

Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2019»

Москва, 29 мая — 1 июня 2019 г.

АННОТИРОВАНИЕ ПАРАЛЛЕЛЬНЫХ ТЕКСТОВ: ПОНЯТИЕ «ДИВЕРГЕНТНЫЙ ПЕРЕВОД»

Инькова О. Ю. (Olga.Inkova@unige.ch)

ИПИ ФИЦ ИУ РАН, Москва, Россия; Женевский
университет, Женева, Швейцария

Ключевые слова: база данных коннекторов, русский язык, корпусная лингвистика, лингвистическое аннотирование, параллельные корпуса, семантика коннекторов, статистические данные

ANNOTATION OF PARALLEL TEXTS: THE CONCEPT OF DIVERGENT TRANSLATION

Inkova Olga Yu. (Olga.Inkova@unige.ch)

Institute of Informatics Problems, FRC CSC RAS, Moscow,
Russia; University of Geneva, Geneva, Switzerland

The annotation of parallel corpora, as well as building of supracorpora databases, challenges linguists with the question of how to define a functional equivalent of the linguistic units that serve as an object of a given study. The paper discusses the concept of divergent translation and whether it is theoretically important for the analysis of logical-semantic relations (LSR). It is shown that relations between states of things can be expressed not only by connectives but also by lexical means (referred to as “alternative lexicalizations” in the works of the Penn Discourse Treebank group) and grammatical tools (syntactic constructions and morphological forms), and by marks of punctuation. While the two latter ways are mentioned in grammars, they are usually not taken into account when the alternative ways of tagging LSR are described, nor are they annotated in corpora or databases. The supracorpora database of connectives, built on the basis of the French and Italian parallel

subcorpora of the Russian National Corpus, introduces new functional capabilities. It stores a representative array of annotations tagged as “divergent translation” (more than 1,250, i. e. 7.7 per cent of the total number), which allows users to collect various statistical data. With these data, one could establish: (1) which LSR tend to be expressed by alternative means and how often they occur compared to connectives, (2) what these alternative means are, (3) which divergent translations may be used to render a given marker of LSR and how often each of them is used, (4) which alternative markers of LSR are specifically employed to convey one or another relation and which of them are able to express several LSR. The conclusive part of the paper suggests that, for the analysis of divergent equivalents, it is central that one and the same alternative means is used by different translators when translating one and the same textual fragment into one and the same language as well as into several languages, which speaks for its productivity. The further development of multi-language and polyvariant parallel corpora and databases would let us find out to what extent the means conveying LSR differ in various languages.

Keywords: corpus linguistics, linguistic annotation, parallel corpora, semantics of connectives, Russian, statistical data, supracorpora database of connectives

1. Введение

Аннотирование параллельных корпусов и создание надкорпусных баз данных требует разработки не только новой информационной технологии (см., в частности [Бунтман и др. 2014]), но и системы терминов [Зализняк и др. 2017]. В данном исследовании проблема разработки терминологии рассмотрена на примере надкорпусной базы данных (НБД) коннекторов, в которую загружены параллельные тексты французского и итальянского подкорпусов Национального корпуса русского языка. Общий объем параллельных текстов составляет примерно 9 млн словоупотреблений. Представительный фрагмент НБД коннекторов доступен по адресу: <http://a179.frccsc.ru/RSCF10004/main.aspx>.

Информация о семантике и функционировании коннекторов хранится в НБД в виде двуязычных аннотаций, *моноэквиваленций* (МЭ), включающих фрагмент текста оригинала с коннектором (столбец 1 на рис. 1) и соответствующий ему фрагмент текста перевода (столбец 3).

