

ВЕРИФИКАЦИЯ СЛОВАРНЫХ ПОМЕТ ИНТЕРНЕТОМ

WWW AS A TOOL FOR VERIFICATION OF STYLISTIC LABELS IN A DICTIONARY

В. И. Беликов (WBelikoph@yandex.ru), *Институт русского языка РАН*

Словарные пометы предлагается разделить на два класса: связаны с типом коммуникативной ситуации и характеризующие языковую компетенцию говорящего. Адекватность вторых сильно зависит от разнообразных знаний лексикографа и Интернет может не только служить большим подспорьем при их верификации, но и помогает найти трудно интерпретируемые словоупотребления, которые требуют корректировки состава помет.

Среди ограничительных лексикографических помет можно выделить те, что связаны с типом коммуникативной ситуации (*груб.*, *пренебр.*, *шутл.*, *вульг.* и т. п.), и те, что в основном характеризуют языковую компетенцию говорящего (*обл.*, *спец.*, *проф.*, *жарг.*, *поэт.*, *высок.* и т. п.). Коммуникативная компетенция заключается в искусстве привязки собственной языковой компетенции к коммуникативной ситуации.

В этой связи русский словарь (лексику и фразеологию) я бы разделил на три группы, границы между которыми несколько размыты. 1. **Общий**, то есть потенциально известный всем носителям языка; не всё отсюда в равной мере попадает в активные словари любого носителя языка. Есть немало тех, кто практически не использует единицы, помечаемые как *вульг.*, *груб.* или даже *разг.-сниж.*, а тех, кто относительно регулярно прибегает к лексике, помечаемой как *книжн.* или *высок.*, вообще немного; попадание в одну из этих категорий — особенность в основном индивидуальная и опосредованно связана с социальными характеристиками говорящего. 2. **Периферийный** словарь активно используется в рамках трудно очерчиваемых и обычно численно крупных социальных общностей типа *молодежь*, *лица с низким образовательным уровнем*, *пожилая интеллигенция*. Открытость этих социальных групп, точнее открытость коммуникативных ситуаций, участниками которых они являются, ведет к тому, что большинство членов общества «экспонируется» их речевым произведениям, тем самым, пассивное владение соответствующими лексическими пластами присуще подавляющему большинству развитых (взрослых) носителей языка. 3. **Узкоспециализированные** словари активно используются лишь в профессиональных (или других узкоспециализированных) контекстах представителями легко очерчиваемых социальных групп (вкл. территориальные); тем, кто с ними не соприкасается, такая лексика даже пассивно оказывается известной лишь случайно (ср.: *волостель*, *комитативный*, *катух*, *нахзац*, *ходить по ништякам*, *давить ливер* — эти единицы, отсутствуют, например, в «Большом толковом словаре русского языка» (БТС), но таковы же и многие, в общие словари попадающие). Редко встречающаяся помета

нов. и стандартная, но не вполне ясная¹ *устар.* в действительности выносят соответствующие единицы за пределы первой, общей группы; обе они наглядно демонстрируют относительность Сосюровской дихотомии: на огромном социально-территориальном пространстве русского языка лексика и появляется, и исчезает далеко не одновременно.

Итак, стандартный носитель русского языка может не знать или знать нетвердо любую единицу, отнесенную к разряду периферийной или узкоспециализированной.

Рядовому пользователю редко приходит в голову выяснять в словаре, что на самом деле значат слова типа *мимоездом* или *минералка* (в БТС — разговорные), а также *здоровущий* или *милаха* (в БТС — фамильярные). Другие общеизвестные слова он если и ищет, то чтобы выяснить, включили их в словарь или «забыли». Иногда находит и узнает, что слова *застукать*, *облаять* (в знач. ‘грубо обругать’), или *срать* — слова хоть и разговорные, но сниженные, *лаять* (в знач. ‘бранить, ругать’), *мордобитие*, *жопа* или *зассать* — уже грубые, *ссать*, *целка* или *хрен* (в эвфемистическом употреблении) — вульгарные, а *блюдство*, *мудозвон* и *засрать* — и не просто вульгарные, но вдобавок еще и грубые. Удивится ли рядовой читатель словаря известной бесстыдности простановки помет у подчеркнутых

¹ Правильно было бы понимать под нею *устаревающее* (и так она расшифровывается, например, в Грамматическом словаре А. А. Зализняка и хороших орфоэпических словарях), но все «общие» толковые словари расшифровывают ее иначе. Но совсем не так, как склонен понимать адресат словаря, в том числе и профессионал: в вопроснике, адресованном 35 лингвистам, сотрудникам филологического факультета МГУ (от не имеющих ученой степени до академика, 19 — с кафедры русского языка), в частности, предлагалось выбрать между расшифровками *устарелое* (как во всех толковых словарях) и *устаревшее*; последнюю выбрали 28 чел., четверо колебались (подчеркнуты либо оба, либо ни одного). Из 33 аспирантов разных языковедческих кафедр МГУ за *устаревший* — 30 чел., двое, подчеркнув было *устаревший*, исправили на *устарелый*, один уклонился от ответа. Итак, «верный» ответ знает 4%, а 85% настаивают на «неправильном», комментируя иногда, что вариант *устарелый* давно устарел.

