

А.С. Буб, З.И. Резанова

СТЕПЕНЬ УСТОЙЧИВОСТИ БИНОМИАЛОВ РУССКОГО ЯЗЫКА В СООТНОШЕНИИ С ПАРАМЕТРАМИ ЧАСТНОСТИ И ПРЕДСКАЗУЕМОСТИ ЭЛЕМЕНТОВ

Результаты были получены в рамках выполнения государственного задания Минобрнауки России, проект № 34.8609.2017/9.10.

Представлены результаты исследования степени устойчивости одного из видов коллокаций русского языка – биномиалов, определено соотношение степени устойчивости, частотности и предсказуемости элементов. По каждому из параметров выделены ядерная, околюдерная и периферийные группы. Выявлены пересечение ядерных групп по составу биномиалов и характер корреляционной связи групп по трем параметрам. Исследование проведено с применением корпусных, психолингвистических, статистических методов.

Ключевые слова: коллокации; биномиалы; устойчивые словосочетания; степень устойчивости; частотность; предсказуемость компонентов.

Термин «коллокация» активно используется в разных направлениях современной лингвистики. В широком смысле под коллокацией понимают «устойчивое словосочетание, близкое расположение компонентов которого обосновано регулярным характером взаимоожидания и является семантически обусловленным, другими словами, сочетание двух или более слов, имеющих тенденцию к совместной встречаемости» [1].

В зарубежной лингвистической традиции термин является «зонтичным», называет любые устойчивые сочетания, обладающие разной степенью семантической спаянности: идиомы, образные выражения, фразеологизмы, дискурсивные маркеры и т.д. (Ш. Балли, Н. Palmer, Т. Herbst, N. Nesselhauf, Т. Shippan и др.).

В российской лингвистической традиции коллокации изучаются в рамках фразеологии, где базовым является термин «фразеологизм», а коллокация выделяется в качестве одного из подвидов фразеологизма. Принимая во внимание высказывание В.Н. Телия об отнесении к «широкому объему» фразеологии всего того, «что воспроизводится в готовом виде, не являясь словом» [2. С. 58], можно оправдать отнесение коллокаций к сфере фразеологии.

Таким образом, в рамках двух традиций исследуется один материал с фокусировкой разных аспектов. Если в российском языкознании акцент ставится на единство семантики, при этом чем дальше от ядра находится единица (устойчивые сочетания с высокой фразеологичностью, дестимологизированные, типа *сесть в калошу, кричать во всю ивановскую*), тем меньше она изучается, то в зарубежной лингвистике основным понятием является коллокабельность, или потенциал к совместной встречаемости. Именно такой потенциал и кладется в основу определения понятия коллокация – это сочетание слов, которые обычно употребляются вместе [3]. Таким образом, основным свойством «коллокаций» является их устойчивость, т.е. регулярная воспроизводимость единиц в готовом виде в речи.

В нашем исследовании мы обращаемся к исследованию свойства устойчивости одного из видов колло-

каций русского языка – биномиалов. Данный термин был образован в англоязычной научной традиции изучения коллокаций как обозначение одного из типов. В настоящее время термин применяется к широкому кругу воспроизводимых единиц, соединенных сочинительной связью и обладающих набором различных дифференциальных признаков; ср., например, определение, данное И.Н. Молодцовой: «Биномиалы, или биномиальные образования (БО), представляют собой экспрессивные, сочинительные сочетания (окказиональные и узуальные), нередко рифмованные, компоненты которых чаще всего соединены такими союзами, как *and, or* и *but*: *part and parcel* (неотъемлемая часть), *near or far* (близко ли далеко), *slowly but surely* (медленно, но верно)» [4. С. 3]. Как представляется, однако, включение в разряд биномиалов окказиональных единиц противоречит их базовому признаку – воспроизводимости, а признак экспрессивности проявляется далеко не во всем составе единиц, относимых другими исследователями данному классу устойчивых, воспроизводимых сочетаний слов.

Как отмечает И.Н. Молодцова, лингвистические исследования, посвященные непосредственно биномиалам английского языка, немногочисленны. Продолжительное время являясь объектом исследований, биномиалы, как правило, не выделяются в отдельный класс, а используются в качестве практического материала изучения таких сочетаний слов, как повторения, клише и др., однако проблематика изучения данных многочисленных в английском языке единиц многоаспектна. В зависимости от практических и теоретических установок авторов в исследованиях биномиалов выделяются и описываются разные аспекты данного класса единиц: «синтаксический (например, рассмотрение сочинительных конструкций), семантический (анализ и классификация семантических связей компонентов посредством компонентного анализа), фонетический (изучение использования эвфонических средств) и ритмический (влияние ритмической организации на порядок следования компонентов), функциональный (изучение коммуникативных функций биномиалов)» [Там же. С. 4–5].

