оценка полезности иллюстративных примеров лингвистом и непрофессионалом может различаться. Интересный подход демонстрирует работа О. А. Митрофановой и С. А. Крылова [2006], отмечающих, что наиболее удачные из модельных примеров «согласуются с узусом, не перегружены элементами частного характера, и в то же время, позволяют демонстрировать спектр семантических / синтаксических валентностей слова». При анализе словаря Ожегова в формате базы данных Starling они констатируют, что с помощью модельных примеров удается выявить 228 синтагматических соседей глагола *говорить*; среди прочих определяются такие «предметно-предикатные» актанты, как *банальности, гадости, двусмысленности, дерзости, колкости, наглости, пакости*.

Однако неспециалист, даже если пользуется оцифрованной версией словаря (а в версии Ожегов—Шведова словарь включен в «Большую энциклопедию Кирилла и Мефодия»), имеет дело не с базой данных, поэтому ему явлены вовсе не типовые актанты глагола, а глагол *говорить* как наиболее тривиальная реализация функции *Oper*, имплицитно содержащаяся уже в соответствующем фрагменте толкования помянутых существительных: ‘избитое выражение’, ‘отвратительные слова’, ‘двусмысленное выражение’, ‘дерзкие слова’, ‘язвительное замечание’, ‘наглое выражение’, ‘отвратительные слова’. Много ли добавляет к семантике *отвратительных слов* и *дерзких выражений* уточнение, что их можно *говорить*?

С точки зрения лингвиста, интересующегося синтагматикой слов *говорить* и *год*, вполне оправдана оценка развернутых модельных примеров *Май холодный — год плодородный* и *Не пугай, говори с толком* как полезных (даны в статьях *май* и *пугать*). Но с точки зрения рядового пользователя оба они ущербны: при отсутствии специального задания и пословицу, и модель управления следует приводить в наиболее распространенных вариантах. Первый пример полезен лишь при иллюстрировании прилагательного *плодородный* (и, разумеется, в словаре пословиц как вариант к *Май холодный — год хлебодородный*), второй — в статьях *говорить* и *толк*, но только при сопоставлении с более нейтральным *говорить толком*⁷.

В отношении цитатных иллюстраций важно определить, что представляет собой «цитатный фонд». Известно, что наиболее насыщенный иллюстративным материалом 17-томный словарь был сильно идеологизирован, но за длительный срок его выпуска (начал он выходить при Сталине, а завершился при Брежнев) идеология менялась, и прессинг на подборку иллюстраций ослабел. Однако, кроме идеологии, вклад цитат из тех или иных авторов и произведений определяется не всегда ясными субъективными причинами, что хорошо видно при анализе размещенного в Интернете оцифрованной версии 2 изд. МАС. Пользуясь расширенным поиском Гугла по адресу расположения словаря (feb-web.ru/feb/mas), можно определить, откуда брались «примеры правильного употребления слова» [МАС 1981: 6], выяснить число словарных статей, где цитируются отдельные произведения. Верхнюю и среднюю часть списка занимают классики вперемешку с лауреатами Сталинских премий, изредка просто орденосносцами⁸, но при этом не ясно, почему цитат из

⁶ В случае редактуры «под стандарт» остается загадкой, почему цыгане *по Бессарабии кочуют*, а не *кочуют по Бессарабии*.

⁷ О распространенности вариантов *говорит* интернет-статистика. На 30.03.06 *Май холодный год хлебодородный* находится Яндексом на 79 сайтах, Гуглом — на 95, ... *год плодородный*, соответственно, на 20 и 16. На *говорить /+1 "+с толком"* Яндекс находит 85 примеров, но если исключить нерелевантные контексты типа *очень, говорит, с толку сбивает, собьет, мягко говоря, с толку* или *с чувством, как говорит, с толком, с расстановкой* (даже сохранив примеры типа *главное говорить с толком и расстановкой*), их окажется всего 45. На *говорить /+1 /толком* — 428 примеров, и, конечно, тоже есть шум. Важнее другое: в первой группе собственно литературных текстов только два: Шекспир (Сон в летнюю ночь, пер. Н. М. Сатина) и Л. Ф. Баум (Дороти и Волшебник в Стране Оз, переводчика не установил). Во второй (*говорить толком*) — тексты Помяловского (Очерки бурсы), А. К. Толстого (Князь Серебряный), Тургенева (Записки охотника), Лескова (Шерамур), Мельникова-Печерского (Бабушкины рассказы), А. Островского (Гроза, Последняя жертва), Чехова (Злоумышленник, Канитель. Радость), Л. Андреева (Савва), Куприна (Чудесный доктор), Ильфа и Петрова (Золотой теленок), Кассиля (Чаша гладиатора), Вампилова (Старший сын) и др.