Контекст коннектора	Коннектор и его признаки	Контекст эквивалента	Эквивалент и его признаки	Признаки МЭ
Голос превосходный, но ведь голос, в конце концов, можно и так слушать, не вступая в брак, не правда ли... Впрочем, неважно.	впрочем <"вопреки ожиданиям"> <иллокутивные> <повествовательное> <начальная > <p CNT q> <CNT>	Sa voix est splendide, mais enfin, une voix, on peut l'écouter sans l'épouser pour autant, pas vrai?... Du reste , c'est sans importance.	du reste <"вопреки ожиданиям"> <иллокутивные> <повествовательное> <начальная > <p CNT q> <CNT>	<ul style="list-style-type: none"> • Cngrn • Dvrg

Рисунок 1. Пример двуязычной аннотации из НБД коннекторов

В процессе аннотирования (подробнее см. [Зацман и др. 2016]; [Inkova, Popkova 2017]) коннектору (столбец 2 на **рис. 1**) и его функциональному эквиваленту (столбец 4) приписываются значения признаков, характеризующих их функционирование в данном контексте. На заключительном этапе заполняется пятый столбец, характеризующий саму МЭ, а именно тип межъязыковой эквивалентности: метки Sngn и Dvrg предназначены для обозначения типа перевода, который считается конгруэнтным (Sngn), если коннектор переведен коннектором, как на **рис. 1**, и дивергентным (Dvrg), если коннектор переведен лексической единицей, не являющейся коннектором, морфологической формой или синтаксической конструкцией. Ср. (1), где *и* | *при этом* переведен в **а.** деепричастием, а в **б.** — относительным придаточным. В таких случаях проставлена метка Dvrg.

- (1) — Да три полсотни с лишком надо будет приложить, — сказал Петрович *и сжал при этом* значительно губы. [Н. В. Гоголь. Шинель (1842)]
- а.** Trois billets de cinquante roubles au bas mot», dit Pétrovitch *en se pinçant les lèvres*. [Trad. H. Mongault]
- б.** Il faudra y mettre trois billets de cinquante roubles au moins, dit Pétrovitch *qui serra les lèvres avec des sous-entendus*. [Trad. B. Kreise]

По мере наполнения НБД появилась необходимость уточнить понятие «дивергентный» перевод, чему посвящен **раздел 2** настоящего исследования. Представительный массив аннотаций с пометой Dvrg (на 31.01.2019 в НБД сформировано 1263 таких МЭ, т. е. 7,7% от общего количества МЭ) позволяет, в свою очередь, сформулировать некоторые гипотезы относительно способов выражения логико-семантических отношений (ЛСО). Эти гипотезы рассматриваются в **разделе 3**. Там же приводятся некоторые статистические данные, а также иллюстрируются возможности, предоставляемые НБД для контрастного исследования способов выражения ЛСО.

2. Понятие дивергентного перевода

Термин «дивергентный перевод» заимствован нами из работы [Johansson 2007], посвященной использованию многоязычных корпусов в контрастивных исследованиях. Тип межъязыкового соответствия характеризуется по трем параметрам: тип текста (оригинал vs. перевод), эксплицитный или имплицитный характер соответствия (наличие эквивалента vs. нулевой перевод) и синтаксическая природа переводного соответствия (конгруэнтный vs. дивергентный эквивалент). Например, соответствие норвеж. *nok* → англ. *probably* охарактеризовано как «перевод, эксплицитный, конгруэнтный», а норвеж. *nok* → англ. *I suppose* как «перевод, эксплицитный, дивергентный» [Там же: 25]. Однако понятие «дивергентный перевод» определено в работе без учета особенностей семантического или функционального класса, к которому принадлежит переводимая языковая единица (ЯЕ).

В случае коннекторов определение соответствия как конгруэнтного или дивергентного вызывает немалые трудности при аннотировании корпусов

параллельных текстов. Но проблема определения, каким языковым средством выражено то или иное логико-семантическое отношение (ЛСО), возникает и при аннотировании одноязычных корпусов, см., например [Taboada 2009]. Первая трудность связана с тем, что функциональный класс коннекторов является морфологически неоднородным и открытым [Инькова-Манзотти 2001: 20–23]. Помимо сочинительных (*а, но* и др.) и подчинительных (*потому что, если* и др.) союзов, а также их «аналогов» (*однако, тем не менее, например*), к этому классу относятся некоторые предлоги, способные сочетаться с именами с пропозитивным содержанием (*после, вместо* и др.), и предикативные структуры, характеризующиеся высокой степенью устойчивости, например, *дело (закljučается/состоит) в том, что* и др.