В то время как варианты научных решений при исследовании биномиалов английского языка еще не исчерпаны, биномиалы русского языка и вовсе оказываются неизученными, поэтому мы обратились к изучению статуса, специфики и места биномиалов в системе русского языка.

На предшествующих этапах исследования при помощи инструментов корпусной лингвистики был выявлен состав биномиалов русского языка (128 единиц), определены ядерные (75) и периферийные (53) единицы данного класса; все включенные в анализ единицы были проранжированы по объективной частотности, определены их субъективная частотность и корреляция двух показателей. Выборку составили биномиалы, отобранные из Национального корпуса русского языка (НКРЯ), корпуса письменных и устных текстов, охватывающих период от середины XVIII до начала XXI в. Объем корпуса составлял на момент обращения 209 198 275 словоупотреблений [5]. Процесс составления выборки описан в статье [6].

Задачами данного этапа являются ранжирование биномиалов по степени устойчивости и установление параметров, соотносимых с данным важнейшим свойством исследуемого класса коллокаций русского языка: частотностью целостной единицы и предсказуемостью второго элемента.

Важным для характеристики биномиальных образований представляется тот факт, что, несмотря на отсутствие подчинительной связи, они характеризуются определенной степенью устойчивости компонентов, или, в терминологии С. Моллин, степенью фиксированности / нефиксированности (далее – (не)фиксированности) порядка слов в биномиале [7].

На ранних этапах исследований ученые фокусировались лишь на так называемых «фризах» (транслитерация английского слова *freezes*, что в переводе означает «замороженные»), т.е. непереверчиваемых единицах, элементы которых могут находиться только в одном фиксированном порядке. В основном лингвистов интересовало, почему предпочтителен только один фиксированный вариант следования слов в биномиале, а не другой. Активно изучались факторы, влияющие на порядок слов в подобных единицах, например рифма, звуковая структура и т.д. [8–11].

Однако в 1959 г. Я. Малкиел в работе *Studies in irreversible binomials* указал, что относительно фиксированности порядка конститuentов все биномиалы английского языка можно расположить на континууме. С одного края континуума будут располагаться единицы с фиксированным порядком составляющих, на другом – единицы со свободным порядком элементов. Между двумя крайними точками расположатся биномиалы с большей или меньшей степенью фиксированности порядка слов, т.е. случаи, когда предпочтение отдается одному или другому порядку следования элементов. Биномиалы, слова в которых можно поменять местами, в терминологии Я. Малкиела получают название *reversible*, или «переворачиваемые» [8]. Таким образом, была сформирована традиция выделения двух видов биномиалов, фикси-

рованных и свободных. Последние, в свою очередь, делятся на абсолютно свободные (порядок слов может свободно меняться) и такие, которые обладают предпочтительным порядком слов. Сам континуум, по мнению Малкиела, является результатом диахронического процесса: свободный биномиал с течением времени начинает более тяготеть к определенному порядку конститuentов и постепенно становится фиксированным.

В своей работе мы следуем данной точке зрения и для определения степени устойчивости биномиалов русского языка применяем методологию, апробированную зарубежными лингвистами при изучении биномиалов английского языка.

На первом этапе определяем степень устойчивости биномиалов в выборке из 128 сочетаний русского языка.

Степень устойчивости определяется на основе выявления коэффициента (не)фиксированности порядка слов в биномиале. Современная корпусная лингвистика, методики поиска по корпусу позволяют эмпирически установить предпочтительный порядок для того или иного биномиала и рассчитать так называемый коэффициент (не)фиксированности (*ir*)reversibility score, в терминологии С. Моллин.

Для подсчета коэффициента (не)фиксированности было предложено несколько методик. Например, П. Рейх (*Reih*, 2005) предлагает выявлять соотношение более частотной версии биномиала и менее частотной. Однако данный метод не может применяться в том случае, если биномиал является абсолютно устойчивым, т.е. частотность одной из его версий равна нулю. С. Моллин, в свою очередь, предлагает следующую формулу для расчета коэффициента устойчивости:

$$(ir)reversibility\ score = \frac{freq}{freq + revfreq} \times 100,$$

где *freq* – частотность более частотного биномиала (т.е. порядок его элементов предпочтителен), а *revfreq* – частотность «перевернутого» биномиала, т.е. биномиала с обратным, менее предпочтительным порядком слов.

Представим пример применения вышеуказанной формулы для биномиала *муж и жена*. Количество вхождений для данного биномиала и его обратного варианта *жена и муж* – 815 и 23 соответственно. Сделав расчеты по исследовательской формуле «количество словоупотреблений биномиала “х” делится на общее количество словоупотреблений и умножается на миллион», получаем показатели частотности 2,87 *ir*m и 0,08 *ir*m для первой и второй единицы соответственно. Подставив в формулу показатели частотности, получаем коэффициент, равный 97.

