

*Д. О. Добровольский, А. А. Кретов, С. А. Шаров*

## **КОРПУС ПАРАЛЛЕЛЬНЫХ ТЕКСТОВ АРХИТЕКТУРА И ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ**

1. Корпус параллельных текстов (**КоПарТ**), входящий в качестве самостоятельного модуля в Национальный корпус русского языка, состоит из множества художественных произведений с их переводами. Корпус включает в себя, с одной стороны, оригинальные русские тексты с их переводом на английский язык, а с другой — английские тексты с их переводом на русский. К моменту на сентябрь 2005 г. КоПарТ включал следующие произведения:

<b>Автор</b>	<b>Название произведения</b>	<b>Кол-во слов в рез. файле</b>
Ч. Диккенс	Записки Пиквикского клуба (гл. 1-15 и 30-43)	300.024
В. Скотт	Айвенго (гл. 1-32)	245.559
Р. Л. Стивенсон	Странная история доктора Джекилла и мистера Хайда	49.529
Р. Л. Стивенсон	Остров сокровищ	130.057
Г. К. Честертон	Неведение отца Брауна	59.600
Т. Драйзер	Сестра Керри	227.885
М. Твен	Приключения Тома Сойера (10 глав)	108.368
Джером К.Джером	Трое в лодке, не считая собаки	130.531
Дж. Лондон	Мартин Иден	104.031
Ф. С. Фицджеральд	Великий Гэтсби	100.000
У. Голдинг	Повелитель мух	120.000
Э. Хемингуэй	Прощай, оружие! (фрагмент)	87.000
Х. Филдинг	Дневник Бриджит Джоунс	142.000
М. М. Додж	Серебряные коньки	164.000
К. Грэм	Ветер в ивах	97.600
Р. Брэдбери	Фаренгейт 451	95.000
А. С. Пушкин	Капитанская дочка	60.840
Н. В. Гоголь	Тарас Бульба	83.810

***Д.О.Добровольский, А.А.Кротова, С.А.Шаров***

Н. В. Гоголь	Мертвые души	174.416
М. Ю. Лермонтов	Герой нашего времени	94.208
А. П. Чехов	На святках	3.199
А. П. Чехов	Крыжовник	7.992
А. П. Чехов	Дама с собачкой	11.879
А. П. Чехов	Мечты	6.464
А. П. Чехов	Егерь	3.361
А. П. Чехов	Счастье	7.351
А. П. Чехов	Красавицы	5.953
А. П. Чехов	О любви	7.229
А. П. Чехов	Человек в футляре	9.526
А. П. Чехов	Мужики	22.609
А. П. Чехов	Пари	5.010
А. П. Чехов	Попрыгунья	17.767
А. П. Чехов	Сапожник и нечистая сила	5.149
А. П. Чехов	В сарае	4.409
А. П. Чехов	Скучная история	42.754
А. П. Чехов	Свирель	6.214
А. П. Чехов	Тоска	4.027
А. П. Чехов	Злоумышленник	2.882

Корпус параллельных текстов постоянно пополняется. Предполагается, что в дальнейшем он будет расширяться как за счет произведений других авторов и жанров, так и за счет других (помимо английского) языков перевода. Имеющийся на сегодня корпус может рассматриваться как пилотный. Выбор включенных в него текстов был мотивирован в первую очередь их высоким литературным качеством, а также наличием переводов на другие языки. Выбор языков пилотного корпуса мотивирован следующими соображениями. Понятно, что на первом этапе было бы нецелесообразно использовать в качестве параллельных русскому сразу несколько языков, так как каждый из них обладает своими особенностями, которые, возможно, могли бы повлиять на процесс выравнивания. То, что в качестве «пилотного» языка был выбран именно английский, объясняется несколькими — прежде всего техническими — причинами. Графика английского языка не содержит диакритических знаков, что делает обработку английских текстов на машинных носителях более легкой. Кроме того, английские тексты оказались наиболее доступными. Поиск электронных версий переводов русской литературы на другие языки часто сопряжен со значительными техническими трудностями.

Все приведенные в таблице тексты были выровнены по методике, описанной в пункте 2. В данный момент различные тексты находятся в разной степени готовности. Определенная часть корпуса доступна для поиска в режиме online: приблизительно 1,1 млн. слов в англо-русской и 0,5 млн. слов в русско-английской.

2. В настоящее время разрабатываются программные средства выравнивания параллельных текстов; кроме того, создается система управления корпусом, призванная удовлетворить запросы пользователей. КоПарТ обеспечен программой **ПарТекс** (идея А. А. Кротова, руководство И. Е. Ворониной, программирование Д. Спесивцева), имеющей на входе два параллельных текста (оригинал и перевод). Выравнивание осуществляется на уровне предложений. Программа выдает на выходе синтезированный текст, в котором последовательно за каждым предложением оригинала следует соответствующее предложение перевода. В таком синтезированном тексте поиск интересующих пользователя слов может осуществляться штатными средствами обычных текстовых редакторов, например, таких, как «Майкрософт Ворд». В программе ПарТекс для поиска может быть «подстрока в строке» и на выход подается текстовый файл, в который входят все пары предложений оригинала и перевода, содержащие заданную последовательность символов. Возможен поиск как английских, так и русских слов или словосочетаний.

Одна из наиболее существенных трудностей выравнивания заключается в том, что авторское членение текста на предложения и абзацы не всегда выдерживается в тексте перевода (см. подробнее раздел 5). Кроме того, в разных языках (а иногда и разных изданиях) приняты различные способы графического оформления, что иногда затрудняет определение границ предложения в автоматическом режиме. Ср., например, различные способы оформления переходов от прямой речи персонажей к авторским ремаркам. В таких случаях результаты автоматического выравнивания нуждаются в коррекции, осуществляемой вручную. Программа позволяет обнаружить в параллельных текстах асимметрию такого рода.

Алгоритм обнаружения асимметрии прост: количество глав оригинала сравнивается с количеством глав перевода, и в случае неравенства выдается сообщение об асимметрии. В случае равенства числа глав осуществляется переход на следующий уровень членения текста. В каждой из глав, начиная с первой, сравнивается

количество абзацев. В случае неравенства выдается сообщение об асимметрии, в случае равенства осуществляется переход к анализу следующей главы. После выравнивания по абзацам осуществляется переход к выравниванию по предложениям. Сравнивается количество предложений в первом абзаце оригинала с количеством предложений в первом абзаце перевода, при несовпадении выдается сообщение об асимметрии, при совпадении количества предложений в первом абзаце оригинала и перевода осуществляется переход к анализу второго абзаца.

Для решения собственно лингвистических задач (в первую очередь в области сопоставительной лексикологии и двуязычной лексикографии) важны отношения типа *слово — слово*, *слово — словосочетание* и *словосочетание — слово*. Такие пары могут быть названы *функционально эквивалентными фрагментами (ФЭФ)*. Типовая поисковая задача пользователя может быть сформулирована так: *Извлечь все случаи употребления слова L (в оригинале) и его ФЭФ (в переводе)*. При решении этой задачи могут быть использованы как минимум две стратегии, назовем их *тотальной* и *имплекативной*.

*Тотальная стратегия* нахождения ФЭФ состоит в следующем. Исходный текст — последовательная цепочка символов, адрес каждого из которых равен числу натурального ряда. Адрес каждого слова в исходном тексте задается интервалом между адресом первого и адресом последнего символа включительно. Если длина исходного текста =  $D^1$ , а длина переводного текста =  $D^2$  то можно составить пропорцию и с ее помощью получить адрес предполагаемого ФЭФ. Например,  $D^1 : D^2 = \text{AdrL} : \text{AdrX}$ , где  $\text{AdrL}$  = длина текста от начала до первого символа искомого слова, а  $\text{AdrX}$  — длина параллельного текста от начала до первого символа предполагаемого ФЭФ.

Рассмотрим это на примере текста повести Джека Лондона «Зов предков», которая в данный момент обрабатывается для включения в параллельный корпус. Длина оригинала ( $D^1$ ) составляет 31850 словоупотреблений, тогда как длина русского перевода ( $D^2$ ) составляет 27963 словоупотреблений. Следовательно, поправочный коэффициент для вычисления адреса слова в тексте ( $K^{\text{adr}}$ ) равен  $D^1 : D^2 = 31850 : 27063 = 1,139005$ . Допустим, нас интересует перевод слова *clash* в контексте *It was inevitable that the clash for leadership should come. Buck wanted it*. Длина текста от начала до слова *clash* включительно ( $\text{AdrL}$ ) равна 9522 словоупотреблениям. Для

того чтобы определить *теоретический адрес ФЭФ в русском переводе* ( $AdrX_{теор}$ ), разделим расстояние до искомого слова оригинала ( $AdrL$ ) на поправочный коэффициент:

$$AdrX_{теор} = AdrL : K^{adr} = 9522 : 1,139005 = 8360.$$

*Действительный адрес интересующего нас ФЭФ* ( $AdrX_{exp}$ ) — 8633, и это слово — *бой*. Ср. *Бой за первенство неизбежно должен был произойти, и Бэк хотел этого*. Таким образом, погрешность в вычислении равна  $8633 - 8360 = 273$  словам, что составляет 3% от величины  $AdrX_{exp}$ .

