

Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2017»

Москва, 31 мая — 3 июня 2017

ПРИНЦИПЫ ОПРЕДЕЛЕНИЯ СТЕПЕНИ ЛИНГВОСПЕЦИФИЧНОСТИ КОННЕКТОРОВ¹

Инькова О. Ю. (Olga.Inkova@unige.ch)

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

PRINCIPLES OF HOW TO DETERMINE THE DEGREE OF LANGUAGE-SPECIFICITY OF CONNECTIVES

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

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

The paper deals with the question of how to determine the degree of language-specificity of connectives. As connectives are non-descriptive linguistic units and their morphological nature may vary, the author suggests to apply functional semantic criteria and distinguishes five types of interlingual isomorphism violation in Russian/French and Russian/Italian.

The author proposes to establish five categories of language-specific connectives:

1. The relation R expressed by the connective K_A of language A is conveyed in language B by a linguistic unit U_B not belonging to the functional class of connectives (for instance, the relation of simultaneousness expressed

¹ Исследование выполнено в ИПИ ФИЦ ИУ РАН в рамках проекта «Логическая структура текста: контрастивный анализ показателей логико-семантических отношений в русском, французском и итальянском языках», поддержанного РФФ (грант № 16-18-10004).

by *pri ètom* in Russian is frequently expressed in French or Italian by means of a present gerund);

2. The relation R expressed by the connective K_A of language A is conveyed in language B by the signal K_B of a relation R' which is semantically contiguous, but nevertheless possesses a more specific equivalent in language A (for instance, the *pri tom* of illocutive simultaneousness is rendered in French with *en outre*, *en plus*, *de plus* disposing in Russian of the more precise equivalents *k tomu že* and *krome togo*);
3. The connective K_A of language A possesses in language B a set of equivalents K_B^i , all occurring with the same frequency (it's the case of Russian *pročëm*, which is mainly translated with *d'ailleurs* and *du reste*);
4. The connective K_A of language A has in language B a systemic equivalent K_B which is the most frequent in translation, but doesn't reproduce all the values and uses of K_A (for instance, the connective *a to* of negative alternative has the French equivalent *sinon*, but the Russian connective is polysemous: it can express both the relations of cause and alternative, as well as a relation of addition — what is not the case of *sinon*);
5. There is in language B a mostly used translation variant K_B of the connective K_A of language A, but this 'equivalent' K_B doesn't possess the combinatory variety of K_A (thus the Russian connective *hotya* introducing a subordinate clause combines in the main sentence with the conjunction *no*, but *hotya ... no* is translated in French either by a single *bien que*, the equivalent of *hotya*, or by a single *mais*, the equivalent of *no*).

The author also describes features helping to determine the degree of language-specificity of connectives that would take into account the translation direction (from Russian to French/Italian and vice versa) and might be subject to a quantitative analysis. These features are: 1) frequency of occurrence of a "zero" functionally equivalent fragment (respectively, translation stimulus); 2) type of functionally equivalent fragment or translation stimulus (congruent or divergent); 3) number of possible translation patterns (respectively, translation stimuli). In conclusion, the author will comment on statistical data confirming the described types of interlingual isomorphism violation in case of the functional class of connectives. The author points out that statistics requires adequate interpretation and should be accompanied by a semantic analysis, as the latter also serves as a tool for the data's review. The study was carried out by using the Russian-French and Russian-Italian supracorpora database of connectives.

