ПРОЕКТЫ

Борис Орехов

ПАРАЛЛЕЛЬНЫЙ КОРПУС ПЕРЕВОДОВ «СЛОВА О ПОЛКУ ИГОРЕВЕ» КАК ПРОЕКТ

В статье описывается предпосылки и история филологического цифрового проекта, поддерживаемого уже более 15 лет. В его основе лежала идея сравнения переводов «Слова о полку Игореве», самого переводимого на русский язык произведения мировой литературы. Однако по мере складывания практик использования ресурса, стало очевидно, что он востребован у пользователей только как библиотека текстов. Корпус действительно содержит множество оцифрованных (иногда даже редких) произведений переводческого искусства. В процессе разработки проекта пришлось отказаться от ряда технических решений, таких, как автоматическое и полуавтоматическое выравнивание. Несмотря на цифровой статус, ручного труда на создание корпуса употреблено больше, чем применения технологий.

Ключевые слова: научный проект, параллельный корпус, слово о полку игореве, художественный перевод

В феврале 2007 года был запущен сайт «Параллельного корпуса переводов “Слова о полку Игореве”». Он был задуман как филологический инструмент для сравнения текстов. Специалисты в области переводоведения хорошо знакомы с ситуацией, когда у некоторого классического литературного текста имеется несколько переводов на один язык. Связано это с естественным процессом старения перевода и желанием обновить звучание статусного текста, а также с желанием предложить собственную переводческую трактовку произведения, имеющего в оригинале широкий спектр интерпретаций. Так, перевод «Фауста», выполненный Н. Холодковским, начинается со строк

Борис Валерьевич Орехов
НИУ «Высшая школа экономики», ИРЛИ РАН
nevmenandr@gmail.com

DOI: 10.31860/cgi-2025-1-60-69

Вы вновь со мной, туманные виденья,
Мне в юности мелькнувшие давно. . .

Конструкция *мелькнуть* + *Dat*, в принципе, вполне естественна для русского языка XIX века: ср. примеры вроде «Во всяком случае, в эту минуту в первый раз, но все еще смутно, *мелькнула мне* мысль, что в „принципах“ известной окраски, если даже они залегли в общее миросозерцание...» [М. Е. Салтыков-Щедрин. Мелочи жизни (1886–1887)]; «С другой стороны, Лизе очень хотелось видеть, какое впечатление произвело бы на него свидание столь неожиданное. . . Вдруг *мелькнула ей* мысль. Она тотчас передала ее Насте; обе обрадовались ей как находке и положили исполнить ее непременно». [А. С. Пушкин. Барышня-крестьянка (1830)]. Однако к концу XX века она уже может считаться устаревшей, при десятках несомненных вхождений этой конструкции в XIX веке, начиная с 1980-х годов ее появление в НКРЯ единичны и требуют оговорок.

Сосуществование нескольких переводов одного текста на некоторый язык принято называть не вполне изящным словом «переперевод» или громоздко, но более естественно «множественный перевод».

Обычно в ситуацию множественного перевода попадают произведения, входящие в мировой канон. Так, десятками исчисляются варианты переводов од Горация (на сайте horatius.ru учтено более 30 переводов 30-й оды III книги од, «Exegi monumentum aere perennius»), сонетов Петрарки, трагедий Шекспира.

«Слово о полку Игореве» находится в особой позиции среди других произведений, часто переводившихся на русский язык. Число переводов «Слова» на русский язык исчисляется сотнями, таким образом, это самое часто переводившееся на русский язык произведение мировой литературы.

Общепринятой методики работы с такой ситуацией множественного перевода нет, но ясно, что общий подход заключается в сравнении этих текстов по значимым для исследовательского вопроса параметрам. В работах М. Л. Гаспарова как основание для сравнения предлагалась мера точности перевода [Гаспаров 2001], и в этом смысле количественный подход как раз демонстрировал свои сильные стороны, так как открывал возможности для сравнения текстов между собой. Но более привычным является качественная работа с таким материалом, как, например, это продемонстрировано в монографии А. Н. Егунова о Гомере в России

[Егунов 2001]. Пример работы подобного рода с переводами «Слова» можно найти в трудах Е. Д. Поварницыной [Поварницына 2020].

