

DOI: 10.31862/2500-2953-2024-3-20-32

УДК 81-114.4

Э.Е. Буров

Иркутский государственный университет,
664003 г. Иркутск, Российская Федерация

О связи между корпусными данными и интуицией носителей языка (на материале сочетаний типа «что-то/кто-то + полное прилагательное»)

Статья посвящена экспериментальному исследованию связи между корпусными данными и интуицией носителей языка. Указанная проблема рассматривается на материале сочетаний типа «что-то/кто-то + полное прилагательное». Гипотеза исследования состоит в том, что на данном участке между корпусными данными и интуицией носителей языка существует прямая связь, и предложения, содержащие высокоустойчивые со статистической точки зрения сочетания, значимо более приемлемы, чем предложения с сочетаниями, не обладающими статистической устойчивостью. Это утверждение проверяется в ходе двух экспериментов на оценку приемлемости. Используется шкала Ликерта от 1 до 7. Первый эксперимент нацелен на исследование сочетаний типа «что-то + полное прилагательное», второй – на исследование сочетаний типа «кто-то + полное прилагательное». Делается вывод о том, что данные обоих экспериментов полностью подтверждают выдвинутую гипотезу. Ставится вопрос о возможных причинах наличия прямой связи между корпусными и экспериментальными данными.

Ключевые слова: эксперимент на оценку приемлемости, мера logDice, языковая интуиция, корпусные данные, неопределенные местоимения

ДЛЯ ЦИТИРОВАНИЯ: Буров Э.Е. О связи между корпусными данными и интуицией носителей языка (на материале сочетаний типа «что-то/кто-то + полное прилагательное») // Рема. Rhema. 2024. № 3. С. 20–32. DOI: 10.31862/2500-2953-2024-3-20-32

© Буров Э.Е., 2024

Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License



E. Burov

Irkutsk State University,
Irkutsk, 664003, Russian Federation

On the relationship between corpus data and native speakers' intuition (on the material of combinations of the type “*čto-to* ‘something’ / *kto-to* ‘somebody’ + full adjective”)

This paper presents an experimental study of the relationship between the corpus data and the intuition of native speakers. This issue is considered using the material of Russian combinations of the type “*čto-to* ‘something’ / *kto-to* ‘somebody’ + full adjective”. The hypothesis of the study is that for this phenomenon there is a direct connection between corpus data and the intuition of native speakers, and sentences containing combinations that are statistically highly stable are significantly more acceptable than sentences with combinations that do not have statistical stability. This statement is verified with the help of two acceptability judgment experiments using 1–7 Likert scale. The first experiment is aimed at studying combinations of the type “*čto-to* ‘something’ + full adjective”, and the second one is aimed at studying combinations of the type “*kto-to* ‘somebody’ + full adjective”. The author concludes that the data from both experiments fully confirm the hypothesis and raise the question about the possible reasons for the existence of a direct connection between the corpus and experimental data.

Key words: acceptability judgment task, logDice measure, patterns of knowledge, patterns of use, indefinite pronouns.

FOR CITATION: Burov E. On the relationship between corpus data and native speakers' intuition (on the material of combinations of the type “*čto-to* ‘something’ / *kto-to* ‘somebody’ + full adjective”). *Rhema*. 2024. No. 3. Pp. 20–32. (In Rus.). DOI: 10.31862/2500-2953-2024-3-20-32

1. Введение

Одним из нетривиальных вопросов современной лингвистики является вопрос о связи между корпусными данными и интуицией носителей языка (см. [Bresnan, 2007; Divjak, 2017; Lau et al., 2017], обзор представлен в [Gerasimova, Lyutikova, 2020]). И.Б. Левонтина отмечает: «Априорно нельзя утверждать, что результаты анализа корпуса (даже если он достаточно большой и претендует на репрезентативность) отражают интуицию носителей языка. Иными словами, языковая интуиция (“patterns of knowledge”) не может быть приравнена к “patterns of use” (как они, в частности, фиксируются в корпусах). <...> Таким образом, может оказаться, что корреляции между корпусными данными и языковой интуицией носителя довольно сложны» [Левонтина, 2022, с. 16].