Key words: database of cross-linguistic equivalences, Russian, French, Italian, contrastive linguistics, corpus linguistics, language-specific connectives, statistical data

1. Введение

Исследования лингвоспецифичных языковых явлений ведутся в отечественной лингвистике начиная с 1990-х годов (см., например, [Зализняк, Левонтина 1996], [Зализняк, Левонтина, Шмелев 2012], [Зализняк 2015], [Шмелев 2015]). С одной стороны, они исходят из гипотезы Сепира-Уорфа, а с другой, опираются на работы А. Вежбицкой, которой принадлежит и сам термин

«лингвоспецифичный» (“language-specific (word-)meanings” [Wierzbicka 1992, 1996]). В [Зализняк 2015: 687] лингвоспецифичной называется «лексическ[ая] единиц[а], заключающ[ая] в себе **уникальную концептуальную конфигурацию**, а именно такую, что во всех существующих ее переводных эквивалентах не хватает какого-то семантического компонента или неустранимым образом присутствует ‘лишний’ компонент». Поэтому лингвоспецифичные единицы являются труднопереводимыми (или вообще непереводаемыми) на другие языки, а сам признак «труднопереводимости» является важнейшим отличительным свойством лингвоспецифичных слов.

Если до последнего времени исследования лингвоспецифичности проводилось на основе сопоставительного семантического анализа языковых единиц, то с развитием корпусной лингвистики, особенно, с созданием параллельных корпусов, открываются новые возможности исследований межъязыковых соответствий, прежде всего, с привлечением статистических данных. Еще один шаг на этом пути был сделан благодаря созданию т.н. «надкорпусных баз данных» (НБД) ([Кружков 2015]). Они представляют собой своего рода надстройку над параллельным корпусом, позволяющей фиксировать реально встретившиеся соответствия языковых единиц и аннотировать их по заданному набору признаков. Отмечу два основных преимущества НБД перед параллельными корпусами: 1) для каждой анализируемой языковой единицы в тексте перевода уже выделен ее функционально-эквивалентный фрагмент (ФЭФ) (термин введен в работе [Добровольский и др. 2005]); 2) анализируемой единице и ее ФЭФ приписывается значение формально охарактеризованных признаков, релевантных для их анализа.

В результате работы с русско-французской НБД лингвоспецифичных единиц в работе [Зализняк 2016: 860] были выделены следующие пять типов отсутствия межъязыкового семантического изоморфизма:

1. Асимметричное членение концептуальной области (например, рус. *рука* — фр. *main* и *bras*).
2. Тип САМОВАР (переводной эквивалент отсутствует; используется заимствованное слово или описательное определение).
3. Тип БАБУШКА (имеется один преимущественный вариант перевода, но он неточный).
4. Тип РОДНОЙ (переводной эквивалент отсутствует; имеется несколько приблизительно равновероятных вариантов перевода, все неточные).
5. Тип РАЗЛУКА (имеется один преимущественный переводной эквивалент, имеющий более точное соответствие в русском языке: *разлука* — *séparation* — *расставание*).

Автором предлагается также ряд характерных признаков, указывающих на вероятную лингвоспецифичность языковой единицы. Эти признаки могут быть программно генерированы в НБД с учетом направления перевода и поддаются количественному анализу.

При переводе с русского языка (Р—Ф):

1. наличие большого количества ФЭФ, в том числе имеющих приблизительно равную частотность;

2. наличие многокомпонентных ФЭФ;
3. слово остается без перевода (нулевой ФЭФ).

При переводе на русский язык (Ф—Р):

4. большое количество стимулов перевода (СП), т.е. тех языковые единицы оригинала, 'реакцией' на употребление которых является употребление русской языковой единицы;
5. наличие многокомпонентных СП;
6. отсутствие какой-либо единицы, которая 'стимулирует' появление анализируемой русской языковой единицы (нулевой СП).

В данной статье предлагается проверить, насколько данные пять категорий и перечисленные признаки лингвоспецифичности, разработанные, прежде всего, для однозначных слов, применимы к анализу функционального класса коннекторов. Такого рода исследование представляет интерес для лексической типологии, лексикографии, теории и практики перевода, а также других областей, связанных с сопоставительным изучением языков. Предлагаемый анализ проводится с использованием русско-французской и русско-итальянской НБД коннекторов общим объемом более 3,5 млн словоупотреблений. Ее концепция, принципы аннотирования и отличия от других существующих НБД изложены в работах [Инькова, Кружков 2016; Зацман и др. 2016].

