

## ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

УДК 811.161.1(045)

DOI: 10.22213/2410-9304-2020-4-54-64

### Статистически устойчивые глагольно-именные сочетания в русских летописях

*Р. А. Верняева*, кандидат филологических наук, ИжГТУ имени М. Т. Калашникова,  
Ижевск, Россия

*С. Р. Зайнуллина*, кандидат филологических наук, доцент, Удмуртский государственный университет,  
Ижевск, Россия

*В работе представлено количественно-статистическое исследование языковых единиц древнейших русских летописей. Инструментом для получения необходимых выборок стал модуль n-грамм ИАС «Манускрипт», позволяющий выявить текстовые сочетания с разным количеством компонентов. Модуль дает возможность провести статистический анализ языковых единиц с помощью мер ассоциации. В рамках данной работы была использована мера T-score, которая позволяет определить, насколько неслучайна сила связанности между словами в словосочетании. В ходе исследования были выявлены статистически устойчивые единицы – коллокации, претендующие в древнерусском языке на статус устойчивых выражений.*

*Предметом исследования стали глагольно-именные сочетания, характерные для грамматической системы древнерусского языка и зачастую входящие в систему устойчивых словесных комплексов.*

*Важным в работе стало выявление статистически устойчивых словесных комплексов, не отраженных в словарях, в связи с тем, что традиционные методы анализа текстов не позволяют объективно оценить тесноту (близость) компонентов сочетаний, зависящую от нескольких количественных параметров.*

*При помощи модулей ИАС «Манускрипт» были выявлены не только устойчивые сочетания, нашедшие отражение в исторических словарях, но и сочетания, чья слитность не зафиксирована в лексикографических источниках.*

*В работе делается вывод о том, что количественно-статистический метод выявления устойчивых сочетаний позволяет получить лингвистически обоснованные результаты.*

**Ключевые слова:** коллокации, модуль n-грамм, лингвистическая статистика, устойчивое сочетание, глагольно-именные сочетания.

#### Введение

Известно, что язык представляет собой сложную систему отдельных единиц. Последние, в свою очередь, как и единицы любой природы, могут иметь количественные характеристики. Посчитать и сопоставить мы можем единицы всех уровней языковой системы.

В последние годы количественный подход при изучении различных лингвистических явлений становится все более распространенным. Он дает возможность более глубокого проникновения в сущность языковых процессов.

В настоящее время существует большое количество работ, посвященных количественно-статистическому анализу языковых единиц в современных текстах (см., например, работы [1–3] и др.), в сравнении с относительно небольшим количеством подобных работ на материале древних текстов, авторитетными среди которых являются исследования В. А. Баранова, Е. Л. Алексеевой, И. В. Азаровой, Д. М. Мирановой и др. [4–7].

#### Актуальность вопроса

В работе «О стандартной форме лингвистических описаний» И. А. Мельчук подчеркивал важность проведения лингвистического описания, исходя из количественных параметров: «простоты, т. е. лингвистического использования правил, и полноты, т. е. максимально возможного охвата языковых единиц различных структурных типов» [8]. Данное утверждение можно отнести также к любому методу изучения устойчивых с лексической точки зрения единиц, изучение становления и функционирования которых является актуальным в истории русского языка. Вопрос о статусе фразеологических сочетаний в древнерусском языке является открытым: можно ли квалифицировать сочетания *сътворити миръ, казнъ сътворити, законъ положити, сила бога* и подобные как устойчивые?

Во вступительной статье к Словарю русского языка 11–17-го вв. (далее: СлРЯ) отмечено, что данное лексикографическое издание содержит

сведения об устойчивых сочетаниях разного типа: 1) *фразеологически связанные сочетания* помещены в конце словарной статьи и отмечены знаком ромба, 2) *терминологические сочетания*, а также типичные, характерные для данного слова лексические сочетания даются под соответствующим значением слова» (СлРЯ, Т.1, с. 10). Устойчивыми в данной работе считаем сочетания, отмеченные двумя указанными способами.

Известно, что язык летописей отличается насыщенностью языковыми средствами, которые имеют признаки фразеологизмов в современном понимании, но данные признаки в совокупности не позволяют нам использовать по отношению к ним термин «фразеологизм», в рамках данного исследования нами используются понятия «коллокация», «устойчивое сочетание».

Определение того или иного сочетания как фразеологически связанного может представлять проблему в том случае, когда данное сочетание имеет неидиоматический характер. Особенно осторожно об идиоматичности следует говорить в том случае, если исследование глагольно-именных сочетаний проводится в диахроническом аспекте: древнерусские устойчивые сочетания отличаются от современных «прежде всего с семантической точки зрения, поскольку реализуют отражающее древнюю концептуальную форму ментальности синкретичное значение» [9].

#### **Цель работы и материал исследования**

В современной лингвистике востребованным инструментом и одновременно материалом для лингвистического исследования все чаще становятся корпуса текстов. Важным при работе с большим массивом данных является наличие инструментальной базы, позволяющей автоматизировать обработку информации и классифицировать результаты выборов.

Материалом для настоящего исследования является полнотекстовый корпус русских летописей, включающий три разновременных списка – Лаврентьевский, Ипатьевский, Радзивилловский, – оснащенный модулем выборов и запросов, модулем статистики, модулем n-грамм и опубликованный на портале ИАС «Манускрипт» (URL: <http://manuscripts.ru/mns/portal.main?p1=23>).

Целью данной работы является: 1) количественно-статистический анализ лексических комплексов в летописных текстах с помощью модуля n-грамм и запросных форм ввода и вывода результатов ИАС «Манускрипт» и 2) выявление статистически устойчивых сочетаний (коллока-

ций), претендующих на статус устойчивых сочетаний или сочетаний, имеющих переносное значение в древнерусском языке.

Ранее в диссертационном исследовании С. Р. Зайнуллиной [10] на материале «Повести временных лет» по Лаврентьевскому списку были определены структурно-семантические признаки и закономерности употребления устойчивых сочетаний в зависимости от жанрово-стилевой принадлежности текстового фрагмента. В работах [11–13] также перечислены критерии выявления устойчивых единиц разных типов в древнерусском тексте, определены возможности варьирования компонентов глагольно-именных устойчивых сочетаний.