Тот же алгоритм может быть применен для вычисления  $AdrX$  в символах (с пробелами и без пробелов).

**Таблица вычисления адреса функционально эквивалентного фрагмента**

«Зов предков»	В словах	В символах с пробелами	В символах без пробелов
Оригинал	31850	176553	143049
Перевод	27963	175589	146249
$K^{adr}$	1,139005114	1,005490093	0,97811951
$AdrL$	9522	52750	42720
$AdrX_{теор}$	8360	52461,97884	43675,6446
$AdrX_{exp}$	8633	54058	44992
ПогрАбс	273	1596,021161	1316,35543
ПогрОтн	0,031631258	0,029524236	0,02925754

Как видим, вычисление  $AdrX_{теор}$  в символах без пробелов дает наименьшую погрешность. Однако независимо от единиц, в которых осуществляется вычисление, точность определения адреса ФЭФ остается в пределах 2,9% — 3,2% истинной длины текста до ФЭФ. Понятно, что по мере удаления от начала текста размер ошибки в абсолютных единицах будет увеличиваться. Именно поэтому практическая польза от тотальной стратегии будет убывать по мере роста величины  $AdrL$ , т. е. пропорционально расстоянию от начала текста до интересующего нас фрагмента текста.

*Импликативная стратегия* поиска ФЭФ ориентирована на границы, задаваемые архитектурой текста. Адрес искомого слова  $L$  задается последовательным сужением зоны поиска, т. е. пошагово:

номер тома, главы в томе, абзаца в главе, предложения в абзаце, слова в предложении. Аналогичным образом ищется предложение, содержащее предполагаемый ФЭФ переводного текста. При этом оптимальный результат достигается в том случае, когда количество единиц на каждом из уровней членения текста совпадает.

Практика показывает, что у каждого из методов есть как свои достоинства, так и свои недостатки. Тотальная стратегия универсальна, но неточна. Ошибка в определении адреса ФЭФ пропорциональна размеру текста. Импликативная стратегия точна, но не универсальна, поскольку предполагает изоморфное членение параллельных текстов. Прежде, чем пользоваться импликативной стратегией, следует убедиться, что у параллельных текстов одинаковое количество абзацев, а в каждом из абзацев — одинаковое количество предложений. Если количество абзацев и предложений не одинаковое, следует либо отказаться от импликативной стратегии, либо восстановить параллелизм текстов в данном отношении. Второе решение предпочтительнее, прежде всего из-за неточности тотальной стратегии.

Таким образом, необходимо создать программу, которая сравнивала бы тексты в их естественных границах, обнаруживала асимметрию в количестве абзацев и предложений и выдавала информацию о местах этой асимметрии пользователю. Имея эту информацию, пользователь может восстановить формальную (по количеству абзацев и предложений) симметрию параллельных текстов, и тем самым обеспечить успешную реализацию импликативной стратегии поиска ФЭФ. После такого предредактирования параллельных текстов становится возможным автоматический поиск ФЭФ для любого содержащегося в тексте оригинала слова. Функциональные роли «исходный текст» и «параллельный текст» задаются самим пользователем, что обеспечивает гибкость программы и возможность анализа как от оригинала к переводу, так и наоборот.

Принципиально возможны и другие способы выравнивания параллельных корпусов. Ср. способ выравнивания Австрийской академии наук в Вене — Austrian Academy Corpus (AAC) — подробнее см. (Fiber, Breiteneder, Dobrovol'skij 2002; Добровольский 2003; 2004). Подлежащие выравниванию тексты предварительно разбивались на предложения по определенным принципам, единым для оригинала и переводов. Такое членение текста предполагает тща-

**Корпус параллельных текстов**

тельное постредктирование результатов автоматической разбивки, позволяющее ориентироваться на определенные синтаксические структуры, а не на употребление знаков препинания: точки, вопросительного и восклицательного знаков. Подобная предварительная обработка текстов, подлежащих выравниванию, существенно повышает качество конечных результатов, получаемых с помощью специальных программ выравнивания в полуавтоматическом режиме.

После выравнивания параллельных текстов КоПарТа по предложениям было проведено морфологическое аннотирование и лемматизация полученного корпуса с использованием тех же средств, что и в Национальном корпусе русского языка в целом (программа *mystem* компании Яндекс). Результат морфологического анализа был представлен в виде таблицы в формате *Corpus WorkBench* (см. Christ 1994):

I		
...		
<seg id=2>		
<s>		
-	-	PUNCT
А	а	CONJ
поворотись	поворотиться	V sov=pov,ed,2-1
-ка	ка	PART
,	,	PUNCT
сын	сын	S muzh,od=ed,im
!	!	PUNCT
</s>		
</seg>		
<seg id=3>		
<s>		
Экой	экий	A =ed,zhen,rod
ты	ты	S ed,od=im
смешной	смешной	A =ed,muzh,im
какой	какой	A =ed,muzh,im
!	!	PUNCT
</s>		
</seg>		

**II**

...
<seg id=2>
<s>
Turn                    NN                    turn
round                   NN                    round
,
my                      PP\$                    my
boy                      NN                    boy
!                        SENT                    !
</s>
</seg>
<seg id=3>
How                    WRB                    how
ridiculous            JJ                      ridiculous
you                     PP                      you
look                    VVP                    look
!                        SENT                    !
</s>
</seg>

Выравнивание между предложениями отражено в идентификаторах сегментов соответствующих текстов. За исключением этих идентификаторов Corpus WorkBench не накладывает ограничений на структуру аннотирования. В английском корпусе использован другой набор морфологических признаков (и даже порядок полей). Механизм поиска в параллельном корпусе позволяет искать по лемме, точной словоформе, морфологическим признакам или их комбинации. Также возможно использование в запросе стандартных регулярных выражений, например, **.\*[а-я]** и т. п.

Поиск по лемме является наиболее простым способом поиска, который позволяет найти все формы одного слова: запрос **стакан** находит все формы (*стакан, стакана, стаканом*). Задание нескольких лемм (в их словарной форме) соответствует поиску фразы, например, запрос **один слово** находит фразы *одним словом, ни одного слова* и т. п. Для поиска точной формы каждую словоформу надо заключать в кавычки, например, запрос **«ни» «одного» «слова»** найдет ровно эти примеры. Морфологические признаки представлены с помощью механизма множеств Corpus WorkBench, поэтому

для их поиска требуется использовать специальный синтаксис запросов Corpus WorkBench, в котором указываются поле запроса и ограничение на его значение. Например, набор признаков слова *словом* выражен как «**n:inan:sg:nom**». Запрос [**gr contains «ins»**] находит все формы в творительном падеже, в то время как [lemma=«слово»&gr contains «ins»] — все формы слова *слово* в творительном падеже (*словом* и *словами*), запрос [**gr contains «pl»&gr contains «ins»**] находит формы множественного творительного.

В результате создатели двуязычных словарей, переводчики, все лингвисты, работающие с параллельными текстами, получают весьма простой и эффективный инструмент сбора материала. Ценность этого инструмента определяется тем, что в лингвистике этап сбора материала является наиболее трудоемким и наименее творческим, а подобные корпуса параллельных текстов позволят сэкономить время и силы для собственно исследовательской работы.

**3.** Создание корпуса параллельных текстов в качестве одного из модулей Национального корпуса русского языка представляется целесообразным как с практической точки зрения, так и с точки зрения развития корпусной лингвистики — одного из наиболее перспективных лингвистических направлений. С помощью параллельных корпусов могут быть также получены интересные результаты в области теоретического языкознания, так как опора на принципиально сопоставимые аутентичные тексты разных языков позволяет выявить часто неожиданные (как квазиуниверсальные, так и специфические) особенности функционирования языковой системы. Корпус параллельных текстов может быть эффективно использован в различных лингвистических исследованиях (в первую очередь, в области контрастивной лексикологии и двуязычной лексикографии), а также в исследованиях по теории перевода, сравнительного литературоведения, культурологии, автоматической обработки текста и др.

Приведем эскизный перечень основных вопросов, при решении которых обращение к корпусам параллельных текстов представляется разумным.

1. Как ведут себя определенные структуры входного языка (L1) и их соответствия выходного языка (L2) в аутентичных контекстах? Насколько системные характеристики этих структур способны предсказать их поведение в реальном дискурсе? Какие типы контекстов оказываются релевантными для выбора

адекватного эквивалента в языке L2? Иными словами, если исследование корпусов показывает, что стандартные «словарные» L2-эквиваленты данной L1-структуры оказываются неприемлемыми в контекстах определенных типов, необходимо выявить релевантные свойства этих контекстов и соответствующим образом переформулировать условия эквивалентности.