2. Параметры квантитативного анализа степени лингвоспецифичности коннекторов

Хорошо известно, что коннекторы, как и многие дискурсивные слова, принадлежат к лингвоспецифичным единицам. Даже такие, казалось бы, семантически пустые коннекторы как англ. *and*, нем. *und*, фр. *et*, рус. *и*, имеют в каждом языке особенности функционирования, которые не позволяют поставить между ними знак равенства. Достаточно привести хорошо известный пример расхождения в употреблении англ. *and*, нем. *und*, фр. *et*, с одной стороны, и русского *и*, с другой, в такой фразе: нем. *Ihre Mutter ist Lehrerin und ihr Vater ist Anwalt*, англ. *Her mother is a teacher and her father is a lawyer*; фр. *Sa mère est enseignante et son père est avocat*; рус. *Ее мать — учительница, а^{/?}и отец — адвокат*. Этот простой пример дает также возможность подчеркнуть три важных положения, которые должны учитываться при анализе лингвоспецифичности. Первое: само свойство лингвоспецифичности является сопоставительным, проявляющимся при анализе межъязыковых соответствий двух или более конкретных языков (это обстоятельство подчеркивалось, в частности, в работе [Зализняк 2015]). Иначе говоря, то, что является лингвоспецифичным для одной пары языков, может не быть таковым в другой. Второе: языковые единицы различаются степенью лингвоспецифичности: она очевидным образом будет выше у рус. *а*, не имеющего системного эквивалента во французском, английском, немецком и многих других языках, и минимальной у *или*, имеющего системные эквиваленты в большинстве европейских языков. Наконец, «труднопереводимость» не является детерминирующим признаком лингвоспецифичности. Как было показано в [Dobrovolskij,

Röppel 2016] на примере *дело в том*, некоторые дискурсивные единицы, не представляя никакого труда для перевода, являются тем не менее лингвоспецифичными, т.е. обладают такой конфигурацией семантических черт, которые не позволяют найти им точного эквивалента в других языках.

Что касается пяти выявленных в [Зализняк 2016] типов отсутствия семантического межъязыкового изоморфизма, разработанного в первую очередь для однозначных слов, то их невозможно механически перенести на функциональный класс коннекторов в силу их семантических и функциональных особенностей. Коннекторы относятся к классу недескриптивных языковых единиц. Их функция — сигнализировать, что между двумя связываемыми ими фрагментами текста существует логико-семантическое отношение (ср. категорию «сигналитивы» в [Иорданская, Мельчук 2007: 347–358]). Говорящий может, однако, никак не выразить это отношение (например, причинное отношение в *Петя не пришел. Он заболел*) или, наоборот, выразить его неоднократно. Так, в высказывании:

Выяснилось, что в бюро иностранцев ни о каком Воланде, а равно также и Фаланде, маге, ровно ничего не слышали. [М. А. Булгаков. Мастер и Маргарита (ч. 1) (1929–1940)]

употреблен коннектор *а равно также и*, хотя можно было бы ограничиться более экономным *а также*. Нарушение межъязыкового изоморфизма может быть связано как раз с тем, что словарный эквивалент языковой единицы в другом языке не может входить в состав таких образований, что прежде всего связано с разнообразной морфологической природой коннекторов. Как известно, в функциональный класс коннекторов входят не только союзы, сочинительные и подчинительные, но и т.н. аналоги союзов наречного типа, а также предикативные конструкции (*из этого следует, дело в том, что и др.*), которые имеют другие синтаксические и сочетаемостные свойства. На этом основании в случае коннекторов мне представляется более адекватным говорить о нарушении функционально-семантического изоморфизма.