Однако хорошо известно, что отношения между ситуациями могут выражаться ЯЕ, не принадлежащими к классу коннекторов. В системе аннотирования, разработанной в рамках Penn Discourse Treebank (PDTB), некоторые из них квалифицируются как «альтернативные лексикализации» [Prasad et al. 2010]. В PDTB зафиксировано 624 таких случая. Этот класс включает i) лексически и синтаксически свободные конструкции (например, *The increase was due mainly to..., These measurements can indicate temperature changes, That is why*, выражающие причинные или следственные отношения¹); ii) конструкции, синтаксически свободные, но с фиксированным лексическим составом (*What's more, To begin with*), iii) некоторые устойчивые наречные выражения, синтаксический статус которых позволяет их отнести к классу коннекторов, но эта функция не является у них основной, а возникает в определенном контексте (*for one thing, as well, too, even, especially*). В НБД коннекторов в тех случаях, когда коннектор переведен одной из перечисленных ЯЕ (2) или, наоборот, такая ЯЕ переводится коннектором (3), межъязыковое соответствие квалифицируется как дивергентное.

- (2) Вот что было-с. *Да еще* слухи о земельной реформе, которую намеревался произвести пан гетман. [Михаил Булгаков. Белая гвардия (1924)]
 Voilà ce qui existait. *Il fallait ajouter à cela* les bruits qui couraient, selon lesquels le seigneur hetman avait l'intention de procéder à une réforme agraire.
 [Trad. C. Ligny]

¹ В [Toldova et al. 2018] речь уже идет не об альтернативных способах выражения причинных отношений, а о коннекторах; ср. приводимые на с. 6 в качестве «discourse connectives for causal relations» глаголы *порождать, позволять, объяснить, давать, изменять, вызывать, приводить* и) и пример: [Неудачно остановившаяся машина стала помехой для быстрых кругов многих гонщиков, включая Фернандо Алонсо.] [и это вызвало расследование FIA.] Такая позиция представляется теоретически уязвимой. Эти глаголы, безусловно выражая семантику причинности, участвуют в описании ситуации, входя в пропозициональное содержание высказывания. Включение в состав «коннектора» демонстратива *это*, осуществляющего резюмитивную анафору, не спасает дела. Если следовать логике исследователей, то коннектором следует признать и предикативные структуры с подлежащим, осуществляющим резюмитивную анафору, типа *эта ситуация, это положение дел* и др. Ср. модификацию приведенного примера: [Неудачно остановившаяся машина стала помехой для быстрых кругов многих гонщиков, включая Фернандо Алонсо,] [и эта ситуация вызвала расследование FIA.] Если говорить о коннекторе в исходном примере, то им является только союз *и* (ср. Миша много гуляет, и это очень хорошо).

- (3) Да приложи к письму какой-нибудь петербургский гостинец... сигар, *что ли*. [И. А. Гончаров. Обломов (1848–1859)]
Ajoute à ta lettre quelque cadeau pétersbourgeois... Des cigares, *par exemple*.
[Trad. L. Jurgenson]

В (2) семантика *да еще*, выражающего аддитивные ЛСО, передается безличной конструкцией франц. *Il fallait ajouter à cela* ‘нужно добавить к этому’. В (3) частица *что ли* употреблена в контексте ЛСО спецификации: первый фрагмент текста содержит наименование множества (*какой-нибудь петербургский гостинец*), а во фрагменте текста со *что ли*, — элемент этого множества (*сигары*). Однако вряд ли можно считать *что ли* выразителем этого ЛСО. Эта частица сохраняет свое модальное значение предположения и может употребляться и в других контекстах, далеких от семантики спецификации; ср. (4), где можно говорить скорее об ЛСО переформулирования, поиска более адекватного наименования, между *спокойный* и *традиционный*.

