

RAFAEL GUZMÁN TIRADO
IRINA A. VOTYAKOVA
(ed.)

TIPOLOGÍA LÉXICA

GRANADA
2013

© LOS AUTORES.
RAFAEL GUZMÁN TIRADO, IRINA A. VOTYAKOVA (ED.).
TIPOLOGÍA LÉXICA.
ISBN: 978-84-938920-4-3. Depósito legal: Gr./ 560-2013
Edita: Jizo Ediciones.
C/ Maestro Vives, nº 8-4º G
Granada 18005
Maqueta: Atticus ediciones, atticusediciones@gmail.com
Imprime: Entorno Gráfico. Atarfe. Granada.

Impreso en España

Printed in Spain

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos www.cedro.org), si necesita fotocopiar o escanear algún fragmento de esta obra.

ПАРАЛЛЕЛЬНЫЕ КОРПУСА В СОСТАВЕ НАЦИОНАЛЬНОГО
КОРПУСА РУССКОГО ЯЗЫКА КАК ИНСТРУМЕНТ
ЛЕКСИКО-ГРАММАТИЧЕСКОЙ ТИПОЛОГИИ¹

Д.В. СИЧИНАВА

ВВЕДЕНИЕ. МОДЕЛИ И СТИМУЛЫ ПЕРЕВОДА

Национальный корпус русского языка (НКРЯ, <http://ruscorpora.ru>) имеет большое число практических применений в области сопоставительной грамматической и лексической семантики. Базой для межъязыковых (типологических и контрастивных) исследований в НКРЯ является прежде всего параллельный корпус — корпус, состоящий из оригинальных и переводных текстов, выровненных по предложениям. В составе Национального корпуса русского языка по состоянию на январь 2013 г. содержится 8 двуязычных параллельных корпусов: англо-русский (и, соответственно, русско-английский; далее наличие текстов, переведённых в обратном направлении, для каждой языковой пары не оговаривается), немецко-русский, французско-русский, испанско-русский, итальянско-русский, польско-русский, украинско-русский и белорусско-русский (см. публикации: (Добровольский, Кретов, Шаров 2005), (Добровольский 2009), (Sitchinava 2012) и др.). Совокупный объём этих корпусов превосходит 37 миллионов словоупотреблений, и развитие их продолжается. На стадии проекта находится ещё ряд двуязычных параллельных корпусов с участием русского (армянский, латышский, венгерский, татарский и другие). Для каждого языка, включенного в пары с русским, Корпус располагает морфологической разметкой, позволяющей осуществлять точный поиск по

1. Статья подготовлена при поддержке программы Президиума РАН № 36-П «Корпусная лингвистика» и гранта РФФИ № 12-06-33038 "Контрастивные корпусные исследования русских и французских глагольных категорий в поливариантных параллельных текстах"

словоформе, поиск по лексеме и/или набору граммем, а также различным сочетаниям этих параметров. Кроме того, имеется многоязычный корпус на базе 25 языков, в основном славянских, содержащий 9 параллельных текстов (но здесь грамматическая разметка имеется уже не для каждого языка); его объём – 5 миллионов словоупотреблений. Планируется создание поливариантных корпусов, в которые входит несколько разных переводов текста на один и тот же язык; в настоящее время ведётся работа над созданием русско-французского поливариантного параллельного корпуса.

Тексты выравниваются по предложениям. В основу выделения отрезков, между которыми устанавливается соответствие, берётся деление на предложения в оригинале (разбиение в переводе нередко бывает другим). В частности, предложения в переводе могут меняться местами, вклиниваться друг в друга и т. п.; в таком случае несколько предложений оригинала и перевода объединяются в пару минимальных соответствующих друг другу отрезков. Важно отметить, что художественный перевод отмечен высокой степенью вольности на разных уровнях, и при составлении и анализе параллельного корпуса это постоянно приходится учитывать. Причины этого разнообразны — общая установка ряда переводчиков, роль автора, участвующего в творческом пересоздании авторизованного перевода (в частности, он может и сам выступать переводчиком), а нередко и цензурные причины. Возможны также просто ошибки перевода, вызванные, в частности, интерференцией.

Для указания наличия той или иной вольности в переводе, – пропусков, добавлений, замены текста с изменением смысла, – в корпусе используется специальный уровень разметки – атрибут **loose**. Например, одно из значений этого атрибута, тег **omit**, используется для значимого пропуска в переводе (пример из «Смерти Ивана Ильича» Л. Н. Толстого во французском переводе М. Оффмана; выделен опущенный в переводе фрагмент)

`<se lang="ru">Было лучшее общество, и Иван Ильич танцевал с княгиней Труфоновой, сестрою той, которая известна учреждением общества «Унеси ты мое горе».</se>`

`<se lang="fr" loose="omit">Ivan Illitch dansa avec la princesse Troufonova, la propre sœur de l'illustre fondatrice de la société « Foin de peines ! ».</se>`

Использование параллельных корпусов для сравнительных исследований — предмет уже обширной исследовательской литературы, назовём, прежде всего, специальный номер журнала STUF под редакцией М. Сисоува и Б. Вельхли (Cysouw, Wälchli (eds.) 2007). При переводе лексики, как и грамматических конструкций (чёткую границу между ними не всегда можно провести), существенную роль играют **модели перевода** и **стимулы перевода**. Обычно для конструкции А в языке В имеется (реально представлено в текстах) несколько неслучайно повторяющихся соответствий – моделей перевода (translation patterns; ср. сборник (Hasselgård, Oksefjell (eds.) 1999). Практически можно утверждать, что, имея дело с переводными текстами и располагая сколько-либо большим корпусом, мы наверняка можем выделить несколько моделей перевода конструкции языка А в языке В.

