

Добровольский Дмитрий Олегович

Институт русского языка им. В. В. Виноградова РАН,
Россия, 119019 Москва, ул. Волхонка, 18/2

Институт языкознания РАН,
Россия, 125009, Москва, Большой Кисловский пер., 1

Стокгольмский университет,
Швеция, SE-10691, Стокгольм, ул. Сванте Аррениуса, 2
dobrovolskij@gmail.com

Корпусный подход к исследованию фразеологии: новые результаты по данным параллельных корпусов*

Для цитирования: Добровольский Д. О. Корпусный подход к исследованию фразеологии: новые результаты по данным параллельных корпусов. *Вестник Санкт-Петербургского университета. Язык и литература*. 2020, 17 (3): 398–411. <https://doi.org/10.21638/spbu09.2020.303>

В статье обсуждаются возможности использования параллельных корпусов в исследовании фразеологии. Результаты проведенного анализа подкрепляют положение о том, что обращение к корпусным данным не только существенно улучшает кросс-лингвистические описания, но и меняет само представление о специфике фразеологии как особого фрагмента лексической системы языка. В частности, материал корпусов параллельных текстов позволяет выявить лингвоспецифические особенности идиом, традиционно считавшихся полностью эквивалентными. В качестве материала анализа используются параллельные подкорпусы Национального корпуса русского языка (НКРЯ): прежде всего немецко-русский, а также русско-немецкий. В отдельных случаях привлекается материал основного корпуса НКРЯ. Помимо задачи выявления сходств и различий между семантически близкими идиомами немецкого и русского языков с помощью корпусного материала, в статье решаются и вопросы общего характера, в частности вопрос о границах понятия межъязыковой эквивалентности и способах его операционализации. Представляется необходимым выделить как минимум два различных типа эквивалентности: эквивалентность на уровне системы языка и эквивалентность на уровне текста. Метод данного исследования может быть охарактеризован как основанное на корпусных данных последовательное сопоставление отдельных пар коррелирующих фразеологизмов разных языков с целью выявления различий между ними, в том числе семантических, синтаксических, прагматических и сочетаемостных. Как правило, речь идет при этом о различиях, относящихся к периферии значения. Часто фразеологизм одного из сопоставляемых языков обнаруживает ограничения на образование грамматических форм; это могут быть, в частности, видо-временные ограничения русских глагольных идиом. К причинам межъязыковых различий относится также степень стертости метафоры, лежащей в основе фразеологизма, и степень его употребительности. Результаты такого анализа могут найти применение в двуязычной лексикографии.

Ключевые слова: фразеология, параллельный корпус, русский язык, немецкий язык

* Статья написана при поддержке РФФИ, проект 17-29-09154 и частично 18-012-90025.

Введение

Основная цель данной работы — показать на конкретных примерах, как использование корпусных данных, в особенности данных параллельных корпусов, может модифицировать и уточнить традиционные представления о межъязыковой эквивалентности в области фразеологии. Мы исходим из основанного на предварительных наблюдениях предположения, что обращение к корпусным данным не только существенно улучшает контрастивные описания, но и меняет само представление о специфике фразеологии — прежде всего идиоматики — как особого фрагмента лексической системы языка.

Достижения Валерия Михайловича Мокиенко в области сопоставительной фразеологии широко известны. Достаточно вспомнить ставшую классической книгу «Славянская фразеология», впервые вышедшую в 1980 г., а затем дополненную и переработанную [Мокиенко 1989]. Надеемся, что предлагаемое вниманию читателей, и прежде всего юбиляра, исследование, концентрирующееся на сравнении фразеологизмов немецкого и русского языков, станет достойным развитием в изучении этой области лингвистики.

Изложение следует такому плану: после описания объекта исследования, его базового метода, а также лежащих в основе метода общих принципов работы с текстовыми корпусами мы обратимся к анализу одной немецкой идиомы, не имеющей в русском языке стандартного системного эквивалента, и сопоставим системные представления об отсутствии эквивалентности с корпусными данными. В следующем разделе статьи анализируются фразеологизмы немецкого и русского языков, которые изначально считались полностью эквивалентными. Задача этого раздела — с помощью изучения корпусных данных выявить лингвоспецифические особенности таких единиц. В заключительном разделе статьи мы остановимся на некоторых вопросах развития теории фразеологии, имеющих и лексикографические приложения.

Объект и метод исследования

Объект предлагаемого исследования — фразеология немецкого и русского языков в сопоставлении. Поскольку в качестве материала исследования используются данные параллельных корпусов Национального корпуса русского языка (НКРЯ), особенность проводимого в работе сопоставительного анализа состоит в его эмпирической направленности. Иными словами, анализ основан не столько на теоретических представлениях о природе межъязыковой эквивалентности в области фразеологии, сколько на данных параллельных корпусов НКРЯ с немецким языком: русско-немецкого и, прежде всего, немецко-русского. В специальной литературе выделяются следующие основные виды работы с корпусом:

- исследования, направляемые корпусом (*corpus-driven*);
- исследования, основанные на корпусе (*corpus-based*);
- исследования, пользующиеся корпусом как источником иллюстративного материала (*corpus-illustrated*, такой принцип использования корпуса называется также *corpus-informed*).