Для исследования переводов «Слова о полку Игореве» был бы полезен вспомогательный инструмент. Владимир Набоков, устами персонажа романа «Бледный огонь», настаивавшего на параллельном чтении двух частей своего текста, предлагал, чтобы обойтись без хлопотного перелистывания назад-вперёд, либо разрезать книгу и скрепить вместе соответствующие страницы произведения, либо купить сразу два экземпляра, которые можно будет положить перед собой:

> I find it wise in such cases as this to eliminate the bother of back-and-forth leafings by either cutting out and clipping together the pages with the text of the thing, or, even more simply, purchasing two copies of the same work which can then be placed in adjacent positions on a comfortable table.
>
> (V. Nabokov “Pale Fire”)

Тексты переводов рассредоточены по отдельным изданиям, которыми можно пользоваться при сравнении двух произведений, но уже работа с тремя текстами создает серьезные затруднения для параллельного чтения, а усложнение исследовательской задачи за счет увеличения числа текстов делает работу едва ли не невозможной. «Жонглировать» четырьмя и более книгами, пытаясь отыскать мельчайшие сходства и различия текстов, крайне затруднительно. Поскольку размышления об этом состоялись еще до глобальной цифровизации, наиболее естественным выходом из ситуации в тот момент было создание «аналогового» «инструмента» сравнения, то есть сшитых в тетрадь листов, где один под другим (а не на разных страницах разворота, как обычно бывает с билингвами) написаны строки переводов. Такое решение не оптимально по многим причинам: во-первых, сравниваемые переводы могут оказаться не на соседних строках, во-вторых, при печатном исполнении такую книгу очень трудно сверстать так, чтобы параллельные части текста оказались точно одна под другой, в-третьих, трудно представить себе объём издания, включающего хотя бы половину общего числа переводов.

Цифровой сервис благодаря свойственной ему динамичности представления контента, противопоставленной статике бумажного справочника, способен разрешить все эти затруднения. Интерфейс, помогающий отбирать переводы для сравнения, позволит ограничиться нужными текстами и не распылять внимание. Электронная

верстка гораздо более гибкая, чем бумажная, и позволяет исправлять ошибки, а цифровой формат нечувствителен к текстовым объемам, с которыми придется иметь дело.

Первоначально идею такого параллельного представления переводов мы с коллегами рассчитывали предложить уже существующим электронным сервисам, у которых, как мы считали, есть ресурсы на то, чтобы воплотить такую идею средствами своих штатных программистов. Однако, переговоры с Фундаментальной электронной библиотекой (ФЭБ) в 2006 году прошли неудачно, коллеги указали на концептуальные и технические сложности реализации идеи. Сейчас, спустя почти 20 лет, понятно, что идея такого корпуса не вполне органична для ФЭБ, и в гораздо большей степени идейно перекликается с Национальным корпусом русского языка (НКРЯ). Действительно, ФЭБ выдерживала курс на максимально точное электронное воспроизведение печатного издания, что во многом чуждо корпусу, концептуальной платформой которого было соединение в поле одного визуального пространства нескольких текстов. Такая ориентация неизбежно приведет к потере точности: не будут воспроизведены типографские строки, пагинация, особенности шрифтового оформления. Особенности исходного издания в предлагаемом проекте не являются обязательными к сохранению.

Зато при проектировании сервиса был учтен опыт параллельного корпуса в составе НКРЯ [Добровольский и др. 2005]. При размещении текстов на таком ресурсе ключевым становится подготовительный процесс *выравнивания*. Выравнивание (alignment) — это способ сопоставить тексты, отыскав и связав между собой интуитивно тождественные отрезки. Даже для прозаических переводов выравнивание часто представляет собой сложности. Переводчики произвольно делят речь оригинала на сегменты, переставляют их местами, меняют их порядок следования в тексте. Еще более сложным образом устроено выравнивание поэтического текста, где на соображения вкуса переводчика накладываются ограничения, связанные со стихотворной формой. Таким образом, если при подготовке текстов для НКРЯ выравнивание производилось полуавтоматически, то для Параллельного корпуса переводов «Слова о полку Игореве», в который попадают и поэтические произведения, альтернатив ручному выравниванию не было. Кроме общих соображений, работу затрудняли и специальные особенности рассматриваемого текста. В Первой публикации есть эпизод (затмение), переставленный относительно хронологической последовательности событий. В науке не решен вопрос, является эта перестановка

следствием эстетического замысла автора или механической порчи текста. Многие исследователи подозревают, что один из листов рукописи с этим эпизодом просто выпал из кодекса и был возвращен не на свое, а на произвольное место. В результате часть переводов воспроизводят композицию Первого издания, а часть гипотетически реконструированную исходную композицию текста. Вторые (например, армянский перевод С. Вауни) имеют в корпусе специальную помету в виде пиктограммы в начале текста 8-го фрагмента.