Представляется значимым провести экспериментальное исследование, направленное на изучение связи между корпусными данными и интуицией носителей языка. Хорошую возможность сделать это дают сочетаемостные свойства неопределенных местоимений-существительных *что-то* и *кто-то*. Как указывает Е.В. Падучева, «*что-то* часто употребляется в контексте прилагательного» [Падучева, 2016]. То же, хотя и в меньшей степени, относится к местоимению *кто-то*. Удобно изучать, с одной стороны, статистическую устойчивость сочетаний типа «*что-то/кто-то* + полное прилагательное» в корпусе, а с другой – приемлемость подобных сочетаний для носителей языка в эксперименте, а затем сопоставлять корпусные и экспериментальные данные.

Корпусный анализ¹ показывает, что одни сочетания типа «*что-то/кто-то* + полное прилагательное» являются высокоустойчивыми со статистической точки зрения, а другие, напротив, неустойчивы. Высокоустойчивые сочетания – это такие сочетания, которые встречаются в текстах намного чаще, чем они встречались бы при случайном соединении их компонентов; неустойчивые, наоборот, встречаются реже, чем при случайном соединении их компонентов. Одним из способов выявления высокоустойчивых и неустойчивых сочетаний является применение метрики \logDice , указывающей на силу синтагматической связи между компонентами словосочетания (см. подробнее [Буров, 2023]), а также корпусные исследования [Захаров, Хохлова, 2014; Хохлова, 2017; Ольховская, Зельникова, 2023], где указанная метрика применяется для анализа сочетаемости русских слов.

Если значение метрики \logDice отрицательное, соответствующее сочетание считается статистически неустойчивым, если положительное –

¹ В настоящем исследовании привлекаются данные основного корпуса Национального корпуса русского языка (НКРЯ).

статистически устойчивым (в той или иной степени). Высокоустойчивым будем считать сочетание с показателем \logDice , превышающим 2.

Например, высокоустойчивы следующие сочетания: *что-то знакомое* ($\logDice = 6,92$), *что-то белое* ($\logDice = 5,40$), *что-то грустное* ($\logDice = 5,17$), *что-то вкусное* ($\logDice = 4,92$), *кто-то чужой* ($\logDice = 5,57$), *кто-то огромный* ($\logDice = 3,63$), *кто-то страшный* ($\logDice = 2,99$), *кто-то умный* ($\logDice = 3,33$). Не обладают статистической устойчивостью следующие сочетания: *что-то популярное* ($\logDice = -0,19$), *что-то бесцветное* ($\logDice = -0,07$), *что-то нередкое* ($\logDice = -0,02$), *что-то сытное* ($\logDice = -0,04$), *кто-то свой* ($\logDice = -0,56$), *кто-то низкий* ($\logDice = -0,43$), *кто-то красивый* ($\logDice = -1,03$), *кто-то глупый* ($\logDice = -0,16$).

Эксперимент позволяет проверить, как соотносятся между собой фиксируемая в корпусе степень статистической устойчивости сочетания типа «*что-то/кто-то* + полное прилагательное» и приемлемость этого сочетания для носителей языка.

Гипотеза состоит в том, что между корпусными данными и интуицией носителей языка существует прямая связь, и предложения с высокоустойчивыми сочетаниями значимо более приемлемы, чем предложения с сочетаниями без статистической устойчивости, несмотря на то что в последних не нарушены какие бы то ни было сочетаемостные ограничения (семантические, лексические, морфо-синтаксические) в смысле [Апресян, 1974, с. 60–62]. Подтверждение этой гипотезы означало бы, что носители языка способны не только отличать противоречащее языковым правилам от не противоречащего им (например, осознавать грамматичность сочетания *что-то новое* и неграмматичность сочетания *что-то новая*), но и определять, что из соответствующего языковым правилам встречается в русской речи нетривиально часто, а что – нетривиально редко.

Было проведено два эксперимента: в первом исследовались сочетания типа «*что-то* + полное прилагательное», во втором – типа «*кто-то* + полное прилагательное».

