

КРИТИКА И БИБЛИОГРАФИЯ

ОБЗОРЫ

© 2010 г. Ю.Л. КУЗНЕЦОВА, Т.В. ВЕЛЕЙШИКОВА

**СОВРЕМЕННЫЕ КОРПУСНЫЕ ИССЛЕДОВАНИЯ ЯЗЫКА:
НОВЫЕ ПОДХОДЫ**

В обзоре рассматриваются новые методы и возможности современной корпусной лингвистики, в том числе связанные с применением статистики и автоматического анализа сочетаемости слов. Наибольшее внимание уделяется работам Ст. Гриса (США) и А. Стефановича (Германия).

В настоящем обзоре будут рассмотрены некоторые новые подходы и методы, в последние годы интенсивно применяемые к анализу данных различных языков. Общим компонентом всех рассматриваемых подходов является прежде всего их корпусный характер: как это и вообще характерно для современной лингвистики, обращение к корпусам различного рода буквально на наших глазах становится практически обязательным для любой серьезной работы, претендующей на надежность и достоверность получаемых результатов.

Авторы выбрали для анализа группу работ, так или иначе связанную с деятельностью двух современных лингвистов, входящих в число наиболее активных сторонников корпусных методов: это Стефан Грис (в настоящее время работает в Калифорнийском университете Санта-Барбары, США) и Анатолий Стефанович (в настоящее время работает в Бременском университете, Германия). Как представляется, многочисленные публикации этих двух авторов (индивидуальные и совместные), а также сборники, выпущенные под их редакцией, дают наиболее адекватное представление о многих важных гранях современной корпусной лингвистики.

Настоящий обзор состоит из двух частей. В первой части (написанной Ю.Л. Кузнецовой) будет охарактеризован метод коллострукционного анализа, предложенный в ряде недавних работ Ст. Гриса и А. Стефановича. Во второй части (написанной Т.В. Велейшиковой) будет дан более подробный анализ сборника статей по когнитивным и корпусным методам изучения языков, выпущенного под редакцией Гриса и Стефановича [Gries, Stefanowitsch (eds.) 2007].

1. КОЛЛОСТРУКЦИОННЫЙ АНАЛИЗ

Коллострукционный анализ – это название, объединяющее несколько методов лингвистического анализа, созданных Ст. Грисом и А. Стефановичем. Название «коллострукция» (англ. *collostruction*) происходит от сложения двух слов *collocation* ‘коллокация’ и *construction* ‘конструкция’. Методы, объединенные под этим названием, находятся на стыке лингвистики и статистики. В качестве лингвистической теории авторы используют терминологию и методы теории грамматики конструкций, созданной Ч. Филлмором [Fillmore 1989; Fillmore, Kay 1995; Goldberg 1995; 2006; Tomasello 2003; Fried, Boas 2005 и др.]. В основе лежит метод, являющийся развитием коллокационного метода

анализа, в рамках которого исследуется, какие слова наиболее частотно встречаются в том или ином контексте. Примером коллокационного анализа может служить разбор различий между *actually* 'действительно' и *in fact* 'на самом деле' в [Oh 2000] или исследование сочетаемостных свойств слов *between* 'между' и *through* 'сквозь' в [Kennedy 1991]. Поиск наиболее частотных вариантов заполнения слота конструкции дает авторам значительные преимущества, в результате которых он становится более «прицельным». Например, использование конструкций позволяет учитывать морфологические ограничения на поиск: так, если известно, что в исследуемой конструкции глагол встречается только в прошедшем времени, это дает возможность отбросить много ненужных контекстов.

Коллострукционный анализ включает в себя три основных метода:

- колексемный анализ, служащий для измерения степени притягивания / отторжения данной леммы относительно слота конструкции;
- различительный колексемный анализ, измеряющий какую из двух схожих функционально конструкций предпочитает конкретная лексема;
- ковариация колексем – метод, позволяющий измерить насколько часто одна лексема в одном из слотов употребляется одновременно с определенными лексемами в других слотах.

Рассмотрим действие каждого из этих методов на конкретных примерах. Сначала опишем поведение английской конструкции *N waiting to happen* с точки зрения коллострукционного анализа и покажем, в чем преимущества этого подхода по сравнению с коллокационным. Обратимся к конкордансу словосочетания *waiting to happen*, составленному традиционным образом. Такой список имеет несколько очевидных недостатков. Все они являются следствием того факта, что линейная структура отнюдь не всегда непосредственно отражает синтаксическую структуру. Например, посмотрев на данные в таблице 1, читатель может подумать, что *business*, *horizon* и *company* занимают именной слот конструкции *N waiting to happen*. Однако, если мы рассмотрим их синтаксическую структуру более подробно, выяснится, что примеры 14, 24 и 28 в таблице 1 покажут нам, что эти существительные являются частью других синтаксических единиц и только случайно находятся непосредственно перед словосочетанием *waiting to happen*. Другую проблему традиционного конкорданса можно проиллюстрировать на примере слов *recovery* и *it*, которые, несмотря на то, что они занимают именной слот конструкции (строки 12 и 28 в таблице 1), не появляются в списке в таблице 2, содержащей коллокации данной конструкции. Это происходит потому, что в соответствующих предложениях эти слова не стоят непосредственно перед *waiting to happen*. Таким образом, можно видеть, что конкорданс дает нам не всегда точное представление о словах, заполняющих слоты конструкции.

А. Стефанович и Ст. Грис предлагают более трудоемкий, но дающий более точные результаты способ: вручную разметить элементы, заполняющие слоты конструкции. Однако, если то или иное слово часто встречается в определенном слоте конструкции, это еще не говорит о том, что конструкция предпочитает это слово. Возможно, частотность слова *accident* очень высока в исследуемом корпусе, и только поэтому оно так часто встречается в конструкции *waiting to happen*. Для того чтобы исключить такую возможность, авторы сравнивают частотность слов в корпусе с частотностью слов в конструкции. Для этого они пользуются тестом Фишера и сравнивают четыре переменных: количество употреблений слова *accident* в конструкции, количество употреблений слова *accident* в корпусе, количество употреблений конструкции *waiting to happen* с другими словами в именном слоте и общее число конструкций в корпусе (ср. таблицу 3). В результате А. Стефанович и Ст. Грис могут сравнить, совпадает ли распределение переменных в именном слоте с распределением тех же переменных в корпусе. Если одна из переменных встречается в конструкции чаще, чем это предсказывает распределение в корпусе, то о такой переменной говорят, что она «притягивается» (*attract*), а если реже, чем предсказывает распределение в корпусе, что она «отталкивается» (*repel*).