С другой стороны, в отличие от лингвоспецифичной лексики, включающей, как правило, однозначные лексемы, многие коннекторы многозначны, в том смысле, что в зависимости от смыслового соотношения соединяемых ими фрагментов текста они могут выражать разные отношения или устанавливать одно отношение, но на разных семантических уровнях. И наоборот, одно отношение может иметь целую парадигму показателей, различия между которыми часто заключаются не столько в семантике, сколько в тех синтаксических условиях, в которых они могут употребляться.

Таким образом, становится понятным, что проблема лингвоспецифичности коннекторов должна решаться на основе функционально-семантических критериев, а предложенные типы нарушения межъязыкового изоморфизма должны быть несколько переформулированы, а некоторые из них не находят себе места при анализе коннекторов. Так, тип 1. «Ассиметричное членение концептуальной области» на первый взгляд присутствует в такой тройке: рус. *скорее... чем и лучше... чем* и ит. *piuttosto che*.

Piuttosto morire, che bere quella medicina cattiva.	Лучше умереть, чем глотать такое ужасное лекарство!
— У меня скорее лапы отсохнут, чем я прикоснусь к чужому, — напыжившись воскликнул кот.	— Mi si paralizzino le zampe piuttosto che toccare la roba altrui, — esclamò il gatto, con sussiego.

Однако в итальянском языке существует также компаратив *meglio* 'лучше', который просто не обладает способностью русского *лучше* выступать в качестве «компонента двухместного союза» (РГ-80, II: 490) для выражения замещения по предпочтению. Поэтому речь здесь идет не столько о различном концептуальном членении, сколько о функциональных возможностях той или иной языковой единицы. В парах русский — французский и русский — итальянский тип 1. пока не был обнаружен.

Тип 2. САМОВАР в том виде, в котором он сформулирован для полнозначной лексики, также не находит себе места при анализе коннекторов. Использование заимствованного коннектора представляется мне мало вероятным явлением. По крайней мере, в паре языков русский — французский таких случаев, насколько мне известно, нет. Употребление в итальянском языке латинского *ergo* как показателя следственных отношений воспринимается как иноязычное вкрапление, имеющее явную книжную стилистическую окраску, поскольку в итальянском языке существуют стилистически нейтральные коннекторы, выражающие это отношение (*dunque, quindi*). Использование описательного определения для выражения логико-семантических отношений также не характерно. Типу 2. будет примерно соответствовать в типологии лингвоспецифичности коннекторов тип 1. Вообще говоря, установить какое-либо прямое соответствие между типами, предлагаемыми Анной А. Зализняк для лингвоспецифичной лексики, и типами, предлагаемыми нами в данной работе для анализа коннекторов, не представляется возможным в силу указанных выше причин.

Предлагается выделять следующие типы отсутствия функционально-семантического межъязыкового изоморфизма для класса коннекторов:

Тип 1. Отношение R, выражаемое коннектором K_A в языке A, выражается в языке B языковой единицей U_B , не принадлежащей к функциональному классу коннекторов. Это может быть синтаксическая конструкция, морфологическая форма, лексическая единица.

Так, отношение сопутствования (определение данного отношения см. в [Инькова в печати]), которое на пропозициональном уровне сигнализирует, что одно действие, представленное как главное, сопровождается другим действием, представленным как участвующее в его реализации, выражается в русском языке коннектором *при этом*, а во французском языке может при общем субъекте обоих действий выражаться формой деепричастия настоящего времени предиката, соответствующего сопровождающему действию.

Маруся дарила сестре платья и кофты. При этом говорилось: — Синюю кофту бери, а зеленую я еще поношу...	Maroussia donnait à sa cousine ses robes et ses blouses en ayant soin d'ajouter: «Prends la blouse bleue, la verte je la porterai encore un peu... »
---	--

Тип 2. Отношение R, выражаемое коннектором K_A в языке A, выражается в языке B показателями K_B близкого по семантике отношения R', имеющего, однако, более точное соответствие в языке A. Например, коннектор *притом* не имеет системного эквивалента во французском языке. Он выражает отношения сопутствования на иллокутивном уровне, сигнализируя, что вводимый им речевой акт 'сопутствует предыдущему', добавляет по поводу сказанного новую информацию, часто с аргументативной функцией (см. [Инькова в печати]), как во втором примере. *Притом* может иметь нулевой эквивалент или переводиться несколькими равночастотными эквивалентами (*en outre, en plus, de plus*), которые имеют более точные соответствия в русском языке (*кроме того, к тому же*).