⁸ Вот выборочный список цитируемых произведений (произведения и авторы-лауреаты Сталинской премии подчеркнуты): Степанов, *Порт-Артур*: 671 цитата; Л. Толстой, *Воскресение*: 574; Пушкин, *Капитанская дочка*: 537; Федин, *Необыкновенное лето*: 506; Николаева, *Жатва*: 481; Тургенев, *Отцы и дети*: 456; Саянов, *Небо и земля*: 399; Казакевич, *Весна на Одере*: 345; Гладков, *Энергия*: 336; А. Н. Толстой, *Сестры*: 305; Тургенев, *Накануне*: 309; Достоевский, *Братья Карамазовы*: 309; Федин, *Первые радости*: 283; Л. Толстой, *Воина и мир*: 275; Бабаевский, *Кавалер Золотой Звезды*: 269; Гладков, *Вольница*: 249; М. Горький, *Трое*: 244; Бубеннов, *Белая береза*: 230; Чехов, *Степь*: 210; Шолохов, *Тихий Дон*: 195; Пушкин, *Евгений Онегин*: 186; Гончаров, *Обломов*: 175; Панова, *Спутники*: 171; Ажаев, *Далеко от Москвы*: 156; Березко, *Мирный город*: 142; Л. Толстой, *Анна Каренина*: 136; Крымов, *Танкер «Дербент»*: 135; Чехов, *Три сестры*: 122; Бунин, *Деревня*: 122; М. Горький, *Коновалов*: 117; М. Алексеев, *Солдаты*: 115; Гладков, *Повесть о детстве*: 108; Березко, *Ночь полководца*: 106; Серафимович, *Город в степи*: 105; Горбатов, *Непокоренные*: 104; Фадеев, *Молодая гвардия*: 102; Задорнов, *Амур-батюшка*: 100; Тургенев, *Бежин луг*: 87; Гладков, *Цемент*: 84; Рыбаков, *Водители*: 78; М. Горький, *Мать*: 69; Павленко, *Счастье*: 66; Чехов, *Вишневый сад*: 65; С. Антонов, *Лена*: 60; М. Горький, *Васса Железнова*: 59; Лидин, *Две жизни*: 41; Лесков, *Соборяне*: 40; Лесков, *Очарованный странник*: 33; М. Горький, *Песня о Соколе*: 22; Василенко, *Звездочка*: 22; Лесков, *Некуда*: 14; Николаева, *Битва в пути*: 13; М. Горький, *На дне*: 9; Бунин, *Жизнь Арсеньева*: 4; Достоевский, *Бесы*: 1.

«Воскресения» в два раза больше, чем из «Войны и мира», почему у Гладкова наиболее известный «Цемент» вместе с удостоенной лавров «Вольницей» в сумме уступают «Энергии», почему Бунинская «Деревня» цитируется в 30 с лишним раз чаще, чем «Жизнь Арсеньева», то же с «Жатвой» и «Битвой в пути» Г. Николаевой; распределение 13 цитат из «Битвы в пути» по томам загадочно: в т. 1 — три цитаты, в томе 3 — одна, в т. 4 — девять, из которых семь — во второй половине тома (*упитать* и далее).

Задачи иллюстрирования в четырехтомном «Словаре русского языка» сформулированы так: «иллюстрации (...) служат дополнительным средством характеристики слова, указывают на сферу его употребления, подтверждают стилистическую характеристику слова, данную при помощи соответствующей пометы» [МАС 1981: 9]. Чуть иначе сказано в последнем фундаментальном издании, которое «продолжает и развивает богатые академические традиции словарного дела в России»: «Цитаты (...) служат для оправдания введения слова в словник Словаря, убеждают читателя в правильности толкования и стилистической классификации слова» [БАС 2004: 3, 29]. Если вдуматься, возложение на иллюстрации функции «дополнительного средства характеристики слова» совершенно неожиданно; буквальное понимание подталкивает к единственной интерпретации: из иллюстрации можно извлечь то, что составитель не стал по каким-то причинам включать в результат своей авторской работы над словарной статьей. В самом деле, толкования и ограничительные пометы — это суждения, порожденные в рамках науки, поэтому они могут быть как истинными, так и ложными, цитаты же представляют собой грамматически стилистически и семантически правильные тексты на русском языке, истинность которых в контексте словаря в общем случае не представляет интереса (я отвлекаюсь от примеров к терминам и подобной им лексике, таковые, впрочем, и иллюстрируются нечасто). Я бы назвал эту важнейшую функцию цитат *к о р р е к ц и о н н о й*; внимательному читателю они помогают оценить истинность предлагаемых лексикографом толкования и стилистических помет и часто убеждают в их правильности. Если толкование неаккуратно, вероятность убедиться в этом для читателя тем выше, чем больше число иллюстративных цитат. В самом деле, если примеры отсутствуют или толкование иллюстрируется лишь интроспективно подобранными автором словарной статьи словосочетаниями, для того, чтобы усмотреть ошибки в толковании, необходимо, как минимум, в точности знать значение слова. При наличии цитат в этом нет нужды.

