Известия Воронежского государственного педагогического университета. 2026. № 3 (312). С. 248–252.
Izvestia Voronezh State Pedagogical University. 2026. (3), 248–252.

Научная статья
УДК 81`23
DOI 10.47438/2309-7078_2026_3_248.

МОДЕЛИРОВАНИЕ РАСПРЕДЕЛЕНИЯ ГРАММАТИЧЕСКИХ ДЕРИВАТОВ ГЛАГОЛОВ *УВИДЕТЬ, УВИДАТЬ, УСМОТРЕТЬ, УГЛЯДЕТЬ, УЗРЕТЬ* В СИСТЕМЕ РЕЧИ

Оксана Тарасовна Косаренко[1], Сергей Викторович Косаренко[2]

Воронежский государственный педагогический университет[1, 2]
Воронеж, Россия

[1]*Кандидат филологических наук, доцент кафедры русского языка, современной русской и зарубежной литературы, ORCID ID: 0000-0002-5657-7023, e-mail: o.gordey@mail.ru*
[2]*Кандидат филологических наук, доцент кафедры теории, истории и методики преподавания русского языка и литературы, ORCID ID: 0000-0003-1338-7534, e-mail: serg.v.kosarenko@gmail.com*

Аннотация. В статье рассматриваются особенности употреблении глаголов зрительного восприятия *увидеть*, *увидать*, *усмотреть*, *углядеть*, *узреть* и их грамматических дериватов, наблюдаемые через систему речи. Результаты статистического анализа указывают на то, что распределение словоформ в узусе реагирует на различие и сходство в семантической организации синонимов и на возможность взаимозаменяться в одном контексте. На материале национального корпуса русского языка анализируется вероятностное отображение речевых закономерностей, доминирующих в каждом кластере на синхронном срезе современного русского языка 1926–2026 гг. В работе впервые выявлены речевые закономерности, положенные в основу моделирования морфологии глаголов и их производных. Исследование позволяет реконструировать матрицу встречаемости грамматических дериватов, принадлежащих к пяти словообразовательным гнездам, с учетом функциональной нагрузки на единицы высокого ранга. В статье доказывается, что совокупность парадигматических и частотных корреляций между формами разных гнезд позволяет распознавать строение языка по строению системы речи, а также прогнозировать развитие системы.

Ключевые слова: *система речи, лексическая прогностика, грамматические дериваты, гнездо, парадигма, словоформа, глагол, синоним, относительная частота употребления, параметрическое ядро.*

Для цитирования: *Косаренко О.Т., Косаренко С.В.* Моделирование распределения грамматических дериватов глаголов *увидеть, увидать, усмотреть, углядеть, узреть* в системе речи // Известия Воронежского государственного педагогического университета. 2026. № 3. С. 248–252. DOI 10.47438/2309-7078_2026_3_248.

Введение

Функциональная близость синонимов, возможность взаимозаменяться в одном контексте усиливают конкуренцию между ними, а также усиливают формальные различия обратно пропорционально степени их семантической общности. Целью исследования является проверка данного теоретического положения о наличии корреляции, асимметрии плана содержания и плана выражения применительно к системе речи. Также решается проблема выявления основных языковых единиц по признаку их статистической устойчивости (регулярности) и функционального веса, поскольку возможна ситуация, когда единица речи одновременно является и единицей системы языка [10]. В этом случае, как показал в своем исследовании Н. Д. Андреев, «данные о системе речи могут быть алгоритмически переведены в данные о языке» [1, с. 26]. С точки зрения лексического прогнозирования также важно определить «малое (параметрическое) ядро» языка по функциональному и синтагматическому параметрам, основываясь на численных характеристиках, не прибегая при этом к анализу семантики [5; 6]. Количественные характеристики глаголов, как предполагают исследователи, являются материальным выражением семантических различий, «степенью этого предопределения» [3, с. 185].