Подробнее об этом см. в первую очередь [Tognini-Bonelli 2001], а также среди прочего [Hunston, Gill 2000; McEnery et al. 2006; Tummers et al. 2005; Копотев 2014]¹. Хотя выделение этих типов признается не всеми исследователями (в том числе и потому, что границы между ними часто нарушаются в конкретных исследованиях), сама идея неоднородности подходов к работе с корпусом представляется разумной и помогает определить метод анализа в каждом конкретном случае. Так, если исследователь использует корпус как источник иллюстративного материала (подход *corpus-illustrated*), его задача состоит прежде всего в отборе контекстов, которые могут претендовать на роль парадигмальных примеров. Данный принцип актуален, в частности, в словарной работе. Не вызывает сомнений, что словарь, в котором иллюстративный материал черпается из корпуса, характеризуется большей объективностью, чем словарь с иллюстративными примерами, основанными на языковой интуиции или случайно оказавшимися в картотеке лексикографа. Это, естественно, не означает, что корпусный принцип работы с иллюстративным материалом освобождает лексикографа от принятия решения в каждом конкретном случае. Преимущество данного принципа скорее в том, что он расширяет возможности выбора и позволяет определить степень употребительности той или иной формы, исходя из объективных данных, снижая тем самым роль интроспекции.

Исследования, направляемые корпусом (*corpus-driven*), могут рассматриваться как противопоставленные способам работы с корпусом как источником материала. При данном подходе сама область исследования задается корпусом. Так, в сфере грамматики сам набор базовых категорий, значимых для исследуемого языка, формируется на основе корпусных данных². Существенно, что при этом подходе рабочая гипотеза не формулируется заранее, а извлекается из корпуса. Анализ полученных данных становится первым шагом исследования, ведущего к теоретическим обобщениям. Иными словами, реализуется принцип «от данных к теории», а не «от теории к ее подтверждению или опровержению с помощью анализа эмпирических данных».

Между этими в известной степени противоположными подходами к корпусному материалу располагается вид работы с корпусом, который в [Tognini-Bonelli 2001] был назван *corpus-based*. В случае если исследование ориентировано на такой подход, исходной точкой является некоторое представление об устройстве изучаемого феномена (часто сформулированное в виде гипотезы или предварительных допущений), а обращение к корпусу позволяет проверить истинность этих представлений на эмпирическом материале. Результатом анализа в нормальном случае является получение нового знания об устройстве изучаемого явления. В основе предлагаемого исследования лежит именно этот вид работы с корпусом, хотя и не в полной мере. Строгое следование принципу *corpus-based* предполагает количественный анализ, что пока — вследствие ограниченного объема параллельных корпусов — плохо применимо к фразеологии, особенно если в фокусе внимания находятся идиомы, а не коллокации или синтаксические фраземы. Объем русско-немецкого корпуса — 5213939 слов, а немецко-русского — 3916879 слов (на мо-

¹ Пользуясь случаем, выражаю искреннюю благодарность Михаилу Вячеславовичу Копотеву за советы и замечания, касающиеся представления видов работы с текстовыми корпусами.

² «The corpus-driven approach <...> aims to derive linguistic categories systematically from the recurrent patterns and the frequency distributions that emerge from language in context» [Tognini-Bonelli 2001: 87].

мент поиска — март 2020). Из порядка 2000 наиболее употребительных в современном немецком языке идиом, описываемых в разрабатываемом словаре «Deutsch-russische Idiome online»³, в немецко-русском параллельном корпусе представлено только 325 единиц. Количество контекстов употребления каждой из них крайне немногочисленно, лишь в редких случаях обнаруживается более десяти вхождений конкретной идиомы. Таким образом, использование приемов количественного анализа в данной работе не имеет смысла, что не мешает, однако, получить результаты, значимые для сопоставительного изучения фразеологии. Это может быть проиллюстрировано на следующем примере.

Немецкая идиома *von Mund zu Mund gehen* имеет в русском языке, казалось бы, идеальное соответствие: 'переходить/передаваться из уст в уста'. И действительно, в немецко-русском корпусе есть контексты, в которых представлена данная пара идиом:

- (1) Tausend freudige Ausrufungen *gingen von Mund zu Mund*. [Novalis. Heinrich von Ofterdingen (1800)].

Тысячи радостных восклицаний *переходили из уст в уста*. [Новалис. Генрих фон Офтердинген (1800) (В. Микушевич, З. Венгерова, С. Ошеров, 1914–1960)].

Встречаются, однако, и случаи, в которых такой способ перевода не используется, причем дело не в том, что переводчик принял решение не переводить *von Mund zu Mund gehen* с помощью «переходить/передаваться из уст в уста», а в том, что это в принципе не допускается нормами русского языка:

- (2) <...> *da ging es von Mund zu Mund*, er habe sich mit einem jungen, bildschönen Fräulein aus zwar armer, aber adliger Familie verlobt. [Gustav Meyrink. Der Golem (1914)].

<...> в тот самый день пронесся слух, что он обручился с одной молодой и необычайно красивой девушкой из бедной, правда, но аристократической семьи. [Густав Майринк. Голем (1914) (Д. Выгодский, 1922)].

Причина невозможности употребить идиому *переходить/передаваться из уст в уста* в русском переводе примера (2) состоит в том, что эта русская идиома, в отличие от ее немецкого коррелята *von Mund zu Mund gehen*, обнаруживает запрет на безличное употребление. Эти и подобные отличия в употреблении сопоставляемых фразеологизмов могут быть обнаружены при обращении к корпусному материалу. Очевидно, что хороший двуязычный словарь должен фиксировать особенности такого рода. Таким образом, применение предлагаемого метода может оказаться полезным не только для получения нового знания о фразеологии, но и для составления фразеологических словарей.