<p>Такое сочетание авантюриности, притом часто бульварной, с идеями, с проблемным диалогом, [...] воспринималось как грубое и ничем не оправданное нарушение «жанровой эстетики».</p>	<p>Cet alliage de l'aventure, souvent boulevardière, avec l'idée, avec le dialogue idéologique, [...] était perçu comme une infraction injustifiée à l'« esthétique de genre ».</p>
<p>Чтоб кончить все это разом, ей оставалось одно: заметить признаки рождающейся любви в Штольце, не дать ей пищи и хода и уехать поскорей. Но она уже потеряла время: это случилось давно, притом надо было ей предвидеть, что чувство разыграется у него в страсть; да это и не Обломов: от него никуда не уедешь.</p>	<p>Pour en finir d'un coup avec tout cela il ne lui restait qu'une chose : ayant deviné chez Stolz un amour naissant, partir au plus vite, sans lui fournir d'aliment. Mais il était trop tard : Stolz couvait son amour depuis quelque temps déjà; de plus il lui aurait fallu prévoir que ce sentiment se développerait chez lui en une passion. D'ailleurs, ce n'était pas Oblomov : on ne pouvait le fuir.</p>

Тип 3. Коннектор K_A языка A имеет несколько равночастотных переводных эквивалентов-коннекторов K_B^i , адекватно отражающих его семантику. Так, *впрочем* в НБД из 285 случаев переводится *d'ailleurs* в 97 случаях и *du reste* — в 75. Ниже приводятся четыре перевода одного фрагмента русского текста, в котором распределение этих двух ФЭФ 50% на 50%.

<p>Нужно знать, что одно значительное лицо недавно сделался значительным лицом, а до того времени он был незначительным лицом. Впрочем, место его и теперь не почиталось значительным в сравнении с другими, еще значительнейшими.</p>	<p>Il faut savoir que le monsieur important n'était devenu un monsieur important que tout récemment, et que, jusqu'alors, il n'avait rien eu d'important. Du reste, sa place, même à présent, n'était considérée pas comme importante si on la comparait à d'autres, bien plus importantes qu'elle.</p>
	<p>Il y a lieu d'indiquer que ce personnage n'était devenu important que depuis peu; son poste, d'ailleurs, n'était pas considéré comme très important, comparé à d'autres, plus importants encore.</p>
	<p>Il faut dire que ledit personnage n'était devenu important que depuis peu; du reste, par rapport à d'autres plus considérables, la place qu'il occupait n'était pas tenue pour bien importante.</p>
	<p>Il faut savoir qu'un certain personnage important était devenu depuis peu un personnage important bien qu'il fût jusque-là un personnage sans importance. D'ailleurs, même maintenant, sa place ne serait pas considérée comme importante par rapport à d'autres beaucoup plus importantes encore.</p>

Тип 4. Коннектор K_A языка А имеет в языке В системный и наиболее частотный переводной эквивалент коннектором K_B , но он не охватывает всех случаев употребления коннектора K_A . Например, русскому *a to* в значении отрицательной альтернативы соответствует фр. *sinon*. Однако, в отличие от *sinon*, *a to* может выражать целый ряд других отношений (причинное, разделительное, соединительное и др.), сводимых тем не менее к единому семантическому ядру, что позволяет говорить о едином союзе *a to* ([Inkova-Manzotti 2005]). В этих случаях *a to* имеет другие ФЭФ.

