

ВАДИМ АЛЕКСЕЕВИЧ БЕЛОВ

кандидат филологических наук, доцент кафедры германской филологии и межкультурных коммуникаций Гуманитарного института
Череповецкий государственный университет (Череповец, Российская Федерация)
belov.vadim.a@gmail.com

ВОЗМОЖНОСТЬ СИНОНИМИЧЕСКИХ ЗАМЕН В УСТОЙЧИВЫХ СОЧЕТАНИЯХ (экспериментальное и корпусное исследование)

В задачи работы входит анализ степени вариативности устойчивых сочетаний и фразеологизмов. Актуальность исследования связана с тем, что в языке существует большое количество сочетаний с разной степенью устойчивости. В лингвистике не разработаны критерии установления степени устойчивости и вариативности сочетания, что определяет проблему нашего исследования. Новизна заключается в том, что в работе используется мера ассоциативности (Mutual Information) – статистический показатель, который демонстрирует количество совместных употреблений единиц с учетом их частотности. Основным методом исследования является проведенный эксперимент по выбору синонима в высказывании. В нем приняло участие 216 человек, испытуемые выбирали синоним, наиболее подходящий контексту. Для эксперимента были отобраны 56 сочетаний с разной степенью устойчивости. Далее с этими сочетаниями были выбраны высказывания из Национального корпуса русского языка. Результаты эксперимента показали, что мера ассоциативности способна предсказывать выбор синонимов в устойчивом сочетании только с учетом других факторов. В исследовании выделены несколько факторов, которые определяют вариативность устойчивых сочетаний и фразеологизмов и которые связаны с семантическими и когнитивными аспектами речевой деятельности: неаддитивность семантики сочетания, категориальное значение (принадлежность лексико-грамматическим разрядам) входящих в сочетание слов, общий смысл высказывания, распространенность сочетаний в современной речи.

Ключевые слова: устойчивое сочетание, фразеологизм, ассоциативности, синонимия, корпусная лингвистика, психолингвистика, идиоматичность

Для цитирования: Белов В. А. Возможность синонимических замен в устойчивых сочетаниях (экспериментальное и корпусное исследование) // Ученые записки Петрозаводского государственного университета. 2020. Т. 42. № 1. С. 48–55. DOI: 10.15393/uchz.art.2020.432

ВВЕДЕНИЕ

В языке существует большое количество несвободных (устойчивых) сочетаний, единицы которых оказываются синтагматически связанны. Свойства устойчивых сочетаний активно обсуждаются в современной лингвистике [6], [7], [10], [12], [14]. Они хранятся в лексиконе и воссоздаются в речи в готовом виде [10], [14], а также имеют ограничения в построении [6].

В отечественной системоцентрической лингвистике для определения устойчивых выражений чаще всего применялся основанный на языковой интуиции анализ. Так, наибольшую известность получила типология В. В. Виноградова. Он выделяет фразеологические сращения (значение которых не может быть выведено из значений компонентов; *попасть впросак*), фразеологические единства (значение которых образно связано с семантикой компонентов; *плыть по течению*), фразеологические сочетания (значение которых мотивировано, но один из компонентов имеет фразеологически связанное значение; *попутить взор*) [2: 140–161]. В таких случаях критерием от-

несения конкретного сочетания к тому или иному типу является интуиция исследователя, поэтому возможны противоречивые результаты.

С развитием информационных технологий открываются широкие перспективы статистической обработки языкового материала. В современной лингвистике разработаны специальные количественные меры для определения степени устойчивости сочетания¹ [13], которые регистрируют случаи совместного употребления единиц в корпусе. Очевидным преимуществом обозначенного подхода может стать объективность результатов и доступ к общирному текстовому материалу [1], [3], [9], [15].

В нашей работе анализируется возможность применения меры ассоциативности для предсказания синонимических замен в устойчивом сочетании. Как представляется, в речи устойчивые выражения ограничены в синонимических заменах, что является важнейшим показателем устойчивости выражения. Например, во фразеологизме *не в коня корм* невозможна синонимическая замена *не в коня пища*.

В научной литературе отмечается вариативность устойчивых сочетаний и фразеологизмы [4], [8]. Например, в словаре А. В. Жукова встречаются фразеологизмы, которые допускают семантические замены: *на языке (устах)*, *на свою голову (шею)*, *под конём (седлом)*, *моя хата (изба) с краю* и другие². В нашей работе изучается зависимость вариативности устойчивого сочетания от статистических показателей степени устойчивости.

МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

Исследование основывается на результатах психолингвистического эксперимента по выбору синонима в высказывании. Для разработки эксперимента использовались данные Национального корпуса русского языка (НКРЯ) и сведения словарей синонимов русского языка. Проведение исследования включало следующие этапы: во-первых, отбор высказываний из НКРЯ с последующим определением меры ассоциативности сочетаний; во-вторых, подбор синонимов для выбранных сочетаний с помощью словарей синонимов под редакцией А. П. Евгеньевой³, Л. Г. Бабенко⁴; в-третьих, проведение эксперимента по выбору синонима в высказывании. В дальнейшем результаты эксперимента были обработаны и обобщены. При обработке результатов оценивалась корреляция (зависимость) меры ассоциативности и возможности синонимических замен в устойчивом сочетании. Для эксперимента были выбраны сочетания с разными значениями меры ассоциативности, часть отобранных сочетаний являлась фразеологизмами. Как отмечалось ранее в работах по корпусной лингвистике, устойчивыми оказываются сочетания, имеющие значение MI 3 и выше [5: 138]. Однако в нашем исследовании будет показано, что часть фразеологизмов, зафиксированных в словарях, имеет значение MI ниже 3.

Эксперимент был разделен на два этапа с использованием разных экспериментальных таблиц. В первом этапе приняло участие 144 испытуемых, им было предъявлено 43 высказывания с разной степенью устойчивости. Во втором этапе участвовало 72 человека; были специально выбраны 13 фразеологизмы из фразеологического словаря русского языка под редакцией А. И. Молоткова⁵. В двух этапах эксперимента в качестве испытуемых выступили в основном студенты Череповецкого государственного университета в возрасте от 18 до 47 лет. Испытуемым предлагались предложения, необходимо было выбрать подходящий для контекста синоним. В примерах исходная форма высказывания была скрыта от участников эксперимента. В случаях, когда пропущенные синонимы отличались по роду / числу, согласуемые слова варьировались по этим грамматическим признакам. В экс-

периментальной таблице высказывания были представлены следующим образом:

Вода сильнее всего, она точит камень и тушит (пламя, огонь, пыл).

Только не говорите мне, что если алкоголизм – это не дурной образ жизни, а заболевание, то он, как всякая/ий/ое прочая/ий/ое (хворь, болезнь, недуг, заболевание), дело приватное.

Экспериментальные задания для первого этапа были оформлены в письменном виде, во втором этапе задания предъявлялись в электронном виде с помощью программы «Google Формы» (<https://docs.google.com/forms>).

Как уже отмечалось, в работе используется показатель *мера ассоциативности*, или *MI (mutual information)*, который рассчитывается на основе данных о частоте совместной встречаемости и частотности входящих в сочетание единиц (см. более подробно см.: [11: 23]). MI рассчитывается с помощью данных НКРЯ. Формула данного показателя представлена следующим образом:

$$MI(x, y) = \log_2 \frac{P(x, y)}{P(x) * P(y)},$$

где x, y – слова, входящие в состав коллокации;
 $P(x, y)$ – количество совместных употреблений слов;
 $P(x)$ – частотность первого слова;
 $P(y)$ – частотность второго слова;
N – объем корпуса, из которого взята частотность слов.

Так, рассчитаем значение показателя для сочетания *желудочная хворь*. Частотность лексемы *желудочный* в основном корпусе НКРЯ составляет 793 употребления, лексемы *хворь* – 325 употреблений, количество совместных употреблений лексем в контексте – 4. Объем основного корпуса составляет 230 млн. Подставляем приведенные значения в формулу и получаем показатель MI – 11,8.

Среди разработанных статистических мер данный показатель демонстрирует наилучшие результаты: «Мера MI, возможно, дает наилучшие усредненные результаты. Она позволяет выделить устойчивые фразеологизированные словосочетания» [5: 141]. В целом все подобные показатели оценивают совместные употребления лексем с учетом их частотности⁶. В данном исследовании мера ассоциативности рассчитывалась не только для исходных высказываний, но и для выражений, полученных после синонимической замены. Например, исходным выражением является сочетание *тушил пламя* (MI – 9), после синонимических замен получаются выражения *тушил огонь* (MI – 9), *тушил пыл* (MI – 0; так как в НКРЯ не встречается таких употреблений).

РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА

Анализ результатов показал, что не наблюдается непосредственной зависимости возможности замены элемента в устойчивом сочетании от меры ассоциативности. В первом этапе эксперимента, где присутствовали

устойчивые и свободные сочетания, регистрируется достаточно слабая корреляция (0,36⁷) меры ассоциативности и результатов экспериментов. Однако для второго этапа эксперимента (с использованием фразеологизмов) общая корреляция значительно выше – 0,69. Более того, зафиксированы примеры, когда корреляция равна 1: например, для фразеологизмов *идти в ногу со временем, валять дурака, кровная обида*.

Таблица 1. Результаты второго этапа эксперимента по выбору синонима в высказывании
Table 1. The results of the second stage of the experiment for the selection of a synonym in the utterance

№	Высказывание	Синоним	Ответы	Процент	MI	Корреляция
1	И вот обычным и ровным голосом он объясняет старшему сыну, что начиная с завтра старший будет заниматься с известным математиком Стрепетовым; если же сын не будет валять дурака и окажется не тупицей, то, поступив в университет, он, возможно, станет на долгие даже годы учеником этого блестящего ученого	дурака	65	94	10,9	1
		глупца	2	3	0	
		болвана	2	3	0	
2	– И меня в лауреаты выдвигали, да вот рожей не вышел...	рожей	31	45	2,02	0,29
		лицом	21	30	0,31	
		рылом	16	23	4,62	
		ряхой	1	1	0	
3	Или если он соберется с духом и на какой-то срок просто выбросит вон из головы эти никчемные гадания: будет – не будет?..	головы	63	91	4,45	0,59
		черепа	4	6	4,01	
		тыковки	2	3	0	
4	Желающих нагреть руки на операциях с землей немало	руки	55	80	2,79	0,99
		кисть	9	13	0	
		пясть	5	7	0	
5	Оксана принялась плакать, потому что перед чемпионатом большую часть их программы надо было еще доводить до ума	ума	60	87	5,04	0,97
		разума	7	10	1,6	
		интеллекта	2	3	0	
6	Он может отведать запретный (зародыш, плод, эмбрион) хотя бы в целях «научного» эксперимента	плод	68	99	10,7	0,93
		эмбрион	1	1	0	
		зародыш	0	0	0	
7	Травить насекомых-вредителей обычными ядами (пестицидами) – гиблое дело как в прямом, так и в переносном смысле	дело	56	81	5,51	0,94
		труд	9	13	2,37	
		поступок	4	6	0	
8	Стараясь идти в (ногу, лапу) со временем , в последние годы Пхёнъян, как предполагают наблюдатели, стал переходить на евро	ногу	66	96	1,97	1
		лапу	3	4	-0,1	
9	А ты все-таки начальник цеха – почти что вольная птица	птица	63	91	6,62	0,36
		птаха	3	4	7,93	
		пичуга	3	4	0	
10	Понятно, разворовали хранилище, затем устроили пожар – и концы в воду	воду	66	96	1,61	0,99
		водицу	3	4	0	
		влагу	0	0	0	
11	Недосказанность, неискренность воспринимались им как кровная обида	обида	59	86	9,06	1
		оскорбление	10	14	7,84	
12	Он тебе, наверное, осточертел своими любовными победами – вот ты его и втолтал в грязь , заставив полюбить свою жену	грязь	66	96	12,6	0,99
		месиво	2	3	0	
		хлябь	1	1	0	
13	Так и не получив признания, а значит, и денег, я повесил нос	нос	65	94	5,86	0,48
		сопелку	4	6	0	
		хобот	0	0	5,68	

Также высокие результаты отмечаются у следующих фразеологизмов:

- *идти в ногу со временем* – см. пример № 8;
- *концы в воду* – см. пример № 10;
- *втоптать в грязь* – см. пример № 12.

Наименее устойчивыми в эксперименте оказались такие выражения, как *рожей не вышел* (пример № 2).