Двум крайним точкам континуума соответствуют коэффициенты 100 и 50 соответственно. Таким образом, можем сделать вывод о том, что порядок слов в биномиале *муж и жена* является скорее фиксированным, нежели свободным. А, например, словосочетания *его и ее* и *ее и его* встречаются в НКРЯ примерно одинаковое количество раз (84 и 52 соответ-

ственно), поэтому порядок слов данного биномиала можно интерпретировать как относительно свободный, поскольку коэффициент (не)фиксированности равен 62.

Таким образом, из ранее составленного списка биномиалов (128 единиц) 23 коллокации интерпретируются как единицы с фиксированным порядком элементов, их коэффициент устойчивости равен 100: *вдоль и поперек, туда и обратно, король и королева, первое и второе, разделяй и властвуй, судить и рядить, мама и тетя, мартышка и очки, пробы и ошибки, зрение и осязание, понедельник и пятница, детский сад и школа, школа и университет, трамвай и троллейбус, воздух и птицы, плавать и бегать, звонить и разговаривать, сахар и специи, сапоги и кроссовки, удивление и грусть, петух и хвост, цирк и зоопарк, князь и крестьянин.*

7 единиц имеют строго свободный порядок и коэффициент устойчивости 50: *яблоки и апельсины, пальто и шарф, свет и сумерки, кожа и шерсть, песок и вода, дверь и потолок, заяц и медведь.*

63 единицы размещаются в континууме ближе к краю с фиксированным порядком (коэффициент (не)фиксированности от 75 до 99): *туда и сюда (99), муж и жена (97), мужчины и женщины (88), до и после (99), вверх и вниз (86), добро и зло (91), читать и писать (90), война и мир (97), жизнь и смерть (92), плоть и кровь (90), здесь и сейчас (93), целиком и полностью (96), мальчики и девочки (80), жених и невеста (94), вкривь и вкось (94), сегодня и завтра (99), глаза и уши (86), дети и внуки (97), коротко и ясно (94), вопросы и ответы (97), теория и практика (96), наука и техника (94), прошлое и будущее (88), север и юг (78), верх и низ (86), гром и молния (81), лето и осень (97), хлеба и зрелищ (98), хлеб и соль (89), солнце и луна (76), весна и лето (98), взлеты и падения (95), соль и перец (88), флора и фауна (82), лук и стрелы (96), мир и спокойствие (86), имя и адрес (81), рога и копыта (97), спрос и предложение (85), правда и ложь (78), солнце и звезды (85), закон и порядок (78), принц и принцесса (90), утро и день (85), положительный и отрицательный (94), акции и облигации (89), боль и страдания (82), совет и любовь (92), кожа и кости (77), семья и друзья (89), крестики и нолики (86), радость и грусть (75), еда и напитки (80), лево и право (76), письмо и конверт (80), утро и ночь (95), обоняние и слух (75), воспитание и природа (80), гайки и болты (94), радость и испуг (83), признаки и симптомы (75), зубы и десны (90), честь и хвала (75).*

Оставшиеся 35 единиц – ближе к «свободному» краю (коэффициент (не)фиксированности от 51 до 74): *зимой и летом (71), небо и земля (68), тут и там (56), золото и серебро (72), восток и запад (60), мамы и папы (53), чай и кофе (72), его и ее (62), черное и белое (59), большой и малый (73), свет и тьма (74), дождь и снег (68), луна и звезды (68), близко и далеко (66), белый и красный (52), надежды и мечты (70), кошка и собака (62), выше и больше (54), стол и кро-*

вать (56), простить и забыть (71), голубой и розовый (67), осень и весна (51), паспорт и билет (64), дверь и окно (67), охи и вздохи (67), скипетр и держава (55), шашки и шахматы (52), цифры и слова (64), издержки и выгоды (67), нитка и иголка (57), кофе и сок (57), горячо и холодно (67), закат и рассвет (67), математика и музыка (60), джинсы и рубашка (67).

Рассчитав коэффициент устойчивости и распределив биномиалы на континууме, соотнесем объективный коэффициент устойчивости с субъективным, а также определим корреляцию данного показателя с другими характеристиками биномиалов, исследуемыми на предыдущих этапах.

Исследование (не)фиксированности порядка слов биномиалов имеет непосредственное отношение к вопросу о воспроизводимости и компонуемости этих единиц в ментальном лексиконе носителя языка. Далеко не всегда данные корпуса могут совпадать с представлениями носителей языка о той или иной единице. Чтобы валидизировать данные корпуса и корпусную методологию, было решено прибегнуть к психолингвистической методологии, методу «субъективных оценок».