Конкорданс конструкции *waiting to happen* (отсортирован по L1)

#	left context	node	right context
1	Stewart said that there was an accident	waiting to happen	and he feared lives would be lost.
2	the horse's knees. It was an accident	waiting to happen."	Recall stewards, dressed in day-glo bibs,
3	you had a cartoon about an accident	waiting to happen.	You could have saved the cartoonists fee
4	Unless, of course, it was an accident	waiting to happen.	That insurer has 1,500 appointed
5	"Why?" "Because Stud's like an accident	waiting to happen,	that's why." "Oh, fuck off, Joey! I'm
6	the site say it was an accident	waiting to happen.	Video-Taped report follows JESSICA
7	the building means it was an accident	waiting to happen.	Unfortunately last night an accident did
8	the horse's knees. It was an accident	waiting to happen."	Blow for "blot on landscape" golf range
9	the return of his body. An accident	waiting to happen.	Charity stunt learn warned you're playing
10	of it. Bands like that are accidents	waiting to happen	in a world where 99 per cent
11	actions which are little more than accidents	waiting to happen.	A little more patience and consideration on
12	yesterday: "I think the recovery has been	waiting to happen	for the last couple of months. It
13	Saturday was an accident that had been	waiting to happen.	I wrote to Sir Bob Reid, the
14	accident at the heart of the company	waiting to happen:	now IBM's signalling of the death of
15	not matter was the real constitutional crisis	waiting to happen,	vindication to all those Euro-sceptics who
16	which Coleman warned him of the "disaster	waiting to happen".	The identity papers seized by the FBI
17	- I'm pulling. "This is a disaster	waiting to happen,"	he added, in a prophecy that would
18	who said that it was "a disaster	waiting to happen".	Our hospitals are so short of cash
19	just had to be one monumental disaster	waiting to happen,	Leith later realised. But to start with,
20	marriage to Mandy Smith was a disaster	waiting to happen.	Urging Jagger to rebuild his marriage with
21	is a graphic example of a disaster	waiting to happen.	Over the weekend all attempts to salvage
22	one of these may be a disaster	waiting to happen.	In Lancashire towns like Oldham, Bolton
23	described in The Independent as" a disaster	waiting to happen".	The management of the economy has
24	- "Well - for a business disaster	waiting to happen,	you seem to have come off remarkably
25	develops this theme, identifying "disasters	waiting to happen"	associated with liquified natural gas, oil and
26	events of this week were an earthquake	waiting to happen.	Historians will argue over what was the
27	the first-half goal rush was an event	waiting to happen.	Young wingers are like young spin bowlers;
28	As if it [sex]'s just over the horizon,	waiting to happen	to me, as weird and wonderful as
29	residents are certain that "an accident is	waiting to happen".	Their fears - which focus on a
30	arguments that a new industrial revolution is	waiting to happen	in space are, for now, unconvincing. The
31	Cause" was a carefully planned invasion just	waiting to happen,	poised at the starting gate for the
32	and I can feel the dream just	waiting to happen,	gathering its energies from somewhere on
33	a graphic illustration of the disaster that's	waiting to happen	out there." Stuck fast: the Bettina Danica
34	in food production. A disaster was	waiting to happen.	Like so many cash crops, sugar is
35	that there may be many more Welkoms	waiting to happen.	and if racial conflict does spread in

В таблице 4 представлены элементы, притягивающиеся к именному слоту конструкции *waiting to happen*.

Нужно отметить, что самый высокий рейтинг притягивания у слов *accident* и *disaster*, лидировавших в конкордансе конструкции *wating to happen*. Таким образом, получается, что простой конкорданс дает для конструкции *waiting to happen* почти такие же результаты, как и коллострукционный анализ. В чем же преимущество данного метода?

Преимущество его станет видно, если мы обратимся не к лексическим, а к грамматическим конструкциям. Например, такой же метод может быть применен к любой грамматической форме. А. Стефанович и Ст. Грис [Stefanowitsch, Gries 2003] применяют свой метод к прогрессиву в английском языке (конструкции *be + V-ing*). Если соста-

Частотность слов, встречающихся в левом и правом контекстах *waiting to happen*

L2	L1	R1	R2	
an	11 accident, disaster	9 in	3 the	2
a	6 accidents, been,	2 and, the, you	2 a, added, at,	1
the	3 is, just	a, associated,	1 could, fears, for,	
disaster	2 company, crisis,	1 blow, charity, for,	he, hospitals,	
accident, are,	1 disasters, earth-	gathering, he,	IBM's, identity,	
business,	quake, event,	historians, I,	if, insurer, its,	
constitutional,	horizon, that's,	Leith, like, now,	Jagger,	
dream, had, has,	was, Welkoms	our, out, over,	Lancashire, last,	
identifying,		poised, recall,	later, little,	
invasion,		that, that's, their,	management,	
monumental,		to, unfortunately,	me, report, seem,	
more,		urging,	so, space,	
revolution, than		video-taped,	stewards, stunt,	
		vindication,	there, to, why,	
		young	will, wingers,	
			with, wrote	

Таблица 3

Кросстабуляция слова *accident* и конструкции *waiting to happen*

	accident	¬ accident	Row totals
[N <i>waiting to happen</i>]	14	21	35
¬ [N <i>waiting to happen</i>]	8,606	10,197,659	10,206,265
Column totals	8,620	10,197,680	10,206,300

Таблица 4

Коллексемы, наиболее сильно притягивающиеся к именному слоту конструкции *waiting to happen*

Collexeme (n)	P _{Fisher exact} (collostruction strenght)
accident (14)	2.12E-34
disaster (12)	1.36E-33
welkom (1)	4.46E-05
earthquake (1)	2.46E-03
invasion (1)	7.10E-03
recovery (1)	1.32E-02
revolution (1)	1.68E-02
crisis (1)	2.21E-02
dream (1)	2.45E-02
it (sex) (1)	2.83E-02
event (1)	6.92E-02

вить обычный конкорданс глаголов в прогрессивной форме, такой список просто отразит частотность глаголов. Интересно понять, однако, какие глаголы *тяготеют* к данной форме, то есть встречаются в этой форме чаще, чем распределение формы в корпусе предсказывало бы. Для этого необходимо узнать, сколько в целом в корпусе встречается глагольных форм и какую долю от них занимают формы прогрессива. Эти цифры необходимо сравнить с количеством форм данного глагола и долей прогрессива среди форм конкретного глагола. Если доля прогрессива для конкретного глагола будет выше среднестатистической, то такой глагол «притягивается» к прогрессивной конструкции, если ниже – «отталкивается».

Таблица 5

Пример таблицы с данными для точного критерия Фишера: глагол *сказать*

	все глаголы	сказать
V	27 616 742	421 203
V, praet	14 348 054	284 115

Коэффициент силы взаимодействия подсчитывается при помощи точного критерия Фишера. Этот критерий используется для оценки взаимодействия двух факторов и применяется к таблице из четырех элементов, такой, как таблица 5. Например, в качестве таких факторов могут служить лексема и грамматическая форма глагола, в таблице 5 приведены данные из Национального корпуса русского языка, характеризующие употребление глагола *сказать* в прошедшем времени. Первая пара цифр показывает, сколько глаголов встречается в корпусе и сколько из них употреблены в форме прошедшего времени. Вторая пара цифр представляет данные о глаголе *сказать*: сколько раз этот глагол встречается в корпусе и сколько раз он встречается в форме прошедшего времени. Сравнив первую пару цифр, мы видим, что в среднем прошедшее время глагола занимает 52 % от всех форм глагола, т. е. чуть больше половины. Если глагол *сказать* употребляется в форме прошедшего времени среднестатистически, то для этого глагола должно быть верно такое же процентное соотношение, то есть формы прошедшего времени должны занимать 52 % от всех форм глагола *сказать*. Данный глагол употребляется в корпусе всего 421 203 раза. 52 % от этого числа – это 218 832. Сравнивая это число с реальным употреблением форм прошедшего времени 284 115, мы видим, что глагол *сказать* употребляется в прошедшем времени чаще, чем среднестатистический глагол в корпусе – в 67 % случаев, то есть глагол *сказать* в терминологии А. Стефановича и Ст. Гриса притягивается к конструкции прошедшего времени. Это легко объяснить – ведь данный глагол в прошедшем времени часто употребляется для оформления прямой речи в тексте.