В названных исследованиях была выработана методика поиска древнерусских устойчивых сочетаний с учетом следующих признаков: 1) закреплённость в исторических словарях, 2) высокая степень употребительности, 3) композиционная обусловленность. Однако исследование показало, что для выявления в тексте устойчивых сочетаний не достаточно данных лексикографических источников, а второй параметр устойчивости нуждается в дополнительной проверке с целью установления степени употребительности компонентов в сочетании друг с другом с помощью специальных инструментов, способных автоматически вычленять в тексте потенциально устойчивые сочетания.

Использование инструментов ИАС «Манускрипт» для подсчета единиц и автоматического анализа их сочетаемости позволяет установить степень спаянности компонентов внутри синтаксической единицы и обуславливает, таким образом, актуальность данной работы.

#### **Количественно-статистические данные исследования**

Количественный фактор имеет большое значение при любых статистических исследованиях устойчивых единиц, например, при определении их частотности.

Для того чтобы выявить сочетания, претендующие на статус устойчивых в летописном корпусе, необходимо установить количество и частотность употребления n-грамм. Получить данные характеристики позволило нам применение различных статистических мер.

Для проведения статистического анализа нами были использованы инструменты ИАС «Манускрипт», в частности запросные формы ввода и вывода результата и модуль n-грамм. Последний создан для выявления текстовых сочетаний с разным количеством компонентов.

Модуль позволяет провести статистический анализ языковых единиц с помощью мер ассоциации (Mutual Information (MI), Pointwise Mutual Information (PMI), T-score, Log-Likelihood, Dice, Inside и др.). Речь идет о статистической ассоциации, основанной на синтаксической и лексической связности. Частота совместной встречаемости лексем может быть важным параметром для нахождения устойчивых сочетаний. Статистические меры учитывают частоту совместной встречаемости единиц, а также частоту каждой отдельной единицы в текстовом корпусе.

В. А. Баранов в работе [14] описал возможности модуля n-грамм «для выявления грамматически и семантически устойчивых словосочетаний, характеризующих тематику текстов» [15]. Автор выявил перечни редких и частотных сочетаний в текстах летописей с помощью мер Mutual Information и T-score и пришел к выводу, что «MI-перечни включают имена собственные, парные именованные, устойчивые библейские и славяно-книжные подчинительные конструкции» [16], а «T-score-перечни дают информацию о событиях, целях, лицах, результатах и о их характеристиках» [17].

В рамках нашего исследования мы анализируем примеры с помощью меры T-score, которая позволяет определить, насколько неслучайна сила связанности между словами в словосочетании. Важным для исследователя является тот факт, что применение данной меры позволяет учитывать «частоту совместной встречаемости ключевого слова и его коллоката» [18] и анализировать «насколько неслучайной (подчеркивание – Р. А., С. Р.) является сила ассоциации между коллокатами» [19].

Для того чтобы оценить эффективность применения статистической меры при квалификации n-грамм как устойчивых и свободных сочетаний, потребовалось сформировать выборку, содержащую ряд запросов.

#### **Выборка беспредложных глагольно-именных сочетаний (биграмм)**

Предметом настоящего исследования являются устойчивые сочетания, состоящие из глагола и имени, связанных между собой предикативной или же подчинительной связью.

Для получения выборки было сформировано два запроса (табл. 1, 2).

*Результаты меры T-score.*

*Запрос глагол – имя существительное.*

Полученная выборка включила в себя сочетания двух типов.

Сочетания первого типа представляют собой предикативные конструкции

(ГЛАГОЛ(128) ИМЯ(83) – 17х,

ПРЕСТАВИТИС(249) КНЯЗЬ(434) – 17х,

ПРИДОША(242) ПОЛОВЦЫ(527) – 17х). Являясь частотными в рассматриваемом корпусе летописных текстов, они не имеют характера статистически устойчивых единиц, т. к. процентное соотношение общего количества употреблений каждого сочетания и количества употреблений каждого из компонентов сочетания слишком мало: например, сочетание

ПРЕСТАВИТИС(249) КНЯЗЬ(434) встречается в корпусе 17 раз. При этом процент употребления первого компонента со вторым составляет 6 %, а употребление второго компонента (он частотнее в корпусе) с первым – всего 3 %. Отсутствие устойчивой сочетаемости между компонентами объясняется также их семантикой. Субъектом в данном случае является лицо, типичное для произведения (Господь, Бог, князь, княгиня, половцы). Предикат же представлен глагольным компонентом, выражающим традиционное, типичное для памятника данной жанрово-стилевой принадлежности действие (компонент *глаголати* в паре со словами *Бъ, Гь, пророк* обычно использовался для отсылки к какому-либо авторитетному источнику, глаго-

лы *преставитис* и *прийти* активно сочетались с наименованиями значимых для повествования действующих лиц). Выборка глагольно-именных сочетаний предикативного типа, таким образом, позволяет получить представление о событийной канве летописного повествования, сделать вывод о том, какие исторические моменты чаще всего оказывались отражены в тексте.

Глагольно-именные сочетания второго типа представлены в выборке значительно более обширно. Они также состоят из глагола и имени, но, в отличие от предикативных конструкций, имеют подчинительный характер: глагол управляет именем. Данные единицы были типичны для текстов древнерусского периода. Называя подобного рода сочетания исходными, В. В. Колесов отмечает их относительно простую структуру (глагол с зависимым словом) и специфическую семантику (глагол раскрывает смысл синкретичного имени) [20].