Итак, параллельные корпуса представляют собой особенным образом подготовленные цифровые текстовые коллекции, в которых тексты на разных языках, являющиеся переводами один другого, сопоставлены. Технически это сопоставление реализовано как разрезание текста на отдельные фрагменты текста (обычно предложения) и нахождение соответствий между этими фрагментами. Например, в книге Rich Karlgaard «Late Bloomers: The Power of Patience in a World Obsessed with Early Achievement» (2019) есть предложение «Lehrer, like Jefferson, could also write cogently». Цель выравнивания в том, чтобы зафиксировать его соотнесенность с предложением из русского перевода (Рич Карлгаард «Поздний расцвет. Как взрослым добиться успеха в мире, одержимом ранним развитием», переводчик Ксения Свешникова, 2020): «Точно так же, как Джефферсон, Лерер обладал убедительным слогом».

Такой подход порождает специфические вопросы, например, как должно производиться деление текста. В условиях выравнивания вольно переведенных поэтических текстов деление на предложения не всегда представляется оптимальным. Кроме того, само деление древнего текста на предложения представляется проблематичным из-за неясности чтения некоторых мест: «Скочи отъ нихъ лютымъ звеpемъ въ плъночи изъ Бела-града, обесися сине мьгле, утръ же воззни стрикусы, отвори врата Нову-граду, разшибе славу Ярославу...» Загадочное слово «стрикусы» вносит изрядную долю неясности в приведенное предложение. Место следует читать то ли как: «Скакнул от них лютым зверем/в полночь из Белгорода,/объятый синей мглой,/добыл он счастье,/с трех попыток отворил ворота Новгорода,/расшиб славу Ярославу...» (Д. Лихачев, «стрикусы» трактуется как «с трех попыток», относится к «отвори врата Нову-граду»); то ли как «Скакнул от них лютым зверем в полночи, из Белгорода,/завесясь синим туманом./Вначале же возник яростным отрядом,/отворил ворота Новгороду, расшиб славу Ярославу...» (Л. Гурченко, «стрикусы» трактуется как «яростный от-

ряд», сохраняет синтаксическую самостоятельность). Однако в традиции филологической науки о «Слове о полку Игореве» уже есть устоявшийся способ членить текст. Впервые он был продемонстрирован в издании «Слова» под редакцией Романа Якобсона [La Geste 1948]. Якобсон разбил текст на 218 фрагментов, и это членение затем многократно воспроизводили в своих изданиях западные слависты (что облегчало выравнивание). Мы решили последовать за авторитетом Р. Якобсона, и искать соответствие в переводах между фрагментами текста, выделенными знаменитым лингвистом.

Так как в коллекцию корпуса должны были быть включены не только переводы на современный русский язык, но и на другие языки, в том числе далекие от европейского стандарта, а выравнивание текстов могло быть только ручным, проблема состояла в том, чтобы найти людей, знающих все требуемые языки, и убедить их помочь с размещением текстов в корпусе. В реальности, конечно, отдельные люди не требовались на каждый новый язык. В группу специалистов, работающих над составлением корпуса, входили славист Е. А. Слободян и романист М. С. Рыбина. Именно они в итоге помогли осуществить выравнивание существенной части текстов на славянских (включая западнославянские польский и чешский) и романских (французский, испанский, португальский) языках.

Однако эта ручная работа имела смысл только при условии создания подходящей технической платформы для размещения вебверсии корпуса. С самого начала планировалось, что корпус «Слова» будет функционировать по модели НКРЯ — как общедоступный бесплатный ресурс с доступом из сети интернет. Во второй половине 2000-х к такому режиму привыкли еще не все университетские люди, и преподавательница в моей alma mater, пользуясь знакомством, спрашивала меня как одного из разработчиков НКРЯ, какую официальную бумагу и куда нужно подать, чтобы поработать с корпусом.

Первая версия корпуса «Слова» была написана на языке программирования PHP, тексты были размещены в базе данных под управлением MySQL. Некоторые технические подробности, ставшие результатом стратегических концептуальных решений, по горячим следам изложены в статье [Орехов 2009]. В дальнейшем подобный стек был признан слишком сложным для такой задачи, и сайт в 2017 году переписан на Python 2, он функционирует через cgi-гейт на обычном виртуальном хостинге, а тексты хранятся в JSON-файле. Объем небольшой поэмы, даже повторенной в сотнях вариантов, не требует мощностей СУБД.

Тогда же интерфейс сервиса получил дополнение в виде визуализации сравнения текстов переводов, реализованной на библиотеке traviz для языка JavaScript.