Для этого была реплицирована процедура, применяемая С. Моллин в ее исследовании. Разработано четыре анкеты, каждая из которых содержала 32 биномиала с прямым и обратным порядком слов, которые испытуемые должны были ранжировать в соответствии с членениями семизначной шкалы Лайкерта (рис. 1).

Всего в опросе приняли участие 40 носителей русского языка, 10 мужчин и 30 женщин. Средний возраст респондентов 22 года. Участникам опроса необходимо было отдать предпочтение одному из представленных в опросном листе вариантов биномиала в соответствии со значениями шкалы (1 и 7 – только этот вариант корректен, 2 и 4 – скорее выберу этот вариант, 3 и 5 – выберу этот вариант, но другой тоже возможен, 4 – оба варианта возможны). Опрос проводился при помощи бумажных анкет, информант мог заполнить несколько анкет.

Всего было получено по 20 реакций на каждый биномиал, т.е. 2 560 реакций. Все реакции учитывались в дальнейшем анализе.

Далее, следуя процедуре С. Моллин, мы сравнили данные корпуса с данными опросников. Для этого все ответы для каждого биномиала были занесены в таблицу excel, где затем при помощи встроенной функции «=срзнач» было определено среднее значение для каждого биномиала. Так, например, среднее значение биномиала *звезды и луна* равен 5 на шкале Лайкерта. Затем каждому делению шкалы был присвоен коэффициент (не)фиксированности: деления 1 и 7 равнялись коэффициенту 100, 4 – 50, деления 3 и 5 соответствовали коэффициенту от 51 до 74, а 2 и 6 приравнялись показателю устойчивости от 75 до 99.

На рис. 2 представлено соотношение объективных показателей и субъективных оценок.

	Только этот	Скорее этот	Этот, но другой возможен	Оба возможны	Этот, но другой возможен	Скорее этот	Только этот	
понедельник и пятница	1	2	3	4	5	6	7	пятница и понедельник
вкось и вкривь	1	2	3	4	5	6	7	вкривь и вкось
кожа и кости	1	2	3	4	5	6	7	кости и кожа
вода и песок	1	2	3	4	5	6	7	песок и вода
практика и теория	1	2	3	4	5	6	7	теория и практика
акции и облигации	1	2	3	4	5	6	7	облигации и акции
петух и хвост	1	2	3	4	5	6	7	хвост и петух
мамы и папы	1	2	3	4	5	6	7	папы и мамы
второе и первое	1	2	3	4	5	6	7	первое и второе
трамвай и троллейбус	1	2	3	4	5	6	7	троллейбус и трамвай
ночь и утро	1	2	3	4	5	6	7	утро и ночь
будущее и прошлое	1	2	3	4	5	6	7	прошлое и будущее
лук и стрелы	1	2	3	4	5	6	7	стрелы и лук
математика и музыка	1	2	3	4	5	6	7	музыка и математика
соль и хлеб	1	2	3	4	5	6	7	хлеб и соль
ложь и правда	1	2	3	4	5	6	7	правда и ложь
мартышка и очки	1	2	3	4	5	6	7	очки и мартышка
князь и крестьянин	1	2	3	4	5	6	7	крестьянин и князь
ошибки и пробы	1	2	3	4	5	6	7	пробы и ошибки
добро и зло	1	2	3	4	5	6	7	зло и добро
голубой и розовый	1	2	3	4	5	6	7	розовый и голубой
выгоды и издержки	1	2	3	4	5	6	7	издержки и выгоды

Рис. 1. Пример анкеты для определения субъективной степени устойчивости биноминала

В анкете (см. рис. 2) X – коэффициент (не)фиксированности, приписываемый на основе статистических подсчетов по данным НКРЯ, S – вариант, который отметили большинство участников опроса, сочетание XS обозначает совпадение субъективных предпочтений и оценки биноминала на основе корпусных данных. В приведенном фрагменте таких биноминалов 5. В целом из 128 биноминалов объективные данные и субъективные оценки совпали у 38 единиц, составивших ядро выборки: *пальто и шарф, заяц и медведь, паспорт и билет, большой и малый, луна и звезды, север и юг, муж и жена, жизнь и смерть, цифры и слова, взлеты и падения, кожа и шерсть, черное и белое, гром и молния, закон и порядок, разделяй и властвуй, спрос и предложение, свет и тьма, стол и кровать, горячо и холодно, соль*

и перец, яблоки и апельсины, дверь и потолок, зимой и летом, кошка и собака, охи и вздохи, вкривь и вкось, кожа и кости, песок и вода, теория и практика, петух и хвост, прошлое и будущее, лук и стрелы, хлеб и соль, правда и ложь, мартышка и очки, добро и зло, надежды и мечты, скипетр и держава.

