

# Вера Сибирцева

---

## Технология использования параллельного подкорпуса Национального корпуса русского языка и коллекций текстов в обучении иностранным языкам

---

Rocznik Instytutu Polsko-Rosyjskiego nr 2, 98-110

---

2013

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej [bazhum.muzhp.pl](http://bazhum.muzhp.pl), gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach  
dozwolonego użytku.

**Вера Сибирцева**

**Технология использования параллельного подкорпуса  
Национального корпуса  
русского языка и коллекций текстов  
в обучении иностранным языкам**

Применение НКРЯ и программ автоматического выравнивания текстов является новым этапом в обучении иностранным языкам, в частности, русскому языку как иностранному. В статье демонстрируются возможности параллельного подкорпуса НКРЯ (на примере русских и польских текстов, а также их переводов) и коллекции выровненных параллельных текстов, созданной при помощи бесплатного программного продукта LF aligner, для последующей разработки творческих лексико-стилистических заданий.

**Ключевые слова:** русский язык как иностранный; польский язык; Национальный корпус русского языка; информационные технологии; компаративный анализ.

## **1. Введение**

Использование текстов художественной литературы в учебных целях, особенно в области обучения иностранным языкам, принято рассматривать как своеобразный феномен, близкий кантовскому «Ding an sich». Многочисленные существующие методики работы с художественным текстом опираются в основном на интуитивный, творческий, «немашинный» анализ. В то же самое время в педагогические и филологические науки все больше проникают статистические и корпусные методы исследования, а многочисленные компьютерные программы (автоматические демонстраторы рифм, семантические референторы текста, креаторы облаков тегов, конкордансеры и т.п.) с успехом могут применяться в обучении иностранным языкам.

Рассмотрим способы применения параллельного подкорпуса НКРЯ и коллекций текстов<sup>1</sup> в обучении, на примере рус-

---

<sup>1</sup> Лингвистическим корпусом называется собрание текстов, существующих на некотором языке в электронной форме, снабженное специальной экстралингвистической и лингвистической информацией (В. Плунгян, <http://www.philology.ru/linguistics2/plungyan-05.htm>). Под коллекцией мы будем понимать неразмеченные (т.е. не снабженные лингвистической информа-

ского и польского языков, и подробно опишем программу, позволяющую создать параллельную коллекцию текстов для дальнейшего лингво-стилистического анализа.

## **2. Национальный корпус русского языка (параллельный подкорпус) в обучении РКИ**

Одним из наиболее интересных и востребованных направлений в современной лингвистике является активно развивающаяся корпусная лингвистика. В сфере сопоставительных исследований и сознательных методов обучения иностранному языку<sup>2</sup> параллельные лингвистические корпуса являются своеобразным банком сопоставленных текстов, при помощи которых можно изучать язык на продвинутом этапе. Если на начальном этапе изучения иностранного языка пальму первенства продолжают удерживать коммуникативные и смешанные методы обучения языку, то для более глубокого знания языка и его особенностей элемент сознательного сопоставления очень важен.

Для лингвистического анализа и выявления особенностей текста на языке оригинала и его перевода обращение к подкорпусу параллельных текстов, содержащемуся в Национальном корпусе русского языка ([www.ruscorgo.ru](http://www.ruscorgo.ru)), представляется наиболее целесообразным. В нем тексту на иностранном языке сопоставлен перевод этого же текста на русский язык, и, соответственно, тексту на русском языке сопоставлен перевод на иностранный. Между элементами оригинального и переводного текста (главами, абзацами, предложениями) установлено соответствие.

Параллельные фрагменты текста располагаются в одинаковом порядке в обеих частях корпуса: оригинальном тексте и переводном. Опорными элементами для сопоставления являются даты, содержащиеся в тексте, имена собственные, а также нередко начало и конец абзаца. Интересный материал для анализа могут представлять собой лексические лакуны – пропуски

---

цией) собрания текстов, в значительной степени уступающие лингвистическим корпусам по объему.

<sup>2</sup> Гальскова Н.Д., Гез Н.И. *Теория обучения иностранным языкам. Лингводидактика и методика.* – М.: Академия, 2007, 336 с.