Таблица 1. Результаты запроса «глагол – имя существительное»

№	N-грамма	F	F(w <sub>1</sub> )...F(w <sub>x</sub> )	T-score
1	НАРЕКОША НМА	28	НАРЕКОША(54) НМА(183)	5.28617
2	ПОСАДН СНА	21	ПОСАДН(110) СНА(476)	4.54992
3	ЗАЛОЖН ЦРКВЬ	20	ЗАЛОЖН(61) ЦРКВЬ(197)	4.46446
4	ПОЛОЖНША ТЪЛО	18	ПОЛОЖНША(157) ТЪЛО(131)	4.22879
5	ГЛТЬ ГЬ	17	ГЛТЬ(128) ГЬ(83)	4.11574
6	ПРЕСТАВНСА КНАЗЬ	17	ПРЕСТАВНСА(249) КНАЗЬ(434)	4.04820
7	ПРНДОША ПОЛОВЦН	17	ПРНДОША(242) ПОЛОВЦН(527)	4.03471
8	ПРНЯ ВЛАСТЬ	14	ПРНЯ(91) ВЛАСТЬ(63)	3.73728
9	НДЕ ЯРОСЛАВЪ	14	НДЕ(564) ЯРОСЛАВЪ(324)	3.60208
10	СТВОРНША МНРЪ	13	СТВОРНША(97) МНРЪ(276)	3.58433
11	ВЛОЖН БЪ	13	ВЛОЖН(47) БЪ(915)	3.57146
12	ПРЕДАСТЬ ДШЮ	8	ПРЕДАСТЬ(29) ДШЮ(49)	2.82699
13	ПРОСЛАВНША БА	8	ПРОСЛАВНША(13) БА(386)	2.82336
14	ПОХВАЛН БА	8	ПОХВАЛН(19) БА(386)	2.82102
15	ПОТРАСЕСА ЗЕМЛА	7	ПОТРАСЕСА(10) ЗЕМЛА(223)	2.64334
16	НМЪТН МНРЪ	7	НМЪТН(91) МНРЪ(276)	2.61862
17	ВЗА ГОРОДЪ	7	ВЗА(131) ГОРОДЪ(235)	2.61250
18	НЗЛОМН КОПЬЕ	6	НЗЛОМН(9) КОПЬЕ(17)	2.44931
19	СКОНУА ЖНВОТЪ	6	СКОНУА(16) ЖНВОТЪ(26)	2.44900
20	ПОЛОЖН ГЛАВОУ	6	ПОЛОЖН(48) ГЛАВОУ(41)	2.44719
21	ГОУБАТЬ ЗЕМЛЮ	6	ГОУБАТЬ(6) ЗЕМЛЮ(607)	2.44524
22	ПРЕСТАВНСА КНАГЪИНН	6	ПРЕСТАВНСА(249) КНАГЪИНН(22)	2.44310
23	НАВОДНТЬ БЪ	6	НАВОДНТЬ(12) БЪ(915)	2.43668
24	ПАДЕ ЖРЕБНН	5	ПАДЕ(15) ЖРЕБНН(19)	2.23570
25	ЯВНСА ЗВЕЗДА	5	ЯВНСА(30) ЗВЕЗДА(17)	2.23542
26	ОУЕНТН ЛЕСТЬЮ	5	ОУЕНТН(84) ЛЕСТЬЮ(28)	2.23306
27	ВЗАША СЕЛА	5	ВЗАША(180) СЕЛА(90)	2.21536
28	СКРОУШН ГЛАВЪ	4	СКРОУШН(5) ГЛАВЪ(33)	1.99976
29	ПОЛОЖНТЬ ДШЮ	4	ПОЛОЖНТЬ(4) ДШЮ(49)	1.99972

Интерес к глагольно-именным устойчивым сочетаниям данного типа обусловлен, прежде всего, тем, что большую их часть модуль позволяет интерпретировать как статистически ус-

тойчивые сочетания (далее: СУС). Опираясь на данные исторических словарей, проанализируем особенности употребления полученных в результате выборки СУС в летописных текстах.

СУС с наибольшим значением меры T-score

является единица **нарекоша**(54) **нма**(183) – 5.28. Общее количество употреблений данной конструкции в летописном корпусе составляет 28х. При этом компонент *нарекоша*(54) употребляется с коллокатом *нма* в 51% от общего употребления первого компонента в целом в корпусе летописей; процент употребления существительного *нма*(183) в сочетании с компонентом *нарекоша* ниже и составляет 15 % от общего употребления именной формы в анализируемых текстах. Статус устойчивого подтверждается данными СДРЯ, где конструкция

*нареци нма* дается в значении ‘дать имя, название’. В тексте ЛЛ встречаем данное сочетание во фрагменте, посвященном основанию города Киева:

н створнша градъ во нма брата своего старѣншаго н нарекоша нма емү кневь (ЛЛ, 4 л.).

Глагол обладает ограниченной сочетаемостью, которая объясняется его значением: ‘назвать (дать имя, название, прозвище)’.

Следующее СУС, функционирующее в летописи, – **предасть**(29) **дшю**(49). *Предати духъ, душу* – устойчивое сочетание со значением ‘умереть’ (СДРЯ, Вып. 10, с. 366). В ИЛ данное сочетание встречаем во фрагменте о кончине Ярослава Мудрого:

а всеволодъ тогда оүѣ шѣа бѣ бо любнмъ шѣмъ·пауе всеа братья·его же нмаше оү себе· ярославү же прнспѣ конѣць жнтья·н предасть·дшю свою·мѣа фебрала вѣ·ѣ (ИЛ, 60 об.).

Несмотря на относительно небольшое количество употреблений данного сочетания (8х) в корпусе, оно претендует на статус устойчивого в силу достаточно высокого процента сочетания глагольного центра с существительным *доуша* – 27 % от общего употребления глагола *предати* в летописном корпусе.

Статистической устойчивостью характеризуется также сочетание **прославнша**(13) **ѣа**(386). Исторические словари не фиксируют его как устойчивое, тем не менее сочетаемость глагольного компонента в значительной мере ограничена: *прославити* – ‘воздать славу, хвалу; прославить, возблагодарить’. Данный коллокат в 61 % от его общего употребления в летописях сочетается только с именем *богъ*. Необходимо отметить, что глагол *прославити* обычно сочетается с одушевленными именами, в тексте летописи славу воздают обычно Богу и Богородице:

н прннесше в новүю црѣвь шверзоша ракү нполннса блгоуханья црѣкы конѣ блгты·вндѣвше же се прославнша ѣа·н мнтрополнта оужасть шенде·(ЛЛ, 61 л.).