Возможно, противоречивые результаты в нашем исследовании получены из-за несовершенства самих статистических показателей и источников данных. Так, в работе Е. В. Ягуновой и Л. М. Пивоваровой отмечается, что мера ассоциативности некорректно работает с низкочастотными единицами [10: 35]. Также наблюдается количественный дисбаланс по наполнению НКРЯ: наиболее объемный корпус – основной корпус – включает литературу с XVIII века, поэтому часть представленных текстов может быть неактуальна для современной языковой личности.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Применение статистических показателей для предсказания вариативности сочетания оправдывается частично. В большей степени мера ассоциативности, используемая в нашей работе, способна предсказывать выбор синонима для фразеологизмов. При этом нельзя определить значения MI, которое предсказывает выбор того или иного синонима. Так, для фразеологизма *концы в воду* рассчитан показатель MI – 1,91, но в 96 % случаев испытуемые восстанавливали его. С другой стороны, для сочетания *гиблое дело* представлено более высокое значение MI – 5,51, но результат эксперимента оказался противоположным: его выбирали в 81 % случаев (см. примеры № 10, 7 из табл. 1).

Далее мы выделим некоторые факторы, которые влияют на возможность синонимических замен. Они связаны с когнитивными и семантическими аспектами.

1. Неаддитивность семантики. Ключевым фактором для восстановления сочетания является не количество совместных употреблений в корпусе, а семантический фактор, связанный с идиоматичностью выражения. Во втором этапе эксперимента фразеологизмы хорошо узнаются испытуемыми, с чем связана высокая корреляция MI и результатов эксперимента.

Однако этот фактор нельзя формализовать с помощью статистических показателей. Сложно выделить значение MI, которое бы регистрировало идиоматичность содержания. Например, устойчивое сочетание *пища для ума* имеет сравнительно невысокое значение MI – 4. Но в высказывании (1) из первого этапа эксперимента оно хорошо восстанавливается испытуемыми (66 реакций, 90 %).

(1) *Но люблю и Дали, и Босха, потому что в их картинах есть пища для ума*⁸.

Таким образом, важным для ограничения вариативности устойчивого сочетания является то, что значение сочетания идиоматично, метафорически переосмысливает значение компонентов. Этот фактор не поддается статистической обработке на данном этапе развития корпусной лингвистики, поэтому в таких случаях приходится опираться на языковую интуицию носителя языка.

2. Категориальное значение входящих в сочетание слов. Результаты эксперимента показывают, что фразеологизмы, включающие абстрактные имена существительные, более вариативны по сравнению с фразеологизмами, содержащими конкретные существительные.

Например, в выражении *кровная обида* (которое толкуется как «Очень тяжелая, глубоко затрагивающая обида человека»⁹) возможна замена на синоним *оскорбление*, что показывают как примеры из корпуса, так и результаты проведенного эксперимента (см. пример № 11 из табл. 1).

В основном корпусе НКРЯ насчитывается 78 сочетаний *кровная обида* и 13 сочетаний *кровное оскорбление*; пример употребления см. в (2):

(2) *Это возмутило всю Фотьянку до глубины души, как самое кровное оскорбление, какого еще не бывало*¹⁰.

Относительно высокая вариативность отмечается для всех сочетаний с абстрактными существительными: *гиблое дело* (пример № 7 из табл. 1), *доводить до ума* (пример № 5 из табл. 1).

Высокая вариативность сочетаний с абстрактными существительными фиксируется не только для фразеологизмов, но и для сочетаний с разной степенью устойчивости. В первом этапе эксперимента были представлены исходные сочетания с абстрактными существительными (см. табл. 2). Как видно из табл. 2, у многих высказываний не отмечается доминирования одного ответа: здесь можно выделить примеры № 7, 9, где ответы испытуемых распределены достаточно равномерно. Однако в примерах № 1, 2, 6 доминирует один ответ, что объясняется влиянием устойчивой конструкции (которое выражается в высоком значении меры ассоциативности).

Отметим, что в эксперименте высокой вариативностью также отличалась часть конкретных существительных – наименований лиц; данная группа конкретных существительных связана с обозначением людей по различным признакам: *врач, друг, ребенок*. Для сочетаний с наименованиями лиц характерно то, что исходные сочетания не доминировали в ответах испытуемых (см. табл. 3). Наиболее равномерно ответы распределены в высказываниях № 2 и № 6.