Однако оценки респондентов еще 52 коллокаций очень близки к коэффициенту объективных данных, т.е. находятся на соседних делениях шкалы, например субъективная оценка устойчивости биноминала *небо и земля* – 6 (коэффициент устойчивости от 75 до 99), а объективный показатель равен 68 (пятое деление шкалы). Данный класс составил околоядерную часть выборки. Таким образом, мы выделяем ядро, околоядерную часть и оставшуюся периферию (38 единиц) по данному параметру.

	Только этот	Скорее этот	Этот, но другой возможен	Оба возможны	Этот, но другой возможен	Скорее этот	Только этот	
воздух и птицы	X	2	3	S	5	6	7	птицы и воздух
пальто и шарф	I	2	3	XS	5	6	7	шарф и пальто
иголка и нитка	1	2	X	S	5	6	7	нитка и иголка
здесь и сейчас	S	X	3	4	5	6	7	сейчас и здесь
кофе и чай	1	2	3	4	X	S	7	чай и кофе
заяц и медведь	1	2	3	XS	5	6	7	медведь и заяц
признаки и симптомы	1	X	3	S	5	6	7	симптомы и признаки
билет и паспорт	1	2	3	4	XS	6	7	паспорт и билет
белый и красный	1	2	X	S	5	6	7	красный и белый
детский сад и школа	X	2	3	S	5	6	7	школа и детский сад
глаза и уши	1	X	S	4	5	6	7	уши и глаза
десны и зубы	1	2	3	4	S	X	7	зубы и десны
король и королева	X	S	3	4	5	6	7	королева и король
большой и малый	1	2	XS	4	5	6	7	малый и большой
дверь и окно	1	2	3	S	X	6	7	окно и дверь
звезды и луна	1	2	3	4	XS	6	7	луна и звезды

Рис. 2. Соотношение объективных и субъективных данных НКРЯ при определении порядка слов в биномиале

Рассмотрим другие характеристики биномиалов, соотносимые со степенью устойчивости.

Частотность биномиала, по мнению исследователей, анализирующих коллокации английского языка, является также характеристикой, соотносящейся со степенью устойчивости. Так, Я. Малкиел предполагает, что для того, чтобы «обрести» более высокую степень устойчивости, биномиал должен «преодолеть» определенный порог частотности [12].

В нашей выборке частотность биномиалов определяется по данным НКРЯ. Показатель частности ipm (item per million) определялся в соответствии с принятой в данной парадигме формулой: «количество словоупотреблений биномиала “х” делится на общее количество словоупотреблений и умножается на миллион» (подробнее см.: [6]). Анализ частотности ядерной части биномиалов, выделенной по признаку субъективной и объективной устойчивости, свидетельствует о значительном разбросе их частотности; по данным, НКРЯ, например, *пальто и шарф* – 0,003 ipm, *заяц и медведь* – 0,003 ipm, *кошка и собака* – 0,056 ipm, *надежды и мечты* – 0,077 ipm, *закон и порядок* – 0,102 ipm, *правда и ложь* – 0,116 ipm, *спрос и предложение* – 0,123 ipm, *лук и стрелы* – 0,172 ipm, *теория и практика* – 0,546 ipm, *вкривь и вкось* – 0,843 ipm, *зимой и летом* – 0,991 ipm, *жизнь и смерть* –

1,510 ipm, *добро и зло* – 1,689 ipm, *муж и жена* – 2,875 ipm. Как видим, ядро, определенное по признаку устойчивости, представлено как высокочастотными, так и низкочастотными биномиалами (частотность от 2, 875 ipm до 0, 003 ipm).

Статистический анализ показателей устойчивости и частотности ядра устойчивости выявил наличие зависимости двух показателей. При p-value = 9,861e–05 корреляция равна 0,5 894 538.

На основе анализа всей выборки биномиалов по параметру частности мы далее также выделили ядро, околядерную часть и периферию, уточнив наш предшествующий анализ. В ядро вошли единицы (16), частотность которых выражается целым числом (от 3,447 до 1,234 ipm), в околядерную часть вошли единицы (47), частотность которых выражается сотыми (от 0,991 до 0,102 ipm), а в периферию выборки были включены биномиалы (65) с частотностью от 0,095 до 0,003 ipm. При сравнении единиц ядра частотности и ядра устойчивости было обнаружено только 3 совпадения: *муж и жена*, *добро и зло*, *жизнь и смерть*: это высокочастотные единицы, характеризующиеся и высокой степенью устойчивости. Однако формула, применяемая при подсчете степени устойчивости, позволяет отнести к ядерной части биномиалы-гапаксы (*пальто и шарф*, *заяц и медведь*, *дверь и по-*

толок, петух и хвост), которые, обладая низкой частотностью ввиду единичной встречаемости в корпусе, являются строго устойчивыми, поскольку их коэффициент (не)фиксированности равен 100.