2. Для произведений, достаточно далеко отстоящих во времени, дополнительно встает вопрос о соотношении представленных в корпусе L1-структур (с их L2-переводами) и норм современного узуса. Насколько изменился узус? В чем причины подобных изменений? Существуют ли более или менее регулярные механизмы, управляющие языковыми изменениями? Если да, можно ли на основе этих данных предсказать дальнейшие изменения языковых норм? Являются ли соответствующие тенденции уникальными для каждого языка или в них просматриваются некоторые общие закономерности?

3. В каких случаях переводчики предлагают нестандартные решения? С чем это связано? Мотивированы ли отклонения от оригинала субъективными факторами или объективными межъязыковыми различиями между L1 и L2, накладывающими определенные ограничения на способы перевода исходных структур? Если одному и тому же месту оригинала в разных переводах соответствуют различные эквиваленты, встает вопрос о семантических отношениях между этими эквивалентами. Являются они в языке L2 квазисинонимами или же речь идет о различных интерпретациях L1-структуры? Если соответствующие выражения L2 квазисинонимичны, в чем состоят их семантические, прагматические и сочетаемостные различия? Если же речь идет о различных интерпретациях, чем мотивированы отклонения от соответствующих структур оригинала? При каких условиях они могут быть признаны допустимыми? Проблемы данного типа решаются особенно эффективно в случае, если параллельный корпус содержит более одного перевода одного и того же оригинального текста (ср. Добровольский 2003 о корпусе параллельных текстов AAC).

4. Какие решения переводчики находят для тех L1-структур, которые не вписываются в современные нормы? Какими средствами языка L2 могут быть переданы особенности подобных структур? Естественно, ответы на такие вопросы могут быть по-

лучены только при использовании текстов, достаточно далеко отстоящих во времени (ср., например, произведения русской классической литературы XIX века).

Из данного каталога вопросов видно, что использование корпусов параллельных текстов дает нетривиальные результаты для целого ряда филологических направлений. Среди них сопоставительная лексикология, двуязычная лексикография, лексическая семантика (в том числе, ее диахронные аспекты), а также теория перевода. Обсуждению чисто лингвистических вопросов посвящен следующий раздел. В заключение этого раздела укажем на те аспекты использования параллельных корпусов, которые относятся скорее к области критики перевода и текстологии.

Корпус параллельных текстов позволяет выявить переводческий брак и редакторский произвол. Во многих переводах имеется неоправданный пропуск значимых с содержательной и эстетической точек зрения фрагментов текста. Это относится и к английскому переводу «Капитанской дочки», и к переводу «Мертвых душ», в которых опущена «Повесть о капитане Копейкине», и к русскому переводу «Дракулы» или «Хижина дяди Тома». Параллельные тексты заставляют также задуматься о границах свободы переводчика в переводе названий произведений. Например, заглавие рассказа А. П. Чехова «На подводе» переведено как «Schoolmistress»; рассказ «Холодная кровь» получил заглавие «The cattle-dealers», а повесть А. С. Пушкина «Капитанская дочка» в одном из английских переводов озаглавлена «Mary[V13]».

4. Современное состояние сравнительной лексикологии и практики составления двуязычных словарей характеризуется ориентацией на сопоставление более или менее изолированных языковых структур. Отрицательным последствием подобной ориентации является недостаточный учет узуса, то есть тех особенностей синтаксического и сочетаемостного поведения единиц языка, которые нельзя объяснить их системными признаками. Так, в принципе известно, что та или иная структура одного языка не может быть во всех контекстах переведена на другой с помощью своего стандартного эквивалента. В определенных контекстах язык L2 традиционно прибегает к другим способам описания соответствующей ситуации. Известно также, что не существует продуктивных правил, по которым можно было бы вывести подобные отклонения от «стандартной эквивалентности» из неких более общих

принципов. Единственный способ описания подобных отклонений — это их тщательная фиксация на аутентичном материале. Только так можно построить исчерпывающие сопоставительные описания и создать словари, удовлетворяющие современным требованиям.

Корпус параллельных текстов представляет собой наиболее адекватный инструмент для выполнения этих задач. Та или иная языковая структура, интересующая исследователя, может быть найдена во всех представленных в корпусе контекстах с их переводами на соответствующий язык. Таким образом, исследователь получает в свое распоряжение набор аутентичных контекстов, представляющих интересующую его структуру в ее естественном окружении, а также самые разнообразные эквиваленты этой структуры в языке-цели. Поскольку эти эквиваленты также оказываются встроенными в естественные контексты, на основе полученных с помощью параллельного корпуса материалов могут быть сделаны выводы о зависимости выбора эквивалента от типа контекста. Подобные результаты практически всегда расходятся с теми сведениями, которые мы можем почерпнуть из существующих словарей, являясь тем самым нетривиальными.

**4.1.** Важным параметром, по которому языки могут различаться между собой, является степень употребительности определенных выражений. Так, некоторое выражение *A* языка *L1* может стандартным образом переводиться на язык *L2* с помощью выражения *B* — вполне корректного с точки зрения норм этого языка. Таким образом, выражения *A* и *B* оказываются эквивалентными в рамках системы соответствующих языков. Тем не менее, их функциональная эквивалентность часто представляется неполной. В частности, это имеет место в случае, когда одно из выражений оказывается в своем языке существенно более употребительным, чем его переводной эквивалент — в своем. Такие случаи хорошо прослеживаются на сопоставлении оригинальных текстов с их переводами на другие языки.

Исследование некоторого языкового явления на основе корпуса параллельных текстов может быть по ряду параметров противопоставлено исследованию этого явления на основе большого корпуса оригинальных текстов. Отличие оригинальных текстов от переводов заключается и в объеме (миллионы слов оригинальных текстов, производимых ежедневно, против относительно небольшого

количества текстов, переводимых с иностранных языков), и в природе авторства (т. е. степени оригинальности и творческой свободы при порождении текста), и в культурном контексте (переводные тексты обычно погружены в культуру исходного языка).

Все эти факторы обеспечивают различия между исходными и переводными текстами по целому ряду параметров. Например, сравнение употребления притяжательных местоимений *my* и *your* в корпусе оригинальных текстов и переводов с русского языка показывает, что сочетания *my heart*, *my soul*, *my lord* встречаются в переводах с русского намного чаще, чем в оригинальных текстах. В то же время *your hands* практически не встречается в переводах с русского, поскольку русские эквиваленты таких частотных английских выражений, как *leave it in your hands* или *come into your hands*, редко используют словосочетание *ваши/твои руки*. Ср., например (1).

(1) If this letter will not *come into your hands* until tomorrow morning.

Возможно, что *вы получите* это письмо только завтра утром.

Таким образом, английскому идиоматическому выражению в данном случае соответствует абсолютно регулярно построенное русское выражение. В результате под влиянием русского оригинала английские переводчики употребляют в подобных ситуациях менее идиоматичные английские конструкции, чем авторы, пишущие на английском языке как на родном. Ср. также известные наблюдения, согласно которым английское слово *well* в функции заполнителя пауз hesitation употребляется в переводах с других языков в несколько раз реже, чем в оригинальных текстах. Параллельный корпус, с одной стороны, предоставляет возможности для выдвижения осмысленных гипотез, касающихся различий в степени употребительности эквивалентных и квазиэквивалентных выражений сравниваемых языков, а с другой — является (при его достаточном объеме) надежным ресурсом для статистической проверки подобных гипотез.

**4.2.** Корпус параллельных текстов дает также материал для выявления и описания тонких семантических и сочетаемостных различий между выражениями L1 и L2, традиционно считающихся полностью эквивалентными. Примером здесь может служить английский модификатор степени проявления признака *absolutely* в сопоставлении с русским словом *абсолютно*. Сразу оговорим, что параллельный корпус, доступный в настоящий момент для поиска, состоит всего из примерно полутора миллиона слов. Соответст-

венно, даже столь распространенные модификаторы, как *absolutely*, *completely*, *крайне*, *совершенно*, *чрезвычайно* встречаются примерно по 30-50 раз каждый. По этой причине параллельный корпус в его сегодняшнем виде позволяет лишь строить определенные гипотезы относительно особенностей употребления подобных слов. Дополнительную информацию о таких словах можно получить путем сравнения немногих примеров их употребления в параллельном корпусе с примерами из большого и представительного корпуса (ср. Шаров 2003), в частности, Национального корпуса русского языка и Британского Национального Корпуса (BNC).

Относительно слов *absolutely* и *абсолютно* интересными представляются следующие наблюдения. По данным BNC, слово *absolutely* употребляется прежде всего в следующих режимах: 1) как интенсификатор оценочного прилагательного, которое уже само по себе выражает высокую степень соответствующего признака (*absolutely appalling / delighted*), 2) как интенсификатор выражений со значением 'необходимости' (*absolutely necessary / vital*) и 3) как интенсификатор выражений с семантическими компонентами типа 'понятность', 'уверенность' (*absolutely clear / sure*); ср. (Altenberg 1991). Русское слово *абсолютно*, будучи, как и *absolutely*, наречием степени с семантикой предельности, обнаруживает некоторые отличия в сочетаемости. Так, интуитивно ясно, что *абсолютно* вряд ли может широко употребляться в режиме 2, плохо сочетаясь со словами типа *должен*, *нужно* (хотя *абсолютно необходимо* не противоречит нормам современного русского языка).