Модуль n-грамм позволил выявить широко известное СУС **нзломн**(9) **копье**(17), устойчивость которого не вызывает сомнений: из 9 употреблений глагола *изломити* в текстовом корпусе бх он использован только с существительным *копье*, что составляет более 66 % от общего употребления глагольного компонента. *Изломити* – ‘сломать, поломать’. И хотя в исторических словарях оно не закреплено как устойчивое, на его традиционный и даже символический характер указывают многие исследователи древнерусской литературы. Так, Д. С. Лихачев отмечает: «Поскольку копье было оружием первой стычки и почти всегда ломалось в ней, нам становится понятным и обычный в летописи термин “изломить копье”, употреблявшийся для обозначения того, что воин первым принял участие в битве» [21]. Именно о таком действии повествует фрагмент, включающий искомое сочетание:

андрѣн же дургевнү възма копье ѣ ѣха на передѣ н съѣхаса переже вснхъ н нзломн копье свое тогда бо доша конь под ннмъ (ИЛ, 158 л.).

Следующее СУС, функционирующее в летописи, – **предасть**(29) **дшю**(49). *Предати духъ, душу* – устойчивое сочетание со значением ‘умереть’ (СДРЯ, Вып. 10, с. 366). В ИЛ данное сочетание встречаем во фрагменте о кончине Ярослава Мудрого:

а всеволодъ тогда оүѣ шѣа бѣ бо любнмъ шѣмъ·пауе всеа братья·его же нмаше оү себе· ярославү же прнспѣ конѣць жнтья·н предасть·дшю свою·мѣа фебрала вѣ·ѣ (ИЛ, 60 об.).

Несмотря на относительно небольшое количество употреблений данного сочетания (8х) в корпусе, оно претендует на статус устойчивого в силу достаточно высокого процента сочетания глагольного центра с существительным *доуша* – 27 % от общего употребления глагола *предати* в летописном корпусе.

Статистической устойчивостью характеризуется также сочетание **прославнша**(13) **ѣа**(386). Исторические словари не фиксируют его как устойчивое, тем не менее сочетаемость глагольного компонента в значительной мере ограничена: *прославити* – ‘воздать славу, хвалу; прославить, возблагодарить’. Данный коллокат в 61 % от его общего употребления в летописях сочетается только с именем *богъ*. Необходимо отметить, что глагол *прославити* обычно сочетается с одушевленными именами, в тексте летописи славу воздают обычно Богу и Богородице:

н прннесше в новүю црѣвь шверзоша ракү нполннса блгоуханья црѣкы конѣ блгты·вндѣвше же се прославнша ѣа·н мнтрополнта оужасть шенде·(ЛЛ, 61 л.).

Модуль n-грамм позволил выявить широко известное СУС **нзломн**(9) **копье**(17), устойчивость которого не вызывает сомнений: из 9 употреблений глагола *изломити* в текстовом корпусе бх он использован только с существительным *копье*, что составляет более 66 % от общего употребления глагольного компонента. *Изломити* – ‘сломать, поломать’. И хотя в исторических словарях оно не закреплено как устойчивое, на его традиционный и даже символический характер указывают многие исследователи древнерусской литературы. Так, Д. С. Лихачев отмечает: «Поскольку копье было оружием первой стычки и почти всегда ломалось в ней, нам становится понятным и обычный в летописи термин “изломить копье”, употреблявшийся для обозначения того, что воин первым принял участие в битве» [21]. Именно о таком действии повествует фрагмент, включающий искомое сочетание:

андрѣн же дургевнү възма копье ѣ ѣха на передѣ н съѣхаса переже вснхъ н нзломн копье свое тогда бо доша конь под ннмъ (ИЛ, 158 л.).

Следующая коллокация – **скончѣ**(16)

**жнвотъ**(26) – зафиксирована в корпусе летописей бх. Несмотря на небольшое абсолютное количество употреблений компонентов друг с другом, относительное их соотношение позволяет квалифицировать данное сочетание как устойчивое: глагольный компонент сочетается с именной формой *живот* в 37 % от общего относительного количества употреблений лексики *с(ъ)кончати* в корпусе текстов. В СлРЯ употреблена устойчивая конструкция *скончати животъ свой* в значении ‘умереть, скончаться’ (СлРЯ, Вып. 24, с. 229). В контексте ИЛ данная коллокация имеет то же значение:

**кнѣвн же прішедшю въ сконі городъ**

**кнѣвъ тѹ н скончѣ жнвотъ сконі н**

**брата ёго щекъ н хорнѣвъ·н сестра**

**нхъ лзѣбѣдъ тѹ скончѣшасѣ** (ИЛ, 5 об.)

Статистической устойчивостью не характеризуются конструкции, компоненты которых имеют достаточно широкую лексическую сочетаемость и не закреплены в исторических словарях как устойчивые (**посади сна,**

**заложн црѣкъвъ, положнша тѣло**). Данные сочетания, подобно рассмотренным выше предикативным сочетаниям, могут представлять интерес с точки зрения отражения в тексте летописи повторяющихся, типичных событий (передача власти, основание церкви, погребение и т. д.).

Нами также были обнаружены устойчивые сочетания, отраженные в лексикографических источниках, но не являющиеся статистически устойчивыми. Одним из таких сочетаний явля-

ется сочетание **створнша**(97) **миръ**(276). Как устойчивое оно зафиксировано в СлРЯ: миръ сотворити – ‘заключить мирный договор’. В МСДРЯ также читаем: миръ творити – ‘заключать мирный договор’. Второе сочетание

**нмѣтн**(91) **миръ**(276) зафиксировано

в СлРЯ в словарной статье глагола имѣти: «С сущ. образует сочетания, обозначающие

действие или состояние по знач. сущ.» (СДРЯ, с. 229). В данном случае актуализировано значение существительного ‘отсутствие войны, со-

стояние мирных отношений’ (СлРЯ, с. 166): **н**

**посла слѣзѣ ко црѣвн·въ дерѣвьстрѣ бо**

**бѣ тѹ·црѣ·ръка снцѣ хоую нмѣтн**

**миръ с тобо твѣрдъ н любовь**·(ЛЛ, 22 л.).

Несмотря на закрепленность данных единиц в исторических словарях, в обоих случаях показатель совместного употребления компонентов является довольно низким притом, что вне отмеченных конструкций глагол и имя функционируют в летописи довольно активно. Отмеченное противоречие, на наш взгляд, связано с широкой сочетаемостью обоих компонентов рассматриваемых единиц, что, в свою очередь, объясняется особым статусом данных глагольно-именных устойчивых комплексов.