Однако трудно себе представить, что какой-нибудь глагол употреблен в корпусе в прошедшем времени в точности в 52 % случаев. Например, глагол *сказать* мог быть употреблен в прошедшем времени в 218 800 случаев (на 32 примера меньше, чем 52 %). Является ли разница в 32 примера достаточным основанием для того, чтобы говорить о серьезном отличии от среднестатистических 52 %? На этот вопрос помогает ответить точный критерий Фишера. Этот критерий предсказывает, какова вероятность того, что наблюдаемое распределение получилось случайно, что и является вероятностью того, что два сравниваемых параметра независимы. Например, если бы в таблице 5 стояло число 218 800, то вероятность независимости факторов была бы 0,94, то есть очень близка к 1. Для таблицы 5, в том виде, как она представлена выше, вероятность независимости факторов равна 0, то есть притягивание глагола *сказать* к конструкции прошедшего времени является статистически значимым¹. Таким образом, мы видим, что

¹ Обычно о статистической значимости говорят, если вероятность того, что такое распределение получилось бы случайным образом, меньше, чем 0,05.

результаты критерия Фишера хорошо соотносятся с нашей интуицией. В случаях, когда нам кажется, что отличие от предсказанной цифры невелико, критерий Фишера дает вероятность независимости близкую к 1. В случае же, если отличие кажется нам существенным – вероятность независимости, согласно критерию Фишера, низкая. Далее мерой силы притягивания и отталкивания будет вероятность независимости факторов, полученная в результате применения точного критерия Фишера. Причем эта вероятность может одинаково характеризовать как силу притягивания, так и силу отталкивания.

Возвращаясь к прогрессивной конструкции в английском языке, можно задуматься о том, какие глаголы должны притягиваться и отталкиваться от прогрессивной конструкции? Эта конструкция, как известно (см. например [Jespersen 1931: 178; Dowty 1979: 145]), обозначает, что ситуация, названная глаголом, находится в процессе развития. Это, в свою очередь, говорит о том, что стативные глаголы, которые означают такие ситуации, которые не могут развиваться, встречаются в данной форме очень редко и только в специальных контекстах, ср. [Lakoff 1970: 121]. В таблице 6 представлены лексемы «притягивающиеся» и «отталкивающиеся» от прогрессивной формы. Можно видеть, что результаты коллострукционного анализа именно таковы, как предсказывает

Таблица 6

Прогрессивная форма: притягивающиеся и отталкивающиеся колексемы

attracted		repelled	
Collexeme (n)	Collostruction strength	Collexeme (n)	Collostruction strength
talk (234)	1.32E-94	be (448)	0
go (640)	1.08E-89	know (31)	1.01E-63
try (282)	8.86E-84	think (160)	4.05E-34
look (371)	4.41E-77	see (72)	6.36E-31
work (250)	2.14E-68	have (247)	1.93E-29
sit (100)	2.55E-57	want (44)	6.51E-21
wait (88)	6.17E-38	mean (15)	7.72E-17
do (539)	2.16E-36	need (5)	1.11E-14
use (264)	3.18E-29	seem (3)	1.02E-10
come (348)	9.65E-26	believe (11)	3.44E-09
run (113)	1.75E-25	call (30)	3.32E-08
move (104)	5.8E-19	put (93)	6.7E-08
live (101)	1.97E-17	remember (12)	9.49E-08
deal (57)	2.19E-16	find (56)	4.58E-07
walk (55)	9.34E-16	include (6)	2.76E-06
watch (46)	2E-15	agree (9)	4.45E-06
wear (48)	3.76E-14	base (2)	2.04E-05
write (123)	1.58E-13	set (34)	3.39E-05
listen (42)	2.18E-12	sound (6)	3.55E-04
seek (48)	8.66E-11	concern (3)	3.92E-04
fight (32)	2.63E-10	imagine (2)	4.97E-04
stand (57)	4.97E-10	let (10)	5.83E-04
study (31)	1.67E-09	mention (8)	1.04E-03
plan (28)	1.87E-09	exist (4)	1.13E-03
increase (54)	2.36E-09	get (294)	1.27E-03
sing (25)	3.54E-09	regard (2)	1.27E-03
approach (25)	5.13E-09	require (12)	1.3E-03
depend (43)	6.21E-09	marry (1)	1.86E-03
speak (71)	1.24E-08	stop (7)	2.13E-03
sell (38)	1.46E-08	indicate (3)	2.29E-03

традиционный анализ. 30 глаголов, которые сильнее всего отталкиваются конструкцией, – это стативы, и лидирует среди них «самый стативный» глагол *be* 'быть'.

В этой таблице и далее число в скобках сразу после лексемы указывает точное количество искомых форм данной лексемы, найденных в корпусе. Можно видеть, что само это число мало связано с притягиванием или отталкиванием, так форма *being* была обнаружена в прогрессивной конструкции 448 раз, однако в сравнении с общим количеством употреблений глагола *be* это так мало, что вероятность случайности такого совпадения равна 0, ср. колонку справа от лексемы, где как раз находится число, которое характеризует силу наблюдаемого взаимодействия (притягивания или отталкивания). В результате глагол *be* охарактеризован как отталкивающийся от прогрессивной конструкции, несмотря на 448 встретившихся примеров употреблений. Коэффициент силы взаимодействия представлен в следующем формате: слева от символа E находится коэффициент, который необходимо умножить на 10 в степени, указанной справа от символа E. Так, например, 1.86E-03 следует читать, как $1.86 \cdot 10^{-3} = 0.00186$. Из таблицы 6 можно видеть, что самыми сильно притягивающимися к прогрессивной конструкции являются лексемы *talk* и *go*. Первая встретила в прогрессивной конструкции 234 раза, а вторая – 640 раз. Вероятность случайности такого результата в обоих случаях очень близка к 0, то есть мы имеем дело со статистически значимым притягиванием.

Другим интересным применением коллокационного анализа является сравнение двух близких конструкций. А. Стефанович и Ст. Грис [Gries, Stefanowitsch 2004] среди других примеров разбирают две известные конструкции английского языка: так называемую дитранзитивную конструкцию (1a) и конструкцию с *to* (1b). Во многих работах утверждается, что эти конструкции различаются только синтаксически (см., например [Burt 1971]) или только на уровне информационной структуры (ср. [Erteschik-Shir 1979: 443, 449; Thompson 1990: 245]), с точки же зрения семантики между этими двумя конструкциями нет различия: и та и другая означает передачу объекта адресату. В других работах отстаивается семантическая природа различий между двумя конструкциями и утверждается, что дитранзитивная конструкция используется при малой дистанции между субъектом и адресатом, а конструкция с *to* при большой дистанции [Thompson, Koide 1987: 400]. В некоторых работах предлагается дальнейшее уточнение этого различия: дитранзитивная конструкция означает 'X заставляет Y получить Z', в то время как конструкция с *to* означает 'X заставляет Z переместиться к Y' (ср. [Goldberg 1995: Section 2.3.1, Chapters 5–7; Pinker 1989; Gropen et al. 1989: 243]).

- (1) a. John sent Mary the book.
b. John sent the book to Mary.
'Джон послал книгу Мэри'.

Если изучить, какие лексемы притягиваются к этим двум конструкциям, можно получить ответ на вопрос, какова природа различия между ними. Если различие между конструкциями только в синтаксической структуре или только в информационной структуре, то набор притягивающихся лексем должен быть одинаковым в обоих случаях. Если же различие между конструкциями семантическое, то лексемы, притягивающиеся к конструкциям, будут разными и будут отражать семантические ограничения каждой из конструкций.