2. Описание первого эксперимента (сочетание «*что-то* + полное прилагательное»)

В ходе первого эксперимента носители языка оценивали предложения, в которых были представлены либо высокоустойчивые сочетания типа «*что-то* + полное прилагательное», либо неустойчивые сочетания того же типа. В эксперименте, таким образом, исследовался фактор «статистическая устойчивость сочетания», который имел 2 уровня:

«высокая степень статистической устойчивости» и «отсутствие статистической устойчивости».

Сочетания для эксперимента отбирались из списка, автоматически сформированного в основном корпусе НКРЯ по запросу «ключ *что-то*, *SPRO*, коллокат *A & plen*, на расстоянии от 1 до 1 от ключа» (поле «Поиск коллокаций»).

Отбор производился с установкой на формирование пар, в которых постпозитивные прилагательные сходны по семантике (*знакомый – популярный, белый – бесцветный, грустный – нерадостный, вкусный – сытный*). Подобных пар, в которых одно прилагательное образует вместе с местоимением *что-то* высокоустойчивое сочетание, а другое прилагательное – сочетание, не обладающее статистической устойчивостью, обнаружилось очень мало. В связи с этим пришлось ограничиться небольшим количеством лексикализаций на условие – четырема: **Он напевал *что-то* знакомое / *что-то* популярное. У него в руке было *что-то* белое / *что-то* бесцветное. Произошло *что-то* грустное / *что-то* нерадостное. Мне дали съесть *что-то* вкусное / *что-то* сытное.**

Предложения распределялись по экспериментальным листам в соответствии с правилом латинского квадрата: 2 листа, т.к. имеется два экспериментальных условия, в каждом листе по 4 собственно экспериментальных, или стимульных, предложения, всего 8 экспериментальных предложений.

В экспериментальных листах содержались как собственно экспериментальные предложения, так и филлеры. Филлеры были грамматически приемлемыми, в них местоимение *что-то* употреблялось без зависимых прилагательных. Филлеры и стимульные предложения находились в соотношении 1 : 1. Предложения 1-го экспериментального листа с комментариями см. в табл. 1.

Участниками эксперимента стали 23 человека – студенты факультета иностранных языков Института филологии, иностранных языков и медиакоммуникации Иркутского государственного университета, а также выпускники факультета теоретической и прикладной филологии того же института. Респонденты распределились по экспериментальным листам следующим образом: 1-й лист – 11 участников, 2-й лист – 12 участников.

Эксперимент распространялся через социальные сети и электронную почту. Опрос проводился на платформе PCIBexFarm (<https://farm.pcibex.net>) с применением методики оценки приемлемости по шкале Ликерта от 1 до 7. Анализ данных был осуществлен в среде разработки RStudio с использованием языка программирования R [R Core Team 2023].

Результаты эксперимента представлены в табл. 2.

**Предложения 1-го опросного листа эксперимента,
направленного на исследование сочетаний типа
«что-то + полное прилагательное», с комментариями
[Sentences of the 1st questionnaire of the experiment
aimed at studying combinations of the type
“*čto-to* ‘something’ + full adjective”, with comments]**

Предложение [Sentence]	Комментарий [Comment]
<i>Он напевал что-то знакомое.</i>	Высокая устойчивость (1-я лексикализация) [High stability (1st lexicalization)]
<i>На столе лежало что-то вроде фотоаппарата.</i>	Филлер [Filler]
<i>У него в руке было что-то бесцветное.</i>	Неустойчивость (2-я лексикализация) [Instability (2nd lexicalization)]
<i>Он что-то достал из кармана.</i>	Филлер [Filler]
<i>Произошло что-то грустное.</i>	Высокая устойчивость (3-я лексикализация) [High stability (3st lexicalization)]
<i>Вдруг я услышал что-то из Бетховена.</i>	Филлер [Филлер]
<i>Мне дали съесть что-то сытное.</i>	Неустойчивость (4-я лексикализация) [Instability (4nd lexicalization)]
<i>С ним что-то случилось.</i>	Филлер [Filler]

Средняя нормализованная оценка предложений с высокоустойчивыми сочетаниями типа «*что-то* + полное прилагательное» составляет 0,19, а средняя нормализованная оценка предложений с неустойчивыми сочетаниями того же типа – минус 0,22. При этом во всех парах сочетаний с семантически сходными прилагательными, кроме пары *что-то грустное* – *что-то нерадостное*, неустойчивое сочетание имеет более низкую оценку, чем высокоустойчивое.