Отмеченные выше сочетания объединяет наличие глагола с максимально обобщенным зна-

чением (*творити, имѣти*). Не случайно составители СлРЯ указывают на характерную особенность подобных составных единиц, в которых основную смысловую нагрузку несет на себе имя. Подобные глаголы способны сочетаться со множеством отвлеченных существи-

тельных (так, например, глагол *имѣти* обычно выступал в паре с существительными *воздер-*

*жание, лѣсть, любовь, наказание, отвѣтъ, постъ, честь* и т. д.). В то же время определенная созависимость компонентов в семантическом плане указывает на неслучайность их связи. Описывая глагольно-именные сочетания данного типа, Н. М. Хоменко отмечает, что они не характеризуются переосмыслением значения и «занимают промежуточное положение между фразеологическими единицами и свободными сочетаниями» [22].

Таким образом, при построении выборки следует учитывать специфику данных устойчивых сочетаний и формировать запрос таким образом, чтобы можно было оценить сочетаемостные возможности семантически ослабленного глагольного компонента.

*Запрос: имя существительное – глагол.*

Таблица 2. Результаты запроса «имя существительное – глагол»

№	N-грамма	F	F(w <sub>1</sub> )...F(w <sub>x</sub> )	T-score
1	<b>БЪ ДАСТЪ</b>	26	<b>БЪ(915) ДАСТЪ(142)</b>	5.02620
3	<b>ПРРКЪ ГЛТЪ</b>	13	<b>ПРРКЪ(57) ГЛТЪ(128)</b>	3.59977
4	<b>СЕЛА ПОЖГОША</b>	12	<b>СЕЛА(90) ПОЖГОША(67)</b>	3.45913
5	<b>ЗНМЪ ПРЕСТАВНСА</b>	11	<b>ЗНМЪ(140) ПРЕСТАВНСА(249)</b>	3.28659
6	<b>ЛЪТА ПРЕСТАВНСА</b>	10	<b>ЛЪТА(154) ПРЕСТАВНСА(249)</b>	3.12762
7	<b>ЪРОСЛАВЪ НДЕ</b>	10	<b>ЪРОСЛАВЪ(324) НДЕ(564)</b>	2.99713
8	<b>СВЪТЪ СТВОРНША</b>	8	<b>СВЪТЪ(93) СТВОРНША(97)</b>	2.81931
9	<b>БЪ НЗБАВН</b>	8	<b>БЪ(915) НЗБАВН(55)</b>	2.77758
10	<b>ЛЪТО ПРЕСТАВНСА</b>	8	<b>ЛЪТО(384) ПРЕСТАВНСА(249)</b>	2.73182
11	<b>МНРЪ СТОНТЪ</b>	7	<b>МНРЪ(276) СТОНТЪ(44)</b>	2.63263
12	<b>ЛЪТЪ ПРЕСТАВНСА</b>	7	<b>ЛЪТЪ(59) ПРЕСТАВНСА(249)</b>	2.62988
13	<b>СВЪТЪ СТВОРН</b>	7	<b>СВЪТЪ(93) СТВОРН(245)</b>	2.62114
14	<b>НЗАСЛАВЪ СЪДЕ</b>	7	<b>НЗАСЛАВЪ(725) СЪДЕ(195)</b>	2.49304
15	<b>ТЛА ТЛНТЪ</b>	6	<b>ТЛА(7) ТЛНТЪ(6)</b>	2.44944
16	<b>ДНЪ СТВОРН</b>	7	<b>ДНЪ(750) СТВОРН(245)</b>	2.44727
17	<b>ДНН НДЕ</b>	6	<b>ДНН(319) НДЕ(564)</b>	2.23958
18	<b>НАСНЛЪЕ ТВОРАХОУ</b>	5	<b>НАСНЛЪЕ(12) ТВОРАХОУ(28)</b>	2.23564
19	<b>НАПАСТЪМН ВИНТН</b>	5	<b>НАПАСТЪМН(11) ВИНТН(31)</b>	2.23563
20	<b>МОУЖЪ ОУДАРНТЪ</b>	5	<b>МОУЖЪ(146) ОУДАРНТЪ(15)</b>	2.23327
21	<b>ПРРКЪ ГЛШЕ</b>	5	<b>ПРРКЪ(57) ГЛШЕ(44)</b>	2.23286
22	<b>ЪРОПОЛКА ХОУЮ</b>	5	<b>ЪРОПОЛКА(112) ХОУЮ(84)</b>	2.22404
23	<b>ДВАЪ ЗАТВОРНСА</b>	5	<b>ДВАЪ(343) ЗАТВОРНСА(50)</b>	2.21415
24	<b>ЛЪТО ПРННДОША</b>	5	<b>ЛЪТО(384) ПРННДОША(108)</b>	2.18306
25	<b>БОЕБОДА БЪ</b>	5	<b>БОЕБОДА(38) БЪ(1151)</b>	2.18017
26	<b>ВРЕМА БАШЕ</b>	5	<b>ВРЕМА(177) БАШЕ(320)</b>	2.16368
27	<b>МНРЪ СТВОРН</b>	5	<b>МНРЪ(276) СТВОРН(245)</b>	2.14964
28	<b>ЗНМЪ НДЕ</b>	5	<b>ЗНМЪ(140) НДЕ(564)</b>	2.13515

Данная выборка также включает в себя сочетание двух указанных типов. Видим устойчивую предикативную конструкцию, подобную рассмотренной выше **ГЛѢТЬ**(128) **ГЪ**(83), –

**ПРѢКЪ ГЛѢТЬ** (ссылка на авторитет), повторяются сочетания с именами собственными ти-

па **НЗАСЛАВЪ СЪДЕ, ДВДЪ ЗАТВОРНСА**

(описание типичных событий). Подобные предикативные сочетания не отмечены в лексикографических источниках.