В английских переводах русских текстов наречие *absolutely* представлено всего тринадцатью словоупотреблениями, многие из которых представляют сочетаемостный режим 2, причем в русском оригинальном тексте слово *абсолютно* в этих фрагментах не встречается. Ср., например, *тебе непременно нужно* — *you absolutely must*. Что касается других (пусть даже принципиально возможных) сочетаний русского модификатора *абсолютно* с выражениями, обозначающими необходимость, оказалось, что в нашем корпусе они реально не представлены. Так, слово *необходимо* представлено в сочетании с модификаторами *жизненно*, *чрезвычайно*, *принципиально*, но не *абсолютно*. Именно сочетания типа *жизненно необходимо* могут рассматриваться как хорошие функциональные эквиваленты узуального английского словосочетания *absolutely vital*.

Что касается сочетаемостного режима 3, в нашем параллельном корпусе представлен лишь один пример с соответствующим английским выражением в исходном тексте:

- (2) And whatever that queer creature understands, he *understood absolutely* his lord's two fixed ideas.

Что бы ни понимало это создание, оно *прекрасно* поняло две навязчивые идеи хозяина.

Словосочетание *understood absolutely*, естественно, не может быть переведено на русский язык дословно. В этом случае в качестве эквивалента модификатора *absolutely* выступает наречие *прекрасно*.

**4.3.** Еще один пример, сочетающий использование небольшого параллельного и представительного одноязычного корпуса, касается перевода глагола *состояться* в значении 'успешно реализовать свои потенциальные (творческие) возможности'; ср. (3).

- (3) Он вполне *состоялся* как актер.

Параллельный корпус содержит всего одно употребление слова *состояться* — в значении, отличном от (3), но, тем не менее, также не вполне тривиальном с точки зрения межъязыковой эквивалентности; ср. (4) с переводом (5).

- (4) И это дикое, бессмысленное пари *состоялось!*

- (5) And this wild, senseless bet was *carried out!*

Двуязычные словари — например, Оксфордский русско-английский словарь или электронный словарь Лингво — приводят в качестве переводных эквивалентов лишь *take place; come about*, что совершенно не отражает значения слова *состояться*, реализуемого в примере (3), и не выглядит естественным в контекстах типа (4). Словарь Ожегова определяет *состояться* как 'произойти, осуществиться' (ср. *заседание не состоялось*), не выделяя других значений. МАС выделяет еще устаревшее подзначение 'быть опубликованным, обнародованным, введенным в действие'. Очевидно, что предлагаемое этими словарями описание семантической структуры данного глагола оказывается заведомо неполным. Могут ли параллельные и одноязычные корпуса помочь в решении задачи адекватного перевода предложений типа (3) и (4)? И если да, то каким образом?

В текущей версии Национального корпуса русского языка слово *состояться* встречается 5551 раз, что дает среднюю частоту около 11 употреблений на миллион слов. Наиболее частые контексты

его употребления представляют собой сочетания с существительными семантического класса МЕРОПРИЯТИЕ: *встреча, выборы, заседание, презентация, премьера, прием, церемония*, обычно в контексте указания времени и/или места проведения мероприятия (что в принципе характерно для этого класса существительных). Такие примеры хорошо поддаются переводу с помощью вышеуказанных эквивалентов, но примеры (3) и (4) им не соответствуют. Для примера (4) параллельный корпус дает удачный контекстуальный эквивалент. Перевод с помощью *take place* в данном случае невозможен, поскольку этот предикат закреплен за существительными класса МЕРОПРИЯТИЕ и фокусирует параметры «время» и «место».

Заметим попутно, что в Британском национальном корпусе существительное *bet* (эквивалент для *пари* в примере (4)) часто встречается также в сочетании с глаголом *take*, например:

- (6) Standing around, waiting for the police to arrive, the resident and the intern *made bets* on whether the case would be a real emergency or just a teenager who had swallowed too much aspirin.

Перевод примера (6) не может быть проанализирован с помощью параллельного корпуса, поскольку подобные контексты в нем не представлены. Очевидно, что решение поставленной здесь задачи с помощью корпуса параллельных текстов возможно лишь в случае, если корпус достаточно велик. В противном случае обращение к параллельному корпусу оказывается эффективным лишь для частотных языковых выражений. Подобные задачи могут, однако, решаться и с помощью одноязычных корпусов.

Возвращаясь к вопросу об адекватном переводе глагола *состояться* в значении ‘успешно реализовать свои потенциальные (творческие) возможности’, обратимся сначала к анализу соответствующих контекстов из Национального корпуса русского языка. Результаты поиска показывают, что *состояться* в этом значении практически всегда употребляется в контекстах, где речь идет о деятелях искусства (художниках, актерах, писателях) или о политических деятелях и движениях; например, (7) и (8).

- (7) А она тем не менее *состоялась*, стала успешным режиссером, публицистом.  
(8) За короткий срок мы набрали немалый политический вес и вполне *состоялись* и как фракция, и как партия.

Оценить способы выражения успешности деятелей искусства в английском на основе корпуса мы можем задав несколько запросов. Например, запрос

as a [lemma=«actor|actress|artist|comedian|dancer|director|musician|producer|writer»], перечисляющий виды творческих профессий, позволяет найти немало примеров, которые могут быть использованы в переводе контекстов типа (3).

- (9) <...> having already *achieved distinction as a writer* <...>.
- (10) <...> *to get established as a writer* <...>.
- (11) He *had subsequently been very successful as a writer*.
- (12) Michael Powell *built a reputation as a director* of energetic quota films.
- (13) <...> and was determined to enable him *to realize his potential as a writer*.

Таким образом, работа с корпусами позволила установить, что в качестве английских эквивалентов глагола *состояться* в значении 'успешно реализовать свои потенциальные (творческие) возможности' могут использоваться такие выражения, как *to achieve distinction (as an X)*, *to get established (as an X)*, *to have been successful (as an X)*, *to build a reputation (as an X)*, *to realize his/her potential (as an X)*.

**4.4.** Использование параллельного корпуса оказывается весьма эффективным при исследовании лексических единиц, обладающих плохо предсказуемой, капризной сочетаемостью. Адекватный перевод таких слов зависит не только от особенностей их семантической структуры, но и от целого ряда (как правило, практически неописанных) узуальных предпочтений. Поскольку узуальные предпочтения такого рода редко оказываются идентичными в разных языках, вычислить правильный L2-эквивалент для данного L1-слова в том или ином типе контекстов оказывается практически невозможным. Понятно, что данные, почерпнутые из представительного корпуса параллельных текстов, оказываются едва ли не единственным надежным эмпирическим фундаментом для исследования природы комбинаторных ограничений такого типа.

Примером подобных случаев могут служить модификаторы степени проявления признака *чрезвычайно* и *крайне*. Русско-английские словари в качестве эквивалентов наречия *чрезвычайно* называют *extraordinarily*, *extremely* и *utterly*. Наши наблюдения за английскими текстами (безотносительно к обсуждаемому здесь корпусу параллельных текстов) позволяют утверждать, что *extraordinarily* в значении модификатора степени используется весьма редко. Ср. пример (14) из романа П. Данкер «Hallucinating Foucault».

- (14) He was *extraordinarily* strong.

Примеры, подобные (14), встречаются существенно реже, чем контексты с другими интенсификаторами экстремально высокой степени (по классификации ван Оса — Os 1989). Так, в параллельном корпусе на сегодняшний день представлены лишь следующие контексты с модификатором *extraordinarily*: (15) и (16) из «Сестры Кэрри» и (17) из «Великого Гэтсби».

(15) Hurstwood realised that he was seeing something *extraordinarily* good.

Герствуд понял, что перед ним *на редкость* хорошая актерская игра.

(16) It was *not extraordinarily* bright and merry this evening.

*Вовсе не* дышал его дом в этот вечер уютом и весельем.

(17) He did *extraordinarily well* in the war.

Военная карьера *удалась* ему.

Интересно, что ни в одном из этих переводов не использовано слово *чрезвычайно*, считающееся наиболее чистым словарным эквивалентом *extraordinarily*. Что касается прочих предлагаемых словарями эквивалентов слова *чрезвычайно*, мы обсудим их несколько ниже. Здесь же укажем, что словарный эквивалент *extremely* предлагается всеми известными словарями в качестве базового (а чаще всего и единственного) эквивалента также и для русского наречия *крайне*. В связи с этим встает ряд вопросов. Различаются ли *чрезвычайно* и *крайне* по семантическим, прагматическим и сочетаемым параметрам? Если да, как передаются эти различия при переводе на английский язык? Обладает ли английское слово *extremely* более широкой сочетаемостью, чем каждое из этих русских наречий, так что оно действительно с одинаковым успехом может выступать в качестве эквивалента как *чрезвычайно*, так и *крайне*? Существуют ли другие, неотмеченные словарями, способы перевода этих наречий? Различаются ли эти английские эквиваленты в зависимости от типа контекста? И наконец, какие английские модификаторы степени переводятся на русский язык с помощью *чрезвычайно* или *крайне*?