Таблица 2. Результаты первого этапа эксперимента по выбору синонима в высказывании (сочетания с абстрактными ИС)**Table 2.** The results of the first stage of the experiment for choosing a synonym in the utterance (abstract nouns)

№	Высказывание	Синоним	Ответы	Процент	MI	Корреляция
1	Фортуна распорядилась так, что Фоменко даже не удалось выступить непосредственно в гонке	доля	0	0	5,04	0,69
		судьба	61	87	7,69	
		фортуна	9	13	7,40	
2	Но ведь Фортуна – продажная девка, и колесо ее всегда вращается	доля	1	2	0	0,96
		судьба	12	17	2,20	
		удел	0	0	0	
		фортуна	57	81	4,73	
3	Неужели и ему повезло наконец? Неужели фортуна к нему лицом повернулась? Но надо теперь быть вдвойне осторожным: жизнь внезапно может опять по башке ударить, если расслабишься...	доля	0	0	0	0,99
		судьба	22	31	2,65	
		фортуна	48	69	8,09	
4	И, чтобы не было скучно, они придумали себе занятие : готовили пельмени	дело	7	10	3,43	0,86
		занятие	55	79	4,77	
		работа	7	10	2,38	
		труд	1	1	1,36	
5	Постарайтесь не ругать ее, а найти интересное для нее занятие	дело	6	9	2,29	0,91
		занятие	43	61	4,52	
		работа	20	29	4,22	
		труд	1	1	1,96	
6	Ностальгия по 80-м – это не только тоска по юности, совпавшей с концом эпохи Брежнева и началом времени политических перемен	грусть	7	10	3,40	0,56
		печаль	1	1	0	
		скука	0	0	2,19	
		тоска	62	89	3,51	
7	Я надеялся, что скука не живет под чеченскими пулями, – напрасно	грусть	8	11	0,82	0,69
		печаль	21	30	3,48	
		скука	18	26	2,82	
		тоска	23	33	1,87	
8	Утром встала здоровая, скука и одиночество испарились, и все со мной в порядке – ни одышки, ничего, словом, дурного	грусть	13	18	0	0,6
		печаль	12	17	0	
		скука	0	0	5,65	
		тоска	47	65	0	
9	Стараясь рассеять печаль работой, я установил под микроскопом новый шлиф рудной породы из Сефидкана	грусть	15	21	7,39	0,33
		печаль	11	15	6,65	
		скука	22	31	7,18	
		тоска	23	32	5,75	

Во втором этапе эксперимента был представлен интересный пример (№ 2 из табл. 1), где реализуется фразеологизм *рожей не вышел*. Данный фразеологизм оказывается вариативным, несмотря на то что он включает конкретное существительное. Вариативность отмечается во фразеологических словарях: так, в словаре А. И. Молоткова представлено выражение *рожей не вышел*¹¹, в словаре А. И. Фёдорова – выражение *рылом не вышел*¹². Эта особенность, скорее всего, связана с эмоциональной окраской фразеологизма и входящих в него компонентов. Более того, в основном корпусе НКРЯ также наблюдается вариативность фразеологизма: 41 употребление сочетаний *рылом не вышел* и 18 употреблений сочетания *рожей не вышел*.

3. Общий смысл высказывания. Важное значение для выбора синонима в сочетании могут иметь семантические признаки, связанные с общим смыслом высказывания. Так, в высказывании (3) были предложены синонимы *корм*, *еда*, *пища*, *харчи*, но 92 % испытуемых (67 из 73 испытуемых) выбрали первое слово. Такой выбор основан на том, что слово *корм* прежде всего употребляется для обозначения пищи животных, а общий смысл высказывания касается животных (ср. толкование слова в МАС: «Корм – пища животных»¹³).

(3) Реклама – она не только на собачий **корм** спрос стимулирует, но и на самих собак¹⁴.

(4) Иногда, отдав **корм** птенцу, ласточка, слегка запрокинувшись, сидела несколько мгновений у края гнезда¹⁵.