Вот два показательных примера. В не содержащем иллюстраций словаре Ефремовой [2000] **анкерок** толкуется как «деревянный бочонок вместимостью от 16 до 50 литров, служащий для хранения пресной воды на шлюпках», а **аул** — как «селение на Северном Кавказе, в Средней Азии и Казахстане» и «жители такого селения». В БАС первое слово толкуется абсолютно идентично (приведены примеры из «Цусимы» Новикова-Прибоя и «Морской души» Л. Соболева, где анкерки с водой используются *на кораблях*), а второе — как «1. Селение на Северном Кавказе и в Средней Азии. (...) [следуют два примера] 2. *В знач собир.* О жителях такого селения. (...) *Этот аул перекочевал далеко за Голодную степь.* Пришв. Кащ. цепь».

В первом случае читатель вправе заподозрить и неточность в толковании, и пропуск пометы *проф.* (детальных помет типа *мор., мед.* или *авиа* в этом издании не предусмотрено), особенно если заглянет в «Словарь морских терминов», приложенный к «Цусиме»: **Анкерок** — бочонок в одно, два, три ведра и больше; употребляется для вина, уксуса и пр., а также для водяного балласта на шлюпках⁹. Однако вторая читательская претензия легко снимается такими примерами с достаточным стилистическим и хронологическим разбросом: *Но водка французская продается от 40 до 60 р. анкерок* (письмо Радищева А. Р. Воронцову из Тобольска об Ирбитской ярмарке, 15.03.1791), *Анкерок 3л, дуб – бук, розница 1500, опт 1400* (размещенный в Интернете «Прайс-лист на продукцию ООО "Русский дом Северо-Запад"»).

Во втором же случае читатель, особенно обратившись к более широкому контексту цитаты, недоумевает: куда же подевалось селение, когда его жители «далеко перекочевали»? В действительности никакого «селения» не было, ибо по-русски это слово не применяется ни к поставленным юртам кочевников, ни к перемещаемым в разобранном виде на лошадях или верблюдах (ср. более подходящее к этому примеру толкование **аул 2** из словаря Ушакова: «административная единица у некоторых кочевых тюркских народов Средней Азии»).

Важнейшая особенность словаря «Языки русских городов» заключается в том, что лексические единицы языка-объекта в общем случае неизвестны составителю, описывать их приходится методами, близкими к полевым, выявляя семантику текстов, признанных соответствующими региональному узусу (маркированному или нет стилистически), путем их анализа либо самим исследователем, либо с помощью информанта. В отличие от ситуации, с которой сталкивается полевой лингвист, составителю словаря язык-объект в целом хорошо известен; плюсы этого очевидны, минус же заключается в том, что при поверхностном знакомстве с новой лексемой¹⁰ исследователь может уподобиться героям известной корейской сказки, пересказанной Н. Г. Гариним-Михайловским¹¹: слишком рано решив, что он «все понял», ошибется и в толковании лексемы, и в ее региональной привязке, и в стилистическом статусе. Само собой разумеется, что модельные примеры и неполно паспортизованные цитаты для иллюстраций в рамках этого проекта непригодны.

Вполне очевидно, что вне Интернета и привязанных к нему оцифрованных текстовых массивов проект «Языки русских городов» был бы неосуществим. Создаваемый в его рамках словарь — продукт обработки

⁹ Можно добавить, что анкерок имеет сплюснутую форму, поскольку оба его днища эллипсоидны.

¹⁰ В определении этого понятия я следую за представителями школы Ю. Д. Аперсяна: «Лексема — слово в одном из его значений» [НОСС 2004: XXXIII].

¹¹ *Жили себе муж и жена, хорошие люди, но только никогда никого до конца не дослушивали и всегда кричали: «Знаем, знаем!».*

материалов, обсуждавшихся на форуме «Городские диалекты», но очередная версия словаря подвергается дальнейшему обсуждению в форуме, и в «готовых» статьях нередко корректируются и толкование, и региональный статус лексемы.