Существенно, что разница между оценками, данными предложениям с высокоустойчивыми сочетаниями, и оценками, данными предложениям с неустойчивыми сочетаниями, статистически значима: $df = 1$; $F = 4,413$; $p = 0,0385$ (однофакторный дисперсионный анализ; зависимость оценки от того, что употребляется в предложении – высокоустойчивое сочетание или сочетание, не обладающее статистической устойчивостью).

Таблица 2

**Средние нормализованные оценки предложений с сочетаниями
типа «что-то + полное прилагательное»
[Average normalized ratings of sentences with combinations
of the type “*что-то* ‘something’ + full adjective”]**

Высокоустойчивое сочетание [Highly stable combination]	Средняя нормализованная оценка предложения, содержащего данное высокоустойчивое сочетание [Average normalized score of a sentence containing a given highly stable combination]	Статистически неустойчивое сочетание [Statistically unstable combination]	Средняя нормализованная оценка предложения, содержащего данное статистически неустойчивое сочетание [The average normalized score of a sentence containing a given statistically unstable combination]
<i>Что-то знакомое</i>	0,53	<i>Что-то популярное</i>	-0,22
<i>Что-то белое</i>	0,94	<i>Что-то бесцветное</i>	0,12
<i>Что-то грустное</i>	-0,35	<i>Что-то нерадостное</i>	-0,19
<i>Что-то вкусное</i>	-0,38	<i>Что-то сытное</i>	-0,60
В среднем [On average]	0,19	В среднем [On average]	-0,22

3. Описание 2-го эксперимента (сочетание «кто-то + полное прилагательное»)

Второй эксперимент был устроен аналогично первому.

На каждое из двух экспериментальных условий было предложено 4 лексикализации: *Что делать, если я встречу кого-то чужого / кого-то своего? На пороге показался кто-то огромный / кто-то низкий. Около двери сидел кто-то страшный / кто-то красивый. Это сказал кто-то умный / кто-то глупый.*

Как и в случае первого эксперимента, в экспериментальных листах содержались, помимо собственно экспериментальных предложений, грамматически приемлемые филлеры. В них местоимение *кто-то* употреблялось без зависимых прилагательных. Филлеры и стимульные предложения находились в соотношении 1 : 1. Предложения 1-го экспериментального листа с комментариями см. в табл. 3.

Таблица 3

**Предложения 1-го опросного листа эксперимента,
направленного на исследование сочетаний типа
«кто-то + полное прилагательное», с комментариями
[Sentences of the 1st questionnaire of the experiment
aimed at studying combinations of the type
“kto-to ‘somebody’ + full adjective”, with comments]**

Предложение [Sentence]	Комментарий [Comment]
<i>Что делать, если я встречу кого-то чужого?</i>	Высокая устойчивость (1-я лексикализация) [High stability (1st lexicalization)]
<i>Ко мне подошел кто-то в сером.</i>	Филлер [Filler]
<i>На пороге показался кто-то низкий.</i>	Неустойчивость (2-я лексикализация) [Instability (2nd lexicalization)]
<i>Кто-то заходил в комнату пять минут назад.</i>	Филлер [Филлер]
<i>Около двери сидел кто-то страшный.</i>	Высокая устойчивость (3-я лексикализация) [High stability (3st lexicalization)]
<i>Он разговаривал с кем-то из съемочной группы.</i>	Филлер [Филлер]
<i>Это сказал кто-то глупый.</i>	Неустойчивость (4-я лексикализация) [Instability (4nd lexicalization)]
<i>Ты, наверное, разговорился с кем-то.</i>	Филлер [Filler]

Таблица 4

**Средние нормализованные оценки предложений с сочетаниями
типа «кто-то + полное прилагательное»
[Average normalized ratings of sentences with combinations
of the type “*kto-to* + full adjective”]**