Таблица 3. Результаты первого этапа эксперимента по выбору синонима в высказывании (сочетания с наименованиями лиц)

Table 3. The results of the first stage of the experiment on the selection of a synonym in the utterance (names of persons)

№	Высказывание	Синоним	Ответы	Процент	MI	Корреляция
1	Я знаю его – это Стас, большой шалун и любитель выпить	кроха	0	0	0	0,43
		малыш	1	1	0,54	
		ребенок	44	64	1,30	
		шалун	24	35	4,67	
2	Если какой-нибудь шалун и получал шлепок, то только в минуты веселья, и на это, конечно, нельзя было обижаться	малыш	14	20	3,59	-0,83
		ребенок	31	45	0,59	
		шалун	26	35	6,86	
3	Мой приятель внес мое имя в лотерею – проводил спортивный магазин, – и я выиграл	друг	37	53	4,71	0,31
		знакомый	8	11	4,58	
		приятель	18	26	6,39	
		товарищ	7	10	3,35	
4	Лекарь считается хорошим, если на десяток вылеченных им страдальцев приходится не больше одного-двух покойников	врач	33	46	-3,78	-0,99
		доктор	20	28	0,88	
		лекарь	16	22	1,29	
		медик	3	4	5,06	
5	И не может ли в этом случае помочь разными медикаментами уездный лекарь ?	врач	8	10	6,70	-0,07
		доктор	35	49	4,17	
		лекарь	27	38	10,02	
		медик	2	3	4,08	
6	Мой товарищ занимает важный пост в органах государственной власти, за рулем при этом всегда сидит сам, но и машина с мигалками ему нужна, потому что на дорогах пробки	друг	14	20	4,70	0,03
		приятель	32	46	6,39	
		товарищ	24	34	3,35	

Сходная закономерность отмечается в высказывании (4), где тоже речь идет о животных, испытуемые чаще всего также выбирают синоним *корм* (36 реакций, 49 %).

Однако подобные факторы достаточно сложно formalизовать, поэтому приходится опираться на языковую интуицию.

4. Распространенность сочетаний в современной речи. В некоторых случаях фразеологизмы с высоким показателем MI оказываются незнакомы носителям языка. Так, фразеологизм *не в коня корм* (MI – 7,5) в высказывании (5) испытуемые узнавали всего 49 раз из 72 ответов (68 %), что связано с тем, что фразеологизм не употребляется в современной речи. В эксперименте были получены реакции *еда* (14 реакций), *харчи* (9 реакций). В газетном корпусе НКРЯ представлено всего 22 употребления слов.

(5) – *Не в коня корм*, – сказал Штрум, и хотя не со всем было ясно, что значат его слова, они вызвали смех¹⁶.

(6) *И не может ли в этом случае помочь разными медикаментами **уездный лекарь**?*¹⁷

Еще меньше испытуемых восстановили устойчивое сочетание *уездный лекарь* в высказывании (6), которое имеет высокое значение MI (10), но оно незнакомо носителям языка. В эксперименте слово *лекарь* в высказывании выбрали только в 37 % случаев. Большая часть реакций

приходится на частотную доминанту *врач* с более широким значением; несмотря на то, что сочетание *уездный врач* имеет меньшее значение MI – 4. По всей видимости, это объясняется тем, что слово *лекарь* постепенно выходит из современной речи: об этом свидетельствует газетный корпус НКРЯ, где представлено всего 3 употребления.

ЗАКЛЮЧЕНИЕ

Результаты проведенного анализа показывают, что на современном этапе развития лингвистики применение статистических показателей для предсказания вариативности устойчивых сочетаний и фразеологизмов не в полной мере оправдано. На выбор синонима в устойчивом сочетании влияет большое количество факторов, связанных с семантическими и когнитивными аспектами речевой деятельности. Таким образом, обращение к языковой интуиции по-прежнему остается основным способом анализа устойчивости сочетаний и фразеологизмов. Важным в этом вопросе оказывается обращение к языковой компетенции носителя языка и языковой ситуации. Вместе с тем выделенные нами факторы позволяют упорядочить представление о степени устойчивости сочетания и вариативности фразеологизмов.