Сам метод может быть охарактеризован как основанное на корпусных данных последовательное сопоставление отдельных пар коррелирующих фразеологизмов разных языков с целью выявления различий между ними, в том числе семантических, синтаксических, прагматических и сочетаемостных.

³ Описание проекта доступно на сайтах http://wvonline.ids-mannheim.de/idiome_russ/index.htm и http://www.europhras.org/index.php?option=com_content&view=category&layout=blog&id=36&Itemid=91&lang=de. Создаваемый корпусный словарь базируется на данных корпуса DeReKo. Наши актуальные работы в области сопоставительной фразеологии немецкого и русского языков ориентированы на словник, положенный в основу этого лексикографического проекта.

Об одной «безэквивалентной» идиоме

В этом разделе мы на примере немецкой идиомы *ein weites Feld* и способов ее перевода на русский язык рассмотрим один достаточно общий вопрос. В исследованиях по фразеологии часто выделяется разряд безэквивалентных единиц, т. е. фразеологизмов, не имеющих в сопоставляемом языке стандартных переводных эквивалентов. Само понятие эквивалентности нуждается в уточнении. Если понимать под эквивалентностью полное совпадение сопоставляемых единиц двух разных языков по всем релевантным для реального употребления параметрам, то «безэквивалентными» окажутся едва ли не все фразеологизмы. Если же относить к примерам отсутствия эквивалентности только те случаи, когда смысл, выражаемый фразеологизмом одного языка в конкретном контексте, не может быть передан на другой язык никакими способами, то безэквивалентных единиц практически не найдется.

Традиционно к безэквивалентным фразеологизмам относят такие выражения исходного языка, которые не имеют в другом языке хорошего фразеологического коррелята (т. е. фразеологизма, максимально близкого по актуальному значению и желательного по внутренней форме), а к полностью эквивалентным — единицы, возводимые к одному источнику, практически тождественные по значению, образной основе и характеризующиеся максимальным сходством компонентного состава, т. е. параллелизмом семантической и лексической структуры. Ср., например, выражение *метать бисер перед свиньями* и его немецкий коррелят *Perlen vor die Säue werfen*. Тем не менее считать данные идиомы эквивалентными можно лишь весьма условно. В исследованиях последних лет показано, что идиома *Perlen vor die Säue werfen* претерпела в своем историческом развитии существенные семантические изменения [Filatkina 2014], так что ее современное значение существенно отличается от значения русской идиомы *метать бисер перед свиньями*.

Таким образом, вопрос, ответ на который должен быть получен с помощью корпусных данных, — это вопрос о границах понятия эквивалентности и способах его операционализации.

Обратимся к материалу. Идиома *ein weites Feld* (букв. «широкое поле») не имеет в русском языке стандартного — фиксируемого нормативными словарями — фразеологического соответствия и, исходя из традиционных представлений, должна быть, видимо, отнесена к разряду безэквивалентных. В «Немецко-русском фразеологическом словаре» данная идиома представлена двумя словарными статьями: «**das ist ein (zu) weites Feld** это слишком обширная область (*сразу всего не охватишь*)» и «**das ist noch ein weites Feld** в этой области предстоит еще многое сделать» [Бинович, Гришин 1975: 171]⁴.

Самое раннее вхождение идиомы в НКРЯ представлено примером (3) и подтверждает словарную информацию об отсутствии в русском языке образного эквивалента.

- (3) „Das ist *ein weites Feld*, von dem ihr da redet“, sagte ich, „und da steht der menschlichen Erkenntnis ein nicht unwichtiger Gegenstand gegenüber“ [Adalbert Stifter. *Der Nachsommer* (1857)].

⁴ В более современном и сопоставимом по объему немецко-русском фразеологическом словаре [Шекасюк 2010] эта идиома не представлена.

— Вы затронули *большую тему*, — сказал я. — Тут перед нами явление очень важное. [Адальберт Штифтер. *Бабье лето* (С. К. Апт, 1999)].

Однако в переводе известного романа Теодора Фонтане «Эффи Брист», где идиома *ein weites Feld* встречается семь раз, она последовательно переводится с помощью русской идиомы *темный лес*. Ср. несколько примеров употребления этой идиомы из НКРЯ (4), где контексты оригинала [Theodor Fontane. *Effi Briest* (1894–1895)] сопоставляются с их русскими переводами [Теодор Фонтане. *Эффи Брист* (Ю. Светланов, Г. Эгерман, 1960)].

- (4) а. <...> jedenfalls wollen wir darüber nicht streiten; *es ist ein weites Feld*. «<...> Во всяком случае, не будем об этом спорить: *это — темный лес*». б. Glaube mir, Effi, *das ist auch ein weites Feld*. «Поверь мне, Эффи, *это темный лес*». в. Aber lassen wir das, oder wie dein Papa immer sagte: „*Das ist ein zu weites Feld*“. «Ну, хорошо, оставим это. Как говорит твой папа, „*это темный лес*“. г. <...> und Briest sagte ruhig: „Ach, Luise, laß... *das ist ein zu weites Feld*“. «<...> а Брист спокойно сказал: „Ах, оставь, Луиза... *Это уже совсем темный лес*“».

В романе эта идиома выступает как элемент языкового портрета одного из персонажей — отца Эффи. Ее многократное употребление несет художественную нагрузку и должно, таким образом, единообразно воспроизводиться в переводе. Эквивалент, предложенный переводчиками, представляется очень удачной находкой. Отметим, что именно роман «Эффи Брист» сделал идиому *ein weites Feld* широко известной. Еще большую популярность она приобрела в связи с выходом в свет в 1995 г. романа нобелевского лауреата Гюнтера Грасса «*Ein weites Feld*». Грасс сознательно отсылает читателя к роману Фонтане — своего любимого писателя. Данную интертекстуальную отсылку можно интерпретировать в том смысле, что роман Грасса начинается там, где кончается «Эффи Брист»: контекст (4г) — это заключительный фрагмент романа Фонтане⁵.