В этой таблице после каждой лексемы в скобках указано количество употреблений этой лексемы в обеих исследуемых конструкциях: сначала количество найденных употреблений данного глагола в дитранзитивной конструкции, а затем количество употреблений этого же глагола в конструкции с *to*. В шапке таблицы приведено количество найденных в корпусе примеров обеих конструкций. Можно видеть, что в исследуемом корпусе было обнаружено 1 035 примеров дативной конструкции и 1 919 примеров конструкции с *to*, то есть среднестатистическое соотношение между конструкциями примерно 1 к 2. Глагол *give* – самый яркий представитель дативной конструкции – имеет значительный «перекосяк» в сторону дативной конструкции: соотношение его примеров – приблизительно 4 к 1. Если бы у глагола *give* было бы среднестатистическое распределение,

Лексемы, притягивающиеся к дитранзитивной и конструкции с *to*

DITRANSITIVE (N = 1,035)		To-DATIVE (N = 1,919)	
Collexeme	Distinctiveness	Collexeme	Distinctiveness
give (461:146)	1.84E-120	bring (7:82)	1.47E-09
tell (128:2)	8.77E-58	play (1:37)	1.46E-06
show (49: 15)	8.32E-12	take (12:63)	0.0002
offer (43: 15)	9.95E-10	pass (2:29)	0.0002
cost (20:1)	9.71E-09	make (3:23)	0.0068
teach (15:1)	1.49E-06	sell (1:14)	0.0139
wish (9:1)	0.0005	do (10:40)	0.0151
ask (12:4)	0.0013	supply (1:12)	0.0291
promise (7:1)	0.0036	read (1:10)	0.0599
deny (8:3)	0.0122	hand (5:21)	0.0636
award (7:3)	0.0260	feed (1:9)	0.0852

то на 461 пример дативной конструкции приходилось бы 855 примеров конструкции с *to*. Так что мы видим, что глагол *give* статистически значимо притягивается к дитранзитивной конструкции².

В таблице 7 представлены результаты коллокационного анализа лексем, притягивающихся к двум интересующим нас конструкциям. Можно видеть, что списки эти совсем не совпадают друг с другом, что свидетельствует о том, что конструкции различаются между собой по семантике. Видно также, что список лексем, предпочитающих дитранзитивную конструкцию согласуется с семантическим анализом, представленным выше. В дитранзитивной конструкции предпочитают глаголы, означающие ситуации непосредственного контакта: *give* 'дать', *tell* 'сказать', *show* 'показать', *offer* 'предложить' и т. д. Напротив, в списке глаголов, предпочитающих дитранзитивную конструкцию лидируют глаголы дистантной передачи объекта, такие как *bring* 'приносить', *pass* 'передавать' и т. п. Таким образом, можно видеть, что списки лексем, «притягивающихся» к двум близким конструкциям, показывают, что различие между этими конструкциями семантического рода. Более того, глаголы в этих списках отражают семантические предпочтения конструкций.

Еще один интересный случай представляют конструкции, в которых семантические ограничения затрагивают несколько слотов одновременно. Рассмотрим каузативную конструкцию с *into* в английском языке (ср. пример (2)), рассматриваемую в работе [Stefanowitch, Gries 2005]. А. Вежбицкая [Wierzbicka 1998] обнаружила, что эту конструкцию можно использовать только в таких ситуациях, в которых объект каузации изначально не хочет выполнять действие, однако каузатор преодолевает его сопротивление, обычно при помощи давления или нечестной уловки.

- (2) Newley had been **tricked into revealing** his hiding place (BNC GUU).
 'Ньюли обманом заставили выдать свое убежище'.

Это описание предсказывает, какие глаголы притягиваются к обоим интересующим нас слотам конструкции. Первый слот будут занимать глаголы, означающие оказание давления или нечестную уловку. Ко второму слоту будут притягиваться глаголы, обозначающие действие, кажущееся исполнителю неприятным или нежелательным. Это

²Сила притягивания следующего за ним глагола *tell* менее значительна только потому, что в корпусе встретилось меньше примеров этого глагола, а чем с меньшим количеством примеров имеет дело тест Фишера, тем выше получается вероятность того, что наблюдаемое распределение получилось в результате случайности.

предсказание во многом оправдывается. Например, список лексем, притягивающихся к первому глагольному слоту, состоит из глаголов уловки (*trick* 'заставить обманом', *fool* 'одурачить', *mislead* 'вводить в заблуждение' и т. д.) и глаголов давления (*coerce* 'вынудить', *force* 'заставить', *bully* 'заставить силой' и т. д.). В списке глаголов, предпочитаемых во втором слоте конструкции, будут присутствовать глаголы, означающие однозначно неприятные действия (*misbehave* 'плохо себя вести', *misrecognize* 'неправильно опознать', *betray* 'предать'), однако в этом списке появятся и глаголы, означающие действия, не всегда ассоциирующиеся с чем-то неприятным, такие как *think* 'думать', *believe* 'верить', *reveal* 'выказать', *prove* 'доказать', *buy* 'купить', *purchase* 'приобрести'. Для того чтобы объяснить, как глаголы, не означающие неприятного события, оказываются частотными во втором слоте каузативной конструкции, необходимо рассмотреть фреймы, в которых они часто используются: вероятно, эти глаголы употребляются в ситуациях, однозначно ассоциирующихся с чем-то нежелательным.

Получается, что два слота конструкции связаны друг с другом, то есть неверно, что в примерах представлено случайное перемножение глаголов, могущих заполнять первый слот и могущих заполнять второй слот. Неверно, что любой глагол из первого списка одинаково часто встречается с любым глаголом из второго списка. Для того чтобы статистически выявить связь, которая может быть между двумя единицами, необходимо измерить взаимное притяжение лексем к этим двум слотам, то есть в терминологии А. Стефановича и Ст. Гриса ковариацию колексем. Для этого необходимо взять вероятность появления глагола₁ в первом слоте конструкции и глагола₂ во втором слоте. Если бы эти два события были бы независимыми, то вероятность появления одновременно глагола₁ в первом слоте и глагола₂ во втором слоте была бы равна произведению этих двух вероятностей. Сравнив это число с реальным количеством пар глагол₁-глагол₂ в конструкции, мы можем сказать, что пара глагол₁-глагол₂ притягивается к конструкции, если реальная вероятность выше предсказанной, и отталкивается, если ниже. Результаты такого анализа представлены в таблице 8.

В этой таблице, как и в предыдущих, справа дается сила притягивания / отталкивания пары лексем по отношению к конструкции, измеренная при помощи точного критерия Фишера. Однако в этом случае справа дается не значение самой вероятности независимости факторов, как в предыдущих случаях, а $\log_{10}$ этого числа. Таким образом, значение 30.06 – силу притягивания компонентов словосочетания *fool into thinking* к конструкции – следует читать как E-30, то есть очень маленькое число, что говорит о статистически значимом притяжении этого словосочетания к конструкции. Заметим,

Таблица 8

Наиболее значимые колексемные комбинации в каузативной конструкции с *into*

Attracted covarying-collexeme pairs in the <i>into</i> -causative		Repelled covarying-collexeme pairs in the <i>into</i> -causative	
<i>fool into thinking</i>	30.06	<i>force into thinking</i>	2.554
<i>mislead into thinking</i>	12.755	<i>coerce into thinking</i>	1.421
<i>mislead into believing</i>	8.355	<i>trick into making</i>	0.945
<i>deceive into thinking</i>	5.651	<i>push into thinking</i>	0.794
<i>trick into parting</i>	5.248	<i>trick into accepting</i>	0.717
<i>encourage into farming</i>	4.652	<i>bully into believing</i>	0.716
<i>dragoon into serving</i>	4.652	<i>talk into believing</i>	0.671
<i>aggravate into producing</i>	4.28	<i>trick into thinking</i>	0.634
<i>panick into seizing</i>	4.078	<i>lead into believing</i>	0.561
<i>seduce into misbehaving</i>	3.966	<i>talk into making</i>	0.536
<i>delude into believing</i>	3.952	<i>force into giving</i>	0.497
<i>torture into revealing</i>	3.75	<i>tempt into thinking</i>	0.42

Лексемы, заполняющие первый слот конструкции при глаголах коммерческих транзакций

RESULT	CAUSE
<i>buy</i>	<i>mislead, hoodwink, lure, entice, boss, pester, diddle, guilt-trip, scare, nag, pressure, steer, tempt, fool</i>
<i>purchase</i>	<i>mislead, lure</i>
<i>pay</i>	<i>con, dupe, harass, intimidate, scare, blackmail, tie, panick, mislead, shame</i>
<i>overpay</i>	<i>dupe</i>
<i>sell</i>	<i>panick, force, entrap, terrify</i>

что в списке отталкивающихся словосочетаний только первые два отталкиваются со статистически значимой силой.