ПРИМЕЧАНИЯ

- ¹ Хохлова М. В. Исследование лексико-сintаксической сочетаемости в русском языке с помощью статистических методов (на базе корпусов текстов): Автoreф. дис. ... канд. филол. наук. СПб., 2011. 26 с.
- ² Жуков В. П., Сидоренко М. И., Шкляров В. Т. Словарь фразеологических синонимов русского языка: Около 730 синонимических рядов / Под ред. В. П. Жукова. М.: Русский язык, 1987. 448 с.
- ³ Словарь синонимов русского языка: В 2 т. / Под ред. А. П. Евгеньевой. Л.: Наука, 1970–1971.
- ⁴ Современный словарь русского языка. Синонимы: более 5000 синонимических рядов: ок. 30 000 слов-синонимов / Под общ. ред. проф. Л. Г. Бабенко. М.: АСТ: Астрель, 2011. 829 с.
- ⁵ Фразеологический словарь русского языка / Под ред. А. И. Молоткова. М.: Сов. энциклопедия, 1968. 543 с.
- ⁶ Также нередко используются показатели T-score, Salience.
- ⁷ Значение показателя, которое демонстрирует полную зависимость, равно 1.
- ⁸ Пример из НКРЯ: Александр Розенбаум, 1987–1998.
- ⁹ Бирик А. К., Мокиенко В. М., Степанова Л. И. Русская фразеология. Историко-этимологический словарь: Около 6000 фразеологизмов / Под ред. В. М. Мокиенко. М.: АСТ, 2005. С. 484.
- ¹⁰ Пример из НКРЯ: Д. Н. Мамин-Сибиряк, 1892.
- ¹¹ Фразеологический словарь русского языка / Под ред. А. И. Молоткова...
- ¹² Федоров А. И. Фразеологический словарь русского литературного языка. М.: Астрель: АСТ, 2008. 828 с.
- ¹³ Словарь русского языка: В 4 т. / Под ред. А. П. Евгеньевой. М.: Русский язык, 1985–1988. Т. II. С. 104.
- ¹⁴ Пример из НКРЯ: Владислав Быков, Ольга Деркач, 2000.
- ¹⁵ Пример из НКРЯ: Фазиль Искандер, 1962.
- ¹⁶ Пример из НКРЯ: Василий Гроссман, 1960.
- ¹⁷ Пример из НКРЯ: Николай Крышкук, журнал «Звезда», 2001.

СПИСОК ЛИТЕРАТУРЫ

1. Борискина О. О. Корпусное исследование языка: мода или необходимость? // Вестник Воронежского государственного университета. Сер.: «Лингвистика и межкультурная коммуникация». 2015. № 3. С. 24–27.
2. Виноградов В. В. Избранные труды. Лексикология и лексикография. М.: Наука, 1977. 310 с.
3. Голубкова Е. Е. Использование лингвистических корпусов при решении семантических проблем // Методы когнитивного анализа семантики слова: компьютерно-корпусный подход / Под ред. В. И. Заботкиной. М.: Языки славянских культур, 2014. С. 34–76.
4. Жуков В. П. Русская фразеология. М.: Высш. шк., 1986. 310 с.
5. Захаров В. П., Хохлова М. В. Анализ эффективности статистических методов выявления коллокаций в текстах на русском языке // Компьютерная лингвистика и интеллектуальные технологии: Труды Международной конференции «Диалог-2010», РГГУ. Вып. 9 (16). М.: РГГУ, 2010. С. 137–143.
6. Иорданская Л. Н., Мельчук И. А. Смысл и сочетаемость в словаре. М.: Языки славянской культуры, 2007. 672 с.
7. Кустова Г. И. Когнитивные модели в семантической деривации и система производных значений // Вопросы языкознания. 2004. № 4. С. 68–84.
8. Мокиенко В. М. Загадки русской фразеологии. М.: Высш. шк., 1990. 160 с.
9. Плунгян В. А. Корпус как инструмент и как идеология: о некоторых уроках современной корпусной лингвистики // Русский язык в научном освещении. 2008. № 2. С. 7–20.
10. Ягунова Е. В., Пивоварова Л. М. Природа коллокаций в русском языке. Опыт автоматического извлечения и классификации на материале новостных текстов // Научно-техническая информация. 2010. Сер. 2. № 6. С. 30–40.
11. Church K., Hanks P. Word association norms, mutual information, and lexicography // Computational Linguistics. 1996. No 16 (1). P. 22–29.
12. Cruse A. Lexical semantics. Cambridge: Cambridge University Press, 1986. 310 p.
13. Evert S. The Statistics of Word Cooccurrences: Word Pairs and Collocations: Ph. D. dissertation. Stuttgart: Universitat Stuttgart, 2004. 353 p.
14. Sinclair J. Corpus, Concordance, Collocation. Oxford: Oxford University Press, 1991. 179 p.
15. Teubert W., Cermakova A. Corpus linguistics: A Short introduction. London: Athenaeum Press, 2008. 154 p.