Тип 5. В языке В имеется один преимущественный вариант перевода коннектора K_A языка А коннектором K_B , который, однако, не обладает сочетаемостными свойствами коннектора K_A . Речь в данном типе идет о возможности многократного маркирования выражаемого коннектором отношения, благодаря использованию нескольких его показателей. Так, *хотя* имеет преимущественный эквивалент *bien que*, адекватно передающий его семантику. Но *хотя* часто входит в состав двухместных образований, в частности, с союзом *но*. *Bien que* не обладает этим свойством. В результате двухместные образования *хотя... но* переводятся эквивалентами их составляющих: либо *bien que*, либо *mais*.

Хотя ты уверял меня, что она простила, но, знаешь... я не был покоен!	Bien que tu m'aies affirmé qu'elle m'avait pardonné... je n'étais pas tranquille!
Она очень обрадовалась Штольцу; хотя глаза ее не зажглись блеском, щеки не запылали румянцем, но по всему лицу разлился ровный, покойный свет и явилась улыбка.	Elle fut ravie de voir Stolz. Ses yeux ne brillèrent pas, ses joues ne devinrent pas vermeilles, mais tout son visage s'irradia d'une lumière étale et paisible. Elle sourit.

Что касается предлагаемых в [Зализняк 2016] параметров для применения количественных методов, то признаки 1., 3.–4. и 6. могут быть использованы без изменений, а 2. и 5. должны быть адаптированы к анализу коннекторов. Поскольку коннекторы являются в своем большинстве многокомпонентными единицами, то критерий многокомпонентности ФЭФ или СП не является релевантным: у однокомпонентного коннектора может быть многокомпонентный ФЭФ с идентичной семантикой (*например* — *par exemple*). Эти критерии предлагается заменить на критерии принадлежности ФЭФ и СП к классу коннекторов. Если коннектору соответствует коннектор, то ФЭФ и СП считаются конгруэнтными, если коннектору соответствует языковая единица, не принадлежащая к классу коннекторов, то ФЭФ и СП считаются дивергентными (термины заимствованы из [Johansson 2007]). Предлагаются, следовательно, такие параметры для применения количественных методов для анализа коннекторов.

При переводе с русского языка (Р–Ф):

1. наличие большого количества ФЭФ, в том числе имеющих приблизительно равную частотность;
2. доля дивергентных ФЭФ;
3. слово остается без перевода (нулевой ФЭФ).

При переводе на русский язык (Ф-Р):

4. большое количество стимулов перевода (СП);
5. доля дивергентных СП;
6. нулевой СП.

Ниже приводятся статистические данные по четырем коннекторам, полученные в результате сплошной выборки случаев их употребления в текстах на русском языке, загруженных в НБД. Они иллюстрируют первые четыре типа нарушения межъязыкового изоморфизма и параметры их квантитативного анализа. Статистика дается для направления Р-Ф, но подтверждается и в направлении Ф-Р (объемы статьи, к сожалению, не позволяют их привести и прокомментировать).

Коннектор	Всего аннотаций	Zero	Zero%	Dvrg	Dvrg%	Models	Models Ratio
<i>при этом</i>	76	17	22,37	25	32,89	35	46,05
<i>притом</i>	58	12	20,69	1	1,72	16	27,59
<i>впрочем</i>	285	12	4,21	11	3,86	36	12,63
<i>а то</i>	109	14	12,84	2	1,83	24	22,02