Высокоустойчивое сочетание [Highly stable combination]	Средняя нормализованная оценка предложения, содержащего данное высокоустойчивое сочетание [Average normalized score of a sentence containing a given highly stable combination]	Статистически неустойчивое сочетание [Statistically unstable combination]	Средняя нормализованная оценка предложения, содержащего данное статистически неустойчивое сочетание [The average normalized score of a sentence containing a given statistically unstable combination]
<i>Кто-то чужой</i>	-0,07	<i>Кто-то свой</i>	-0,14
<i>Кто-то огромный</i>	-0,18	<i>Кто-то низкий</i>	-0,74
<i>Кто-то страшный</i>	0,09	<i>Кто-то красивый</i>	-0,88
<i>Кто-то умный</i>	-0,40	<i>Кто-то глупый</i>	-0,52
В среднем [On average]	-0,15	В среднем [On average]	-0,57

В эксперименте принял участие 21 человек. Респонденты распределились по экспериментальным листам следующим образом: 1-й лист – 10 участников, 2-й лист – 11 участников.

Результаты второго эксперимента представлены в табл. 4.

Предложения с высокоустойчивыми сочетаниями типа «кто-то + полное прилагательное» получили среднюю нормализованную оценку $-0,15$, а предложения с неустойчивыми сочетаниями того же типа – минус $0,57$. В каждой из 4 пар сочетаний с семантически сходными прилагательными неустойчивое сочетание имеет более низкую оценку, чем высокоустойчивое.

Разница между оценками предложений с высокоустойчивыми сочетаниями и оценками предложений с неустойчивыми сочетаниями, как и в случае первого эксперимента, статистически значима: $df = 1$; $F = 4,268$; $p = 0,042$ (однофакторный дисперсионный анализ; зависимость оценки от того, что употребляется в предложении – высокоустойчивое сочетание или сочетание, не обладающее статистической устойчивостью).

4. Обсуждение

Экспериментальные данные полностью подтверждают исходную гипотезу. Носители языка действительно воспринимают как менее приемлемые те сочетания, которые менее характерны для русской речи, несмотря на то что никакие сочетаемостные ограничения в этих сочетаниях не нарушены. Наблюдается прямая связь между корпусными данными и языковой интуицией носителей.

В завершение следует сказать об одном дополнительном наблюдении.

Обращает на себя внимание, что предложения первого эксперимента получили в среднем более высокие оценки, чем предложения второго. Средняя нормализованная оценка предложений с сочетанием «что-то + полное прилагательное» составляет $-0,02$, средняя нормализованная оценка предложений с сочетанием «кто-то + полное прилагательное» – минус $0,36$.

Если объединить результаты первого и второго эксперимента и принять за зависимую переменную оценку предложения, а за независимую – тип сочетания с точки зрения употребленного в нем местоимения («что-то + полное прилагательное» или «кто-то + полное прилагательное»), то однофакторный дисперсионный анализ покажет, что разница между оценками предложений с сочетанием «что-то + полное прилагательное» и оценками предложений с сочетанием «кто-то + полное прилагательное» статистически значима: $df = 1$; $F = 5,621$; $p = 0,0188$ (зависимость оценки от типа сочетания – «что-то + полное прилагательное» / «кто-то + полное прилагательное»).

Показательно также, что предложения с сочетанием «*кто-то* + полное прилагательное» имеют в среднем значимо более низкие оценки, чем грамматичные филлеры с местоимением *кто-то* ($df = 1$; $F = 24,35$; $p = 1,94 \cdot 10^{-6}$), в то время как предложения с сочетанием «*что-то* + полное прилагательное» значимо не отличаются по оценкам от грамматичных филлеров с местоимением *что-то* (хотя и оцениваются в среднем ниже, чем грамматичные филлеры).

Это коррелирует с тем фактом, что в текстах существенно большей склонностью употребляться с постпозитивными полными прилагательными обладает местоимение *что-то*: согласно данным основного корпуса НКРЯ, показатель \logDice для сочетания «*что-то* + полное прилагательное» составляет 5,31, а для сочетания «*кто-то* + полное прилагательное» – 1,25.

Можно, таким образом, констатировать, что в данном случае, как и в случае основных двух экспериментов, обнаруживается прямая связь между экспериментальными и корпусными данными.