Очевидно, что в рамках данной статьи ответить на все эти вопросы не удастся. Напомним, что доступный в настоящее время для поиска корпус параллельных текстов относительно невелик и не позволяет исчерпывающе описать нетривиальные особенности межъязыковой эквивалентности. Здесь целесообразно скорее наметить возможные пути анализа лексических единиц такого рода с помощью параллельных корпусов.

Остановимся на вопросе о семантических различиях между *чрезвычайно* и *крайне*. Достаточно очевидно, что асертивная часть значения модификаторов, относящихся к одной и той же степени интенсивности, практически идентична. Их семантические различия выводятся в первую очередь из внутренней формы, которая в некоторых случаях может накладывать определенные ограничения на сочетаемость, но чаще скорее способствует профилированию различных аспектов описываемой ситуации. Для *чрезвычайно* и *крайне* могут быть предложены следующие предварительные толкования.

*чрезвычайно*  $P \approx$  'P в очень высокой степени; говорящий считает, что степень P настолько высока, что она выходит за пределы области обычного'

Под «пределами области обычного» здесь понимаются привычные пределы вариативности признака P, т. е. пределы его вариативности, внутри которых степень P воспринимается как не нарушающая стандартных ожиданий. В эту область входит и высокая степень выраженности признака, т. е. значение 'очень P'.

*крайне*  $P \approx$  'P в очень высокой степени; говорящий считает, что степень P настолько высока, что она близка к пределу, возможному для P'

Из анализа сочетаемости этих наречий становится очевидно, что собственно семантические особенности в незначительной степени влияют на сочетаемостные ограничения. Подобная зависимость наблюдается скорее у «молодых» модификаторов с яркой внутренней формой типа *чудовищно*. Ср. запреты на сочетаемость типа *??мне чудовищно приятно*. Яркая внутренняя форма подобных модификаторов обеспечивает связь с другими (первичными) значениями этих слов. Наличие таких семантических мостов, отвечающих за внутреннее единство слова, отличает подобные слова от прочих наречий степени, см. подробнее (Филипенко 2003: 73).

Анализ корпусов современных текстов<sup>1</sup> показывает, что для современного употребления слова *чрезвычайно* характерны следующие тенденции (которые, естественно, нельзя интерпретировать как жесткие правила).

---

<sup>1</sup> Здесь был использован Национальный корпус русского языка (<http://www.ruscorpora.ru>), корпусы русских текстов Тюбингского университета (Германия) — Tübinger Russische Korpora: <http://www.sfb441.uni-tuebingen.de/b1/korpora.html>, текстовая база данных Машинного фонда русского языка и корпусы Отдела экспериментальной лексикографии Института русского языка РАН.

1. В сфере действия показателя степени признака *чрезвычайно* чаще находятся прилагательные или наречия; сочетания с глаголами в личных формах и инфинитиве встречаются существенно реже<sup>2</sup> (*чрезвычайно интересный / заинтересован* лучше, чем *чрезвычайно интересуюсь*).

2. Слова, находящиеся в сфере действия *чрезвычайно*, чаще обозначают «нефизические» сущности (*чрезвычайно тяжелое положение* лучше, чем *чрезвычайно тяжелый чемодан*).

3. Эти слова обозначают свойства, тяготеющие скорее к положительному, чем к отрицательному полюсу различных шкал (*чрезвычайно широкий* <одаренный, энергичный> встречается чаще, чем *чрезвычайно узкий* <бездарный, вялый>)<sup>3</sup>. См. подробнее (Dobvol'skij 2004; Добровольский в печати).

Что касается сочетаемости слова *крайне*, то она отличается от сочетаемости *чрезвычайно* прежде всего по параметру 3. Еще в (Червенкова 1975: 17) было отмечено, что «тенденции к положительному» (т. е. к сочетанию с обозначениями желательных для говорящего свойств), характерной для *чрезвычайно*, противопоставлена «тенденция к отрицательному», характерная для *крайне*. Эти тенденции, естественно, отнюдь не абсолютны. Можно найти достаточно примеров, в которых *чрезвычайно* модифицирует обозначения отрицательно оцениваемых (*чрезвычайно трудно/тяжело*), а *крайне* — обозначения положительно оцениваемых признаков (*крайне оптимистично, крайне популярен*). Соответствующие различия между этими модификаторами носят характер статистически верифицируемых предпочтений, а не абсолютных запретов. Понятно, что опора на большие корпуса оказывается здесь главным средством проверки подобных гипотез.

Не исключено, что признак 'область обычного' до известной степени ответствен за особенности сочетаемости слова *чрезвычай-*

---

<sup>2</sup> Здесь, видимо, в первую очередь важна синтаксическая позиция — *чрезвычайно* неохотно занимает позицию члена предложения, непосредственно зависящего от глагола. Это становится очевидным при сопоставлении глаголов с соответствующими причастиями. Поскольку причастия занимают в предложении позиции, аналогичные позициям прилагательных, они сочетаются с *чрезвычайно* намного чаще, чем соответствующие глаголы.

<sup>3</sup> Понятно, что из всех этих «правил» есть исключения. Ср. совершенно приемлемое с точки зрения современного словоупотребления сочетание *чрезвычайно исхудал*, нарушающее сразу все три «правила» (пример Ю. Д. Апресяна).

но. В силу относительно прозрачной морфологической структуры этого слова его «этимологическая память», фиксирующая семантические связи с ОБЛАСТЬЮ ОБЫЧНОГО, обращается не к «норме» как нулевой отметке на некоторой шкале, а к идее границ стандартных ожиданий. Связь наречия *чрезвычайно* с идеей выхода за пределы обычного поддерживается также семантикой прилагательного *чрезвычайный*, причем даже не столько в коррелирующем значении (*чрезвычайно удивлен — чрезвычайное удивление*), сколько в значении, реализуемом в таких сочетаниях, как *чрезвычайное положение/происшествие, чрезвычайные меры*.

В том или ином понимании идея нормы значима для семантики всех модификаторов степени. Отличие *чрезвычайно* от, например, *очень* видится в том, что для *очень* НОРМА предстает в виде точки на шкале. Все проявления признака Р, которые находятся между этой точкой N и полюсом интенсивности I, могут быть обозначены как *очень Р*. Для *чрезвычайно* НОРМА — не точка, а целый отрезок шкалы между N и I, на котором располагаются все значения признака Р, не выходящие за рамки привычного. Лишь когда степень проявления признака Р оказывается за пределами этого отрезка, приближаясь к полюсу I, проявление признака Р может быть охарактеризовано как *чрезвычайно Р*. Это соответствует интуитивному представлению о том, что *чрезвычайно Р* сильнее, чем просто *очень Р*. А, например, *крайне* отличается от своих квазисинонимов тем, что фокусирует идею «края», максимальной близости к полюсу I<sup>4</sup>.

Анализ данных из Национального корпуса русского языка позволяет утверждать, что *чрезвычайно* в целом тяготеет к положительному полюсу различных шкал: «хорошо», «много», «большой», а *крайне* — к соответствующим отрицательным полюсам: «плохо», «мало», «маленький». Например, сочетание *чрезвычайно много* (и его (квази)синонимы) довольно частотно, при том что *чрезвычайно*

<sup>4</sup> Традиционные представления о том, что различия между такими модификаторами, как *чрезвычайно* и *крайне*, сводятся прежде всего к противопоставлению «сильнее-слабее» (ср., например, Ховалкина 1995: 118, 121), не подтверждаются материалом. Дело в том, что они необязательно привязаны к одной шкале; «о единой шкале мер уместно было бы говорить только в том случае, если бы разные ее элементы прилагались бы к одним и тем же объектам (были бы определены — в математическом смысле слова — на одном и том же множестве объектов)» (Филипенко 2003: 69). В целом, слово *чрезвычайно* обнаруживает более широкую сочетаемость, чем *крайне* (Janus 1981: 189).

*мало* (и его синонимы) встречается реже, хотя говорить, что это сочетание ненормативно или странно, было бы неверным. Сочетание *крайне много* не встретилось в корпусе вообще ни разу, при том что *крайне мало* (с синонимами) оказывается достаточно частотным. Аналогично ведут себя словосочетания *чрезвычайно низкие* vs. *крайне низкие* (*доходы, прибыли* и т. п.) vs. *чрезвычайно высокие* vs. *крайне высокие*, где наиболее частотными оказываются сочетания *чрезвычайно высокие* и *крайне низкие*. Сочетания типа *крайне редко* вполне узуальны, а сочетание *крайне часто*, видимо, вообще не соответствует современным нормам. Ср. также редкие сочетания *чрезвычайно неохотно/неудачно/неудовлетворительно* и частотные *крайне неохотно/неудачно/неудовлетворительно* при отсутствии контекстов на *крайне охотно/удачно/удовлетворительно*. Подобные примеры можно было бы продолжить. В целом складывается впечатление, что соответствующие тенденции несимметричны. Для *крайне* тенденция к отторжению обозначений, близких к положительному полюсу, проявляется сильнее (вплоть до запретов на сочетаемость), чем для *чрезвычайно* — тенденция к отторжению отрицательного.