Для констатации нейтрального статуса лексемы важным является функционирование ее в официальных контекстах, использование в научном и публицистическом стилях, авторском тексте художественных произведений, регулярное появление в периодике без кавычек. Ср. такие примеры:

садить ‘сажать’ в разн. знач.: *Капустник — огород, где сажают капусту* (статья по диалектологии, издано в Кемерове, авторский текст); *Если человек <...> выдал много зарплатой или заработал много прибыли <...> [он] много, много, много работал. И такого социально активного господина мы садим [в тюрьму]? Где логика?* (главный федеральный инспектор по Пермской области, выпускник филологического факультета Пермского университета).

спадывать ‘спадать, падать, отваливаться, соскакивать’: *Вы знаете, кем становится и что думает о себе человек, лишенный ремня и звания, когда у него спадывают брюки?; У одной моей туфельки отрывается пуговица. Туфелька начинает спадывать с ноги, хлястик громко шлепает* (повести из екатеринбургского журнала «Урал», оба автора — местные уроженцы).

садоогород ‘садовоогородное товарищество и отдельный участок в нем’: *Комитет защиты садоогородов* (офиц. название общественной организации в г. Ижевске); *Администрация города сообщает об открытии сезонных пригородных маршрутов для поездки на садоогороды* («Удмуртская правда», заголовок).

вехотка ‘мочалка’: *Возможна контактно-бытовая передача вируса гепатитов в семейных очагах при несоблюдении гигиенических правил (использование в семье одной зубной щетки, вехотки, бритвенного прибора, маникюрных ножниц и т. д.)* (Сайт департамента охраны здоровья населения Кемеровской области); *Благодаря Н. К. Владимировой иркутяне познакомились именно на выставках с люфой, которая не только съедобна и хороша на вкус, но и при полном созревании и выделке является вехоткой* («Восточно-Сибирская правда»).

Для большинства типов лексики наиболее надежным источником иллюстраций является региональная периодика, однако в местных СМИ достаточно широко представлены идентичные тексты, в одних случаях это закономерность (например, издания ИД «Провинция» с частично общими материалами выходят в 40 регионах России от Калининграда до Хабаровска и от Архангельска до Астрахани и Ростов-на-Дону), в других — плагиат. Так, при региональной привязке лексемы **бокал** ‘высокая вместительная чайная чашка’ выявился идентичный текст, происхождение которого вряд ли установимо; фраза *Перед тем как идти в гости и для того, чтобы не захмелеть, надо выпить бокал хорошо заваренного с мятой зеленого или черного чая* присутствует в периодике Оренбурга (1998), Екатеринбурга и Калуги (2000), Костромы (2001), Ростова-на-Дону (2003).

Для иллюстрирования сниженной или детской лексики более пригодны форумы и сетевая литература, ср. названия игр типа салочки: *пятнашки* (Петербург), *ловитки* (Крым), *догоняшки* (Красноярск, Новосибирск), *догонялки* (Ульяновск), *крысы* (Саратов) и т. п.

В целом же иллюстрирование упомянутого словаря, как и вообще его создание, немыслимо без интернет-диалога с многочисленными участниками форума.

Список литературы

1. БАС: Большой академический словарь русского языка. Т. 1. А—Бишь. М., Т. 2. Благо—Внять. СПб: Наука, 2004—2005.
2. Ефремова Т.В. Новый словарь русского языка. Толково-словообразовательный. Т. 1—2. М.: Рус. яз., 2000.
3. МАС: Словарь русского языка: В 4 т. / Под ред. А.П. Евгеньевой. — 2-е изд., испр. и доп. — М.: Рус. яз., 1981—1984. — Т. 1—4.
4. Митрофанова О.А., Крылов С.А. «Типовой» контекст: случайность или закономерность? // Настоящее издание, 2006.
5. НОСС: Новый объяснительный словарь синонимов русского языка / Под общ. рук. Ю. Д. Апресяна. — 2-е изд., испр. и доп. — М.: Языки славянской культуры; Вена: Венский славистический альманах, 2004.
6. Семенова С.Ю. Примеры в компьютерном семантическом словаре: некоторые наблюдения над процессом подбора // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции Диалог’ 2003. М., 2003. С. 593—598.
7. Толковый словарь русского языка / Под ред. проф. Д.Н. Ушакова. — М.: ОГИЗ, 1935—1940. — Т. 1—4.