

Секция «Иностранные языки и регионоведение»

Электронные языковые корпуса как источник новой информации о фразеологических единицах

Кротова Елена Борисовна

Аспирант

Московский государственный университет имени М.В. Ломоносова, Факультет иностранных языков и регионоведения, Москва, Россия

E-mail: elena_krotova@inbox.ru

С давних времён для преодоления языковых барьеров между разными народами составляются глоссарии и словари. Столь же давно существует проблема полноты и достоверности информации, содержащейся в них, её соответствия реальной языковой практике. С недавних пор появился новый, исключительный по важности для лексикографов ресурс, работа с которым помогает вывести словари на новый уровень: электронные языковые корпуса. О них и пойдёт речь в данном докладе. Задачей доклада является показать на примере одной идиомы, какие проблемы могут возникать при работе с корпусами, а также рассмотреть, какую новую информацию о данной идиоме они помогают получить.

Для начала рассмотрим, какая информация об идиоме *wie vom wilden Affen gebissen sein* = *снижен. тронуться; свихнуться; сбрендить (грубо)* [3], (букв. «быть укушенным дикой обезьяной») имеется в некоторых из существующих словарей.

В словаре [1] идиома объясняется следующим образом:

wie vom wilden Affen gebissen (salopp; ganz von Sinnen, verrückt) = фам. сумасшедший, неменяемый.

В онлайн-словаре [2], который составляется на основе имеющихся словарей с возможностью внесения изменений и дополнений носителями языка, даётся следующее объяснение: *ugs., salopp unberechenbar / leicht verrückt sein = фам., разг. непредсказуемый, сумасшедший, рехнувшийся.* В отличие от первого определения второе выражает меньшую степень признака (*leicht verrückt*).

В немецко-русских словарях идиома встречается в несколько другой форме. В словаре [3] идиома даётся в форме *wie vom wilden Affen gebissen sein = фам. тронуться, свихнуться; сбрендить.* В словаре [4] идиома, не отмеченная как таковая, встречается в двух примерах: 1) *ich bin doch nicht vom wilden [blauen] Affen gebissen — я пока ещё не сошёл с ума; меня на кривой не обведешь;* 2) *vom blauen Affen gebissen sein — быть не в своём уме.* В [1] и в [2] рассматриваемые варианты не указаны. Поиск данных вариантов идиомы в корпусах DeReKo [5] результатов не дал. Поиск в Google по форме "vom wilden blauen Affen gebissen" даёт 4 попадания, три из которых представляют собой цитаты из немецко-русских словарей. На запрос "vom blauen affen gebissen" около 240 попаданий, в то время как поиск в Google (на 21.12.2011) по форме "vom wilden Affen gebissen" даёт 198000 результатов. Можно сделать вывод, что при создании словаря идиом, ориентированного на активное употребление, варианты *vom blauen Affen gebissen sein* и *vom wilden blauen Affen gebissen* можно не учитывать.

Поиск данной идиомы по форме "vom wilden Affen gebissen" в DWDS-Kernkorpus [6] даёт только одно попадание, причём за 1927 г. В DWDS Zeitungskorpus [7] – два контекста, за 2001 и 2005 из одной и той же газеты (Der Tagesspiegel).

Понятно, что какие-либо выводы об особенностях синтаксического поведения идиомы, об её семантических и прагматических свойствах на основании 3 контекстов делать трудно.

Посмотрим, какие данные об этой идиоме можно получить из корпусов DeReKo [5]. Попаданий на эту идиому немного, всего лишь 21, включая немецкие, австрийские и швейцарские источники. Делать вывод о неупотребительности идиомы представляется, однако, неверным. Можно предположить, что стилистическая окраска идиомы накладывает ограничения на её употребление в публицистике (т.к. корпуса DeReKo в основном состоят из архивов газет).

Подробный анализ контекстов показал следующее. Было найдено 5 контекстов с языковой игрой (в проанализированных контекстах языковая игра возникает не за счёт модификаций идиомы, а за счёт специфического контекста, препятствующего реализации актуального значения идиомы), 4 контекста, в которых идиома реализуется в буквальном прочтении, 1 контекст с метатекстовой функцией, в двух контекстах нет достаточно информации, чтобы определить, в каком значении употребляется идиома. Т.о. получается, что для лексикографического анализа (в случае если не учитывается языковая игра) непригодны 12 контекстов, т.е. 57%, из которых 23% - контексты с языковой игрой.

С другой стороны, анализ оставшихся 12 контекстов (9 из DeReKo и 3 из DWDS) предоставляет, среди прочего, следующую информацию об идиоме:

- возможность употребления идиомы в сослагательном наклонении, часто в сочетании с глаголом *müssen* и/или частицей *ja* (ср. *Eine Strafe sei es nicht gewesen, "dann müssten wir ja vom wilden Affen gebissen sein"*. (Rhein-Zeitung, 03.03.2005);
- возможность употребления идиомы в качестве обособленного причастного оборота в форме *wie vom wilden Affen gebissen*;
- неупотребительность или значительно меньшая степень распространённости вариантов идиомы *vom wilden blauen Affen gebissen*“ и *vom blauen Affen gebissen*“.

Из данного краткого анализа видны некоторые недостатки и достоинства работы с корпусами. К недостаткам, в первую очередь, относится то обстоятельство, что корпуса (пригодные для фразеографических исследований) несбалансированны. Сбалансированные корпуса (напр., [6]) слишком малы для исследования идиом. Большая часть контекстов берётся из публицистики, хотя многие идиомы характерны скорее для устной речи. В результате невозможно получить относительно точные сведения о реальной употребительности идиомы. Особенности публицистического стиля также можно объяснить большое количество контекстов с языковой игрой. Тем не менее, корпуса предоставляют новые сведения об исследуемом языковом элементе, помогающие более точно описать его семантические, синтаксические и прагматические особенности.

Литература

1. Das grosse Wörterbuch der deutschen Sprache / Duden, 2000.
2. Redensarten-Index: <http://redensarten-index.de/suche.php>

3. Немецко-русский словарь живых идиом / Д.О. Добровольский. М.: Метатекст, 1997. – 201 с.
4. Большой немецко-русский словарь по общей лексике / Под общ. рук. Москальской О.И. «Русский язык-Медиа», 2004.
5. The Mannheim German Reference Corpus (DeReKo): <https://cosmas2.ids-mannheim.de/cosmasweb/>
6. DWDS Kernkorpus: <http://www.dwds.de/resource/referenzkorpora/>
7. DWDS Zeitungskorpora: <http://retro.dwds.de/?corpus=1&qu=Recherche&cc=TSP&sort=1&cp=1>