Первые два коннектора — *при этом* и *притом* — принадлежат к типам нарушения межъязыкового изоморфизма, где в сопоставляемом языке нет специализированного коннектора для соответствующего логико-семантического отношения: *при этом* — Тип 1, *притом* — Тип 2. Различия между двумя типами отражаются в количественных показателях. Если количество нулевых эквивалентов у них примерно одинаковое (чуть больше 20%), то доля дивергентных ФЭФ достигает у *при этом* почти 33%, поскольку выражаемое им отношение сопутствования на пропозициональном уровне часто передается морфологической формой деепричастия настоящего времени. Однако, *при этом* может выражать и отношение сопутствования между двумя речевыми актами, для которого переводчики используют коннекторы близкой семантики добавления к сказанному, а таких коннекторов во французском языке целая парадигма (*du reste, pour le reste, par ailleurs, de surcroît, de plus, en plus, en outre* и др.). Этим объясняется большое количество моделей перевода *при этом*. Поскольку *притом* может выражать только отношение сопутствования на уровне речевых актов, доля дивергентных переводов у него не достигает 2%, но так как он не имеет системного эквивалента в паре русский — французский, то количество моделей переводов у него достаточно высоко (27, 59%).

Впрочем относится к Типу 3. Все три показателя у него невысокие. Если и фиксируется значительное количество моделей перевода, то все они принадлежат к одному семантическому классу, а их количество в два раза меньше, чем у *притом*, и в почти в четыре, чем у *при этом*. Помимо *du reste* и *d'ailleurs*, *впрочем* может быть переведен *au demeurant* (26 случаев) с такой же семантикой или семантически близкими противительно-уступительными коннекторами: *cependant* (13), *mais* (7), *pourtant* (6) и т.д. Принадлежащее к Типу 4. *а то* не только имеет в два раза больше, чем *впрочем*, моделей перевода, приближаясь по этому параметру к *притом*, но они, кроме того, не являются

синонимичными: помимо *sinon*, на которое приходится более 50%, в качестве ФЭФ могут выступать причинные коннекторы (*car, parce que*), противительные (*tandis que*) или разделительные (*ou*). При интерпретации этих данных нужно, однако, иметь в виду, что даже при наличии системного эквивалента с минимальным различием в семантике, как например рус. *или* и фр. *ou*, этот эквивалент используется менее, чем в 80% случаев (для *или* в 76,22%, для *если* на долю *si* приходится 64,29%).

Что касается Типа 5., то двухместные образования *хотя... но* передаются в 50% одним *bien que*, в 18,18% одним *mais*, остальное приходится на другие коннекторы уступительной семантики.

3. Заключение

Предлагаемые пять типов нарушения межъязыкового семантико-функционального изоморфизма в функциональном классе коннекторов находят, таким образом, подтверждение при количественном анализе по шести признакам, указывающим на вероятную лингвоспецифичность коннектора. Не вызывает сомнения, что статистические данные требуют адекватной интерпретации и должны сопровождаться собственно семантическим анализом, являясь одновременно средством его проверки, что и было сделано в настоящем исследовании. Не вызывает сомнения и тот факт, что для надежности статистических данных необходимо увеличить объем доступных для обработки языковых фактов. Однако представляется очевидным, что использование количественных методов анализа с привлечением верифицируемых и программно генерируемых данных повышает надежность лингвистического описания.

Литература

1. Dobrovol'skiy D. O., Kretov A. A., Sharov S. A. (2005), Korpus parallel'nykh tekstov: Arkhitektura i vozmozhnosti ispol'zovaniya [Corpus of parallel texts: Architecture and applications]. Natsional'nyy korpus russkogo yazyka: 2003–2005 [Russian National Corpus: 2003–2005], Moscow, Indrik, pp. 263–296.
2. Dobrovol'skij D., Pöppel L. (2016), The Discursive Construction *дело в том, что* and its Parallels in other Languages: a Contrastive Corpus Study, Computational Linguistics and Intellectual Technologies. Proceeding of the Annual Conference “Dialogue” (2016), Vol. 15 (22), Moscow, RGGU, pp. 134–145.
3. Inkova-Manzotti O. (2005), Encore sur la conjonction russe *a to* [Again about the Russian conjunction *a to*], Revue des études slaves [Review of the Slavic Studies], Vol. 76/4, pp. 485–497.
4. Inkova O. (in press), Konnektory s formantom *pri*: korpusnoe issledovanie [The connectors of Russian with the element *pri*: corpus-based study], Russian Linguistics, 29 pp.
5. Inkova O., Kruzhhkov M. (2016), Nadkorpusnye russko-frantsuzskie basy danykh glagol'nykh form i konnektorov [Supracorpora databases of Russian and