Среди особенностей первой версии сайта — наличие многоязычного интерфейса. Наряду с русским на сайте присутствовали также английский, хорватский и украинский языки. Украинский появился в этом ряду благодаря тому, что это один из лидеров после русского по числу переводов. В корпусе их насчитывается более 20. На все выбранные языки существовало больше одного перевода «Слова». Однако дальнейшее развитие корпуса и изменение интерфейса заставили отказаться от многоязычности: консультации с экспертами (О. Давыдова, Р. Божич, Д. Паньков) не успевали за нововведениями, и многоязычность стала тормозить проект.

Кроме технических задач и задач выравнивания, при создании корпуса приходилось решать и самые затратные вопросы — оцифровки. После того, как были обработаны для корпуса все, имеющиеся в цифровом виде в интернете переводы, были затребованы в библиотеках и переведены в электронный формат десятки переводов на десятки языков мира, включая грузинский, арабский, китайский. На эти языки приходится в первую очередь обращать внимание, потому что в тот момент для них не существовало общедоступных средств OCR. В контексте проекта это означало, что эксперты не только выравнивали текст, но и с высокой вероятностью набирали его вручную. Самым ценным библиографическим источником о существовании переводов являлись соответствующие статьи в Энциклопедии «Слова о полку Игореве» [Алексеев и др. 1995; Творогов 1995]. Неоценимую помощь проекту в оцифровке, выравнивании и сопутствующих процессах оказали энтузиастически включившиеся в дело Н. Мухаметдинова, Е. Шаульский, А. Полян, С. Оплатек, А. Оськина и другие.

Такая масштабная деятельность, далеко выходящая за рамки волонтерской, требовала финансовой поддержки, оказанной проекту в 2008 году Российским гуманитарным научным фондом в рамках гранта № 08-04-12104в.

На текущий момент на ресурсе сосредоточено 247 текстов «Слова», включая 12 непереводных реконструкций исходного текста, 112 переводов на современный русский язык, 52 перевода на другие славянские языки, 5 переводов на языки народов России, 16 на, условно говоря, восточные языки. Некоторые любопытные закономерности, связанные с этими переводами, изложены в научно-популярной публикации [Орехов 2010].

До сих пор не разрешенным является вопрос авторских прав. За 18 лет существования проекта создатели не получили ни одной претензии, касающейся копирайта. По всей видимости, на спокойствие создателей работает небольшие масштаб и известность проекта. Вряд ли при аналогичных трудностях таким отсутствием забот могли бы похвастаться более заметные и масштабные проекты вроде *Слово Толстого* и *Пушкин <цифровой>*.

Однако претензии от пользователей поступали, но касались они не правовых вопросов, а недостатков интерфейса и оцифровки. Профессиональные пользователи были недовольны выбором разбивки Р. О. Якобсона и настаивали на альтернативных способах выравнивания. Этот запрос удовлетворить невозможно, так как членение текста, предложенное американским славистом русского происхождения, является одной из конструктивных особенностей сервиса, и отказ от него подразумевал бы не только глобальную техническую перестройку ресурса, но и работу по подготовке всех текстов корпуса с нуля. Часть запросов поступала от непрофессиональных переводчиков «Слова», которые видели корпус в качестве авторитетной публикационной площадки. Заметим, что этот запрос создатели корпуса старались удовлетворять, не выстраивая внутрикультурных иерархий. Часть переводов на сайте имеют помету «Текст прислан переводчиком».

Тем не менее, длинная история сервиса позволяет замерить его востребованность на освоенной временной дистанции, и прийти к неутешительным выводам. Корпус в итоге не стал тем, для чего планировался — инструментом сравнения текстов. Отклики в научной периодике и известные нам практики использования сайта за прошедшие 18 лет свидетельствуют, что ресурс представляет собой для пользователя скорее библиотеку текстов, рассматриваемых каждый в отдельности, а не в сравнительном аспекте. Вероятно, это ожидаемо, так как запрос на сравнение текстов — запрос весьма специфический, но даже в тех случаях, когда он возникает в научном поле, корпус может быть не задействован, как в [Поварницына 2020].

Это не значит, что само по себе поле сравнения переводов бедное и малоперспективное. Здесь еще остается целый ряд нерешенных вопросов. Например, что наименее охотно воспроизводят переводчики, пропуская фрагменты исходного текста? Какие фрагменты имеют тенденцию к тому, чтобы быть пропущенными? Кажется, что таких мест в «Слове» больше ближе к концу произведения. Если это так, то в этом проявляет себя известный древнерусским тек-

стологам эффект «усталости» переписчиков, которые совершают больше ошибок ближе к концу рукописи, или там сосредоточены риторически наименее выразительные или сложнопереводимые части? Как связана непереводимость с испорченностью и трудночитаемостью текста?