в переводе. Корпусная лингвистика позволяет использовать точные количественные данные о частотности различных грамматических конструкций, а также рассматривать синонимичную лексику в контексте. Так, глаголы «молвить» и «сказать» могут переводиться на польский язык одинаково:

– Довершайте начатое, – сказал он, – не останавливайтесь на полдороге. [Н.А. Островский. Как закалялась сталь (ч. 2) (1930–1934)]

„Doprowadźcie do końca rozpoczęte dzieło – powiedział – nie zatrzymujcie się w połowie drogi” [Nikołaj Ostrowski. Jak hartowała się stal (cz 2)]

– молвил Иван, не зная, что делать с этим цилиндром дальше. [М.А. Булгаков. Мастер и Маргарита (ч. 1) (1929–1940)]

– powiedział Iwan nie wiedząc, co teraz z tym cylindrem począć. [Michaił Bułhakow. Mistrz i Małgorzata (cz 1)]

При этом количественная выдача результатов свидетельствует о том, что глагол «молвить» является малоупотребительным, устаревшим.

Так, своеобразие выражения модального значения в русском языке можно показать при помощи вводного слова «авось». Приведём несколько результатов поиска оригинального текста и перевода по слову «авось»:

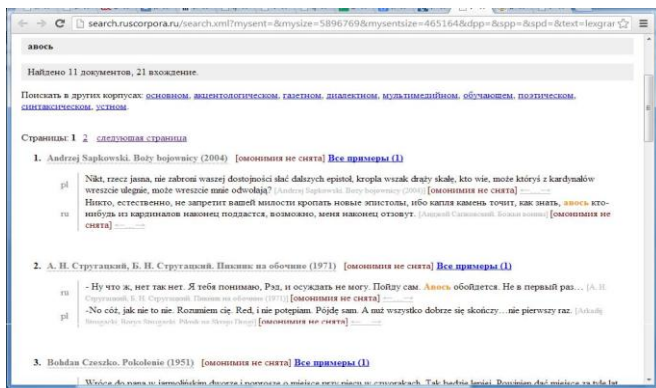


Рис. 1. Результаты поиска «авось» в НКРЯ

– Ну что ж, нет так нет. Я тебя понимаю, Рэд, и осуждать не могу. Пойду сам. Авось обойдётся. Не в первый раз... [А.Н. Стругацкий, Б.Н. Стругацкий. Пикник на обочине (1971)]

– No cóż, jak nie to nie. Rozumiem cię. Red, i nie potępiam. Pójdę sam. A nuż wszystko dobrze się skończy...nie pierwszy raz. [Arkadij Strugacki, Borys Strugacki. Piknik na Skraju Drogi]

Pan sądzi, że można każdej kobiecie, która jest w teatrze, śmiało mówić o swojej miłości, ot tak sobie...a może się uda!... [Władysław Reymont. Komediantka (1896)]

Вы полагаете, каждой женщине, если она в театре, можно без стеснения говорить о своей любви. Авось удастся! [Владислав Реймонт. Кomediantka]

– Widzisz, synku! nie brykaj! –szeptał sufler, tak mocno trzymając go, że Glas nie mógł się poruszyć. –Położyłeś się na amen!... Malowałeś Dobka, umalował ciebie Dobek... [Władysław Reymont. Komediantka (1896)]

– Смотри, сынок, не брыкайся! — шептал Добек, вцепившись в ногу врага так, что тот не мог пошевелиться. — Понадеялся на авось! Размалевал Добека, и Добек в долгу не останется. [Владислав Реймонт. Кomediantka]

Авось дадим отпор Пугачёву. Господь не выдаст, свинья не съест! [А.С. Пушкин. Капитанская дочка (1836)]

A nuż oprzemy się Pugaczowowi? Jak Pan Bóg dopuści, to i z kija wypuści! [Aleksander Puszkín. *Córka kapitana*].

Как видим из примеров, слову «авось» соответствуют различные эквиваленты. Причем при обратном переводе данные эквиваленты также могут переводиться различно.