- (4) И хочется чего-то более спокойного, традиционного, *что ли*. [Вероника Стрельникова. Опять акробатика, милый? // «Даша», 2004]

Словари справедливо не указывают для *что ли* значения показателя спецификации. Во французском переводе, напротив, использован специализированный показатель этого ЛСО — *par exemple*.

Помимо ЯЕ, относящихся к «альтернативным лексикализациям», в НБД зафиксированы и другие функциональные эквиваленты коннекторов. Их можно разделить на две группы: грамматические (синтаксические конструкции, морфологические формы), примеры которых мы видели в (1), и лексические. Так, отношение результативной аналогии, при которой длящегося в момент наблюдения действие или состояние субъекта, описанное в главной части, непосредственно продолжает то же действие или состояние субъекта, описанное в придаточной части, или является его результатом, часто выражается лексемами с семантикой неизменности [Кобозева, Инькова 2018: 206–210]: в (5)–итал. *rimanere immobile* ‘остаться неподвижным’ и франц. *s’immobiliser* ‘застыть в неподвижности’:

- (5) Катерина Ивановна как стояла на месте, так и осталась, точно громом пораженная. [Ф. М. Достоевский. Преступление и наказание (1866)]
а. Katerina Ivanovna rimase lì immobile, come colpita da un fulmine.
[Trad. G. Kraiski]
б. Catherine Ivanovna s’immobilisa sur place comme frappée par la foudre.
[Trad. E. Guertik]

Наконец, семантика коннектора может быть выражена семантически насыщенным знаком препинания: двоеточием или тире, причем как в оригинале (6), так и переводе (7).

- (6) Конечно, от оперуполномоченного многое зависит, и таинственны, туманны дорожки к высотам жизни — зав. баней, хлеборез.
[Василий Гроссман. Жизнь и судьба (1959)]

Bien sûr, bien des choses dépendaient de l'oper, bien sûr, les voies qui mènent vers les sommets de la vie, *par exemple* être responsable du bain ou des rations de pain, sont sombres et mystérieuses. [Trad. A. Berelowitch]

- (7) Подлинная многопланность разрушила бы драму, *ибо* драматическое действие, опирающееся на единство мира, не могло бы уже связать и разрешить ее. [М. М. Бахтин. Проблемы поэтики Достоевского (1963)]
Une véritable multiplicité de plans serait préjudiciable à la pièce: l'action dramatique qui s'appuie normalement sur l'unité de l'univers représenté serait alors incapable de servir de lien et d'apporter des solutions. [Trad. I. Kolutcheff]

Эти знаки препинания, особенно двоеточие, совместимы с ограниченным числом ЛСО (как правило, разного рода причинными ЛСО и ЛСО спецификации), поэтому, в отличие от других знаков препинания (точка, точка с запятой и запятая), они могут служить сигналом того, что между двумя фрагментами текста существует некоторое ЛСО.

Во всех перечисленных выше случаях в НБД для МЭ проставляется признак Dvrg.

В процессе аннотирования были выявлены, однако, такие случаи перевода, когда двух меток — Sngn и Dvrg — оказалось недостаточно для характеристики переводного соответствия, поскольку речевая реализация (PP), т. е. та форма, в которой коннектор встретился в конкретном контексте и которая является единицей аннотирования в НБД [Инькова 2018а], может быть переведена сочетанием коннектора и ЯЕ, не являющейся коннектором. Это могут быть как PP, выражающие одно ЛСО, так PP, представляющие собой сочетание коннекторов, выражающих разные ЛСО:

- (8) Звали его полностью Лазарь Рувимович Цехновицер, он был худой, длинноносый, курчавый, *а также* учился играть на скрипке.
[Сергей Довлатов. Иностранка (1986)]
Son nom exact était Lazare Rouvimovitch Tsekhnovitser, il était maigre, avait le cheveu frisé, le nez long *et* apprenait *même* le violon.
[Trad. J. Michaut-Paternò]
- (9) Соня дала свой адрес *и при этом* покраснела.
[Ф. М. Достоевский. Преступление и наказание (1866)]
Sonia donna son adresse *et* rougit *en le faisant*. [Trad. É. Guertik]

В (8) показатель аддитивных ЛСО *а также* переведен сочетанием союза *и* и градационной частицы *même* 'даже'; в (9) PP *и при этом*, сочетающая показатель соединительных ЛСО и ЛСО сопутствования, переведена союзом *et* и деепричастием глагола *faire* 'делать' с анафорическим местоимением *le* 'это', передающими одновременность двух действий.