Материалы и методы

Рассмотрим грамматические дериваты (ГД) глаголов *УВИДЕТЬ* (√v), *УВИДАТЬ* (√w), *УСМОТРЕТЬ* (√s), *УГЛЯДЕТЬ* (√g), *УЗРЕТЬ* (√z), принадлежащих к пяти

© Косаренко О.Т., Косаренко С.В., 2026

словообразовательным гнездам. Статистические данные за последнее столетие (1926–2026 гг.), полученные на базе корпуса НКРЯ [7], позволяют вскрыть систему, которая упорядочивает распределение синонимичных форм, конкурирующих за преимущественное выражение того или иного грамматического значения. Появление каждой формы имеет свою вероятность, а предсказуемость их появления обусловлена тем, что имеются крупные отклонения «условных» вероятностей (т. е. при условии появления единиц в определенной роли) от «независимых» вероятностей (т. е. относительной частоты единиц, наблюдаемой вне условий) [1, с. 24, 165–166]. Соответственно и речевые закономерности, выполняющиеся в парадигматическим или синтагматическом планах, имеют вероятностный характер.

Результаты и обсуждение

Моделирование по морфологическим признакам ГД глаголов. Каждое гнездо имеет свою «внутреннюю» меру упорядоченности ГД (кластер 1), что индивидуализирует гнезда и противопоставляет парадигмы ГД. Совокупность всех единиц пяти гнезд (корней) составляет кластер 2, который можно представить в виде матрицы. Матрица встречаемости словоформ в речи является результатом формализации грамматических закономерностей. Сопоставим кластеры 1 и 2, предварительно исключив малоупотребительные и не входящие в активной словарь носителей русского языка ГД. Определенный квантитативный параметр маркирует отобранные единицы как ядерные высокого ранга либо как периферийные для языка явления.

Модель распределения в речи ГД в кластере 1. Чтобы определить языковой центр гнезда по формам, имеющим статистический вес, введем условия: при пороге отсечения 1% доли в каждом гнезде получаем для первого кластера ряд единиц, составляющих «большое ядро», совокупно охватывающее 45 словоформ всех гнезд. Регистрация рангов по функциональной нагрузке ведет к выявлению объективных взаимоотношений между гнездами [4]. Вероятностное распределение элементов системы речи следующее. Если в речи встречается словоформа с корнем s, то с вероятностью 36% это будет инфинитив (инф.) *усмотреть* или форма прошедшего времени (прош. вр.) (23%). Если в тексте встречается словоформа с корнем √z, то с вероятностью 26,84% это будет форма прош. вр. *узрел* или с вероятностью 21% инфинитив. Если в тексте встречается словоформа с корнем √g, то с вероятностью 39% это будет форма прош. вр. *углядел* или с вероятностью 18% инфинитив. Если в тексте встречается словоформа гнезда √v, то с вероятностью 36% это будет форма прош. вр. *увидел* или инфинитив 17%. Если в тексте встречается словоформа инф. √w, то с вероятностью 46% это будет форма прош. вр. *увидал* или деепричастие *увидав* 18%. В целом отмечается доминирование прош. вр. м. р. (кроме √s). Роль инфинитива определяется тем, что в гнездах √s и √v инфинитив входит в топ 2 по частоте, в других гнездах он менее значим. Количественная доля деепричастия заметна в гнезде √z (11%) и w (18%), а для остальных гнезд умеренна. Формы буд. вр. наиболее разнообразны и частотны в √v и √z, а в √w отсутствуют вовсе. Краткие причастия страдательного залога встречаются только у s. Как видим, наиболее сбалансированное распределение форм гнезда у √v, а наиболее «смещенное» – у √w и √g.

По объему набора ГД позволяет судить о внутреннем разнообразии (богатстве) гнезда [2]. Лидером является ГД «увидеть» с максимальным количеством словоформ, зафиксированных в корпусах текстов, включая и архаичные образования. Данный показатель коррелирует с показателем максимальной активности гнезда, форм, востребованных в речи. В перспективе развития системы эти взаимообусловливающие факторы обеспечат √v в группе глаголов зрительного восприятия лидирующие позиции.