Сила коллокационной связи является также одной из значимых характеристик бинмиала. Бинмиал не только интерпретируется в научной традиции как «характерные, часто встречающиеся сочетания слов», но отмечается также, что взаимное появление таких сочетаний «основано на регулярном характере взаимного ожидания и обусловлено семантическими факторами» [1]. Регулярный характер взаимного ожидания подразумевает наличие некоей связи между словами, благодаря которой они складываются в одну коллокацию. Такая связь также может различаться по своей силе. На наш взгляд, сила коллокационной связи коррелирует с понятием предсказуемости, т.е. способности одного слова предсказывать появление другого. Высокая степень предсказуемости, в свою очередь, будет доказывать целостное хранение бинмиалов в ментальном лексиконе человека.

Для определения степени предсказуемости бинмиалов в нашем исследовании было разработано 8 анкет, каждая из которых содержала по 32 слова, после полнозначного слова следовал союз *и*, открывающий связь с потенциальным следующим компонентом, например *судить и...*

Четыре анкеты включали первые слова бинмиалов, например: *мужчины и ...*, *судить и...* и т.д., и еще четыре анкеты содержали вторые слова бинмиалов, например: *женщины и ...*, *рядить и ...*

В эксперименте приняли участие 163 респондента, 38 мужчин, 125 женщин в возрасте от 17 до 25 лет. В качестве задания им предлагалось написать, какое слово должно идти после союза. Всего было получено 6 526 реакций (некоторые респонденты заполняли несколько анкет). В последующем анализе учитывались все анкеты.

Ответы были разделены на два типа: 1) ответы, в которых в качестве второго компонента респонденты называли слово – второй компонент бинмиала; 2) ответы, в которых в качестве второго компонента были избраны слова, не входящие в бинмиал. Далее на основе общего количества ответов и ответов, полученных для каждого бинмиала, был определен коэффициент «правильных» ответов, т.е. таких, в которых в качестве ответа называлось второе слово бинмиала. Для этого общее количество ответов умножали на 100 и делили на количество «правильных ответов». Коэффициент рассчитывался как для анкет, в которых содержалось первое слово бинмиала, так и для анкет, которые включали в себя второе слово бинмиала. Чем больше количество правильных ответов, тем выше коэффициент (от 0 до 100).

Коэффициент 0 был выведен для 21 единицы из первого типа анкет и для 31 единицы из анкет второго типа. Для остальных единиц коэффициент «правильных» ответов в анкетах первого типа составил от 3 (например, *звонить и разговаривать* – из 33 ответов лишь 1 совпал с продолжением бинмиала) до 100 (например, *лево и право* – из 26 ответов совпали все). Коэффициент правильных ответов для анкет второго

типа составил от 5 (например, *сок и кофе* – из 22 ответов совпал 1) до 100 (*отрицательный и положительный* – из 22 ответов совпали все).

На данном этапе в дальнейшем анализе использовались лишь анкеты первого типа. Исходя из полученных результатов, по аналогии с показателями устойчивости и частотности были выделены ядро, околоядерная часть и периферия, для этого все бинмиалы также были размещены на континууме предсказуемости от 0 до 100. Ядро выборки по параметру «предсказуемость» составили 52 бинмиала коэффициентом предсказуемости от 100 до 57: *вопросы и ответы, до и после, крестики и нолики, лево и право, тут и там, война и мир, зимой и летом, флора и фауна, его и ее, правда и ложь, взлеты и падения, вдоль и поперек, гром и молния, разделяй и властвуй, добро и зло, вверх и вниз, жених и невеста, мальчики и девочки, мужчины и женщины, плоть и кровь, положительный и отрицательный, свет и тьма, целиком и полностью, муж и жена, север и юг, вкривь и вкось, близко и далеко, рога и копыта, черное и белое, восток и запад, нитка и иголка, мамы и папы, теория и практика, золото и серебро, мама и тетя, верх и низ, горячо и холодно, небо и земля, сегодня и завтра, детский сад и школа, шашки и шахматы, кожа и кости, первое и второе, спрос и предложение, кошка и собака, туда и сюда, жизнь и смерть, здесь и сейчас, чай и кофе, мартышка и очки, коротко и ясно, прошлое и будущее. Околоядерную часть образовали 55 единицы с коэффициентом от 50 до 3. Бинмиалы с нулевым коэффициентом предсказуемости (21 единица) были отнесены к периферийной части выборки: *воздух и птицы, заяц и медведь, признаки и симптомы, мир и спокойствие, воспитание и природа, имя и адрес, радость и испуг, кожа и шерсть, солнце и звезды, сахар и специи, честь и хвала, петух и хвост, математика и музыка, князь и крестьянин, удивление и грусть, простить и забыть, зрение и осязание, кофе и сок, утро и день, дети и внуки, дверь и потолок...* Одна единица имела коэффициент предсказуемости 100 – *вопросы и ответы*, показатель предсказуемости остальных бинмиалов варьировал от 3 до 97.*

Анализ совпадающих единиц ядерных частей двух выборок по коэффициентам устойчивости и предсказуемости выявил совпадение 19 единиц. Таким образом, мы получили гораздо более объемное пересечение ядерных частей выборок по данным параметрам, нежели в предыдущем сравнении.