Отмеченные здесь тенденции, регулирующие сочетаемость слова *чрезвычайно*, видны особенно рельефно на фоне норм XIX века, допуская гораздо более широкую и свободную сочетаемость этого слова; см. подробнее (Dobrovol'skij 2004; Добровольский в печати). Старые нормы сочетаемости прослеживаются в приведенных ниже примерах (21) из «Тараса Бульбы» Гоголя и (23) из «Злоумышленника» Чехова, где нарушается тенденция 2, что придает этим контекстам несколько архаичный характер.

Перейдем к вопросу об английских эквивалентах этих русских модификаторов. В нашем параллельном корпусе встретились четыре различных эквивалента наречия *чрезвычайно*, причем только один из них — *extremely* — укладывается в рамки «классического» словарного описания. Прочие эквиваленты — *exceedingly, greatly, immense* (с соответствующими преобразованиями синтаксической структуры высказывания) — обычно словарями не фиксируются. А *utterly*, будучи одним из типичных словарных эквивалентов *чрезвычайно*, вообще не встретилось в корпусе в этом качестве (что, естественно, ни о чем не говорит, так как объем корпуса ограничен)<sup>5</sup>.

---

<sup>5</sup> С другой стороны, анализ английских корпусов показывает, что для *utterly* характерны предельные употребления, т. е. *utterly* эквивалентно скорее русским словам *совершенно, полностью, окончательно, совсем*, чем словам типа *чрезвычайно*.

Перевод с помощью наречия *extremely* или — при соответствующих синтаксических изменениях — прилагательного *extreme* представляется наиболее стандартным; ср. контексты (18)-(21).

- (18) <...> неизбежно вырастает в целую задачу, сложную *чрезвычайно* <...>.  
<...> inevitably grows into a regular problem of *extreme* intricacy <...>.
- (19) <...> ему стало *чрезвычайно* горько и обидно на беспорядок, который в природе <...>.  
<...> he felt *extremely* bitter and resentful of the impropriety in the conduct of nature <...>.
- (20) <...> не стараясь связать их в мотив, но тем не менее в его писке слышалось что-то суровое и *чрезвычайно* тоскливое.  
<...> with no attempt at forming a tune, and yet there was something harsh and *extremely* dreary in the sound of the piping.
- (21) <...> Тарас был *чрезвычайно* тяжел и толст.  
<...> Taras was *extremely* stout and heavy.

Прочие способы перевода представлены единичными контекстами; ср. (22)-(25).

- (22) Остапу и Андрию казалось *чрезвычайно* странным, что при них же приходила на Сечь гибель народа <...>.  
It seemed *exceedingly* strange to Ostap and Andrii that, although a crowd of people had come to the Setch with them <...>.
- (23) Перед судебным следователем стоит маленький, *чрезвычайно* тощий мужичонка в пестрядинной рубашке и латаных портах.  
An *exceedingly* lean little peasant, in a striped hempen shirt and patched drawers, stands facing the investigating magistrate.
- (24) Эта улица *чрезвычайно* походила на вывороченную внутренность заднего двора.  
This street *greatly* resembled a back-yard turned wrong side out.
- (25) В последние два года заточения узник читал *чрезвычайно* много, без всякого разбора.  
In the last two years of his confinement the prisoner read an *immense* quantity of books quite indiscriminately.

Для того чтобы дать точные ответы на вопрос о причинах выбора того или иного конкретного модификатора в английских переводах, потребуется большее количество контекстов, достаточное для выделения определенных типов сочетаемости. Пока можно лишь сказать, что употребление более формального и редкого *exceedingly* (по сравнению с *extremely*) объясняется, видимо, стилистическими причинами (22-23), употребление наречия *greatly* в сочетании *greatly resembled* (24) — тем, что *extremely* еще в меньшей степени, чем *чрезвычайно* переносит соседство с глаголами, а упот-

ребление *immense* в сочетании *read an immense quantity of books* (25) — неузуальностью буквального перевода *читал чрезвычайно много* как *?read extremely much* или *?read extremely many books*. Английское наречие *extremely* в принципе плохо сочетается с *much* и с *many*. В некоторых случаях выбор модификатора мотивирован узуальными предпочтениями английского языка или вкусом переводчика, например: *Бродяга был чрезвычайно доволен моим подарком — As for the vagabond, he was well pleased with my present*. Иногда модификатор вообще остается непереуведенным (ср. с *видом чрезвычайно встревоженным — with a terrified air*), что, однако, вряд ли отражается на степени адекватности перевода в целом, поскольку минимальной смысловой единицей при переводе является не слово, а конкретный аспект ситуации.

Эти примеры показывают, что даже единичные контексты из параллельного корпуса позволяют расширить наши представления об особенностях сочетаемости слов в разных языках. Это особенно значимо, когда речь идет о, казалось бы, полностью эквивалентных словах, которые обнаруживают, однако, существенные расхождения в комбинаторном потенциале.

Слово *крайне* во всех контекстах корпуса переводится с помощью *extremely*. Ср. примеры с (26) по (28).

(26) Это маленький тщедушный человек, слабосильный и болезненный, с мелкими, бесцветными и *крайне* неопределенными чертами лица.

He was a frail little man, weak and sickly-looking, with small, colourless, and *extremely* indefinite features.

(27) В последнее время Ольга Ивановна вела себя *крайне* неосторожно.

Olga Ivanovna had been *extremely* imprudent in her conduct of late.

(28) Она и другая невестка, Фекла, которая теперь сидела поодаль и слушала, — обе были *крайне* неразвиты и ничего не могли понять.

Both she and Fyokla, the other sister-in-law, who was sitting a little way off listening, were *extremely* ignorant and could understand nothing.

Очевидно, что *extremely* действительно представляет собой едва ли не универсальный перевод русского *крайне*. Не в последнюю очередь этому способствует параллелизм внутренней формы. Очевидно также, что обратное неверно: не всякий контекст, содержащий *extremely*, может быть переведен на русский язык с помощью *крайне*. Причина этого заключается, видимо, в том, что в современном русском языке модификатор *крайне* обнаруживает довольно ограниченную сочетаемость, при том что английское *extremely*

обладает достаточно широкими комбинаторными возможностями. Косвенным подтверждением этой гипотезы служит зафиксированный на материале корпуса факт, что с помощью *extremely* на английский язык переводится как *крайне*, так и (чаще всего) *чрезвычайно*, хотя в современном русском языке они отнюдь не полные синонимы.

Перейдем к поиску ответа на вопрос, какие английские модификаторы степени переводятся на русский язык с помощью *чрезвычайно* или *крайне*? Чисто умозрительно можно было бы предположить, что, если и *крайне* и (чаще всего) *чрезвычайно* переводятся с помощью *extremely*, то и в английских текстах там, где в русском переводе появляются *чрезвычайно* и *крайне*, должно стоять наречие *extremely* или его близкие синонимы типа *exceedingly*. Однако это не так. Подобные переводные соответствия встречаются скорее редко: *some of these queer faces amused Ben exceedingly* — *некоторые из этих диких голов чрезвычайно забавляли Бена*; *placed me in such an extremely awkward situation* — *поставила меня в чрезвычайно неловкое положение*.

В большинстве случаев появления в переводах слова *чрезвычайно* в английском оригинале какой бы то ни было модификатор отсутствует вообще, ср. (29-32).

(29) <...> and my reasons, though private, are honourable and compelling.

<...> для этого у меня есть основания, хотя и личные, но вполне добропорядочные и *чрезвычайно* веские.

(30) The night after the funeral, at which he had been sadly affected <...>.

Вечером после похорон, *чрезвычайно* его расстроивших <...>.

(31) For two months, however, I was true to my determination; for two months, I led a life of such severity as I had never before attained to, and enjoyed the compensations of an approving conscience.

Однако два месяца я свято соблюдал свое решение, два месяца я вел *чрезвычайно* строгую жизнь, о какой и мечтать не мог прежде, и был вознагражден за это блаженным спокойствием совести.

(32) The door was very strong, the lock excellent <...>.

Дверь оказалась на редкость крепкой, а замок — *чрезвычайно* хитрым <...>.

В контексте (31) присутствует указательный элемент в метафорической функции (*I led a life of such severity as I had never before attained to*), который в каком-то смысле может быть сопоставлен с модификатором степени, но в любом случае им не является.