Можно видеть, что результаты коллострукционного анализа в этом случае указывают на несколько важных для этой конструкции фреймов. Первые четыре примера иллюстрируют соотношение между уловкой в первом слоте конструкции и знанием / верой во втором слоте конструкции. Нужно заметить, что другой класс глаголов, часто появляющихся в первом слоте, – глаголов давления – не участвует в этом фрейме. Можно видеть, что пары *force into thinking* ‘заставить думать’, *coerce into thinking* ‘вынудить думать’, *bully into believing* ‘заставить верить силой’ отталкиваются от конструкции. Оба эти явления отражают культурные представления говорящих о том, как одни люди влияют на мысли других людей: это часто происходит при помощи уловок, но редко при помощи насилия.

Другой интересный фрейм, который можно видеть в рамках этой конструкции, – это фрейм, представленный парой *torture into revealing* ‘вынудить рассказать под пытками’³. В этом фрейме, напротив, в первом слоте преобладают глаголы физического давления: именно физическое давление используется для того, чтобы вынудить человека рассказать информацию, которую он хочет скрыть. В качестве результата в этом фрейме часто встречаются другие глаголы сообщения информации: *admit* ‘признать’, *confess* ‘признаться’, *prove* ‘доказать’.

Еще один фрейм, выявляемый коллострукционным анализом в каузативной конструкции, связывает между собой глаголы уловки и глаголы коммерческих транзакций. Проанализировав более расширенный вариант таблицы 8, А. Стефанович и Ст. Грис выявили частотные употребления каузативной конструкции, в которых первый слот занимали глаголы коммерческих транзакций: *buy* ‘покупать’, *purchase* ‘приобретать’, *overpay* ‘переплачивать’, *pay* ‘заплатить’, *sell* ‘продать’. Результаты такого анализа представлены в таблице 9. Рассмотрим такие притягивающиеся к данной конструкции пары, в которых второй слот заполнен одним из этих глаголов. Посмотрим, какие глаголы при этом будут заполнять первый глагольный слот конструкции. Результат такого поиска можно видеть в колонке «CAUSE» таблицы 9. Легко видеть, что этот фрейм выявляет культурное знание о действиях продавца: нередко покупатель оказывается вовлечен в ситуацию покупки вопреки собственной воле под влиянием нечестной уловки со стороны торгового агента.

Таким образом, можно видеть, что изучение ковариации колексем – взаимного распределения слов, заполняющих разные слоты конструкции, выявляет, что в рамках одной конструкции, например, такой как каузативная конструкция с семантикой каузации

³ Интересно, что для выражения этого фрейма в русском языке используется другая конструкция: **пытать рассказать*, аналогичное вполне приемлемому в русском *вынудить считать*, не является грамматичным.

неприятного действия, могут скрываться несколько культурно значимых фреймов, рассказывающих о том, в каких случаях одно лицо может заставлять другое выполнять неприятные действия. В рамках каждого из таких фреймов имеются дополнительные семантические ограничения на переменные. Например, во фрейме «нечестной продажи» участвуют в первую очередь глаголы уловки в первом слоте и глаголы коммерческих транзакций во втором слоте. Коллострукционный анализ позволяет легко увидеть такие взаимные предпочтения, что, в свою очередь, помогает обнаружить фреймы, важные для данной конструкции.

Приведенные здесь примеры дают представление о методах и возможностях коллострукционного анализа и о той пользе, которую он может принести для решения самых разных лингвистических задач, устанавливая условия притяжения, отталкивания и взаимного распределения языковых единиц в контексте конструкций разной структуры.

И еще об одном преимуществе данного метода: его независимости от исходных теоретических предпосылок. Действительно, как правило, статистические методы существенно опираются на изначальные предположения исследователя – в этом их слабость: если какой-то важный параметр почему-либо оказался не включен в изначальный список параметров, следы его влияния невозможно обнаружить в результатах исследования. Коллострукционный анализ не имеет такого недостатка. Действие этого метода целиком базируется на корпусной и статистической информации и не зависит от заранее сформулированных гипотез. Поэтому для лингвистов коллострукционный метод может стать по-настоящему действенным методом независимой проверки предположений относительно изучаемой конструкции и одновременно послужить инструментом пропаганды теории грамматики конструкций среди сторонников других теорий, ведь возможность такой независимой проверки служит аргументом в пользу истинности постулатов теории грамматики конструкций.

2. КОРПУСА И КОГНИТИВНАЯ ЛИНГВИСТИКА

Современные лингвистические исследования в области синтаксиса и лексики все в большей степени базируются на методах, предполагающих обработку значительных по объему массивов данных и тем самым эффективно решающих проблему репрезентативности материала. В современных условиях подобный массив данных представляют так называемые электронные корпуса текстов. Использование этих корпусов формирует целое направление в лингвистике – корпусную лингвистику, которая разрабатывает общие принципы построения корпусов текстов с учетом современных компьютерных технологий и методику сбора письменных и устных текстов, а также способов их хранения и анализа [Баранов 2001; McEnergy, Wilson 2004]. Работа с корпусами текстов помогает исследователям избежать субъективности и дает возможность автоматического поиска информации и сплошного анализа материала. Исследования, основанные на корпусах текстов, как правило, используют функциональный подход к семантике и сочетаемости языковых единиц; для современной корпусной лингвистики характерен интерес не только к новым технологиям, но и к новым теоретическим концепциям языка, соответствующим новым возможностям обработки данных.

Все исследования сборника [Stefanowitsch, Gries 2007] так или иначе основаны на корпусах текстов и количественных методах. Используются данные английского, немецкого, нидерландского и русского языков. Основной идеей сборника, таким образом, можно считать демонстрацию возможностей, которые открывают современные электронные ресурсы. Авторы надеются, что их книга поможет сформировать новую методологическую парадигму, основанную на использовании более объективных и достоверных данных.

В статье Д. Дивьяк «Способы выражения намерения: разграничение и структурирование неточных синонимов» анализируются проблемы точной и неточной синонимии. Работа состоит из двух частей. Первая часть посвящена выявлению неточных синонимов. Д. Дивьяк использует корпусную базу данных лексических признаков, которая

предполагает корреляцию между дистрибутивным и семантическим сходством, а также кластерный анализ для определения уровня сходства неточных синонимов. Работа основана на данных Национального корпуса русского языка, в котором были выявлены конструкции [V + Inf], выражающие концепт «намерение осуществить действие». В зону данного концепта вошли пять русских синонимических глаголов, таких как «намереваться», «собираться», «предполагать», «думать» и «хотеть». В русском языке существует еще один глагол, который по своим семантическим признакам должен был бы войти в данную группу, а именно «планировать», но автор исключил его из своего списка, поскольку у глагола отсутствуют структурные свойства, характерные для остальных глаголов этого типа.