Для таких случаев в новой версии НБД введена метка Sngn + Dvrg. На 01.02.2019 эта метка проставлена для 89 МЭ (учитываются все направления перевода).

3. О чем нам говорит дивергентный перевод?

В работах, выполненных в рамках проекта РФФ «Логическая структура текста: контрастивный анализ показателей логико-семантических отношений в русском, французском и итальянском языках», наличие большого количества дивергентных эквивалентов или стимулов перевода используется как один из параметров лингвоспецифичности коннектора [Инькова 2018б]. Однако функциональные возможности НБД позволяют получить и другую статистику по признаку Dvrg. Эта статистика позволяет, например, увидеть, для каких ЛСО дивергентный перевод показателя является распространенным явлением, а для каких — скорее редкость².

Таблица 1. Доля дивергентных переводов для ЛСО (выборочно) в направлении русский-французский

ЛСО	Dvrg	Всего	%
сопутствование	62	348	17,8 %
спецификация	86	726	11,8 %
сравнительные	29	288	10,1 %
аналогия	10	102	9,8 %
переформулирование	81	864	9,4 %
сопоставительные	24	282	8,5 %
причина	32	398	8 %
условные	39	580	6,7 %
аддитивные	43	683	6,3 %
уступительные	43	730	5,9 %
альтернатива	9	420	2,1 %
отрицательная альтернатива	3	170	1,7 %
соединительные	1	263	0,4 %
«вопреки ожиданиям»	40	1346	0,3 %

В НБД самая высокая доля дивергентных переводов — у показателей отношения сопутствования: почти 18%, а наименьшая (приближающаяся к нулю) — у показателей соединительных ЛСО и ЛСО «вопреки ожиданиям». У ЛСО сравнения, аналогии, причины и переформулирования доля дивергентных переводов составляет около 10%. Чуть больше их у ЛСО спецификации (почти 12%).

Важно также узнать, какой из показателей того или иного ЛСО отличается наибольшей долей дивергентных переводов и каких именно. В Таблице 2 приведены данные для ЛСО переформулирования и причины в направлении перевода русский — французский.

² Подчеркнем, что статистику, приводимую в работе, можно получить непосредственно в НБД, задав соответствующий запрос. Кроме того, абсолютные цифры в таблицах являются в НБД гиперссылками, отсылающими к аннотациям, отвечающим критериям запроса, что позволяет пользователю сразу получить доступ к корпусу примеров и проводить их дальнейший семантический анализ.

Таблица 2. Распределение дивергентных переводов по показателям ЛСО в НБД

Переформулирование	Всего МЭ	Dvrg, %
<i>то есть</i>	633	11,22
<i>одним словом</i>	68	5,88
<i>словом</i>	84	0

Причина	Всего МЭ	Dvrg, %
<i>потому что</i>	180	12,22
<i>ибо</i>	50	10
<i>так как</i> <i>то</i>	31	3,23

Мы видим, что, если у обоих ЛСО доля дивергентных переводов сопоставима (9,4% для переформулирования и 8% для причины), распределение этих переводов по показателям каждого из ЛСО существенно различается. В группе показателей переформулирования лидирует *то есть* (11,22%), тогда как у *словом* пока не зафиксировано ни одного дивергентного перевода. В группе причинных ЛСО лидирует *потому что* с 12,22%, чуть меньше таких переводов у *ибо* (10%).