В выданных результатах встречаются также примеры лексических лакун:

– Ja tu nie będę czekał, aż mnie pochwyć —zawołał. — Spróbuję ucieczki; tak czy owak ginać, lepiej choć się pokusić o wyzwolenie. [Józef Ignacy Kraszewski. Kunigas (1881)]

– Я не стану ждать, пока меня поймают, — закричал он, — попробую бежать, все равно ведь погибать; авось спасусь! [Юзеф Игнацы Крашевский. Кунигас]

Wróć do pana w jarmolińskim dworze i poproszę o miejsce przy piecu w czworakach. Tak będzie lepiej. Powinien dać

miejsce za tyle lat roboty... [Bohdan Czeszko. Pokolenie (1951)]

Вернусь к помещику, авось даст угол за печкой где-нибудь в бараках. Не откажет после стольких лет работы... [Богдан Чешко. Поколение].

Одной из форм работы с иностранным языком на продвинутом этапе может стать подбор примеров для создания упражнений. Такой тип работы не только углубляет знания языка, но и способствует самоорганизации, позволяет взглянуть на текст глазами исследователя и педагога, а не только обучающегося-потребителя. Продемонстрируем некоторые примеры обращения к подкорпусу параллельных текстов Национального корпуса русского языка и опишем технические шаги, необходимые для вывода искомых примеров на экран компьютера или ноутбука. НКРЯ снабжён дружественным пользовательским интерфейсом и удобной навигацией, освоить которую можно без обращения к подробной, но, к сожалению, очень объёмной инструкции пользователя.

Для подготовки упражнений и заданий по русскому языку как иностранному удобно пользоваться поиском в параллельном корпусе, задав определённые «грамматические признаки». В данном случае нас интересуют местоимение «весь» и имя прилагательное «целый» (с морфологической точки зрения). Поскольку основное количество примеров выдаётся с неснятой омонимией, в «лексико-грамматическом поиске» необходимо для каждого слова задать второе слово, указав только его грамматический признак – имя существительное, чтобы исключить попадание в список результатов грамматических омонимов («достигнуть определённых *целей*», «сказал *все*»).

Так, для различения значения слов «весь» и «целый» можно предложить обучающимся найти примеры использования этих слов и на основе найденных предложений составить подстановочные упражнения.

I było ich mnóstwo, cała chmara, rój. [Andrzej Sapkowski. Narrenturm (2002)].

И было их множество, целая туча, рой. [Анджей Сапковский. Башня шутов]

Kiedy w Europie całej, popatrzcie jeno, ciągiem ninie zgiełk bitewny? [Andrzej Sapkowski. Lux Perpetua (2006)].

Когда во всей Европе, посмотрите только, беспрестанно шум битвы. [Андрей Сапковский. Свет вечный]

W takiej chwili- chciałby cały świat nabić dynamitem i rozsadzić. [Bolesław Prus. Lalka (t. 2) (1887-1890)]

В эту минуту он готов был весь мир начинить динамитом и взорвать. [Болеслав Прус. Кукла (т. 2)]

Za drzewami owocowymi ciągną się niedostępne zarośla, gaje wiktów i rokitnicy. [Stefan Żeromski. Popioły (1902)]

За плодовыми деревьями тянутся непроходимые заросли, целые рощи ив и раkitника. [Стефан Жеромский. Пепел]

Другим видом работы являются мини-исследования: например, поиск слова «наконец» как вводного или в качестве обстоятельства. Оформление вводных слов в русском языке отличается от польского, поэтому данный вид работы не только расширяет словарный запас, но и развивает контекстную и пунктуационную грамотность. Для того, чтобы искомое слово являлось вводным, необходимо на странице поиска указать «дополнительные признаки»: слово «после запятой» и «перед запятой»:

Когда же слухи об этом дошли, наконец, и до мужа, Анна, не задумываясь, покинула его дом и переехала к своему любовнику. [В.Я. Брюсов. Через пятнадцать лет (1909)]

Kiedy zaś pogłoski o tym dotarły wreszcie i do męża, Anna, niewiele myśląc, porzuciła jego dom i przeniosła się do swego kochanka. [Walery Briusow. Po piętnastu latach].