Однако проведенный корреляционный анализ зависимости между степенью устойчивости бинмиала и степенью предсказуемости его элементов с использованием критерия хи-квадрат Пирсона показал отсутствие корреляционной зависимости коэффициента устойчивости и коэффициента предсказуемости ядерной части выборки по параметру устойчивости, $p\text{-value} = 0,0006634$, она составила 0,5180451.

На следующем этапе анализа был установлен характер корреляции частотности всего бинмиала и силы коллокационной связи с применением критерия хи-квадрат Пирсона. Анализ показал, что корреляция для объективной частотности бинмиала и количества совпавших ответов статистически значима как для первого,

так и для второго слова. При $p\text{-value} = 4,159\text{e}-08$ она составляет 0,4 615 912 и 0,4 188 926 соответственно.

Таким образом, анализ соотносимых признаков класса воспроизводимых в речи сочетаний слов, связанных синтаксической связью – устойчивости порядка слов, частотности использования в речи, предсказуемости второго элемента при актуализации первого – выявил принцип актуализации признаков, формирующих биномиальный комплекс, который

может быть назван «плавающим». Рассмотренные признаки проявляют себя как относительно независимые друг от друга, восприятию устойчивости биномиального комплекса может способствовать выдвигание на первый план одного из признаков в речевом функционировании. У биномиала же как инвариантной, внетекстовой единицы ни один из рассмотренных признаков не может быть признан доминирующим.

ЛИТЕРАТУРА

1. Firth J.R. *Papers in linguistics 1934–1951*. London : Oxford University Press, 1957. 232 p.
2. Телия В.Н. Русская фразеология. Семантический, прагматический и лингвокультурологический аспекты. М. : Языки русской культуры, 1996. 288 с.
3. Jackson H. *Words and their Meaning*. London and New York : Longman, 1995. 288 p.
4. Молодцова И.Н. Английские биномиалы: статус, сущность, функции : автореф. дис. ... канд. филол. наук. Белгород, 2002. 22 с.
5. Национальный корпус русского языка. URL: <http://www.ruscorgpora.ru/> (дата обращения 12.08.2018).
6. Резанова З., Буб А. Коллокации-биномиалы в русской речи: семантические типы, объективная и субъективная частотность // *Quaestio Rossica*. 2017. Т. 5, № 4. С. 1164–1177. DOI 10.15826/qr.2017.4.273.
7. Mollin S. The (Ir)reversibility of English Binomials: Corpus, constraints, developments. John Benjamins B.V., 2014. 264 p.
8. Jespersen O. *Growth and Structure of the English Language*. Leipzig : Teubner, 1905. 259 p.
9. Abraham R.D. Fixed order of coordinates: a study in comparative lexicography // *Modern Language Journal*. 1950. № 34 (4). P. 276–287.
10. Cooper W.E., Ross J.R. World order // *Papers from the Parasession on Functionalism*. Chicago, IL : Chicago Linguistic Society, 1975. P. 63–111.
11. Benor S.B., Levy R. The chicken or the egg? A probabilistic analysis of English binomials // *Language*. 2006. № 82 (2). P. 233–277. DOI: 10.1353/lan.2006.0077.
12. Malkiel Ya. Studies in irreversible binomials // *Lingua*. 1959. № 8. P. 113–160. DOI: 10.1016/0024-3841(59)90018-X.

Статья представлена научной редакцией «Филология» 28 декабря 2018 г.

RUSSIAN BINOMIALS: THE (IR)REVERSIBILITY SCORE AND ITS RELATION TO THE FREQUENCY AND PREDICTABILITY OF THE COMPONENTS

Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal, 2018, 437, 15–22.

DOI: 10.17223/15617793/437/2

Aleksandra S. Bub, Tomsk State University (Tomsk, Russian Federation). E-mail: aleksandrabub@yandex.ru

Zoya I. Rezanova, Tomsk State University (Tomsk, Russian Federation). E-mail: rezanovazi@mail.ru

Keywords: collocations; binomials; fixed collocations; (ir)reversibility score; frequency; predictability of components.