Встречаются случаи, в которых русскому *чрезвычайно* соответствуют его более или менее отдаленные квазисинонимы, обычно не фигурирующие в роли словарных эквивалентов: *completely weird* — *чрезвычайно странный*, *incredibly nice and attractive* — *чрезвычайно приятный и обаятельный*, *fearfully grim and forbidding* — *чрезвычайно мрачным и неприступным*, *greatly excited* — *чрезвычайно возбужденно*, *considerably astonished* — *чрезвычайно удивленный*, *immensely rich* — *чрезвычайно богат*, *especially dirty faces* — *чрезвычайно грязные физиономии*, *that will rather surprise you* — *что тебя чрезвычайно удивит*, *I am deeply interested* — *меня чрезвычайно заинтересовал*, *a very spirited and rampacious animal* — *чрезвычайно горячее и резвое животное*, *particularly fat* — *чрезвычайно толстый*, *unusually precarious* — *чрезвычайно непрочное*, *singularly romantic* — *чрезвычайно живописно*, *most faithful* — *чрезвычайно преданный*, *very much satisfied* — *чрезвычайно довольна*, *much touched* — *чрезвычайно расстрогана*.

Часто бывает трудно сказать, как переводится тот или иной конкретный модификатор, поскольку он является частью особой синтаксической конструкции со своим относительно стандартным лексическим наполнением. Ср.: *they were in a high state of excitement* — *они были чрезвычайно взволнованы*, *in the most delightful manner* — *чрезвычайно приятно*, *he was a most particularly short man* — *он был чрезвычайно маленького роста*, *looking extremely pleased* — *с чрезвычайно довольным видом*, *I should esteem it highly if..* — *я была бы чрезвычайно признательна, если бы...*, *in well-affected ecstasies of joy* — *искусно притворяясь чрезвычайно обрадованной*, *they wore a good deal of hair* — *у них были чрезвычайно пышные волосы*. В некоторых случаях подобные конструкции оказываются лексикализованными в столь высокой степени, что могут рассматриваться как устойчивые коллокации. Наиболее ярким примером таких выражений может служить русское словосочетание *чрезвычайно важный/важно*, которое в силу своей частой повторяемости в дискурсе и достаточно высокой стабильности используется при переводе самых разнообразных английских выражений, передающих тот же смысл. Это в первую очередь такие выражения, как *highly important, so important, very important*. Ср. также *with an air of great importance* — *с чрезвычайно важным видом*, *on a matter of considerable importance* — *по чрезвычайно важному делу*, *will be considered significant* — *будет считаться чрезвычайно важным*, *a very vital matter* — *чрезвычайно важный момент*.

Что касается модификатора *крайне* из русских переводов, то ему в ряде случаев тоже не сопоставляется какой бы то ни было модификатор оригинала; ср. (33) и (34).

(33) <...> he gave an impression of deformity without any nameable malformation, he had a displeasing smile <...>.

<...> он производил впечатление уродства, хотя никакого явного уродства в нем заметно не было, улыбался он *крайне* неприятно <...>.

(34) O, we must be careful.

Мы должны быть *крайне* осторожны.

Есть контексты, в которых коррелятами *крайне* являются слова *very* и *most*; ср. (35) и (36).

(35) I am painfully situated, Utterson; my position is a *very* strange — a *very* strange one.

Вы не понимаете, в каком нахожусь положении, — сбивчиво ответил доктор. — Оно *крайне* щекотливо, Аттерсон, *крайне* щекотливо и странно, очень странно.

(36) My two natures had memory in common, but all other faculties were *most* unequally shared between them.

Мои две натуры обладали общей памятью, но все остальные их свойства распределялись между ними *крайне* неравномерно.

В одном случае перевод сопровождается полной перестройкой синтаксической структуры предложения: *крайне P => P can hardly be exaggerated*.

(37) The importance of this to Dr. J. can hardly be exaggerated.

Доктору Джекилу *крайне* важно получить этот препарат.

Проанализированные примеры показывают, что между конкретными модификаторами степени в разных языках нет взаимно-однозначных соответствий. В принципе это касается и многих других семантических классов слов. Каждое конкретное слово обладает своим уникальным набором сочетаемостных ограничений и предпочтений. Эта идея относится к числу традиционных теоретических положений лингвистики. То новое, что вносит обращение к корпусам параллельных текстов в решение подобных задач, — это возможность эмпирической проверки соответствующих гипотез на представительном материале. Кроме того, данные корпусов параллельных текстов позволяют ясно продемонстрировать, в сколь сильной степени выбор переводного эквивалента зависит от конкретной конструкции, компонентом которой является данная лексическая единица<sup>6</sup>.

<sup>6</sup> Результаты подобных наблюдений коррелируют с идеями Е. В. Падучевой (2004: 114, 121-122, 128 и далее) о модификациях значения слова под влиянием

5. Как уже было отмечено выше, одна из наиболее существенных трудностей выравнивания заключается в том, что авторское членение текста на предложения не всегда выдерживается в тексте перевода. В связи с этим возникает, казалось бы, наивный вопрос: зачем переводчики меняют границы предложений? Не проще ли было бы сохранить авторское членение текста? Пытаясь ответить на этот вопрос, мы обнаружили, что этот вид переводческих трансформаций, несмотря на частоту его использования в переводческой практике, является малоизученным. В процессе изучения данной проблемы были получены результаты, которые, возможно, помогут систематизировать наши знания об этих переводческих приемах, а также позволят получить предварительное количественное описание данного явления.

В качестве материала был выбран роман Т. Драйзера «Сестра Кэрри» и его перевод, выполненный М. Волосовым. С помощью программы Course.exe, позволяющей обнаружить все случаи асимметричного перевода предложений, в тексте был зафиксирован 951 случай такого рода: 493 примера членения одного предложения на два и более и 458 примеров объединения двух (и более) предложений в одно. Попытка систематизировать и классифицировать эти случаи привела к следующим результатам (см. подробнее Гурина, Добровольский, Кретов 2005). Кратко остановимся на некоторых из них.

**5.1.** Начнем с анализа случаев, в которых одно английское предложение было при переводе на русский язык разделено на два самостоятельных предложения. Самой многочисленной группой примеров (236 случай) оказались предложения, являющиеся частями диалога. Эти случаи делятся на две подгруппы. В первой подгруппе новая реплика при переводе оформляется как отдельное предложение (ср. (38)), а во второй подгруппе слова автора разделяются на два предложения, причем второе предложение начинается новый абзац (ср. (39)).

(38) «Oh», he answered, in a very pleasing way and with an assumed air of mistake, «I thought you did.»

О! — приятно улыбнулся он, как бы изумляясь своей ошибке. — Очевидно, я ослышался.

---

контекста, требующих перестройки толкования, в частности, изменения акцентного статуса его элементов.

(39) «It's from Hurstwood», said Drouet, noting the superscription as he tore it open.

– От Герствуда, — сказал Друэ, тотчас узнав почерк приятеля.

Он вскрыл конверт и принялся читать.

Большое количество подобных примеров подтверждает систематичность преобразований этого типа. Тем не менее, вряд ли имеет смысл предполагать, что такие преобразования каким-то образом влияют на смысл высказывания. Скорее речь идет о традиционных различиях оформления диалога в русском и английском языках.

В следующую группу вошли сложносочиненные английские предложения, в которых в позиции подлежащего находятся разные субъекты и которые были разделены на два или более предложений в русском тексте; ср. (40). Всего было обнаружено 85 таких примеров.

(40) *They* filed out, and *he* affected to take no notice of her.

Пассажиры один за другим стали покидать вагон. Друэ сделал вид, будто не обращает на Кэрри внимания.

В этом примере обращает на себя внимание еще один переводческий прием «прояснения» смысла текста — раскрытие анафоры: *they* vs. *пассажиры*, *he* vs. *Друэ*. Ср. также *she* vs. *Каролина* в контексте (41).

Далее были выявлены 53 случая разделения предложения, содержащего однородные члены или параллельные конструкции. Ср. (41).

(41) *She* was eighteen years of age, bright, timid, and full of the illusions of ignorance and youth.

Каролине только что исполнилось восемнадцать лет. Девушка она быламышленная, но застенчивая, преисполненная иллюзий, свойственных неведению и молодости.

Также и этот пример показывает наличие — свойственной по крайней мере анализируемому переводу — тенденции к «ориентации на смысл» в ущерб формальной точности, когда переводчик с помощью различных синтаксических преобразований и лексических замен проясняет те места текста, которые, по его мнению, могут восприниматься читателем как недостаточно «прозрачные» (о параметрах различных стратегий художественного перевода см. Добровольский 2004). Контекст (41) иллюстрирует целый ряд таких приемов: это членение предложения-оригинала (видимо, оцененного как синтаксически перегруженное) на два самостоятель-

ных русских предложения, раскрытие анафоры, а также введение в русский текст противительного союза *но*, при том что английский оригинал не содержит каких бы то ни было адверсативных элементов.

Следующую группу контекстов с непараллельным членением составляют предложения, содержащие эмфазу. Ср. (42).

(42) *It's a second New York — great.*

Это в полном смысле слова второй Нью-Йорк. *Изумительный город!*

Интересно, что все эмфатические элементы в английском тексте выделены тире. Тем самым и в английском оригинале элементы эмфазы оформляются как самостоятельные высказывания. Различия между соответствующими английскими контекстами и их русскими переводами объясняются, по-видимому, ориентацией на различные традиции пунктуационного выделения.