Однако характер наблюдаемой связи, как кажется, здесь другой.

В случае основных двух экспериментов причину корреляции между корпусными и экспериментальными данными можно усматривать в том, что узус влияет на языковую интуицию: носители языка, нетривиально часто встречаясь в текстах с сочетаниями типа *что-то знакомое*, *что-то белое*, *кто-то страшный*, начинают воспринимать их как более приемлемые, чем сочетания типа *что-то популярное*, *что-то бесцветное*, *кто-то красивый*, которые, наоборот, встречаются в текстах нетривиально редко. При этом объяснить, почему сочетание *что-то знакомое* высокоустойчиво, а сочетание *что-то популярное* неустойчиво, затруднительно. По-видимому, это связано с частотностью прилагательных *знакомый* и *популярный*: *знакомый* частотнее, чем *популярный* (см. [Ляшевская, Шаров, 2009]). Но сама эта частотность представляет собой речевую данность, которую сложно каким-либо образом объяснить.

В случае же с дополнительным наблюдением связь между корпусными и экспериментальными данными имеет другой характер, т.к. и те, и другие данные указывают на одну и ту же особенность семантики местоимений *что-то* и *кто-то*: «понятие “предмета”, которое стоит за *что*-местоимениями..., гораздо абстрактнее понятия “лица”, стоящего за *кто*-местоимениями» [Кузьмина, 1989, с. 202].

Итак, местоимение *что-то* обладает более неопределенной семантикой, чем *кто-то*, а значит, ему более естественно сочетаться с постпозитивными полными прилагательными, которые в некоторой мере компенсируют неопределенность. Большая естественность конструкции «*что-то* + полное прилагательное» находит отражение, с одной

стороны, в большей устойчивости этой конструкции в корпусе, а с другой – в большей приемлемости этой конструкции для носителей языка. В результате корпусные и экспериментальные данные и оказываются напрямую связаны.

Можно заключить, что экспериментальный анализ сочетаний типа «что-то / кто-то + полное прилагательное» позволяет не только констатировать наличие на этом участке прямой связи между корпусными данными и интуицией носителей языка, но и выявить, что природа прямой корреляции между корпусными и экспериментальными данными может быть различной.

Библиографический список / References

Апресян, 1974 – Апресян Ю.Д. Лексическая семантика. Синонимические средства языка. М., 1974. [Apresyan Yu.D. Leksicheskaya semantika. Sinonimicheskie sredstva yazyka [Lexical semantics. Synonymous means of language]. Moscow, 1974.]

Буров, 2023 – Буров Э.Е. Выявление сочетаемостных свойств слова с помощью статистических мер ассоциации // «Цифра» в социально-гуманитарных исследованиях: метод, поле, реальность: материалы конференции молодых ученых. Иркутск, 14–16 ноября 2022 г. Иркутск, 2023. С. 8–11. [Burov E.E. Identification of words' syntagmatic properties by means of statistical measures of association. "Tsifra" v sotsialno-gumanitarnykh issledovaniyakh: metod, pole, realnost. Irkutsk, 2023. Pp. 8–11. (In Rus.)]

Захаров, Хохлова, 2014 – Захаров В.П., Хохлова М.В. Выделение терминологических словосочетаний из специальных текстов на основе различных мер ассоциации // Технологии информационного общества в науке, образовании и культуре: сборник научных статей. Труды XVII Всероссийской объединенной конференции «Интернет и современное общество» (IMS–2014). СПб., 2014. С. 290–293. [Zakharov V.P., Khokhlova M.V. Extracting terminological phrases by different association measures. Tekhnologii informatsionnogo obshchestva v nauke, obrazovanii i kulture. Trudy XVII Vserossiiskoi obedinennoi konferentsii «Internet i sovremennoe obshchestvo» (IMS–2014). St. Petersburg, 2014. Pp. 290–293. (In Rus.)]

Кузьмина, 1989 – Кузьмина С.М. Семантика и стилистика неопределенных местоимений // Грамматические исследования. Функционально-стилистический аспект / отв. ред. Д.Н. Шмелев. М., 1989. С. 158–231. [Kuzmina S.M. Semantics and stylistics of indefinite pronouns. Grammaticheskie issledovaniya. Funktsionalno-stilisticheskii aspekt. D.N. Shmelev (ed.). Moscow, 1989. Pp. 158–231. (In Rus.)]