Для поиска обстоятельства в «дополнительных признаках» для слова «перед запятой» необходимо поставить минус:

Udawałam szalone nim zainteresowanie wyłącznie w tym celu, żeby mąż się wreszcie zaniepokoił. [Joanna Chmielewska. Wszyscy jesteśmy podejrzani (1966)]

Я притворялась безумно заинтересованной им с единственной целью – чтобы муж наконец забеспокоился. [Иоанна Хмелевская. Мы все под подозрением]

Подведём некоторые итоги использования параллельных корпусов в обучении.

Билингвальность как основополагающее дидактическое свойство параллельных корпусов позволяет успешно формировать лексические навыки у обучающихся иностранным языкам. Благодаря контексту корпуса параллельных текстов способствуют более точному определению значения слова и выбору синонима из ряда возможных. Продуманная навигация по параллельному подкорпусу Национального корпуса русского языка не только сокращает время нахождения искомым слов, но и позволяет производить поиск в различных сферах функционирования языка, от повседневного общения до деловой и профессиональной коммуникации.

Особенности словоупотребления могут быть показаны как в узком, так и в расширенном контексте. Выдача результатов поиска в формате KWIC (конкордансный список примеров) даёт возможность просмотреть несколько вариантов употребления слова в контексте на языке оригинала.

### **3. Использование компьютерной программы LF aligner для создания учебных коллекций параллельных текстов**

Как видим, применение национальных (генеральных) лингвистических корпусов в обучении иностранным языкам имеет очень широкий прагматический спектр. Но возможности параллельных подкорпусов ограничены тем количеством текстов, которые в них содержатся. Стиль переводчика может в большой степени влиять на текст, и часто, работая с готовым корпусом, исследователь анализирует прежде всего стиль конкретного переводчика, а не сам текст. Кроме того, в переводе часто наблюдается искажённый, калькированный или упрощённый вариант языка.

Джарл Эбелинг (Jarle Ebeling) сформулировал основные недостатки параллельных корпусов так:

- язык перевода искажается под влиянием языка источника;
- язык перевода отличается от естественного языка;
- переводчики делают ошибки;
- перевод зависит от индивидуального стиля переводчика;



- перевод непредсказуем и обусловлен только текстом-источником и его особенностями»<sup>3</sup>.

При исследовании различных переводов одного и того же произведения можно проследить диахронные и стилистические изменения в лексике. Для таких целей собирается электронная коллекция текстов. «Сложные» художественные тексты, как правило, содержат большое количество пропусков или случаев асимметричного перевода и дают наиболее богатый материал для филологического анализа и интерпретации. Помимо несохранения авторской разбивки текста на абзацы и предложения, обучающийся может столкнуться с иным визуальным оформлением текста: в разных языках (иногда также в хронологически разных изданиях произведения) могут использоваться различные способы графического оформления прямой речи и диалогов.

Компенсировать эти недостатки позволяет использование в исследованиях большого числа переводов разных авторов, а также переводов в разных направлениях (с языка и на язык). Для учебных целей может быть создана коллекция параллельных текстов при помощи бесплатной программы для некоммерческого использования LF aligner.

Программа LF aligner, применяемая для выравнивания параллельных текстов (собственно, относящаяся к вспомогательным инструментам переводчика) и проведения статистических подсчетов в лингвистических исследованиях, интересна в целях обучения иностранному языку по нескольким причинам. Во-первых, визуальное представление результатов выровненных текстов удобно для анализа. Во-вторых, возможности программы не ограничиваются двумя текстами, при необходимости можно добавить четыре и более текстов на различных языках. И, наконец, интерфейс программы удобен для работы филолога, простые команды не требуют специфических программистских навыков, а весь процесс обработки текстов занимает несколько минут. Для корректной работы LF Aligner не-

---

<sup>3</sup> *Corpus linguistics: an international handbook* / edited by Anke Ludeling and Merja Kyto. – Berlin, 2008. 776 p.