Особый интерес среди полученных предикативных сочетаний представляет сочетание

**БЪ ДАСТЬ**. Запрос демонстрирует статистическую устойчивость и активность данной конструкции (26х). Употребление глагола *дати* с существительным *богъ* в летописях составляет 18 % от общего количества употреблений данного глагола. Анализируемая единица закреплена в СлРЯ как устойчивая: *богъ даль (дсть)* – о получении, приобретении чего-л., о прибавлении семьи (СлРЯ, Вып. 1, с. 255). Наличие данного сочетания свидетельствует о том, что русский человек эпохи Средневековья осмыслял те или иные события как свидетельство проявления

Божьей воли: **ОУЕНТИ МА ХОТАУЕ**

**ДА БОУДН СО МНОЮ БЪ Н СЛА**

**ЖНВОТВОРАЩАГО ХА А КАКО МН БЪ**

**ДАСТЬ Н ТО РЕКЪ ПОВЕРЖЕ НМЪ**

**ГРАМОТЫ** (ИЛ, 127 об.). Сочетание *Бог даст* знакомо и современному русскому языку, данный фразеологизм имеет разговорную окраску и употребляется с целью выражения ‘надежды на благополучный исход дела’ (ФСРЛЯ).

Кроме того, данная выборка, подобно предыдущей, содержит подчинительные глагольно-именные сочетания с неполнозначными глаго-

лами (см. например, **СВѢТЬ СТВОРНША**

и **МНРЪ СТВОРН**), однако таких единиц по данному запросу значительно меньше, поскольку при подчинении на первом месте обычно стоит глагол, управляющий именем, а не наоборот.

## Выводы

1. Проанализированные статистические данные и контексты позволили сделать вывод о том, что устойчивыми сочетаниями можно считать сочетания, характеризующиеся разной степенью слитности семантики компонентов.

2. Выявлено, что количественно-статистический метод определения устойчивых сочетаний позволяет получить лингвистически обоснованные результаты.

3. Устойчивость анализируемых единиц в летописном тексте доказывается не только закрепленностью в исторических словарях, но и стабильностью употребления – статистической устойчивостью.

4. В результате выборки по запросу «глагол – имя существительное» были выявлены частотные предикативные сочетания, обозначающие типичные для летописного текста события, а также статистически устойчивые подчинительные сочетания.

5. Выборка по запросу «имя существительное – глагол» подтвердила выводы, полученные в результате выборки «глагол – имя существительное», и продемонстрировала возможность выявления СУС предикативного типа (*бог даст*).

6. При помощи модулей ИАС «Манускрипт» были выявлены не только устойчивые сочетания, нашедшие отражение в исторических словарях, но и сочетания, чья слитность не зафиксирована в лексикографических источниках. Такой результат оказался возможен благодаря тому, что модуль n-грамм ИАС «Манускрипт» способен определить неслучайность связи компонентов сочетания.

## Список источников и сокращений

### Словари

МСДЯ – Срезневский И. И. Материалы для словаря древнерусского языка по письменным памятникам / И. И. Срезневский. Репринт. изд.: в 3 т. М.: Наука, 1989. 996 с.

СлРЯ – Словарь русского языка (XI–XVII вв.) [Электронный ресурс] // Этимология и история слов русского языка. URL: <http://etymolog.ruslang.ru/index.php?act=xi-xvii>, свободный.

ФСРЛЯ – Фразеологический словарь русского литературного языка [Электронный ресурс]. М.: Астрель, АСТ. А. И. Фёдоров. 2008. URL: <https://phraseology.academic.ru/>, свободный.

### Электронный корпус русских летописей

ЛЛ – Лаврентьевская летопись 1377 г. (РНБ, Ф.п.IV.2) [Электронный ресурс] // Портал «Манускрипт: Славянское письменное наследие». URL:



[http://manuscripts.ru/mns/main?p\\_text=32500902](http://manuscripts.ru/mns/main?p_text=32500902), требуется регистрация.

ИЛ – Ипатьевская летопись перв. пол. (сер.?) XV в. (БАН, 16.4.4) [Электронный ресурс] // Портал «Манускрипт: Славянское письменное наследие». URL: [http://manuscripts.ru/mns/main?p\\_text=32151080](http://manuscripts.ru/mns/main?p_text=32151080), требуется регистрация.

РЛ – Радзивилловская летопись кон. XV в. (БАН, 34.5.30) [Электронный ресурс] // Портал «Манускрипт: Славянское письменное наследие». URL: [http://manuscripts.ru/mns/main?p\\_text=43296853](http://manuscripts.ru/mns/main?p_text=43296853), требуется регистрация. ПВЛ – Повесть временных лет об. – оборотный лист л. – лицевой лист

### Библиографические ссылки

1. Хохлова М. В. Исследование лексико-синтаксической сочетаемости в русском языке с помощью статистических методов (на базе корпусов текстов) : автореф. дис. ... канд. филол. наук. Санкт-Петербург, 2010. 27 с. [Электронный ресурс]. URL: <https://clck.ru/RfVmT> (дата обращения: 20.07.2020 г.).

2. Борисова Е. Г. Коллокации. Что это такое и как их изучать. М. : Филология, 1995. 49 с. [Электронный ресурс]. URL: <https://search.rsl.ru/ru/record/01001721026> (дата обращения: 20.07.2020 г.).

3. Влавацкая М. В. Комбинаторно-семантический анализ коллокаций как метод лингвистического исследования (на примере колоративных коллокаций, образованных по адъективному типу) // Научный диалог. 2020. № 4. С. 40–57.

4. Баранов В. А. Опыт создания модуля n-грамм системы «Манускрипт» и оценки эффективности его использования для поиска коллокаций в корпусе М. В. Ломоносова // Интеллектуальные системы в производстве. 2016. № 4 (31). С. 124–131. ISSN 1813–7911.

5. Баранов В. А. Корреляционный анализ биграмм русских евангельских списков XI–XIV веков // Интеллектуальные системы в производстве. 2018. Т. 16, № 4. С. 65–74. ISSN 1813–7911.

6. Баранов В. А., Жолобов О. Ф. Древнерусские рукописи как объект статистического анализа // Интеллектуальные системы в производстве. 2019. Т. 17, № 4. С. 63–77. ISSN 1813–7911. DOI 10.22213/2410-9304-2019-4-63-77.

7. Алексеева И. В., Азарова И. В., Миронова Д. М. Кластеризация рукописей на базе совпадения различий как основа публикации славянской традиции // Материалы XLIV междунар. филол. научн. конф.: секция прикладной и математической лингвистики, 11–15 марта 2014 г. / [отв. ред. М. В. Хохлова]. СПб. : Филол. фак. СПбГУ, 2014. С. 10–22.