Насколько нам известно, роман «*Ein weites Feld*» на русский язык пока не переведен. Возможно, это не в последнюю очередь связано с трудностями, которые вызывает перевод его названия. Дословный перевод «Широкое поле» вряд ли уместен (хотя именно так роман представлен в русскоязычной «Википедии»), так как отсутствие в русском языке такого фразеологизма не позволит читателю соотнести название романа с его содержанием, не говоря уже о потере столь значимой отсылки к «Эффи Брист». Можно было бы воспользоваться в качестве названия идиомой *темный лес*, и это было бы более приемлемо с литературоведческой точки зрения, так как позволило бы сохранить столь важные для автора интертекстуальные связи. Однако такое решение было бы оправданно только при условии, что от русского читателя ожидается хорошее знание немецкой литературы XIX века. Только если читатель знает, что «Эффи Брист» в переводе Ю. Светланова и Г. Эгерман кончается словами «это уже совсем темный лес», он в состоянии в полной мере понять намерение автора и оценить место романа Грасса в истории немецкой литературы.

Какие выводы можно сделать из этого наблюдения относительно понятия межъязыковой эквивалентности? Во-первых, идиомы, не имеющие в другом языке хороших эквивалентов на уровне языковой системы (т.е. единиц, практи-

⁵ Имеется в виду, конечно, не продолжение сюжета, а определенные имплицитные автором параллели между Германией конца XIX и конца XX вв.

чески идентичных по значению и образной основе), могут тем не менее обнаруживать весьма удачные эквиваленты на уровне конкретного текста; ср. *ein weites Feld* и *темный лес*. Таким образом, о наличии или же отсутствии эквивалентности уместно говорить не вообще, а лишь применительно к перспективе, в которой осуществляется сопоставление. Это может быть как системно ориентированная, так и текстоцентрическая перспектива. Данные, которые предоставляет в распоряжение исследователя корпус, однозначно направлены на текст. Во-вторых, удачное соответствие в конкретном тексте совсем не обязательно окажется столь же удачным во всевозможных контекстах употребления и, соответственно, не всегда оказывается пригодным для включения в двуязычный словарь.

Так, при обращении к другим контекстам становится очевидным, что идиомы *ein weites Feld* и *темный лес* могут выступать в качестве коррелятов далеко не во всех случаях. Это объясняется спецификой значения каждой из этих идиом, мотивированной различиями в их внутренней форме. Поскольку образная составляющая является частью актуального значения, различные метафорические основы этих идиом не могут не сказываться на их функционировании. Представление о чистом поле, которое лишено ориентиров и потому вызывает у человека определенную неуверенность, и представление о темном лесе весьма различны. По этой причине идея неясности и невозможности предсказать будущее, в которой идиомы *ein weites Feld* и *темный лес* совпадают, мотивированы различными факторами: открытость и, следовательно, неопределенность будущего, фиксированная во фразеологизме *ein weites Feld*, непрозрачность, а значит и неизвестность для выражения *темный лес*. Эти различия становятся очевидными при анализе корпусных примеров. Первые контексты с идиомой *темный лес* в основном корпусе НКРЯ относятся к середине XIX в.

- (5) Антон также глядел на дряблую избушку свою, бог весть, о чем он думал: чужая голова — *темный лес* [Д. В. Григорович. Антон-горемыка (1847)].

Употребить в переводе этого примера идиому *ein weites Feld* невозможно.

Итак, эквивалентность на уровне системы представляет собой значимый параметр сопоставительного описания состава и структуры фразеологии. Эквивалентность на уровне текста значима для литературоведения (в том числе для изучения интертекстуальных связей) и переводческих штудий⁶. Хорошему двуязычному словарю важно найти некоторое равновесие между этими понятиями и выработать функционально ориентированный подход, т. е. подход, учитывающий как системные факторы, так и способы перевода в тех или иных контекстных условиях; подробнее об этом см. [Добровольский 2011].

Квазиэквиваленты во фразеологии

Данный раздел посвящен выявлению лингвоспецифических особенностей квазиэквивалентных идиом немецкого и русского языков, т. е. тех, которые близки и по актуальному значению, и по внутренней форме, и по компонентному составу. На отдельных примерах будет проверена следующая гипотеза: различия между

⁶ Из последних работ об использовании параллельных корпусов в переводоведении см. [Doval, Sánchez Nieto 2019].

структурно и семантически близкими фраземами, в том числе восходящими к одному источнику, могут быть выявлены на уровне отдельных (часто маргинальных) семантических признаков, узуса, встроенности в контекст, степени профилированности метафоры и пр.

Обратимся к материалу немецко-русского параллельного корпуса НКРЯ. Идиома *aus Leibeskräften* имеет в русском языке целый набор стандартных эквивалентов: *что есть/было мочи, что было сил, изо всех сил, изо всей силы*. Действительно, в ряде контекстов из корпуса это немецкое выражение переводится с помощью одного из этих эквивалентов.

- (6) <...> dort *aus Leibeskräften* die Glocke zieht und, wenn die Alte herauskommt, mit falscher Freundlichkeit fragt, ob hier vielleicht Herr und Madame Spucknapf wohnen, worauf man mit großem Gekreisch davonrennt... [Thomas Mann. Buddenbrooks (1896–1900)].