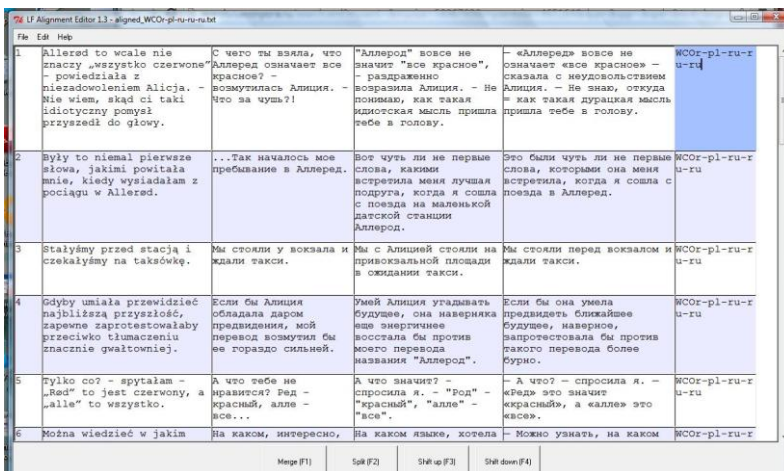
обходимо использовать только латиницу, как в местах хранения файлов, так и в их названиях.

Обрабатываемые тексты могут быть в привычных текстовых форматах: обычный текст (сохранённый в UTF-8), документ Microsoft Word или защищённый документ.

Кроме того, сопоставляться могут локальные html-файлы, страницы сайта (достаточно указать ссылки на страницы файлов оригинала и перевода), а также документы Еврокомиссии и Европарламента. Таким образом, в зависимости от специализации студентов, изучающих иностранный язык, могут привлекаться не только художественные, но и специальные (технические, экономические и т.д.) тексты.

Опишем пример использования программы для сопоставления трёх вариантов перевода фрагментов романа И. Хмелевской «Все красное» (В. Селиванова, М. Кронгауз, О. Кузнецова для курса обучения иностранным языкам по методу Ильи Франка). Программа LF Aligner наименее требовательна к пользовательским навыкам.

Работа с обычным текстом в кодировке Юникод представляется нам наиболее удобной. Анализировались начальный фрагмент романа, где обсуждается название города Аллерод, а также диалоги с участием сотрудника полиции господина Мульдгорта. Выдаваемый результат удобен для визуальной оценки качества выравнивания, так как представляет тексты в виде таблицы из нескольких столбцов (в нашем случае пять): с предложениями из текста-оригинала, трёх текстов-переводов, и столбцом, содержащим метаинформацию об источниках текста. Для анализа можно выбрать различное членение текста (два варианта выравнивания) программой LF Aligner: с применением сегментации на предложения, и без неё. В отдельном окне программой выдаётся информация о количестве сегментов в тексте-оригинале и переводах, в случае с текстами И. Хмелевской их было 25 (начало романа). Просмотр и правка сегментированного текста осуществляется при помощи встроенного редактора.



1	2	3	4	5	6
Allerrod to wcale nie znaczy „wszystko sztywne” – powiedziała z niezadowolaniem Alicja. – Nie wiem, skąd ci taki idiotyczny pomysł przyszedł do głowy.	С чего ты взяла, что Allerrod означает все красное? – возмутилась Алиция. – Что за чушь?!	„Allerrod” вовсе не значит “все красное”, а разравжимо Алиция. – Не понимаю, как такая idiotская мысль пришла тебе в голову.	– „Allerred” вовсе не означает «все красное» – сказала с неудовольствием Алиция. – Не знаю, откуда – как такая дурацкая мысль пришла тебе в голову.	MSOR-pl-ru-r u-ru	
Były to niemal pierwsze słowa, jakimi powitała mnie, kiedy wysiadałam z pociągu w Allerred.	...так началось мое пребывание в Allerred.	Все чуть ли не первые слова, какими встретила меня лучшая подруга, когда я сошла с поезда на маленькой датской станции Allerrod.	Это были чуть ли не первые слова, которыми она меня встретила, когда я сошла с поезда в Allerred.	MSOR-pl-ru-r u-ru	
Staliśmy przed stacją i czekaliśmy na taksówkę.	Мы стояли у вокзала и ждали такси.	Мы с Алицией стояли на привокзальной площади в ожидании такси.	Мы стояли перед вокзалом и ждали такси.	MSOR-pl-ru-r u-ru	
Gdyby umiała przewidzieć najbliższą przyszłość, zapewne zaprotestowałaby przeciwko tłumaczeniu znacznie gwałtowniej.	Если бы Алиция обладала даром предвидения, мой перевод возмутил бы ее гораздо сильнее.	Умей Алиция угадывать будущее, она наверняка еще энергичнее восстала бы против моего перевода названия “Allerrod”.	Если бы она умела предвидеть ближайшее будущее, наверняка zaprotestowała бы против такого перевода более бурно.	MSOR-pl-ru-r u-ru	
Tylko co? – spytałam – „Red” to jest czerwony, a „alle” to wszystko.	А что тебе не нравится? Red – красный, alle – все...	А что значит? – спросила я. – “Red” – “красный”, “alle” – “все”.	– А что? – спросила я. – “где” это значит “красный”, а “alle” это “все”.	MSOR-pl-ru-r u-ru	
Można wiedzieć w jakim	На каком, интересно,	На каком языке, хотела	– Можно узнать, на каком	MSOR-pl-ru-r	