НБД позволяет также узнать, каковы эти дивергентные переводы. Например, для *потому что* наиболее частотным из 7 зафиксированных в НБД дивергентных переводных эквивалентов является двоеточие (40,1%), как в (10), где два из четырех переводчиков прибегают к такому решению; за ним следуют различные причастные конструкции в прошедшем времени, со вспомогательным глаголом (18,2%) и без него (4,6%), и в настоящем времени (18,2%), а затем еще один знак препинания: тире (9,1%), как в (11), где оба переводчика выбрали это решение.

- (10) <...> коллежский асессор Ковалев не мог слышать запаха, *потому что* закрылся платком *и потому что* самый нос его находился бог знает в каких местах. [Н. В. Гоголь. Нос (1832–1833)]
 <...> l'assesseur de collège Kovaliov ne pouvait pas s'en rendre compte: il avait caché son visage sous un mouchoir, et d'ailleurs son nez se trouvait en cet instant Dieu sait où. [Trad. B. de Schloezer]
 <...> le major Kovaliov ne s'en trouvait point incommodé: il tenait son mouchoir sur son visage, et d'ailleurs son nez se promenait... Dieu sait où. [Trad. H. Mongault]
- (11) Он по необходимости сидел в классе прямо, слушал, что говорили учителя, *потому что* другого ничего делать было нельзя. [И. А. Гончаров. Обломов (1848–1859)]
 Il lui fallut se bien tenir en classe, écouter ce que disaient les professeurs — il y était bien obligé. [Trad. A. Adamov]
 Se pliant à la nécessité, il se tenait droit en classe, écoutait ce que disaient les professeurs — on ne pouvait rien faire d'autre [Trad. L. Jurgenson]

Два других дивергентных эквивалента: конструкция *Adj+que+fin_être* (12) и выделительная конструкция *c'est N qui/que* зарегистрированы с единичными употреблениями.

- (12) Ничто не нарушало однообразия этой жизни, и сами обломовцы не тяготились ею, *потому что* и не представляли себе другого житья-бытья. [И. А. Гончаров. Обломов (1848–1859)]
 Rien ne troublait la monotonie de cette vie, qui ne pesait point aux habitants d'Oblomovka, *incapables qu'ils étaient même de s'imaginer un train-train différent*. [Trad. L. Jurgenson]

У *ибо* зафиксировано пока три дивергентных эквивалента: двоеточие (7) и причастие прошедшего времени без вспомогательного глагола (по 40%), которые мы уже видели для *потому что*, и относительное придаточное (20%). Заметим, что его выбирают не только французские (13а), но и 3 из 15 итальянских переводчиков этого контекста; один из таких переводов приведен в (13б).

- (13) <...> вскрикнул, может быть, в первый раз от роду, *ибо* отличался всегда тихостью голоса. [Н. В. Гоголь. Шинель (1842)]
 а. <...> s'exclama, pour la première fois de sa vie, sans doute, le malheureux Akaki Akakiévitch, *qui d'ordinaire parlait à voix très basse*. [Trad. H. Mongault]
 б. <...> e fu forse questo il primo urlo che avesse emesso dacchè era nato, *lui che parlava sempre sottovoce*. [Trad. Duchessa D'Andria]

Иная семантика дивергентных переводов у *то есть*. Хотя здесь тоже встречаются двоеточие и тире, но на их долю приходится в совокупности всего 5,6%, а самые частотные дивергентные переводы — глагольное выражение *vouloir dire* 'хотеть сказать' (39,5%), как в (14), и глагол *entendre* 'понимать что-то под чем-л.' (15,5%). Используются также глаголы *signifier* 'означать', *dire* 'сказать', *s'expliquer* 'объясниться', *appeler* 'называть'.

- (14) Между нами уже было не так, как раньше, *то есть* мы дружили, но все больше я на нее покрикивала [Светлана Алексиевич. Время секунд хэнд (2013)]
 Ce n'était déjà plus comme avant entre nous, *je veux dire*, on s'entendait bien, mais je lui criais de plus en plus souvent dessus. [Trad. S. Benech]

Заметим, что если некоторые переводы, и не только дивергентные, являются контекстуально обусловленными, что часто приводит к значительной перестройке контекста оригинала (см. на примере *хотя* [Нуриев 2018]), и зафиксированы с единичными вхождениями, то другие могут рассматриваться как регулярные модели перевода. Для сопоставительного *a* — это местоименный повтор (15) [Зализняк, Микаэля 2018], для *при этом* — деепричастие настоящего времени (1)³.