Для определения границ группы неточных синонимов автор использует данные «конструкционных сетей», которые характеризуют глаголы намерения. Для каждого глагола создается «сочетаемостный профиль», на основе которого определяется сходство между глаголами. Д. Дивьяк отмечает, что нельзя руководствоваться только методиками анализа значения, а следует учитывать более «объективную» сочетаемость лексемы. Во второй части работы автор дает более детальную характеристику четырем глаголам из пяти, а именно «намереваться», «собираться», «думать» и «хотеть», основываясь на 47 параметрах (или, как их назвала Б. Аткинс, ID-признаки – синтаксические или лексические маркеры в предложениях, которые указывают на отдельное словарное значение слова), которые Д. Дивьяк условно разделил на две группы. Первая группа включает в основном грамматические признаки глаголов, тип предложения, падежное оформление актантов и т. д. Вторая группа отражает сочетаемостную информацию, в том числе о наличии отрицания, наречий, частиц, коннекторов и др.

Кластерный анализ дает различные результаты для каждой из предложенных групп ID-признаков. Однако, если учитывать сразу все свойства, кластеризация является практически идентичной формальной классификации. Иерархическое объединение показывает, например, что при учете всех формальных реализаций глаголы «думать» и «хотеть» больше всего схожи между собой. Однако, если основываться на семантических парафразах, можно обнаружить, что больше всего схожи между собой глаголы «думать» и «намереваться».

Таким образом, автор делает вывод, что каждый сочетаемостный профиль содержит определенный набор семантических признаков. Если две лексемы описываются одним и тем же набором признаков, то можно говорить о полных синонимах. Сочетаемостный профиль также дает возможность подсчитать уровень сходства лексем при помощи стандартных кластерных методик.

Статья Ст. Гриса «Корпусные методы и когнитивная семантика: несколько значений глагола *to run*» посвящена анализу семантики английской лексемы *run*. Автор работает на стыке двух направлений – когнитивной и корпусной лингвистики. По его мнению, использование корпусных методов способствует решению некоторых трудных вопросов когнитивной лингвистики – таких, как определение прототипов, разграничение значений слова, автоматическое разрешение неоднозначности и др.

В начале статьи кратко описаны методы исследования. Анализ значений слова делается в работе с точки зрения как когнитивной, так и корпусной лексикографии. Это обусловлено тем, что, по мнению автора, современные когнитивные исследования рассматривают полисемию в основном как систему взаимосвязанных смыслов, в то время как корпусная лингвистика фокусируется, прежде всего, на дистрибутивных особенностях различных значений слова. Автор указывает, что исследование многозначности лексем и фразеологизмов остается одной из центральных проблем когнитивной лингвистики. Он описывает несколько подходов к решению данного вопроса, используемых в современной лингвистике, например, метод полной спецификации (*full-specification approach*) и метод принципиальной многозначности (*principled-polysemy approach*) – и указывает на недостатки каждого из них. Наиболее подробно Ст. Грис останавливается на корпусном методе, предложенном Дж. Кишнером и Р. Гиббсом [Kishner, Gibbs 1996]. В качестве примера автор приводит два следующих подхода. Первый подход Ст. Грис

рассматривает на примере работы Б. Аткинса [Atkins 1987], посвященной исследованию ID-признаков, т. е. синтаксических или лексических маркеров, которые в контекстных употреблениях указывают на определенный смысл слова. На основании этого Б. Аткинс анализирует 441 контекстное окружение лексемы *danger* и получает несколько семантически важных ID-признаков. Второй лексикографический подход предложен П. Хэнксом [Hanks 1996], который исходит из того, что семантика глагола определяется моделью его управления, в частности, количеством дополнений, которые глагол может иметь. В связи с этим, П. Хэнкс предлагает опираться на сочетаемостный профиль глагола (*behavioral profile*), что перекликается с идеями Б. Аткинса.

Основная часть статьи посвящена когнитивному анализу многозначного глагола *to run*. Базой для исследования послужили 815 примеров употребления названного глагола, взятые из Британского компонента Международного корпуса английского языка (*the British component of International corpus of English*) и из Брауновского корпуса (известный корпус американского варианта английского языка Брауновского университета). На основе анализа этого материала автор выделяет 53 значения глагола *to run* и разрабатывает специальную схему, демонстрирующую различного рода связи между выявленными значениями.

В заключительной части автор сочетает метод, предложенный Б. Аткинсом (анализ ID-признаков) с методом П. Хэнкса (сочетаемостный профиль). Проанализировав примеры употребления глагола *to run* и выявив его значения, Ст. Грис выделяет ID-признаки, характерные для данной единицы. В результате он получает сочетаемостный профиль глагола *to run*, основывающийся на 252 различных ID-признаках.

Следующий важный момент в работе Ст. Гриса – это определение прототипического значения глагола *to run*. Исходя из постулатов, выдвинутых в работе А. Тайлер и В. Эванс [Tyler, Evans 2001], автор указывает, что для данного глагола естественно считать прототипическим значение 'движение'. Однако данные корпуса и сочетаемостный профиль анализируемого глагола показывают, что таким значением необходимо признать 'быстрое пешее движение' (*fast pedestrian motion*), поскольку именно оно является наиболее частотным (и в этимологических словарях дается как одно из первичных).

Определив прототипическое значение глагола *to run*, Ст. Грис возвращается еще к одной важной теме когнитивных исследований – полисемии. При анализе значений слова достаточно спорным является вопрос о том, разграничивать ли два значения как самостоятельные смыслы или же рассматривать их как одно более общее значение. На основе сочетаемостного профиля, когнитивного анализа и корпусных данных автор исследует некоторые достаточно близкие значения глагола *to run* (например, «течь», «избегать чего-л. неприятного, опасного» и т. д.) и доказывает, что их следует считать отдельными самостоятельными значениями.

Подобные ситуации показывают, что сочетаемостные профили являются многообещающей методикой для анализа языковых данных в области лексической семантики. Однако оборотная сторона этой методики – ее сложность, а также то, что корпуса текстов не всегда предоставляют полный объем необходимой информации.

В статье С. Вульффа «Конструкции *Go-V* vs. *go-and-V* в английском языке: случай конструкционной синонимии?» разбираются примеры двух внешне схожих моделей:

1. *Go find the books and show me* букв. 'Пойди найди книги и покажи мне';

2. *Now, just keep polishing those glasses while I go and check the drinks* букв. 'Ну, ты пока протирай стаканы, а я пойду и поищу чего-нибудь выпить'.

Данные примеры образуют конструкции '*go-and-V*' и '*go-V*', которые в лингвистике традиционно считаются одним и тем же явлением (в обоих случаях глагол *go* занимает ведущее положение в конструкции, которая по своей сути асимметрична).

В отличие от генеративно-трансформационного подхода, при котором конструкция *go-V* просто считается редуцированным вариантом конструкции *go-and-V*, С. Вульфф доказывает, что обе данные конструкции следует считать самостоятельными. В статье используются следующие методы исследования:

– коллокационный подсчет совпадений (*collocational overlap estimation*);

- конструкционный анализ устойчивых словосочетаний (collostructional analysis);
- дифференциальный лексемный анализ устойчивых словосочетаний (distinctive collexeme analysis).

Автором было обработано 5 320 примеров на каждую конструкцию, вручную отобранных из Британского национального корпуса (BNC). Анализ этих примеров приводит к выводу, что существует некоторое количество общих ограничений в слоте V2, характерных только для конструкции *go-V* (например, подлежащее в ней всегда должно быть агентивным, ср. *The fumes go and bother the people upstairs*, но не *The fumes go bother the people upstairs*).

Для анализа конструкций используется подход, который не подразумевает прямых деривационных отношений между конструкциями. Однако данный подход не может дать объяснение таким случаям, когда обе конструкции появляются в одних и тех же предложениях и передают одно и то же значение. С. Вульфф предполагает, что подобные особенности конструкций могли сформироваться диахронически.