тых слов — неизвестно. Если удивится, то и к остальным пометам не будет относиться с полным доверием.

И будет прав. В вышеперечисленных случаях составители словаря характеризовали лишь коммуникативные ситуации, в которых реально могут использоваться достаточно широко известные слова, но для простановки иного рода помет требовалось знание и функционирования, и семантических тонкостей лексики. Интроспекция тут не всегда поможет. Как известно, *многие вещи нам непонятны не потому, что наши понятия слабы; но потому, что сии вещи не входят в круг наших понятий*. Лексикограф не исключение. Как верифицировать собственное понимание того, что не входят в круг твоих понятий. Раньше был единственный путь — работать с информантом. Теперь можно работать с Интернетом.

Верификация Интернетом помет, которые я отнес к числу характеризующих языковую компетенцию говорящего, легко убеждает, что в БТС многие такого рода пометы расставлены неверно. Многочисленные и разнородные примеры приводятся в докладе. Но анализ реального использования лексики ставит и сложные теоретические вопросы. Вот два примера.

Глагол **возопить** БТС характеризует как книжный. Анализ текстов показывает, что эта помета справедлива лишь отчасти. Поиск Яндексом дает «сайтов — не менее 4451», что не принципиально ниже чем у (ср. *завопить*: «не менее 4823») — значит, глагол достаточно частотен. В выдаче, признанной наиболее релевантной (первые 50 сайтов), действительно преобладают «книжные» словоупотребления: 5 — метаязыковые контексты, 21 — религиозные, 4 — навеяны религиозными (при этом стилистика может быть не вполне высокой: *Ну подставил [щеку] раз, потом два, ну а потом уже возопил "Еще раз дотронешься - убью суку". И хорошо если еще возопил, а не сразу пришел*). Но 20 вхождений книжными никак не назовешь, ср. из сетевой беллетристики: — *Ты что начальник, охерел!? — возопил Татарин <...>; "Машика, ангидрит твою!!" — грозно возопил Подшивалов <...>*.

Как быть? В сознании многих слово явно стилистически не маркировано, но доля «высоких текстов» при его употреблении на порядок выше, чем у «нормальной» немаркированной лексики. Семантика при «обычном» и «высоком» узусе идентична. Вряд ли это слово уникально с точки зрения распределения в современных текстах; при его описании требуется какое-то новое лексикографическое решение.

Лубок² (‘твёрдая накладка в повязке на место костного перелома’) имеет помету *мед.*, что априори довольно неожиданно, учитывая, например, польское *tubek* с тем же значением. Не восстанавливать же праславянскую специализированную медицинскую терминологию. У *шины* в том же значении ограничительных помет нет.

Поиск Яндексом на *шина + перелом* из первых 50 сайтов (выдача по релевантности) дает 48 медицинских (включая любые тексты, посвященные помощи травмированным людям) и 2 ветеринарных. При аналогичном поиске на *лубок + перелом* — 26 документов касаются ветеринарии, 12 — медицины, 4 — беллетристика (К. Чапек «Большая докторская сказка» — 3 сайта и В. Шаламов «Сухим пайком»), 3 — метаязыковые (словарные) сайты, по одному разу представлены сайты, где речь идет о травмах человека в археологическом и этнографическом контекстах и еще 3 сайта представляют собой информационный шум (*лубок* — изобразительный, а *перелом* никак с ним не связан).

Направив поиск в ветеринарную сторону, получаем: *лубок + (перелом/вывих) + (собака/кошка)* дает 428 сайтов, а *шина + (перелом/вывих) + (собака/кошка)* — всего 145. Вроде бы ветеринары предпочитают лубок шине. Но в «Ветеринарном энциклопедическом словаре» есть только *шина*, слово *лубок* вообще не используется. (поиск велся на портале www.rubricon.com), нет этого слова и в текстах, адресованных профессионалам.

Поиск Яндексом на *шина + перелом + фаланга* дает 265 сайтов (подавляющее большинство — медицинские), а на *лубок + перелом + фаланга* — 32, собственно медицинских нет.

Вывод ясен: лубок — если и термин, то ветеринарный. То есть это ветеринарный термин «для народа»? Но народ его не очень-то знает: опрос 27 студентов-лингвистов показал, что приспособление для иммобилизации при травме все спонтанно называют *шиной*, а *лубок* пассивно известен лишь восьми. Слово явно устаревает в среде молодого поколения носителей русского языка, но по текстам сделать такой вывод невозможно; при этом похоже что его представительность в текстах привязана в основном к жанру популярной литературы по ветеринарии.

Литература

Большой толковый словарь русского языка. — СПб.: Норинт, 1998.

Ветеринарный энциклопедический словарь. — М.: Советская энциклопедия, 1981.