<...> *изо всей силы* дергать колокольчик, с притворной учтивостью спрашивать у отворившей дверь старушки, здесь ли проживают господин и госпожа Плевок, и затем с гиканьем улепетывать [Томас Манн. Будденброки (Н. Ман, 1953)].

- (7) Sagst du, daß sie mit den Beinen strampelte? *Aus Leibeskräften!* Aber es nützte ihr nichts, das Sterbesakrament mußte sie haben [Thomas Mann. Der Zauberberg (1924)].
— Ты говоришь, она брыкалась? — *Изо всех сил!* Но все было напрасно, она вынуждена была причаститься святых тайн [Томас Манн. Волшебная гора (В. Станевич, 1959)].

Встречаются, однако, примеры, в которых ни один из этих эквивалентов не используется.

- (8) Sie nahm singend vom Lokal Besitz, begriff nicht gleich die Lage, sang *aus Leibeskräften* weiter, verstummte dann verlegen [Friedrich Dürrenmatt. Justiz (1985)].

Студенты начали с пением разбредаться по залу, не сразу сообразили что к чему, *продолжали драть глотку* и лишь потом замолкли смущенно [Фридрих Дюрренматт. Правосудие (С. Фридлянд, 1988)].

- (9) Sämtliche „Suitiers“, darunter Doktor Gieseke und Peter Döhlmann, verabschiedeten ihn am Bahnhofe und überbrachten ihm Blumen und sogar Zigarren, wobei sie *aus Leibeskräften* lachten... in der Erinnerung ohne Zweifel an all die Geschichten, die Christian ihnen erzählt hatte [Thomas Mann. Buddenbrooks (1896–1900)].

Местные suitiers в полном составе, среди них доктор Гизеке и Петер Дельман, явились на вокзал, поднесли Христиану цветы и даже сигареты, причем все хохотали *до упаду*, видимо, вспоминая истории, которые он им рассказывал [Томас Манн. Будденброки (Н. Ман, 1953)].

Такой выбор способов перевода не случаен. Интенсификаторы *что есть/было мочи, что было сил, изо всех сил, изо всей силы* сочетаются преимущественно с глаголами физического действия, в то время как немецкое выражение *aus Leibeskräften* отличается гораздо более широкой сочетаемостью. Сочетания русских интенсификаторов *что есть/было мочи, что было сил* и подобных с глаголами типа *петь, смеяться* (а также *ругаться* — ср. пример 10 из русско-немецкого параллельного корпуса) нежелательны.

- (10) — А я так слышал, что третьего дня у Катерины Ивановны он *отдельывал* меня на чем свет *стоит*, — вот до чего интересовался вашим покорным слугой [Ф. М. Достоевский. Братья Карамазовы (ч. 1–2) (1878)].

Ich aber habe gehört, daß er mich vorgestern bei Katerina Iwanowna *aus Leibeskräften schlechtgemacht hat*. Da sieht man, wie sehr er sich doch für meine Wenigkeit interessiert [Die Brüder Karamasow (1.–2. Teile). Fjodor Dostojewski (Hermann Röhl, 1924)].

В русском языке выбор интенсификатора зависит от семантического класса глагола; ср. *петь во все горло, смеяться/хохотать до упаду, ругать на чем свет стоит*. В корпусе встречаются единичные контрпримеры, т. е. сочетания *смеяться изо всех сил* и *хохотать что было мочи*.

- (11) Над князем она, говорят, *смеется изо всех сил*, с утра до ночи... [Ф. М. Достоевский. Идиот (1868–1869)].

Allerdings *macht sie sich*, wie mir gesagt wird, über den Fürsten vom Morgen bis zum Abend *aus Leibeskräften lustig*... [Fjodor Dostojewski. Der Idiot (Hermann Röhl, 1981)].

- (12) Офицер, не помня себя, бросился на нее; около Настасьи Филипповны уже не было ее свиты; приличный господин средних лет уже успел стушеваться совершенно, а господин навеселе стоял в стороне и *хохотал что было мочи* [Ф. М. Достоевский. Идиот (1868–1869)].

Der Offizier, vor Wut außer sich, stürzte sich auf sie; ihr Gefolge hatte Nastasja Filippowna nicht mehr um sich; der anständige Herr in mittleren Jahren war bereits völlig verschwunden, und der angeheiterte Herr stand ein wenig abseits und *lachte aus Leibeskräften* [Fjodor Dostojewski. Der Idiot (Hermann Röhl, 1981)].

Эта особенность может объясняться как определенными чертами индивидуального стиля Достоевского, так и спецификой языка эпохи. В целом для русского языка такого рода словосочетания нехарактерны. Ср. количественные показатели, полученные в результате поиска в основном корпусе НКРЯ:

- *смеяться изо всех сил* — 3 вхождения (в том числе приведенный выше пример 11 из «Идиота»);
- *смеяться что было сил* — 0;
- *смеяться что было мочи* — 0;
- *смеяться что есть мочи* — 0;
- *хохотать изо всех сил* — 1;
- *хохотать что было сил* — 0;
- *хохотать что было мочи* — 1 (приведенный выше пример 12);
- *хохотать что есть мочи* — 0.