- French verbal forms and connectors], O. Inkova, A. Trovesi eds, *Lingue slave a confronto* [Slavic languages in comparison], Bergamo, Bergamo University Press, pp. 365–392.
6. *Iordanskaya L. N., Mel'chuk I. A.* (2007), *Smysl i cochetaemost' v slovare* [The sense and the combinatory in dictionary], Moskva, Yazyki slavyanskoy kul'tury.
 7. *Johansson S.* (2007), *Seeing through Multilingual Corpora. On the use of corpora in contrastive studies*, Amsterdam, John Benjamins.
 8. *Kruzhkov M. G.* (2015), *Informacionnye resursy kontrastivnykh lingvisticheskikh issledovaniy: elektronnye korpusa tekstov* [Information resources for contrastive studies: electronic text corpora], *Sistemy i sredstva informatiki* [Systems and Means of Informatics], Vol. 25, N2, pp. 140–159.
 9. RG-80 — *Russkaya grammatika* [Russian grammar]. Shvedova N. Yu. (ed.). Moscow, Nauka, 1980.
 10. *Shmelev A. D.* (2015), *Russkie lingvospetsifichnye leksicheskie edinitiy v parallel'nykh korpusakh: vozmozhnosti issledovaniya i "podvodnye kamni"* [Russian language-specific lexical units in parallel corpora: prospects of investigation and "pitfalls"], *Computational Linguistics and Intellectual Technologies. Proceeding of the Annual Conference "Dialogue"* (2015), Vol. 14 (21), Moscow, RGGU, pp. 584–594.
 11. *Wierzbicka A.* (1992). *Semantics, culture, and cognition. Universal human concepts in culture-specific configurations*, N. Y., Oxford, Oxford University Press.
 12. *Wierzbicka A.* (1996). *Semantics: primes and universals*, Oxford, Oxford University Press.
 13. *Zaliznyak Anna A., Levontina I. B., Shmelev A. D.* (2012), *Konstanty i peremennye russkoj jazykovoj kartiny mira* [Constants and Variables of the Russian language picture of the world], Moskva, Yazyki slavyanskoy kul'tury.
 14. *Zaliznyak, Anna A.* (2015), *Lingvospetsifichnye edinicy russkogo yazyka v svete kontrastivnogo korpusnogo analiza* [Russian language-specific words as an object of contrastive corpus analysis], *Computational Linguistics and Intellectual Technologies. Proceeding of the Annual Conference "Dialogue"* (2015), Vol. 14 (21), Moscow, RGGU, pp. 683–695.
 15. *Zaliznyak, Anna A.* (2016), *Baza dannykh mezh"yazykovykh ekvivalentsiy kak instrument lingvisticheskogo analiza* [A Database of cross-linguistic equivalences as an instrument of linguistic analysis]. *Computational Linguistics and Intellectual Technologies. Proceeding of the Annual Conference "Dialogue"* (2016), Vol. 15 (22), Moscow, RGGU, pp. 763–775.
 16. *Zaliznyak, Anna A., Levontina I. B.* (1996), *Otrazhenie natsional'nogo kharaktera v leksike russkogo yazyka* [The reflection of the national character in the vocabulary of the Russian language], *Russian Linguistics*, Vol. 20, pp. 237–264.
 17. *Zatsman I. M., Inkova O. Yu., Kruzhkov M. G., Popkova H. A.* (2016), *Predstavlenie krossyazykovykh znaniy o konnektorakh v nadkorpusnykh bazakh dannykh* [Representation of cross-lingual knowledge about connectors in supracorpora databases], *Informatika i ee primeneniya* [Informatics and Applications], Vol. 10/1, pp. 106–118.