Наиболее существенные общие выводы автора таковы:

1. Основываясь на методе коллокационного подсчета совпадений, предложенном Ст. Грисом [Gries 2001], автор показывает, что конструкция '*go-V*' не представляет собой подкласс конструкции '*go-and-V*', хотя между ними можно обнаружить большое количество общих свойств. Подсчет совпадений в данном случае очень важен, поскольку обе конструкции являются лексически и дистрибутивно связанными, но, в свою очередь, показывают высокую степень независимости.

2. Конструкция '*go-V*' предпочитает глаголы, обозначающие непредельные процессы и события, и намного меньше – глаголы, обозначающие предельные процессы. В этом случае был использован конструкционный анализ устойчивых словосочетаний, поскольку он подсчитывает степень лексической ассоциации глаголов в слоте V2 с учетом каждой конструкции и затем сравнивает акциональный класс глагола для получаемых списков лексем.

Статья Б. Хампе и Д. Шёнефельд «Синтаксические прыжки или лексические изменения? – Еще раз о “Продуктивном синтаксисе”» основана на поиске в Британском национальном корпусе (BNC) всех случаев использования таких английских глаголов как *encourage*, *fear*, *support* и *bore* в комплексно-транзитивных аргументных структурах, которые часто ассоциируются с другими каузативными глаголами. Авторы продолжают исследование типа продуктивности, которое было начато в работе по «синтаксическим слияниям» (syntactic blends) [Hampe, Schoenefeld 2003].

Было исследовано 766 примеров из Британского национального корпуса, а также Международного корпуса английского языка (ICE-GB). Данные примеры содержат 45 глаголов с различной аргументной структурой. Для их анализа были использованы следующие методы: сравнение распределений глаголов в различных конструкциях, сопоставление конструкционных распределений при употреблении одного глагола, а также лексемный анализ устойчивых словосочетаний.

Авторы рассматривают продуктивное использование глаголов в нетипичных моделях, например, в предложениях типа *He supported them through the entrance door* или *She bore them stupid*. Подобные случаи конструкций 'движения, вызванного какой-либо причиной' или так называемых «результативных» конструкций были подробно исследованы в грамматике конструкций, где была предложена модель, согласно которой необычные аргументные свойства глагола объяснялись через свойства конструкции, в которой он был употреблен [Goldberg 1995].

Для анализа языкового материала Б. Хэмп и Д. Шёнефельд исследовали конструкции, в которых у глаголов появлялся дополнительный аргумент. Глаголы *encourage*, *support* и *fear* только в 1 % всех употреблений используются в каузативном или атрибутивном значении и в результате приобретают аргументный слот, отсутствующий в большинстве типичных употреблений этих глаголов. Иначе обстоит дело с глаголом *bore*, у которого результативные употребления, а также употребления в конструкции каузированного движения, хотя и не являются центральными, но относительно более

часты (около 7 %). Подобные случаи должны трактоваться как идиомы. Таким образом, авторы статьи утверждают, что «аргументная продуктивность» является не только синтаксической по своей природе – разные глаголы в разной степени употребительны в конструкции с дополнительным аргументом. Исследованные различия в поведении предикатов показывают, что лексическое значение предиката также должно быть включено в рассмотрение в «фузионной модели» А. Голдберг при взаимодействии глагола и конструкции.

Статья Г. Гилквин «Место прототипов в корпусной лингвистике: в центре – каузация» посвящена исследованию взаимоотношений когнитивных прототипов и частот встречаемости; задачей автора является определить, в какой степени перифрастические каузативные конструкции (типа *get your father to run us out*) могут считаться реализациями одной из трех когнитивных моделей прототипической каузации. В основе исследования лежит теория прототипов, которая утверждает, что для смежных концептов характерна кластеризация важнейших прототипических употреблений (эта идея пришла из психологии). Когнитивисты склонны считать прототип когнитивно наиболее значимым элементом, тогда как корпусная лингвистика часто отождествляет прототип с элементом, наиболее частотным в корпусе. Таким образом, лингвисты часто ставят знак равенства между такими понятиями, как «частотный», «выделенный (салиентный)» и «прототипический». Однако, по мнению Г. Гилквин, языковая частотность не совпадает с когнитивной салиентностью, по крайней мере в области прототипической каузации.

Автор сравнивает три модели каузальности:

- 1) иконическая последовательность участников (прототипический порядок событий);
- 2) модель «бильярдного шара» (прототипические участники);
- 3) прямая манипуляция (прототипический порядок и участники).

Согласно первой модели порядок аргументов в каузативной конструкции отражает порядок, в котором соответствующие персонажи участвуют в ситуации каузирования. Вторую модель автор соотносит с лангакеровской моделью «бильярдного шара» в каузации с акцентом на причинном движении. Третья модель была предложена Дж. Лакоффом и М. Джонсоном – это модель прямой манипуляции, где прототипическая манипуляция определяется на основе некоего набора общепринятых свойств.

Для сравнения была привлечена база данных Британского национального корпуса, из которой была произведена выборка всех конструкций с основными перифрастическими каузативными глаголами (*get, cause, have, make*). Всего было обработано 3 574 конструкции.

С учетом классификации элементов конструкции сделан вывод, что наиболее типичным видом каузации является цепочка действий, где энергия передается от виновника действия к причине действия и далее к пациенсу. Таким образом выбирается первая модель – модель иконической последовательности участников.

Статья Дж. Ньюмана и С. Райс «Схемы переходности в английских словах *eat* и *drink* в Британском национальном корпусе» посвящена анализу синтаксических свойств английских глаголов *eat* и *drink*. Предполагается, что переходность является сложным феноменом, который далек от того, чтобы быть четким бинарным различительным признаком. Базой данных для исследования опять-таки послужил Британский национальный корпус (проанализировано 934 примера употребления глагола *drink* в устной речи и 2 000 примеров в письменной; для глагола *eat*, соответственно, 2 623 и 2 000 примеров). Все примеры были обработаны вручную, а также размечены для исследования различных морфосинтаксических свойств.

Обнаружилось, что глаголы *drink* и *eat* в основном используются как переходные глаголы, причем намного чаще в устной речи, чем в письменной. Глагол *drink* встречается преимущественно в таких контекстах, которые требуют переходной формы глагола (особенно, когда данный глагол употребляется в выражении *drink an alcoholic beverage*). По сравнению с лексемой *drink* лексема *eat* намного чаще появляется как в устной, так и в письменной речи, а в словосочетаниях обычно ставится на первое ме-

сто (*ate and drank*, вместо *drank and ate*). По мнению авторов, подобная особенность может говорить о когнитивной выделенности: когда мы едим и пьем, питье является как бы дополнением к еде, но не наоборот. Другие выводы авторов касаются характера дополнений, которые обычно используются в переходной конструкции. Например, со словом *eat* наиболее часто используется дополнение *food*. Среди 20 наиболее часто используемых лексем многие обозначают определенные виды еды, такие как *breakfast, lunch, dinner*. Таким образом, данные этого исследования отчасти противоречат результатам предыдущих работ, которые свидетельствовали о незначительной доли переходных употреблений в устной речи.

Статья М. Лемменса «Экспериментальные модели, выявленные в результате работы с корпусом» посвящена семантике позиционных глаголов в нидерландском языке: *zitten* 'сидеть', *liggen* 'лежать', *staan* 'стоять' и их каузативных коррелятов: *zetten* 'сажать', *leggen* 'класть', *steken / stoppen* 'ставить, вставлять'. В исследовании также рассматривается глагол *doen* 'делать', который в некоторых контекстах может использоваться как глагол местоположения. Автор проанализировал 7 550 примеров, взятых из корпуса INL (письменный корпус нидерландского языка), и осуществил их разметку на базе различных морфосинтаксических и семантических свойств. Было обнаружено, что во многих случаях отсутствует прямая связь между каузативными и некаузативными глаголами (впрочем, для разных пар в разной степени). Интересно, что употребление позиционных глаголов применительно к человеку в корпусе встречается довольно редко (например, менее 1 % всех случаев для глаголов *leggen* и *zetten*). Исключения бывают двух типов: это ситуации, когда люди не могут самостоятельно контролировать свое положение (маленькие дети, больные и т. п.), а также ситуации, когда люди предстают пассивными жертвами манипуляции (например, при изгнании из дома или страны).