Еще один сходный по сути пример. Немецкая идиома *etw. an den Nagel hängen* и русская *повесить на гвоздь что-л.* не только совпадают по актуальному значению и внутренней форме, но и обнаруживают параллели в компонентном составе. Тем не менее они различаются по сочетаемостным потенциям. В корпусе встречаются примеры, в которых *etw. an den Nagel hängen* переводится с помощью *повесить на гвоздь что-л.*:

- (13) Er pflegte auch zu sagen, er verstehe nun gerade genug von der Kunst, um ein anständiger Liebhaber und Sammler zu sein, wenn das Glück ihn reich machen wollte, und

um diesen Preis würde er sofort seine Palette *an den Nagel hängen* [Gottfried Keller. Der grüne Heinrich (zweite Fassung) (1879–1880)].

Часто он говаривал, что смыслит в искусстве ровно столько, чтобы стать личным любителем или коллекционером, если фортуна обогатит его, и за эту цену он готов немедленно *повесить* палитру на гвоздь [Готфрид Келлер. Зеленый Генрих (Ю. Афонькин, Г. Снимщикова, Д. Горфинкель, Н. Бутова, 1972)].

Бросается в глаза, что в примере (14) из того же произведения, переведенного теми же переводчиками, вместо идиомы *повесить на гвоздь* употребляется глагол *бросить*. Особенно интересно, что речь в (14) идет о том же самом персонаже, что и в (13), и о том же роде деятельности.

(14) <...> er beabsichtige gleich nach der Hochzeit seine verzwickte Malerei *an den Nagel zu hängen* [Gottfried Keller. Der grüne Heinrich (zweite Fassung) (1879–1880)].

<...> намерен сейчас же после свадьбы *бросить* свое замысловатое художество [Готфрид Келлер. Зеленый Генрих (Ю. Афонькин, Г. Снимщикова, Д. Горфинкель, Н. Бутова, 1972)].

Причина нежелательности употребления выражения *повесить на гвоздь* в (14) в том, что данная идиома не сочетается с абстрактными существительными. На семантическом уровне это связано с тем, что в русской идиоме сильнее профилирована метафора, чем в ее немецком корреляте. Ср. также подтверждающие это сочетаемостное ограничение примеры (15) и (16).

(15) Ein wenig Geld hatte ich übrig, also *hing* ich meine literarischen Verpflichtungen *an den Nagel* und richtete mich ein, beim ersten Frühlingsbeginn zu packen und abzureisen [Hermann Hesse. Peter Camenzind (1904)].

У меня еще оставалось немного денег, и я, *махнув рукой* на свои сочинительские обязательства, решил при первых же признаках весны собрать вещи и немедленно уехать [Герман Гессе. Петер Каменцинд (Р. Эйвадис, 1995)].

(16) Aber jetzt... Irgendwann, wenn ich genug verdient hab, *häng* ich meinen Beruf *an den Nagel* und mach was anderes [Michael Ende. Momo (1973)].

Но сейчас... Когда-нибудь, когда я накоплю денег, я *плюну* на свою профессию и начну что-нибудь другое [Михаэль Энде. Момо (Ю. И. Коринец, 1982)].

Еще один тип межъязыковых различий связан с ограничениями на образование видо-временных форм русских глагольных идиом, и шире — любых грамматических форм фразеологизмов одного из сопоставляемых языков. В [Добровольский 2011] обсуждался пример употребления немецкой идиомы *jmdn. an der Nase herumführen* и ее русского квазиэквивалента *водить за нос кого-л.* В корпусе DeReKo встречается весьма значительное количество контекстов, в которых идиома *jmdn. an der Nase herumführen* употреблена в результативном значении. Перевод таких контекстов на русский язык с использованием идиомы *водить за нос кого-л.* оказывается нежелательным, поскольку форма совершенного вида *проводить за нос* допустима только в контекстах снятой утвердительности. Ср. примеры (17) из основного корпуса НКРЯ.

(17) а. <...> а народ не дурак, *за нос* его так просто *не проведешь*. [Владимир Шаров. Воскрешение Лазаря (1997–2002)]; б. *За нос* такого *провести нетрудно*: прилаской — и верти им, любому слову поверит [Михаил Веллер. Колечко (1983)].

Кроме того, — и это тоже важный фактор в определении степени эквивалентности сопоставляемых фразеологизмов — идиомы *jmdn. an der Nase herumführen* и *водить за нос кого-л.* различаются по степени употребительности. Это наблюдение подтверждается примерами из русско-немецкого корпуса. Поиск на *an der Nase herumführen* дает 11 вхождений, из них только шесть с идиомой *водить за нос* в русском оригинале. Ср., например:

- (18) Совсем не такой подлец! По крайней мере, долго себя не *морочил*, разом до последних столбов дошел [Ф. М. Достоевский. Преступление и наказание (1866)].
Sie sind gar nicht solch ein Schuft! Sie *haben sich wenigstens nicht lange an der Nase herumgeführt*, sind gleich bis zu der letzten Grenze gegangen [Fjodor Dostojewski. Verbrechen und Strafe (Alexander Eliasberg, 1924)].
- (19) «Энергия необходима, — говаривал он тогда, — а со всем тем он обыкновенно оставался в дураках и всякий несколько опытный чиновник *садился* на него *верхом*» [И. С. Тургенев. Отцы и дети (1860–1861)].
„Energie ist unerlässlich, — sagte er, — sie ist die vornehmste Eigenschaft eines Staatsmanns. Trotz dieser stolzen Sprache aber ward er fast immer *düpiert*, und jeder nur etwas erfahrene Beamte *führte ihn an der Nase herum*“ [Iwan Turgenew. Väter und Söhne (Friedrich von Bodenstedt, 1860–1880)].
- (20) И много приходило ему в голову того, что так часто уносит человека от скучной настоящей минуты, *теребит*, *дразнит*, шевелит его и бывает ему любо даже и тогда, когда уверен он сам, что это никогда не сбудется [Н. В. Гоголь. Мертвые души (1835–1852)].
Was kommt einem Menschen bei einem Spaziergang nicht alles in den Kopf, was ihn der langweiligen Minuten der Gegenwart vergessen läßt, die Phantasie beschäftigt, *an der Nase herumführt*, reizt und ihm auch dann gefällt, wenn er selbst überzeugt ist, daß nichts davon eintreffen wird! [Nicolaj Gogol. Die toten Seelen (Michael Pfeiffer, 1978)].