The article presents the results of a research of the degree of fixedness of one of the collocation types of the Russian language, binomials, which are currently poorly researched by Russian linguists. At the previous stages of the study, a set of 128 Russian binomials was composed by the authors of the publication according to the Russian National Corpus. All units were ranked according to their objective and subjective frequency. The research tasks described in the article are the following: to rank binomials according to the degree of fixedness and to find out the parameters correlated with this important property of the class of collocations of the Russian language under study: the frequency of the whole unit and the predictability of the second element. The degree of fixedness is determined on the basis of the (ir)reversibility score, which is calculated by the formula: the number of occurrences in the corpus (data from the Russian National Corpus) of a binomial with the direct word order divided by the sum of occurrences of direct word order binomials and reverse word order units, and multiplied by 100. Statistical objective data analysis was accompanied by the psycholinguistic method of obtaining subjective estimates using the scaling method. A comparison of the results of the objective and subjective estimations of the degree of fixedness of binomials allowed the authors to distinguish the nuclear part (38 units), the near-nuclear part (52), and the peripheral part (38) of the sampling. The frequency of a binomial was determined using the data of the Russian National Corpus in accordance with the formula adopted in this paradigm: the number of uses of the binomial is divided by the total number of words, and multiplied by a million. According to this criterion, 16 units (frequency from 3.447 ipm to 1.234 ipm) were allocated into the nuclear part of the sampling, 47 into the near-nuclear part (from 0.991 ipm to 0.102 ipm), and the periphery consisted of 65 (0.095 ipm to 0.003 ipm). During the comparison of the units of the nucleus of frequency and the nucleus of fixedness, only three coincidences were found: *muzh* [husband] and *zhena* [wife], *dobro* [good] and *zlo* [evil], *zhizn'* [life] and *smert'* [death]. However, the analysis using Pearson's chi-square test revealed the statistical significance of the correlation for the objective frequency of a binomial and the (ir)reversibility score of the nucleus of this sample. At $p\text{-value} = 9.861\text{e}-05$, it was 0.5894538. To determine the degree of predictability of binomials, eight questionnaires were developed, each containing 32 words. Each word was followed by a conjunction *i* [and], revealing the link with the potential next component, for example, *sudit' i . . .* [judge and . . .]. The analysis of the coincident units of the nuclear parts of the two samples based on the fixedness and predictability coefficients revealed a coincidence of 19 units. However, statistical analysis revealed the absence of a correlation dependence of the stability factor and the predictability coefficient of the nuclear part of the sample with respect to the stability parameter, $p\text{-value} < 0.05$. Thus, the analysis revealed that the features examined are relatively independent of each other, the perception of the fixedness of the binomial complex can be facilitated by highlighting one of the features in speech functioning. However, in a binomial as an invariant, an extra-textual unit, none of the characteristics can be recognised as dominant.

REFERENCES

1. Firth, J.R. (1957) *Papers in linguistics 1934–1951*. London: Oxford University Press.
2. Teliya, V.N. (1996) *Russkaya frazeologiya. Semanticheskii, pragmaticheskii i lingvokul'turologicheskii aspekty* [Russian phraseology. Semantic, pragmatic and linguocultural aspects]. Moscow: Yazyki russkoy kul'tury.

3. Jackson, H. (1995) *Words and their Meaning*. London; New York: Longman.
4. Molodtsova, I.N. (2002) *Angliyskie binomialy: status, sushchnost', funktsii* [English binomials: status, nature, functions]. Abstract of Philology Cand. Dis. Belgorod.
5. Russian National Corpus. [Online] Available from: <http://www.ruscorpora.ru/>. (Accessed: 12.08.2018). (In Russian).
6. Rezanova, Z. & Bub, A. (2017) Binomials in Russian speech: semantic types and objective and subjective frequency. *Quaestio Rossica*. 5(4). pp. 1164–1177. (In Russian). DOI: 10.15826/qr.2017.4.273.
7. Mollin, S. (2014) *The (Ir)reversibility of English Binomials: Corpus, constraints, developments*. John Benjamins B.V.
8. Jespersen, O. (1905) *Growth and Structure of the English Language*. Leipzig: Teubner.
9. Abraham, R.D. (1950) Fixed order of coordinates: A study in comparative lexicography. *Modern Language Journal*. 34 (4). pp. 276–287.
10. Cooper, W.E. & Ross, J.R. (1975) World order. In: Grossman, R.E. et al. (eds) *Papers from the Parasession on Functionalism*. Chicago, IL: Chicago Linguistic Society.
11. Benor, S.B. & Levy, R. (2006) The chicken or the egg? A probabilistic analysis of English binomials. *Language*. 82 (2). pp. 233–277. DOI: 10.1353/lan.2006.0077.
12. Malkiel, Ya. (1959) Studies in irreversible binomials. *Lingua*. 8. pp. 113–160. DOI: 10.1016/0024-3841(59)90018-X.

Received: 28 December 2018