Стимулом перевода назовем конструкции в языке оригинальных текстов, соответствующие определённой конструкции в языке перевода («что стимулирует появление конструкции X в переводе на язык В?»)

СООТВЕТСТВИЯ ГРАММАТИЧЕСКИХ КОНСТРУКЦИЙ

В работе (Сичинава 2011) представлены результаты статистического исследования стимулов перевода в английских и немецких текстах для конструкции с частицей *было*, типа *пошёл было, да вернулся*. Выявились любопытные устойчивые соответствия, причём не только и не столько грамматические, сколько лексические. Например, в немецком — частицы *eben* и *gleich* ‘равно’, в английском — наречия непродолжительного действия типа *for a while* ‘некоторое время’. Эти соответствия немало говорят о семантике русской конструкции (не имеющей точного соответствия в английском и немецком языках и в большинстве случаев лишённой какого-либо лексического стимула при переводе). То же верно и для соответствий иноязычных грамматических категорий в русском. Так, перфект прогрессива в английском языке (*I have been doing*) передаётся русским несовершенным видом (настоящего либо прошедшего времени) в актуально-длительном значении. Вместе с тем отмечено такое лексическое средство для передачи перфекта прогрессива, как глаголы типа *начать, решить*, указывающие на актуальность для текущего момента начального предела ситуации:

Perhaps people *have been celebrating* Bonfire Night early — it's not until next week, folks! [Joanne Kathleen Rowling. Harry Potter and the Sorcerer's Stone (1997)] [омонимия не снята]

Кажется, народ уже *начал праздновать* день Порохового Заговора — рановато, господа, он будет только на следующей неделе! [Дж. К. Роулинг. Гарри Поттер и Волшебный камень / М. Спивак].

Well, it's a new thing the boss *has been trying*. [Stephen King. The Lawnmower Man (1975)] Это новая методика, которую *решил испробовать* наш босс. [Стивен Кинг. Газонокосильщик / пер. не указан]

Для соответствий французского времени *passé antérieur* в русских текстах характерны такие дополнительные лексические средства, как *лишь, немедленно, сразу* (как в зависимом, так и в главном предложении):

Quand Passepartout *eut fini*, il se trouva calme et comme soulagé. [Jules Verne. Le tour du monde en quatre-vingt jours (1872)]

Отведя душу, Паспарту *сразу* успокоился и почувствовал большое облегчение. [Жюль Верн. Вокруг света за восемьдесят дней / Н. Габинский, Я. Лесюк]

Плюсквамперфект в испанском языке имеет в ряде употреблений значение аннулированного результата. При его помощи может переводиться не только русское *было*, но и прошедшее время от русского глагола с противоположным значением:

Напряжение сразу вдруг разрядилось. Все *вспомнили* о павианах. [А. Н. Стругацкий, Б. Н. Стругацкий. Град обречённый (1972)]

La tensión desapareció al momento. *Habían olvidado* ('забыли') a los babuinos <...> [Arkady Strugatsky, Boris Strugatsky. Ciudad maldita / Justo E. Vasco]

ЛЕКСИЧЕСКИЕ СООТВЕТСТВИЯ

Исследование моделей и стимулов перевода в реальном корпусе является, в частности, инструментом верификации утверждений о «языковой специфичности» тех или иных лексем, их «переводимости» и т. п. Повидимому, наиболее «специфичной» с семантической точки зрения лексике каждого языка будет отвечать наибольший разброс моделей и стимулов перевода – поскольку какое-либо «главное», подходящее ко всем или почти всем контекстам соответствие в другом языке отсутствует. Так, лексемы, относимые некоторыми исследователями к «концептам» так называемой «русской языковой картины мира» (что бы это ни значило), такие, как *тоска*, *душа* или *удаль*, демонстрируют весьма широкий разброс моделей при переводе на английский язык (и, соответственно, стимулов при переводе с английского языка).