Итак, научный ресурс, о котором шла речь, существует уже больше 15 лет. Для электронной филологии срок немалый. Появившись в сети в феврале 2007 года, он не стал первым ни в ряду корпусов русского языка (НКРЯ появился тремя годами ранее), ни в ряду русскоязычных сайтов, специально ориентированных на интересы филологов (в этом смысле первенцем можно назвать Русскую виртуальную библиотеку, основанную 1 декабря 1999 года, о её специфической ориентации по сравнению с другими электронными собраниями текстов см. в статье Евгения Горного[1]). Тем не менее, в 2007 году набор профессиональных веб-приложений, которые могли бы помочь в исследовании литературных текстов, был весьма скуден, а само сообщество, по большому счету, не готово к их использованию. Сейчас этот проект уже не смотрится экзотично, а напротив, органично встраивается в растущий конгломерат аналогичных ресурсов, объединяемых зонтичным термином Digital Humanities.

Параллельный корпус переводов «Слова о полку Игореве» был задуман по аналогии с особым подвидом лингвистических корпусов, параллельными корпусами, как инструмент, облегчающий сравнение между собой текстов специальной природы — «интуитивно равнозначных (синонимичных)» [Мельчук 1999 С. 5]. Уже во вторую очередь корпус мог быть сетевой коллекцией труднодоступных произведений, но именно в этом качестве оказался востребован.

Поэтому остается сожалеть, что в корпусе всё ещё нет многих интересных текстов: русского перевода С. К. Шамбинаго, вьетнамского перевода Тхеу Тоан Дыха (Ханой, 1987), древнегреческого перевода П. И. Нищинского (информация о том, что такой текст существует, взята из [Новый 1916. Стб. 659]), но и сейчас он выполняет вмененные ему историей и пользователями задачи.

Примечания

[1] Горный Е. Интернет и филология. Субъективные заметки к годовщине РВБ. http://rvb.ru/about/meta/inter-phil.html

Литература

Исследования

Алексеев и др. 1995 — Алексеев А., Павлик (Оденсе) Я., Айвазян К., Творогов О., Тульчинский А., Буланин Д., Бегунов Ю., Белоброва О., Артико Д., Камбарбаева Г., Старшова Т., Булахов М., Штурм Г., Келер С., Гарней Ю., Рымарович С., Бельдиян В., Водолазкин Е. Гребнева Э., Воинов С., Накамура Ё. Переводы «Слова» и литература о нем на языках народов мира // Энциклопедия «Слова о полку Игореве»: В 5 т. СПб.: Дмитрий Буланин, 1995. Т. 4. П–Слово. 1995. С. 23–76.

Гаспаров 2001 — Гаспаров М. Л. Подстрочник и мера точности // Гаспаров М. О русской поэзии: Анализы. Интерпретации. Характеристики. СПб.: Азбука, 2001. С. 361–372.

Добровольский и др. 2005 — Добровольский Д. Кретов А. Шаров С. А. Корпус параллельных текстов: архитектура и возможности использования // Национальный корпус русского языка: 2003–2005. Результаты и перспективы. М.: Индрик, 2005. С. 263–296.

Егунов 2001 — Егунов А. Н. Гомер в русских переводах XVIII-XIX веков. М: Индрик, 2001. 400 с.

Мельчук 1999 — Мельчук И. А. Опыт теории лингвистических моделей «СМЫСЛ ⇔ ТЕКСТ» М.: Школа «Языки русской культуры», 1999. XXII, 346 с.

Новый 1916 — Новый энциклопедический словарь. Пг., 1916. Т. 28.

Орехов 2010 — Орехов Б. В. Переводы «Слова о полку Игореве»: мифы и реальность // Нева. 2010. № 1. С. 155–161.

Орехов 2009 — Орехов Б. В. Параллельный корпус переводов «Слова о полку Игореве»: итоги и перспективы // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009. С. 462–473.

Поварницына 2020 — Поварницына Е. О лексической вариативности в переводах с мёртвого языка (флора и фауна в «Слове о полку Игореве» и его переводах на французский язык) // Изв. Сарат. ун-та. Нов. сер. Сер. Филология. Журналистика. 2020. Т. 20, вып. 1. С. 15–20. DOI: 10.18500/1817-7115-2020-20-1-15-20

Творогов 1995 — Творогов О. В. Переводы «Слова» на современный русский язык // Энциклопедия «Слова о полку Игореве»: В 5 т. СПб.: Дмитрий Буланин, 1995. Т. 4. П–Слово. 1995. С. 76–78.

La Geste 1948 — La Geste du prince Igor'. New-York, 1948.