Статья Д. Шёнефельд «От концептуализации к лингвистической экспрессии: источник языковых различий» также посвящена анализу английских, немецких и русских позиционных глаголов *сидеть, стоять, лежать* с учетом корпусных данных. Целью исследования является описание механизмов влияния когнитивных образов-схем на интерпретацию позиционного сценария. Анализ семантики позиционных глаголов демонстрирует, что их использование с неодушевленными существительными свидетельствует о расширении значений глаголов; при этом механизмы расширения не совпадают в исследуемых языках. Так, английской фразе *salt and pepper sit on the table* в немецком и в русском языках соответствует фраза с глаголом «стоять», а не «сидеть»: нем. *Salz und Pfeffer stehen auf dem Tisch*, русск. *соль и перец стоят на столе*.

По мнению автора, при буквальном использовании глаголов местоположения выбор того или иного глагола зависит от того, какая образ-схема в сознании говорящего определяется как наиболее значимая; исследуемые языки различаются в том, что в них выбор глагола местоположения в идентичных ситуациях не совпадает; в разных языках в похожих ситуациях пространственные отношения между предметом и его положением трактуются неодинаково, что приводит к несовпадению в использовании предлогов места; в некоторых случаях в сопоставимых ситуациях в одном языке используется глагол местоположения, в других – совершенно иной глагол.

СПИСОК ЛИТЕРАТУРЫ

- Баранов 2001 – А.Н. Баранов. Введение в прикладную лингвистику. М., 2001.
 Atkins 1987 – Beryl T. Sue Atkins. Semantic ID tags: corpus evidence for dictionary senses // Proceedings of the Third annual conference of the UW Centre for the New Oxford English dictionary. 1987.
 Burt 1971 – M.K. Burt. From deep to surface structure. New York, 1971.
 Croft, Cruse 2004 – W. Croft, D.A. Cruse. Cognitive linguistics. Cambridge, 2004.
 Dowty 1979 – D.R. Dowty. Word meaning and Montague grammar : the semantics of verbs and times in generative semantics and in Montague's PTQ. Dordrecht, 1979.
 Erteschik-Shir 1979 – N. Erteschik-Shir. Discourse constraints on dative movement // T. Givón (ed.). Discourse and syntax: syntax and semantics 12. New York; San Diego; London:, 1979.

- Fillmore 1988 – *Ch.J. Fillmore*. The mechanisms of Construction Grammar // Sh. Axmaker, A. Jaisser, H. Singmaster (eds.). Proceedings of the Fourteenth annual meeting of the Berkeley linguistics society. University of California, Berkeley, 1988.
- Fillmore 1988 – *Ch.J. Fillmore*. The mechanisms of Construction grammar // BLS. 1988. 14.
- Fillmore, Kay 1995 – *Ch.J. Fillmore, P. Kay*. Construction grammar. Unpublished manuscript. University of California at Berkeley, 1995.
- Fried, Boas (eds.) 2005 – *M. Fried, H.C. Boas* (eds.). Grammatical constructions: Back to the roots. Amsterdam; Philadelphia, 2005.
- Goldberg 1995 – *A.E. Goldberg*. Constructions: A construction grammar approach to argument structure. Chicago, 1995.
- Goldberg 2006 – *A.E. Goldberg*. Constructions at work: the nature of generalization in language. Oxford, 2006.
- Gries 2001 – *S.Th. Gries*. A multifactorial analysis of syntactic variation: particle movement revisited // Journal of quantitative linguistics. 2001. 8(1).
- Gries, Stefanowitsch 2004 – *S.Th. Gries, A. Stefanowitsch*. Co-varying collexemes in the into-causative // M. Achard, S. Kemmer (eds.). Language, culture, and mind. Stanford (CA), 2004.
- Gries, Stefanowitsch 2007 – *S.Th. Gries, A. Stefanowitsch* (eds.). Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis. Berlin, 2007.
- Gropen et al. 1989 – *J. Gropen, S. Pinker, M. Hollander, R. Goldberg, R. Wilson*. The learnability and acquisition of the dative alternation in English // Language. 1989. 65(2).
- Hampe, Schoenefeld 2003 – *B. Hampe, D. Schoenefeld*. Creative syntax. Iconic principles within the symbolic // W.G. Mueller, O. Fischer (eds.). From sign to signing. Proceeding of the 3d symposium on iconicity in language and literature. Amsterdam; Philadelphia, 2003.
- Hanks 1996 – *P. Hanks*. Contextual dependancy and lexical sets // International journal of corpus linguistics. 1996. 1(1).
- Jespersen 1949 – *O. Jespersen*. A modern English grammar on historical principles. V. 7: Syntax. Copenhagen, 1949.
- Kennedy 1991 – *G. Kennedy*. Between and through: The company they keep and the functions they serve // K. Aijmer, B. Altenberg (eds.). English corpus linguistics. London, 1991.
- Kishner, Gibbs 1996 – *J.M. Kishner, R.W. Jr. Gibbs*. How just gets its meanings: Polysemy and context in psychological semantics // Language and speech. 1996. 39(1).
- Lakoff 1987 – *G. Lakoff*. Women, fire, and dangerous things. What categories reveal about the mind. Chicago; London, 1987.
- Lakoff 1990 – *G. Lakoff*. The invariance hypothesis: Is abstract reason based in image schemas? // Cognitive linguistics. 1990. 1(1).
- Lakoff 1991 – *G. Lakoff*. Cognitive linguistics versus generative linguistics: How commitments influence results // Language and communication. 1991. 11(1/2).
- McEnery, Wilson 2004 – *T. McEnery, A. Wilson*. Corpus linguistics. 2nd ed. Edinburgh, 2004.
- Oh 2000 – *S.-Y. Oh*. *Actually and in fact* in American English: A data-based analysis // English language and linguistics. 2000. 4(2).
- Pinker 1989 – *S. Pinker*. Learnability and cognition: The acquisition of argument structure. Cambridge (MA), 1989.
- Stefanowitsch, Gries 2003 – *A. Stefanowitsch, S.Th. Gries*. Collostructions: investigating the interaction between words and constructions // International journal of corpus linguistics. 2003. 8(2).
- Stefanowitsch, Gries 2005 – *A. Stefanowitsch, S.Th. Gries*. Covarying collexemes // Corpus linguistics and linguistic theory. 2005. 1(1).
- Thompson 1990 – *S.A. Thompson*. Information flow and dative shift in English discourse // J.A. Edmondson, C. Feagin, P. Mühlhäusler (eds.). Development and diversity: Language variation across time and space. Dallas, 1990.
- Thompson, Koide 1987 – *S.A. Thompson, Y. Koide*. Iconicity and 'indirect objects' in English // Journal of pragmatics. 1987. 11(3).
- Tomasello 2003 – *M. Tomasello*. Constructing a language: a usage based theory of language acquisition. Cambridge, 2003.
- Tyler, Evans 2001 – *A. Tyler, V. Evans*. Reconsidering prepositional polysemy networks: The case of over // Language. 2001. 77(4).
- Vendler 1967 – *Z. Vendler*. Verbs and times. Linguistics in philosophy. New York, 1967.
- Wierzbicka 1998 – *A. Wierzbicka*. The semantics of English causative constructions in a universal-typological perspective // M. Tomasello (ed.). The new psychology of language. Hillsdale (NJ), 1998.