В 43 предложениях причиной членения предложения стал *ing*-clause, который в принципе может переводиться на русский как причастный или деепричастный оборот, но в рассматриваемых случаях, тем не менее, был вынесен в отдельное предложение. Ср. (43).

(43) *When six o'clock came she hurried eagerly away, her arms aching and her limbs stiff from sitting in one position.*

Когда наконец пробило шесть, она порывисто встала и поспешно вышла. *Руки у нее болели*, все тело ломало от работы в одном положении.

**5.2.** Обратимся к противоположному переводческому приему — объединению предложений. Всего в анализируемом тексте было найдено 458 примеров.

Самую большую группу составляют предложения с разными подлежащими (113 случаев); например, (44).

(44) *She could scarcely toss her head gracefully. Her hands were almost ineffectual. The feet, though small, were set flatly.*

*Она* не умела игриво откидывать назад головку, часто не знала, куда девать руки, и хотя ножки у нее были маленькие, ступала *она* тяжело.

В случаях, подобных (44), синтаксические структуры оригинала преобразуются в переводе таким образом, что у соответствующих предложений оказывается идентичное подлежащее. Понятно, что объединение простых предложений с идентичными подлежащими в более сложные структуры — вполне естественное решение. Это может быть предложение с однородными сказуемыми, объединенными одним подлежащим (ср. первые два предложения оригинала в (44) и русский перевод до слов *и хоть*), или сложносочиненное

предложение (части которого могут включать придаточные) — ср. с контексте (44) *она не умела... и ступала она тяжело*.

В 67 случаях были объединены предложения с референтно идентичными подлежащими, выраженными местоимениями. Ср., например, (45).

(45) *That is, I live at Columbia City. I have never been through here, though.*

Вернее, я живу в Колумбия-сити, а здесь *мне* еще не приходилось бывать.

В отдельную группу (36 примеров) выделяются контексты, в которых подлежащее первого предложения выражено существительным (в том числе именем собственным), а подлежащее второго — местоимением; например, (46).

(46) *Drouet was not a drinker in excess. He was not a moneyed man.*

*Друэ* не был пьяницей и не был человеком денежным.

Обратимся к предложениям, входящим в состав диалогов. Был обнаружен 31 пример объединения предложений при диалоге. Здесь можно выделить две подгруппы, иллюстрируемые контекстами (47) и (48) соответственно.

(47) *He leaned forward to put his elbows upon the back of her seat and proceeded to make himself volubly agreeable.*

«Yes, that is a great resort for Chicago people. <...>»

Слегка наклонившись вперед, он положил локти на спинку ее сиденья и заговорил, желая показать себя приятным спутником:

— Да, прекрасный уголок, отличные отели.

(48) «Oh, I don't know», *returned Hurstwood*. «He's got the money, all right», *and a little twinkle passed over his eyes*.

— Право, не знаю, имеет ли он с ними дело или нет, но денежки у него водятся, — *отозвался Герствуд, и в глазах его блеснул алчный огонек*.

В первом случае (47) перевод сохраняет структуру оригинала. Об «объединении» предложений здесь можно говорить лишь в чисто техническом смысле. В английском языке граница между авторской речью и речью персонажа маркируется точкой, а в русском — двоеточием. Иными словами, здесь мы имеем дело с различными традициями пунктуационного выделения. Во втором случае (48) имеет место перестройка частей высказывания в переводе. В русском тексте две авторские ремарки оригинала *returned Hurstwood* и *and a little twinkle passed over his eyes* объединены в одну.

Восемь раз формальная асимметрия возникла при переводе эллиптических конструкций. Ср. (40). Очевидно, что и здесь речь идет о чисто графических различиях.

(49) Then he rang the bell. *No answer.*

Он позвонил — *ответа не последовало.*

И наконец, асимметрия в членении текста на предложения возникает при переводе парцелированной темы, оформленной как самостоятельное восклицательное предложение (6 примеров). Ср., например, (50).

(51) *The police!* They would be after him.

*А полиция, которая тотчас же бросится по его следам!*

В оригинале смысловой субъект высказывания в результате парцелляции вынесен за пределы предложения и формально образует отдельное предложение, т. е. речь идет о фокусированной топикализации. В переводе этот способ коммуникативного выделения не сохранен.

**5.3.** Таким образом, на проанализированном материале (который, естественно, позволяет сделать лишь весьма предварительные выводы) было установлено, что при переводе с английского языка на русский прием изменения границ между предложениями используется достаточно часто. С приблизительно одинаковой частотой встречаются случаи членения одного предложения на два и более и случаи объединения двух и более предложений в одно. Среди случаев разделения одного предложения оригинала на два и более самостоятельных были выделены следующие типы предложений:

- предложения, являющиеся частями диалога;
- предложения, в которых в позиции подлежащего находятся разные субъекты;
- предложения, содержащие однородные члены или параллельные конструкции;
- предложения, содержащие эмфазу;
- предложения, содержащие *ing*-clause.

Среди случаев объединения двух и более предложений в одно выделяются:

- предложения с разными подлежащими, находящимися в определенных семантических отношениях друг с другом (например, «часть-целое»);
- предложения с референтно идентичными подлежащими, выраженными местоимениями;
- группы предложений, в которых подлежащее первого предложения выражено существительным (в том числе, именем собственным), а подлежащее второго — местоимением;

- предложения, входящие в состав диалога;
- эллиптические конструкции;
- предложения с парцелированной темой, оформленной как самостоятельное восклицательное предложение.

**6.** Намеченные в этой статье пути и способы использования создаваемого корпуса параллельных текстов для решения различных лингвистических задач не исчерпывают всех возможностей работы с подобными корпусами. При расширении эмпирической базы — в особенности при включении в корпус различных переводов одного и того же оригинала, выполненных в разное время, или его переводов на различные языки — появится возможность ставить и другие вопросы, такие, как границы свободы переводчика в выборе функциональных эквивалентов, статус различных переводческих стратегий в разные эпохи, ограничения, накладываемые типом языка и соответствующими культурно-семиотическими параметрами, и пр.

### **Литература**

- Гурина Е. Е., Добровольский Д. О., Кретов А. А. Асимметрия членения текста в параллельном корпусе (на материале романа Т. Драйзера «Сестра Кэрри») // Вестник Воронежского гос. университета. Серия «Лингвистика и межкультурная коммуникация», 2005, № 1, с. 67-74.
- Добровольский Д. О. Корпус параллельных текстов и литературный перевод // НТИ сер. 2, 2003, № 10, с. 13-18.
- Добровольский Д. О. Лингвистические аспекты перевода художественной прозы (на материале романа Ф. М. Достоевского «Идиот») // Языки в современном мире. Материалы конференции. Том I. М.: МГУ, 2004, с. 30-45.
- Добровольский Д. О. Факторы сочетаемости: семантика, прагматика, узус // Русский язык в научном освещении, в печати.
- Добровольский Д. О., Кретов А. А., Шаров С. А. Корпус параллельных текстов // НТИ сер.2, 2005, № 6, с. 16-27.
- Падучева Е. В. Динамические модели в семантике лексики. М.: Языки славянской культуры, 2004.
- Филипенко М. В. Семантика наречий и адвербиальных выражений. М.: Азбуковник, 2003.
- Ховалкина А. А. Лексическое выражение степени величины признака в современном русском языке. Симферополь, 1995.
- Червенкова И. В. Общие адвербиальные показатели меры признака (в современном русском литературном языке). АКД. М., 1975.

**Д.О.Добровольский, А.А.Кротова, С.А.Шаров**

- Шаров С. А. Представительный корпус русского языка в контексте мирового опыта // НТИ сер. 2, 2003, № 6, с. 8-18.
- Altenberg B. Amplifier collocations in spoken English // S. Johansson, A. B. Stenstrom (eds). English Computer Corpora. Berlin: Mouton de Gruyter, 1991.
- Biber H., Breiteneder E., Dobrovol'skij D. Corpus-based study of collocations in the AAC // Proceedings of the Tenth EURALEX International Congress, Vol. 1. Copenhagen: CST, 2002, p. 85-95.
- Christ O. A modular and flexible architecture for an integrated corpus query system // Proceedings COMPLEX'94. Budapest, 1994. Also available from the URL: <http://www.ims.uni-stuttgart.de/projekte/CorpusWorkbench/Papers/christ:complex94.ps.gz>
- Dobrovol'skij D. Lexical semantics and combinatorial profile: a corpus-based approach. In: Geoffrey Williams and Sandra Vessier (eds.). Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004. Lorient: UBS, 2004, p. 787-796.
- Janus E. Wykładniki intensywności cechy (na materiale polskim i rosyjskim). Wrocław, Warszawa, Kraków, Gdańsk: Zakład Narodowy im. Ossolińskich — Wydawnictwo Polskiej akademii nauk, 1981.
- Os C. van. Aspekte der Intensivierung im Deutschen. Tübingen: Narr, 1989.