Заключение

Проанализированный материал позволяет сделать некоторые выводы. Во-первых, говоря о межъязыковой эквивалентности, мы должны различать разные ее аспекты и разные уровни рассмотрения. Изолированное сопоставление фразеологизмов разных языков по параметрам структуры и принадлежности к тому или иному семантическому классу позволяет лишь сделать некоторые утверждения по поводу эквивалентности на уровне системы, которые вписываются в общие вопросы устройства фразеологической системы того или иного языка, механизма заимствований и происхождения. Изучение фразеологизмов, встроенных в конкретный контекст, дает совершенно иные результаты. При сопоставлении оригинальных контекстов и их переводов на другие языки мы получаем результаты, касающиеся эквивалентности на уровне текста. Изучение фразеологии на основе материалов корпусов способствует прежде всего развитию этого направления. Если для системной перспективы важно само устройство фразеологической системы, то при перспективе, ориентированной на текст, в фокусе внимания оказывается приемлемость и удачность перевода.

Во-вторых, далеко не все фразеологизмы, которые представляются полностью эквивалентными при изолированном предъявлении, действительно совпадают по

всем релевантным для употребления параметрам. Такое положение связано с понятием лингвоспецифичности единиц языка.

Вопрос, насколько лингвоспецифична идиоматика, давно обсуждается в специальной литературе и может рассматриваться как один из базовых вопросов теории фразеологии. На эту тему существует несколько мнений. Наиболее экстремальными из них являются, с одной стороны, утверждение (часто высказываемое вскользь как нечто само собой разумеющееся), что идиомы уникальны по определению, а фразеология отражает характер народа, и, с другой — тезис о доминировании сходства над различиями. Этот тезис базируется на серьезных эмпирических исследованиях и подкрепляется этимологически: большинство идиом разных языков (по крайней мере в европейском культурном ареале) восходит к единым источникам. См. подробнее [Piirainen 2012].

Как это часто бывает, истина лежит где-то посередине. В разных языках (в частности, в русском и немецком) встречается некоторое количество идиом, соотносимых не только по ядерному значению и лексической структуре, но и по всем значимым для употребления признакам. По крайней мере, проанализированный на сегодня материал не позволил выявить никаких релевантных различий между ними. Ср. (*jmd.*) *würde sich im Grab umdrehen* «(кто-л.) перевернулся бы в гробу»; *wie Hund und Katze* «как кошка с собакой»; *mit offenem Mund dastehen [dasitzen]* «стоять [сидеть] разинув рот (от удивления)». В то же время есть и идиомы, не имеющие в другом языке стандартных эквивалентов. Ср. обсуждаемый выше пример *ein weites Feld*.

Наконец, выделяется весьма значительная группа единиц, которые, обнаруживая существенное семантическое и формальное сходство, отличаются в реальном употреблении. Различия могут быть обусловлены разными факторами. Основные факторы таковы:

- неполное совпадение семантики (как правило, речь идет о семантических признаках, относящихся к периферии значения);
- степень стертости метафоры и место образной составляющей в плане содержания идиомы;
- сочетаемостные предпочтения и ограничения;
- ограничения на образование грамматических форм, свойственные единицам одного из сопоставляемых языков;
- степень употребительности идиомы.

Выявление и тщательное описание этих различий существенно для развития двуязычной лексикографии.

Источники

НКРЯ — Национальный корпус русского языка. <http://www.ruscorpora.ru>.

DeReKo — Das Deutsche Referenzkorpus des IDS Mannheim im Portal COSMAS II (Corpus Search, Management and Analysis System). <https://cosmas2.ids-mannheim.de/cosmas2-web>.

Словари и справочные издания

Бинович, Гришин 1975 — Бинович Л. Э., Гришин Н. Н. *Немецко-русский фразеологический словарь*. Под ред. д-ра Малиге-Клаппенбах и К. Агрикола. 2-е изд., испр. и доп. М.: Русский язык, 1975. 656 с.

- Копотев 2014 — Копотев М. *Введение в корпусную лингвистику. Электронное учебное пособие для студентов филологических и лингвистических специальностей университетов*. Praha: Animedia, 2014. 194 с.
- Шекасюк 2010 — Шекасюк Б. П. *Новый немецко-русский фразеологический словарь*. 2-е изд., перераб. и доп. М.: Либроком, 2010. 864 с.
- Deutsch russische Idiome online — http://wvonline.ids-mannheim.de/idiome_russ/index.htm и http://www.europhras.org/index.php?option=com_content&view=category&layout=blog&id=36&Itemid=91&lang=de.