Рис. 2. Результат сопоставления текстов в LF aligner

Уже первый фрагмент даёт представление о различных стратегиях переводчиков. Перевод Максима Кронгауза содержит значительные лакуны, опускаются различные детали, недостаток которых не влияет на общий ход сюжета, но лишает текст оригинала шарма, свойственного женскому разговору. Учебный, буквальный перевод Ольги Кузнецовой наиболее близок оригиналу, в данном случае автором перевода сознательно выбиралось следование букве исходного текста, при этом ироничный стиль Иоанны Хмелевской не сохранился. Перевод Веры Селивановой общепризнанно считается наиболее удачным, хорошим русским языком передаётся весь заряд иронии и юмора оригинального текста и авторская любовь к деталям.

Фрагменты романа, содержащие диалоги с господином Мульдгортом, интересны для рассмотрения способов передачи своеобразного польского языка датского полицейского. Приведём только один пример, в котором переводчики обыгрывают неправильную форму множественного числа в вопросе полицейского:

– Co robiły one)? – spytał pan Muldgaard. – Dlaczego tylko my? – zaprotestowała Zosia z oburzeniem i pretensją, w przekonaniu, iż pytanie odnosi się wyłącznie do kobiet.

– Во оноже время, – сформулировал следователь свой следующий вопрос, – што она твориху? – Кто „она”? И почему он спрашивает только о женщинах? – возмутилась Зося [В. Селиванова].

– Кто и что делали особы? – Почему только мы? – вознегодвала Зося, считая, что вопрос относится к женщинам [М. Кронгауз].

В учебном переводе Ольги Кузнецовой приходится прибегать к пояснениям:

– Что делали они (one – они, по отношению только к лицам женского пола)? – спросил пан Мульгор. – Почему только мы? – запротестовала Зося с возмущением и обидой, убеждённая, что вопрос относится только к женщинам.

Разумеется, стилистические особенности текста можно анализировать только на достаточно продвинутом уровне изучения иностранного языка, но нашей задачей в данном случае являлась демонстрация возможностей программы LF aligner в учебных целях. Сопоставление скриншота окна программы и последовательно приведённых примеров наглядно визуализируют разницу представления параллельных текстов.

#### **4. Выводы**

Как видим, корпусные технологии с успехом могут применяться в обучении иностранному языку, в нашем конкретном случае – русскому языку как иностранному. Использование корпусных технологий позволяет получить актуальную информацию о современном функционировании языка, в то время как традиционные учебники не успевают своевременно отражать активные изменения на уровне лексики и грамматики, а также не всегда содержат исчерпывающее количество поясняющего материала.

Знакомство с компьютерными программами и умение ими пользоваться не только в исследовательских, но и в обучающих целях во многом способствует нетривиальному, творческому построению обучения русскому языку как иностранному. Обращение к корпусам и коллекциям текстов расширяет

и углубляет знания о языке, формирует у обучающихся адекватные профессиональные компетенции.