³ Подробный семантический анализ дивергентных переводных эквивалентов не входит в задачу настоящего исследования, да и невозможен в силу ограниченности его объема. Примеры такого анализа см. в посвященной этому вопросу монографии *Семантика коннекторов: контрастивное исследование* под ред. О. Ю. Иньковой.

(15) Обломов задумался, а Алексеев барабанил пальцами по столу.

[И. А. Гончаров. Обломов (1848–1859)]

Oblovov devint pensif, *Alexéev, lui*, pianotait sur la table. [Trad. L. Jurgenson]

Кроме того, как альтернативный способ выражения ЛСО могут использоваться ЯЕ, имеющиеся также в языке оригинала, как в приведенных выше примерах, или же не существующие в системе языка оригинала, как в (16), где для перевода условных отношений, выраженных в оригинале коннектором, используется отсутствующая в русском языке синтаксическая конструкция с двумя предикатами в ирреалисе (и с обязательной инверсией в первом), соединенными асемантическим союзом *que* 'что'.

(16) <...> и не представляли себе другого житья-бытья; а если б и смогли представить, то с ужасом отвернулись бы от него [И. А. Гончаров. Обломов (1848–1859)]

<...> incapables qu'ils étaient même de s'imaginer un train-train différent; en auraient-ils été capables qu'ils l'auraient condamné avec horreur.

[Trad. L. Jurgenson]

Наконец, мы можем узнать, является ли дивергентный перевод «специализированным» средством выражения некоторого ЛСО или может обслуживать несколько ЛСО. К первым можно отнести безличный глагол франц. *il s'ensuit* 'из этого следует', который может служить эквивалентом только для коннекторов ЛСО следствия. Ко вторым — относительное придаточное, использующееся при переводе коннекторов сопоставления (20%), спецификации (17,14%), аддитивности (14,29%), переформулирования (11,43%), сопутствования (11,43%), причины (8,58%), уступительных (5,72%) и др.

4. Заключительные замечания

Теоретический интерес понятия «дивергентный перевод» для описания средств выражения ЛСО не вызывает, на наш взгляд, сомнения. Основное преимущество параллельного корпуса состоит в том, что он служит средством выявления различий в системах языков, особенно в тех случаях, когда они не фиксируются словарями и грамматиками. Использование НБД коннекторов, созданной на основе параллельных текстов предоставляет исследователям принципиально новые возможности и позволяет установить: 1) для каких ЛСО характерны альтернативные способы выражения и какова их частотность по сравнению с коннекторами; 2) каковы эти альтернативные способы выражения; 3) какие дивергентные переводы может иметь тот или иной показатель ЛСО и какова частотность каждого из них; 4) какие альтернативные средства выражения ЛСО являются специфическими для того или иного отношения, а какие могут обслуживать несколько ЛСО.

Важным при таком анализе представляется тот факт, что одни и те же модели дивергентного перевода зафиксированы в НБД для показателей одного и того же ЛСО в разных языках, в разных направлениях перевода и у разных

переводчиков одного и того же контекста оригинала. Это позволяет сделать обобщения относительно регулярности того или иного альтернативного способа выражения ЛСО и установить, каков набор этих способов в каждом из сопоставляемых языков, а также чем этот набор отличается в разных языках.