Левонтина, 2022 – Левонтина И.Б. Частицы речи. М., 2022. [Levontina I.B. Chastitsy rechi [Particles of speech]. Moscow, 2022.]

Ляшевская, Шаров 2009 – Ляшевская О.Н., Шаров С.А. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М., 2009. [Lyashevskaya O.N., Sharov S.A. Chastotnyi slovar sovremennogo russkogo yazyka (na materialakh Natsionalnogo korpusa russkogo yazyka) [Frequency dictionary of modern Russian based on the Russian National Corpus]. Moscow, 2009.]

Ольховская, Зельникова, 2023 – Ольховская А.И., Зельникова А.А. Математические инструменты измерения коллокаций в аспекте изучения фразеологии // III Костомаровский форум. Славянская культура: истоки, традиции,

взаимодействие: материалы Международной научно-практической конференции. М., 2023. С. 418–423. [Olkhovskaya A.I., Zelnikova A.A. Mathematical tools for measuring collocations in the aspect of phraseology research. III Kostomarovskii forum. Slavyanskaya kultura: istoki, traditsii, vzaimodeistvie. Moscow, 2023. Pp. 418–423. (In Rus.)]

Падучева, 2016 – Падучева Е.В. Местоимения неизвестности. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru>). На правах рукописи. М., 2016. [Paducheva E.V. Mestoimeniya neizvestnosti. Materialy dlya proekta korpusnogo opisaniya russkoi grammatiki (<http://rusgram.ru>) (Pronouns of the Unknown. Materials for the project of corpus description of Russian grammar (<http://rusgram.ru>)). On the praach of the manuscript. Moscow, 2016.]

Хохлова, 2017 – Хохлова М.В. Сопоставительный анализ статистических мер на примере частеречных предпочтений сочетаемости существительных // Компьютерная лингвистика и вычислительные онтологии. 2017. Вып. 1. С. 166–171. [Khokhlova M.V. A comparative analysis of the association measures based on the analysis of part-of-speech collocational preferences of the Russian nouns. *Kompyuternaya lingvistika i vychislitelnye ontologii*. 2017. Iss. 1. Pp. 166–171. (In Rus.)]

Bresnan, 2007 – Bresnan J. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. *Roots: Linguistics in Search of its Evidential base*. Vol. 96. S. Featherston, W. Sternefeld (eds.). Berlin, 2007. Pp. 77–96.

Divjak, 2017 – Divjak D. The role of lexical frequency in the acceptability of syntactic variants: Evidence from *that* clauses in Polish. *Cognitive Science*. 2017. No. 4. Pp. 354–382. DOI: 10.1111/cogs.12335

Gerasimova, Lyutikova, 2020 – Gerasimova A., Lyutikova E. Intralingual variation in acceptability judgments and production: Three case studies in Russian grammar. *Frontiers in Psychology*. 2020. No. 11. 348. DOI: 10.3389/fpsyg.2020.00348

Lau et al., 2017 – Lau J.H., Clark A., Lappin S. Grammaticality, acceptability, and probability: A probabilistic view of linguistic knowledge. *Cognitive Science*. 2017. No. 41. Pp. 1201–1241. DOI: 10.1111/cogs.12414

R Core Team, 2023 – R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2023.

Статья поступила в редакцию 12.08.2024

The article was received on 12.08.2024

Сведения об авторе / About the author

Буров Эдуард Евгеньевич – аспирант кафедры русского языка и общего языкознания факультета теоретической и прикладной филологии, Иркутский государственный университет; преподаватель кафедры русского языка и общего языкознания факультета теоретической и прикладной филологии, Иркутский государственный университет

Eduard E. Burov – postgraduate at the Department of Russian Language and General Linguistics of the Faculty of Theoretical and Applied Philology, Irkutsk State University; lecturer at the Department of Russian Language and General Linguistics of the Faculty of Theoretical and Applied Philology, Irkutsk State University, Irkutsk, Russian Federation

E-mail: zhenyabur@mail.ru