Дирк Ноэль писал, что параллельные корпуса – это «собрание мнений информантов о значении лингвистических форм в исходном тексте, с дополнительным достоинством, состоящим в том, что они доступны исследователю без необходимости организации эксперимента»<sup>4</sup>. [Corpus linguistics: an international handbook 2008]. Эти слова в полной мере можно отнести к учебным коллекциям параллельных текстов, создаваемых в учебных и в исследовательских филологических целях.

Рассмотренные методы анализа текста в обучении русскому языку как иностранному с неменьшим успехом могут быть перенесены на обучение другим языкам. Два фактора: постоянное развитие национальных корпусов и существование бесплатного контента в Интернете – все больше привлекают внимание филологов и педагогов к точным и современным методам исследования.

Последовательное описание применения существующих программ, тестирование их на художественном материале и сопоставление полученных данных с существующими традиционными исследованиями может стать новым этапом анализа художественного текста и обучения иностранным языкам в век цифровых технологий.

В данной научной работе использованы результаты, полученные в ходе выполнения проекта № 12-01-0129 «Автоматическая обработка данных и их филологическая интерпретация: So einfach wie möglich. Aber nicht einfacher», выполненного в рамках Программы «Научный фонд НИУ ВШЭ» в 2013 г.

**TECHNIKA WYKORZYSTANIA RÓWNOLEGŁEGO PODKORPUSU NARODOWEGO KORPUSU JĘZYKA ROSYJSKIEGO I ZBIORÓW TEKSTÓW W NAUCZANIU JĘZYKA OBCEGO**

Wykorzystanie Narodowego Korpusu Języka Rosyjskiego (NKJR) otwiera nowe możliwości w nauczaniu języków obcych, zwłaszcza języka rosyjskiego jako Obcego. W artykule zaprezentowano możliwości równoległego podkorpusu NKJR na

---

<sup>4</sup> *Corpus linguistics: an international handbook* / edited by Anke Ludeling and Merja Kyto. – Berlin, 2008. 776 p.

bazie tekstów polskich i rosyjskich oraz ich tłumaczeń jak i zbioru tekstów stworzonego przy pomocy programu LF aligner.

Słowa kluczowe: język rosyjski jako obcy; język polski; Narodowego Korpusu Języka Rosyjskiego; technologie informatyczne; analiza porównawcza

#### **TECHNOLOGIES OF USING PARALLEL SUBCORPORA OF RNC AND TEXT COLLECTIONS IN TEACHING OF FOREIGN LANGUAGES**

Philological research, especially in the field of literature, is usually considered a "thing-in-itself"; the intrinsic value of this phenomenon involves extremely intuitive, creative, "human-readable" analysis. Meanwhile, modern variety of computer programs (semantic text referentors, tag clouds, concordancers, etc.), created also for the humanities, such as sociology, psychology, management, cannot but draw a philologist's attention. The steps, how to work with a parallel subcorpus in Russian National Corpus, described in detail. Reviewed freeware LR aligner (for non-commercial use), compares translations in Russian the novel "All Red" by J. Chmielewska. As examples of lexical items selected the modal word "avos", the word "nakonets" as an introductory and the circumstances of the word "ves" and "tsely". The Program LF aligner treated three translations of the novel, the authors are M. Krongauz, V. Selivanova, O. Kuznetsova. Consistent description of the existing programs, testing them on art material and comparison of the received data with the existing traditional research, especially in the field of philology and foreign language teaching, is a new step of a text analysis.

Key words: Russian as a foreign language; Polish as a foreign language; Russian National Corpus; information technologies; comparative analysis.

**Сибирцева Вера Григорьевна**, кандидат филологических наук, доцент департамента прикладной лингвистики и межкультурной коммуникации факультета гуманитарных наук.

Национальный исследовательский университет – Высшая школа экономики в Нижнем Новгороде (Россия, Нижний Новгород).

Область научных интересов: преподавание русского языка как иностранного, польский язык, лингвистические вопросы преподавания языков, корпусная лингвистика, информационно-коммуникативные технологии в преподавании языков, немецкоязычная литература.

E-mail: vsibirtseva@hse.ru