Литература

1. *Buntman N. V., Zalizniak Anna A., Zatsman I. M., Krushkov M. G., Loshchilova E. Yu., Sitchinava D. V.* (2014) *Informatsionnye tekhnologii korpusnykh issledovaniy: printsipy postroeniya krosslingvisticheskikh baz dannkh* [Information technologies for corpus studies: underpinnings for cross-linguistic database creation]. *Informatics and applications*. 2014. Vol. 8. No. 2. Pp. 98–110.
2. *Inkova-Manzotti O. Yu.* (2001) *Konnektory protivopostavleniya vo frantsuzskom i russkom yazykakh: Sopostavitel'noe issledovanie* [Connectors of opposition in French and Russian: A comparative study]. Moscow: Informelektro. 434 p.
3. *Inkova O., Popkova N.* (2017). Statistical data as information source for linguistic analysis of Russian connectors. *Informatics and applications*. 2017. Vol. 11, No. 3. Pp. 123–131.
4. *Johansson S.* (2007) *Seeing through multilingual corpora: On the use of corpora in contrastive studies*. Amsterdam: John Benjamins. 377 p.
5. *Inkova O. Yu.* (2018a) *Nadkorpusnaja baza dannykh kak instrument izutcheniya formal'noj variativnosti konnektorov* [Supracorpora database as an instrument of the study of the formal variability of connectives]. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2018"*. Moscow, May 30–June 2, 2018. <http://www.dialog-21.ru/media/4299/inkovaoyu.pdf>.
6. *Inkova O. Yu.* (2018b) *Lingvospetsifichnost' konnektorov: metody i parametry opisaniya* [The language-specificity of connectives: methods and parameters of description]. *Semantika konnektorov: kontrastivnoe issledovanie* [Semantics of connectives: a contrastive study], O. Inkova (ed.). Moscow: TORUS PRESS. Pp. 5–23.
7. *Inkova O. Yu. ed* (2018в) *Semantika konnektorov: kontrastivnoe issledovanie* [Semantics of connectives: a contrastive study]. Moscow: TORUS PRESS.
8. *Kobozeva I. M., Inkova O. Yu.* (2018) *Kak i ego dvukhmestnye varianty* [Kak and its bi-places variantes]. *Semantika konnektorov: kontrastivnoe issledovanie* [Semantics of connectives: a contrastive study], O. Inkova (ed.). Moscow: TORUS PRESS. Pp. 168–239.
9. *Nuriev V. A.* (2018) *Khotya. Semantika konnektorov: kontrastivnoe issledovanie* [Semantics of connectives: a contrastive study], O. Inkova (ed.). Moscow: TORUS PRESS. Pp. 269–300.
10. *Prasad R., Joshi A., Webber B.* (2010) *Realization of Discourse Relations by Other Means: Alternative Lexicalizations*. *Proceedings of the 23rd International Conference on Computational Linguistics (Beijing, China — August 23–27, 2010): Posters*. Pp. 1023–1031.

11. *Taboada M.* (2009) Implicit and explicit coherence relations. *Discourse, of Course. An overview of research in discourse studies*, J. Renkema (ed.). Amsterdam/Philadelphia: John Benjamins. Pp. 127–140.
12. *Toldova S., Pisarevskaya D., Kobozeva M., Vasilyeva M.* (2018) The cues for rhetorical relations in Russian: “cause–effect” relation in Russian rhetorical structure treebank. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2018”*. Moscow, May 30–June 2, 2018. <http://www.dialog-21.ru/media/4338/toldovas.pdf>.
13. *Zalizniak Anna A., Mikaelyan I. L.* (2018) Soyuz a [La conjunction a]. *Semantika konnektorov: kontrastivnoe issledovanie [Semantics of connectives: a contrastive study]*, O. Inkova (ed.). Moscow: TORUS PRESS. Pp. 24–79.
14. *Zalizniak Anna A., Zatsman I. M., Inkova O. Yu.* (2017) Nadkorpurnaja baza dannykh konnektorov: postroenie systemy terminov [Supracorpora database on connectives: term system development]. *Informatics and applications*. 2017. Vol. 11, No. 1. Pp. 101–109.
15. *Zatsman I. M., Inkova O. Yu., Kruzhekov M. G., Popkova N. A.* (2016) Predstavlenie krosslingvisticheskikh znanij o konnektorakh v nadkorpurnykh bazakh dannykh [Representation of cross-lingual knowledge about connectors in supracorpora databases]. *Informatics and applications*. 2016. V. 10. No. 1. Pp. 106–118.