Литература

- Добровольский 2011 — Добровольский Д. О. Сопоставительная фразеология: межъязыковая эквивалентность и проблемы перевода идиом. *Русский язык в научном освещении*. 2011. 22(2): 219–246.
- Мокиенко 1989 — Мокиенко В. М. *Славянская фразеология*. 2-е изд., испр. и доп. М: Высшая школа, 1989. 287 с.
- Doval, Sánchez Nieto 2019 — Doval I., Sánchez Nieto M. T. (eds.). *Parallel Corpora for Contrastive and Translation Studies: New Resources and Applications*. Amsterdam, Philadelphia: John Benjamins, 2019. ix+301 pp.
- Filatkina 2014 — Filatkina N. Constructionalization, Konstruktionswandel und figurative Sprache (sprach) historisch betrachtet. In: *Figurative Sprache — Figurative Language — Langage figuré*. Dalmas M., Piirainen E. (Hg.). Tübingen: Stauffenburg, 2014. S. 41–57.
- Hunston, Gill 2000 — Hunston S., Gill F. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam, Philadelphia: John Benjamins, 2000. 288 pp.
- McEnery et al. 2006 — McEnery T., Xiao R., Tono Yu. *Corpus-based Language Studies: An Advanced Resource Book*. London: Taylor & Francis, 2006. 386 pp.
- Piirainen 2012 — Piirainen E. *Widespread Idioms in Europe and Beyond. Toward a Lexicon of Common Figurative Units*. New York etc.: Peter Lang, 2012. 591 pp.
- Tognini-Bonelli 2001 — Tognini-Bonelli E. *Corpus Linguistics at Work*. Amsterdam, Philadelphia: John Benjamins, 2001. 224 pp.
- Tummers et al. 2005 — Tummers J., Kris H., Geeraerts D. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory*. 2005. 1(2): 225–261.

Статья поступила в редакцию 31 марта 2020 г.

Статья рекомендована в печать 8 июня 2020 г.

Dmitrij Olegovich Dobrovol'skij

Russian Academy of Sciences, Vinogradov Russian Language Institute,
18/2, ul. Volkhonka, Moscow, 119019, Russia

Russian Academy of Sciences, Institute of Linguistics,
1, Bolshoi Kislovsky per., Moscow, 125009, Russia

Stockholm University,
26, Svante Arrhenius väg, Stockholm, SE-10691, Sweden
dobrovol'skij@gmail.com

Corpus-based approach to phraseology research: New evidence from parallel corpora*

For citation: Dobrovol'skij D. O. Corpus-based approach to phraseology research: New evidence from parallel corpora. *Vestnik of Saint Petersburg University. Language and Literature*. 2020, 17 (3): 398–411. <https://doi.org/10.21638/spbu09.2020.303> (In Russian)

* This paper is based on the work supported by the RFFI (Russian Foundation for Basic Research) under Grant 17-29-09154.

The present paper discusses the possibility of using parallel corpora in the study of phraseology. The results of the analysis support the assumption that the use of corpus data not only significantly improves cross-linguistic descriptions, but also changes the very idea of the specifics of phraseology as a subsystem of the lexicon. Parallel corpus data make it possible to identify the language-specific features of idioms that were traditionally considered completely equivalent. As the empirical bases, parallel subcorpora of the Russian National Corpus (RNC) are used: Russian-German and primarily German-Russian. In some cases, the material of the main corpus of the RNC is involved. In addition to the task of identifying similarities and differences between semantically similar idioms of German and Russian, the paper also addresses general issues, in particular the question of the boundaries of the concept of cross-linguistic equivalence and ways to operationalize it. At least two different types of equivalence have to be distinguished: equivalence at the level of the language system and equivalence at the text level. The proposed method includes a subsequent comparison of individual pairs of correlating idioms of different languages based on corpus data. It aims at identifying differences between them, including semantic, syntactic, pragmatic and combinatorial ones. The results of the performed analysis may find application in bilingual lexicography.

Keywords: phraseology, parallel corpus, Russian, German.

References

- Добровольский 2011 — Dobrovol'skij D.O. Contrastive phraseology: cross-linguistic equivalence and translation of idioms. *Russkii iazyk v nauchnom osveshchenii*. 2011. 22 (2): 219–246. (In Russian)
- Мокиенко 1989 — Mokienko V.M. *Slavic Phraseology*. 2nd ed., revised and expanded. Moscow: Vysshaya shkola Publ., 1989. 287 p. (In Russian)
- Doval, Sánchez Nieto 2019 — Doval I., Sánchez Nieto M.T. (eds.). *Parallel Corpora for Contrastive and Translation Studies: New Resources and Applications*. Amsterdam: Philadelphia: John Benjamins, 2019. 301 p.
- Filatkina 2014 — Filatkina N. Constructionalization, Konstruktionswandel und figurative Sprache (sprach) historisch betrachtet. In: *Figurative Sprache — Figurative Language — Langage figuré*. Dalmas M., Piirainen E. (Hg.). Tübingen: Stauffenburg, 2014. P.41–57.
- Hunston, Gill 2000 — Hunston S., Gill F. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: Philadelphia: John Benjamins, 2000. 288 p.
- McEnery et al. 2006 — McEnery T., Xiao R., Tono Yu. *Corpus-based Language Studies: An Advanced Resource Book*. London: Taylor & Francis, 2006. 386 p.
- Piirainen 2012 — Piirainen E. *Widespread Idioms in Europe and Beyond. Toward a Lexicon of Common Figurative Units*. New York etc.: Peter Lang, 2012. 591 p.
- Tognini-Bonelli 2001 — Tognini-Bonelli E. *Corpus Linguistics at Work*. Amsterdam, Philadelphia: John Benjamins, 2001. 224 p.
- Tummers et al. 2005 — Tummers J., Kris H., Geeraerts D. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory*. 2005. 1 (2): 225–261.

Received: March 31, 2020

Accepted: June 8, 2020