Поступила в редакцию 23.05.2019

Vadim A. Belov, PhD in Philology, Cherepovets State University
(Cherepovets, Russian Federation)
belov.vadim.a@gmail.com

POSSIBILITY OF SYNONYMIC REPLACEMENT IN COLLOCATIONS (EXPERIMENTAL AND CORPUS RESEARCH)

The purpose of the paper is to analyze the degree of variation of collocations and idioms (phraseological units). The relevance of the study is that there are many collocations with different degrees of stability, but there are no criteria of stability/variation of collocations and idioms. Originality is connected with using mutual information statistical indicator for predicting the variability (stability) of collocations. Mutual information is a special statistical tool for measuring the collocation strength. The research method is based on the psycholinguistic experiment and the Russian National Corpus data. Two hundred and sixteen subjects took part in the experiment. The subjects' task was to choose a synonym for a given context. For the experiment we selected 56 collocations with different

degrees of stability, and then selected entries with these collocations from the Russian National Corpus. The results demonstrated no direct correlation between mutual information and the results of experiment. Therefore, it was assumed that the selection of synonyms in an utterance is affected by certain cognitive and pragmatic features: the idiomatic meaning of an utterance, the categorical meanings of the words in a collocation, the general meaning of an utterance, and the frequency of a collocation in modern speech.

Keywords: collocation, idiom, mutual information, synonymy, corpus linguistics, psycholinguistics, idiomatic meaning

Cite this article as: Belov V. A. Possibility of synonymous replacement in collocations (experimental and corpus research). *Proceedings of Petrozavodsk State University*. 2020. Vol. 42. No 1. P. 48–55. DOI: 10.15393/uchz.art.2020.432

REFERENCES

1. Boriskina O. O. Corpus study of language: fashion or necessity? *Proceedings of Voronezh State University. Series: Linguistics and intercultural communication*. 2015. No 3. P. 24–27. (In Russ.).
2. Vinogradov V. V. Selected works. Lexicology and lexicography. Moscow, 1977. 310 p. (In Russ.).
3. Golubkova E. E. Using linguistic corpora for solving semantic problems. *Metody kognitivnogo analiza semantiki slova: kompyuterno-korpusnyy podkhod* (V. I. Zabotkina, Ed.). Moscow, 2014. P. 34–76. (In Russ.).
4. Zhukov V. P. Russian phraseology. Moscow, 1986. 310 p. (In Russ.).
5. Zakharov V. P., Khokhlova M. V. Analysis of the effectiveness of statistical methods for detecting collocations in Russian texts. *Proceedings of the International Conference “Dialogue 2010” on Computational Linguistics and Intelligent Technologies, Russian State University for the Humanities*. Issue 9 (16). Moscow, 2010. P. 137–143. (In Russ.).
6. Jordanskaya L. N., Mel'chuk I. A. Sense and compatibility in a dictionary. Moscow, 2007. 672 p. (In Russ.).
7. Kustova G. I. Cognitive models in semantic derivation and a system of derived values. *Voprosy yazykoznaniya*. 2004. No 4. P. 68–84. (In Russ.).
8. Mokiyenko V. M. Mysteries of Russian phraseology. Moscow, 1990. 160 p. (In Russ.).
9. Plungyan V. A. Corpus as a tool and an ideology: some lessons of modern corpus linguistics. *Russkiy jazyk v nauchnom osveshchenii*. 2008. No 2. P. 7–20. (In Russ.).
10. Yagunova E. V., Pivovarova L. M. The nature of collocations in the Russian language. The experience of automatic extraction and classification of the material of news texts. *Nauchno-tehnicheskaya informatsiya*. 2010. Series 2. No 6. P. 30–40. (In Russ.).
11. Church K., Hanks P. Word association norms, mutual information, and lexicography. *Computational Linguistics*. 1996. No 16 (1). P. 22–29.
12. Cruse A. Lexical semantics. Cambridge, 1986. 310 p.
13. Evert S. The Statistics of word cooccurrences: word pairs and collocations: Ph. D. dissertation. Stuttgart, 2004. 353 p.
14. Sinclair J. Corpus, concordance, collocation. Oxford, 1991. 179 p.
15. Teubert W., Cermakova A. Corpus linguistics: a short introduction. London, 2008. 154 p.

Received: 23 May, 2019