8. Мельчук И. А. О стандартной форме лингвистических описаний // Вопросы языкознания. 1963. № 1. С. 8.

9. Пименова М. В. Красотою украси : выражение эстетической оценки в древнерусском тексте. СПб. : СПбГУ; Владимир : ВГПУ, 2007. С. 53.

10. Зайнуллина С. Р. Традиционные формулы русских летописей: структура и семантика (на мате-

риале «Повести временных лет» по Лаврентьевскому списку) : дис. ... канд. филол. наук. Ижевск, 2015. 160 с.

11. Зайнуллина С. Р. Жанрово-стилевая обусловленность функционирования устойчивых сочетаний в «Повести временных лет» по Лаврентьевскому списку // Языковые категории и единицы: синтагматический аспект : материалы XII Междунар. науч. конф., посвящ. 65-летию каф. рус. яз. (г. Владимир, 26–28 сент.) : отв. ред. М. В. Пименова. Владимир, 2017. С. 149–150.

12. Зайнуллина С. Р. Проблемы изучения устойчивых сочетаний русского языка в диахроническом аспекте // Слово, высказывание, текст в когнитивном, прагматическом и культурологическом аспектах : материалы IX Междунар. науч. конф., Челябинск, 18–20 апр. 2018 г. / Мин-во образования и науки РФ, ФГБОУ ВО «Челябин. гос. ун-т» ; отв. ред. Л. А. Нефедова. Челябинск, 2018. Т. 1. С. 22–25.

13. Зайнуллина С. Р. Специфика функционирования устойчивого сочетания ротъ ходити в «Повести временных лет» по Лаврентьевскому списку // И. А. Бодуэн де Куртенэ и мировая лингвистика : междунар. конф. (VII Бодуэновские чтения) : Казан. федер. ун-т, 28–31 окт. 2019 г. : тр. и материалы : в 2 т. / под общ. ред.: К. Р. Галиуллина, Е. А. Горобец, Э. А. Исламовой. Казань, 2019. Т. 1. С. 75–79.

14. Баранов В. А. Корпусные исследования средневековых славянских рукописей: статистически значимые n-граммы (коллокации) древнерусских летописей // Электронный научно-образовательный журнал «История». 2020. Т. 11, вып. 3(89): Цифровое содружество наук: опыт применения информационных технологий в истории и смежных дисциплинах / ред. А. Ю. Володин, Артур Пол, Клива Клер [Электронный ресурс]. URL: <https://history.jes.su/s207987840009440-3-1/>. – DOI: 10.18254/S207987840009440-3. – ISSN: 2079-8784.

15. Указ. соч.

16. Указ. соч.

17. Указ. соч.

18. Хохлова М. В. Исследование лексико-синтаксической сочетаемости в русском языке с помощью статистических методов (на базе корпусов текстов) : автореф. дис. ... канд. филол. наук. Санкт-Петербург, 2010. С. 12 [Электронный ресурс]. URL: <https://clck.ru/RfVmT> (дата обращения: 20.07.2020 г.).

19. Указ. соч., с. 12.

20. Колесов В. В. Древнерусский литературный язык. Л. : Изд-во Ленингр. ун-та, 1989. С. 156.

21. Лихачев Д. С. Устные истоки художественной системы «Слова о полку Игореве» // «Слово о полку Игореве». Сборник исследований и статей. М.-Л., 1950. С. 74.

22. Хоменко Н. М. Устойчивые словосочетания в древнерусском языке XII–XV вв. (на материале произведений Кирилла Туровского и Елифания Премудрого) : автореф. дисс. ... канд. филол. наук. Ленинград, 1990. С. 9.