Например, каждый из перечисленных английских переводов слова *удаль* в параллельном корпусе пока встретился по одному разу: *abandonment*, *assurance*, *violence*, *effrontery*, *jockeying*, *bravado*, *daring*, *fearless confidence*. При этом характерно, что в переводах с английского на русский слово *удаль* и его производные встречаются исчезающе редко. Английское слово *soul* переводится практически однозначно как русское *душа* (кроме фразеологизмов типа *poor soul*), в то время как русская *душа* передается значительно разнообразнее: так, в контекстах типа *в глубине души* на месте этого слова выступает английское *mind*. Интересен пример с обратным распределением (*soul* в английском тексте и *сознание* в русском):

And, all through that silent drive back to Green Street, the *souls* of both of them revolved a single thought: 'Why, oh! why should I have to expose my misfortune to the public like this? [John Galsworthy. In Chancery (1920)]

И все время, пока они молча ехали на Грин-стрит, у обоих *в глубине сознания* неотступно вертелась одна и та же мысль: "Почему, ах, почему мне приходится вот так выставлять напоказ мои невзгоды? [Джон Голсуорси. В петле / М. Богословская]

Разнообразие переводных моделей и стимулов для слова *тоска* чрезвычайно велико (перечисляем не только переводы с помощью существительного): *gloom*, *dismay*, *grief*, *agony*, *excruciating feelings*, *distracted moods*, *anguish*,

ache, despair, woe, loneliness, distress, frustration, desperation, blue, misery, nuisance, sadness, nostalgia, yearning, longing, lust, missing, glumly, sulkily, wistfully. В то же время параллельный «спектр» английского слова *yearning*, предлагаемого, например, А. Вежицкой в качестве наиболее приемлемого, хотя всё же и приблизительного, соответствия русской *тоске*, на проверку существенно смещён в сторону слов с иной семантикой, скорее страстного желания как такового (*желание, жажда, жадность, тянуть, томление, томный, томилло желание*).

При переводе с русского на французский язык слово *тоска*, помимо ряда устойчивых субстантивных соответствий (прежде всего *tristesse* и *angoisse*, а также *ennui, désespoir, mélancolie, chagrin, tourment, nostalgie*) нечасто, но систематически передаётся также перестройкой конструкции (*в тоске и слезах – en larmes de douleur, ужас и тоска – terreur mortelle, туная тоска – stupeur obtuse et chagrine*) или сочетаниями синонимов (*от тоски – d'ennui, de tristesse, тоска без конца – tristesse et nostalgie sans fin*).

Корпус является также эффективным средством исследования распределения близких синонимов. Например, русские *влажный* и *мокрый* и их английские объекты, по данным корпуса, распределяются следующим образом в соответствии с типом объекта:

(Ладони, руки) *damp, wet, moist* (от пота) => *влажный*

(Трава, земля, песок) *damp, humid, wet, soaking* => *влажный, мокрый*

(Одежда) *damp, humid, soaking* => *мокрый*.

Единичные отклонения от этой схемы связаны с авторским стилем переводчика: *Trees, forced by the damp heat, found too little soil for full growth, fell early and decayed: creepers cradled them, and new saplings searched away up*. [William Golding. *Lord of the Flies* (1954)]

Мокрой жарой деревья толкало в рост, но на хилом слое почвы они не заживались, рано валились и гнили; их обнимали лианы, сквозь них пробивались новые ростки. [Уильям Голдинг. Повелитель мух / Е. Суриц]

В русском языке *мокрая жара* — сочетание на грани приемлемости; в 230-миллионном одноязычном корпусе НКРЯ оно встретилось только один раз в контексте *ню-йоркская «мокрая жара»*, где оно оформлено как элемент «чужой речи» и, очевидно, также является дословной передачей такого же или схожего английского выражения.

Таким образом, параллельный корпус позволяет быстро получать большое число примеров реальных переводческих решений, принятых носителями языка при создании перевода, и анализировать выявленные соответствия в лексике и грамматике. Задача их интерпретации (оценка неслучайности соответствий, интерференции и ошибок, анализ контекста, выделение различных значений и т. п.), разумеется, остаётся делом исследователя; тем не менее объём предварительной черновой работы при помощи параллельного корпуса может быть существенно сокращён.

БИБЛИОГРАФИЯ

- ДОБРОВОЛЬСКИЙ, Д.О., КРЕТОВ, А.А., ШАРОВ, С.А. (2005): Корпус параллельных текстов: архитектура и возможности использования. // *Национальный корпус русского языка: 2003-2005*. М., Индрик, С. 263—296.
- ДОБРОВОЛЬСКИЙ, Д.О. (2009): Корпус параллельных текстов в исследовании культурно-специфичной лексики. // *Национальный корпус русского языка: 2006—2008. Новые результаты и перспективы*. Спб., Нестор-История, С. 383—401.
- СИЧИНАВА, Д.В. (2011): Комплексное исследование одноязычного и параллельного корпусов в грамматических исследованиях // *Труды Международной конференции «Корпусная лингвистика—2011»*, СПб., СПбГУ, С. 316—332
- CYSOUW, M. & WÄLCHLI, B. (eds.) (2007): *Parallel Texts. Using Translational Equivalents in Linguistic Typology. Theme issue in Sprachtypologie & Universalienforschung STUF 60.2*.
- HASSELGÅRD, H. & OKSEFJELL, S. (eds.) (1999): *Out of Corpora: Studies in Honour of Stig Johansson*. Amsterdam — Atlanta, GA: Rodopi.
- SITCHINAVA, D. (2012): Parallel corpora within the Russian National Copus // *Prace Filologiczne*, LXIII, 2012, S. 271—278