## References

1. Khokhlova M.V. *Issledovanie leksiko-sintaksicheskoi sochetaemosti v russkom yazyke s pomoshch'yu statisticheskikh metodov (na baze korpusov tekstov)* [Research of lexical-syntactic collocation in Russian using statistical methods (based on text corpora)]: PhD thesis. Sankt-Peterburg, 2010, 27 p. (in Russ.). Available at: <https://clck.ru/RfVmT> (accessed 20.07.2020 г.).
2. Borisova E.G. *Kollokatsii. Chto eto takoe i kak ikh izuchat'* [Collocations. What are they and how to study them]. Moscow, Filologiya, 1995, 49 p. (in Russ.). Available at: <https://search.rsl.ru/ru/record/01001721026> (accessed 20.07.2020).
3. Vlavatskaya M.V. [Combinatorial-semantic analysis of collocations as a method of linguistic research (on the example of colorative collocations formed according to the adjective type)]. *Nauchnyi dialog*, 2020, no 4, pp. 40-57. (in Russ.).
4. Baranov V.A. [The experience of creating an n-gram module of the "Manuscript" system and evaluating the effectiveness of its use to search for collocations in M. V. Lomonosov's building]. *Intellektual'nye sistemy v proizvodstve*, 2016, no. 4 (31). pp. 124–131. ISSN 1813–7911 (in Russ.).
5. Baranov V.A. [Correlation analysis of the bigrams of the Russian Gospel lists of the XI-XIV centuries]. *Intellektual'nye sistemy v proizvodstve*, 2018, vol. 16, no. 4. pp. 65–74. ISSN 1813–7911 (in Russ.).
6. Baranov V.A., Zholobov O.F. [Old Russian manuscripts as an object of statistical analysis]. *Intellektual'nye sistemy v proizvodstve*, 2019, vol. 17, no. 4, pp. 63-77 (in Russ.). ISSN 1813–7911. DOI 10.22213/2410-9304-2019-4-63-77.
7. Alekseeva I.V., Azarova I.V., Mironova D.M. *Klasterizatsiya rukopisei na baze sovpadeniya raznochtenii kak osnova publikatsii slavyanskoi traditsii* [Clustering of manuscripts based on coincidence of discrepancies as the basis for the publication of the Slavic tradition] // *Materialy XLIV mezhdunar. filol. nauchn. konf.: sektsiya prikladnoi i matematicheskoi lingvistikoi, 11-15 marta 2014 g.* [Materials of XLIV International Philological Scientific Conference: Section of Applied and Mathematical Linguistics, March 11-15, 2014] (eds. Khokhlova M. V.). SPb.: Filol. fak. SPbGU, 2014, pp. 10-22 (in Russ.).
8. Mel'chuk I.A. [About the standard form of linguistic descriptions]. *Voprosy yazykoznaniiya*, 1963, no. 1. p. 8 (in Russ.).
9. Pimenova M.V. *Krasotoyu ukrasi : vyrazhenie esteticheskoi otsenki v drevnerusskom tekste* [Beautify with beauty: the expression of aesthetic appreciation in the Old Russian text]. SPb. : SPbGU; Vladimir : VGPU, 2007, p. 53 (in Russ.).
10. Zaynullina S.R. *Traditsionnye formuly russkikh letopisei: struktura i semantika (na materiale «Povesti vremennykh let» po Lavrent'evskomu spisku)* [Traditional formulas of Russian chronicles: structure and semantics (based on the "Tale of Bygone Years" according to the Laurentian list)] :PhD thesis, Izhevsk, 2015. 160 p. (in Russ.).
11. Zaynullina S.R. *Zhanrovo-stilevaya obuslovlennost' funkcionirovaniya ustoichivyykh sochetanii v «Povesti vremennykh let» po Lavrent'evskomu spisku* [Genre and style conditionality of the functioning of stable combinations in the "Tale of Bygone Years" according to the Laurentian list]. *Yazykovye kategorii i edinitsy: sigmaticheskii aspekt : Materialy 12 mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Language categories and units: syntagmatic aspect: materials of the 12th international scientific conference]. Vladimir, 2017. pp. 149–150. (in Russ.).
12. Zaynullina S.R. *Problemy izucheniya ustoichivyykh sochetanii russkogo yazyka v diakhronicheskom aspekte* [Problems of Studying Stable Combinations of the Russian Language in the Diachronic Aspect]. *Slovo, vyskazyvanie, tekst v kognitivnom, pragmaticheskom i kul'turologicheskom aspektakh : materialy IX Mezhdunarodnoi nauchnoi konferentsii* [Word, utterance, text in cognitive, pragmatic and cultural aspects: materials of the IX International Scientific Conference]. Chelyabinsk, 2018. Vol. 1. pp. 22–25.
13. Zaynullina S.R. *Spetsifika funkcionirovaniya ustoichivogo sochetaniya rotz khoditi v «Povesti vremennykh let» po Lavrent'evskomu spisku* [The specifics of the functioning of a stable combination of « ротъ ходити » in the "Tale of Bygone Years" according to the Laurentian list]. *I. A. Boduen de Kurtene i mirovaya lingvistika : mezhdunarodnaya konferentsiya (VII Boduenovskie chteniya)* [I. A. Baudouin de Courtenay and world linguistics: international conference (VII Baudouin readings)]. Kazan', 2019. Vol. 1. pp. 75–79.
14. Baranov V.A. [Corpus studies of medieval Slavic manuscripts: statistically significant n-grams (collocations) of ancient Russian chronicles] // *Elektronnyy nauchno-obrazovatel'nyy zhurnal «Istoriya»*, 2020, T. 11, Vol. 3(89). URL: <https://history.jes.su/s207987840009440-3-1/>. – DOI: 10.18254/S207987840009440-3. – ISSN: 2079-8784.
15. *The specified composition.*
16. *The specified composition.*
17. *The specified composition.*
18. Khokhlova M.V. *Issledovanie leksiko-sintaksicheskoi sochetaemosti v russkom yazyke s pomoshch'yu statisticheskikh metodov (na baze korpusov tekstov)* [Research of lexical-syntactic collocation in Russian using statistical methods (based on text corpora)]: PhD thesis. Sankt-Peterburg, 2010, p. 12 (in Russ.). Available at: <https://clck.ru/RfVmT> (accessed 20.07.2020 г.).
19. *The specified composition*, p. 12.
20. Kolesov V.V. *Drevnerusskii literaturnyi yazyk* [Old Russian literary language]. L. : Izd-vo Leningr. unta, 1989. p. 156 (in Russ.).
21. Likhachev D.S. *Ustnye istoki khudozhestvennoi sistemy «Slova o polku Igoreve»* [Oral origins of the artistic system "Words about Igor's Host"]. «Slovo o polku Igoreve». *Sbornik issledovaniy i statei* ["Words about Igor's Host". Collection of studies and articles]. M.-L., 1950, p. 74 (in Russ.).

22. Khomenko N.M. *Ustoichivye slovosochetaniya v drevnerusskom yazyke XII-XV vv (na materiale proizvedenii Kirilla Turovskogo i Epifaniya Premudrogo)* [Stable phrases in the Old Russian language of the XII-XV centuries (based on the works of Cyril Turovsky and Epiphanius the Wise)]: PhD thesis. Leningrad, 1990, p. 9 (in Russ.).

\*\*\*

### Statistically Stable Verb-Noun Combinations in Russian Chronicals

R. A. Vernyaeva, PhD in Philology, Associate Professor, Kalashnikov ISTU, Izhevsk, Russia

S. R. Zaynullina, PhD in Philology, Associate Professor, Udmurt State University, Izhevsk, Russia

*The paper presents a quantitative and statistical study of the linguistic units in the oldest Russian Chronicals. The tool for producing the necessary samples was the n-gram module of the information analysis system "Manuscript" which allows identifying text combinations with a varied number of components. The module allows conducting a statistical analysis of linguistic units using a list of association measures. In this research, we used the T-score measure, which allows you to determine how non-random is the strength of connectivity between words in a collocation. The study revealed statistically stable combinations – collocations that claim the status of set expressions in the old Russian language.*

*The research subject is verb-noun combinations that were widely used in the old Russian language and are often included in the system of stable word complexes.*

*It is important that we identify statistically stable word complexes that are not reflected in dictionaries due to the fact that traditional methods of text analysis do not allow us to objectively determine the closeness (proximity) of components of combinations that depend on several quantitative parameters.*

*The modules of the "Manuscript" made it possible to identify not only stable combinations that are recorded in historical dictionaries but also combinations, the merging of which is not recorded in lexicographic sources.*

*The paper concludes that the quantitative and statistical method for identifying stable combinations contributes to obtaining linguistically valid results.*

**Keywords:** collocations, n-gram module, linguistic statistics, fixed expression, verb-noun combinations.

Получено: 20.10.2020