Востребованные ГД образуют матрицу (табл. 1), эмпирически состоящую из 75 ячеек (15 позиций для репрезентации грамматических категорий), хотя не все позиции в матрице грамматических категорий остаются занятыми представителями всех гнезд.

Таблица 1. Матрица встречаемости грамматических дериватов всех типов с указанием места по рангу доли в гнезде (метрика freq) и рангу относительной частотности в речи (метрика ipm)

Грамматическая категория	**ГД s**		**ГД z**		**ГД g**		**ГД v**		**ГД w**	
	freq	*ipm*	*freq*	*ipm*	*freq*	*ipm*	*freq*	*ipm*	*freq*	*ipm*
Инфинитив	1	15	2	21	2	23	2	2	5	17
Прошедшее время, м. р., ед. ч.	2	18	1	20	1	19	1	1	1	11
Прошедшее время, ж. р., ед. ч.	5	27	8	35	3	25	3	3	3	14
Прошедшее время, ср. р., ед. ч.	10	42								
Прошедшее время, мн. ч.	3	22	4	29	4	28	5	5	4	16
Деепричастие	4	26	3	24	6	34	4	4	2	13
Будущее время, 1-е л., ед. ч.			11	45			10	10		
Будущее время, 2-е л., ед. ч.	8	40	6	32	5	31	6	6		

Грамматическая категория	**ГД s**		**ГД z**		**ГД g**		**ГД v**		**ГД w**	
	freq	*ipm*	*freq*	*ipm*	*freq*	*ipm*	*freq*	*ipm*	*freq*	*ipm*
Будущее время, 3-е л., ед. ч.	6	36	7	33	7	39	7	7		
Будущее время, 1-е л., мн. ч.			9	37			8	8		
Будущее время*, 2-е л., мн. ч.	9	41	10	44			9	9		
Будущее время, 3-е л., мн. ч.			5	30			11	12		
Императив* 2-е л., мн. ч.	9	41	10	44						
Кр. страд. прич. прош. вр. ед.ч. ср.р.	7	38								
Кр. страд. прич. прош. вр. ед.ч. ж.р.	11	15								

*Омонимия 2 л. мн. ч. будущего времени и императива в НКРЯ не снята.

Модель распределения в речи ГД в кластере 2. Если во втором кластере ограничить выборку только высокочастотными речевыми единицами с показателем ipm ≥ 1,0, то из них преодолевают пороговое значение 24, образуя «малое ядро». Нормой речи является доминирование единиц гнезда v, которые перекрывают редкие варианты в десятки раз. Доминирование обусловлено стилистически нейтральной окраской ГД v, тогда как ГД √w и ГД √s закреплены за разговорной или книжной речью и используются значительно реже.

Мера перепада, т. е. отношение частотностей двух языковых элементов, стоящих рядом в списке по убывающей частоте, также показывает некоторую упорядоченность системы речи. Мера перепада показывает, что морфологическая парадигма глагола несимметрична: есть явный «центр» (прошедшее время, инфинитив) и «периферия» (архаизмы, страдательные причастия). В целях экономии усилий и для повышения эргономичности языка его носители предпочитают использовать в речи ограниченное число форм, а остальные используются лишь в узких стилистических целях. Резкий перепад частот воспринимается интуитивно, оттеняя наиболее выразительные узлы системы речи и составляя узус. Так, в сводном списке имеется три заметных, резких перехода: 1) между 1-м и 2-м местом отмечается падение частоты почти в 2 раза: *увидел* (ipm 166,75) → *увидеть* (ipm 77,77); 2) в зоне средних (продуктивных) частот ощутимый «обрыв» в 2,3 раза есть между 5-й и 6-й строками: *увидели* (38,73) → *увидишь* (16,88); 3) при переходе из «средней» зоны в зону «низкой частоты» между 12-й и 13-й строками есть падение почти в 2 раза: *увидят* (6,6) → *увидав* (3,13).

Для измерения меры упорядоченности системы речи определим, как меняется объем набора элементов, сравнивая их ранги в кластере 1 и кластере 2. Наблюдаемое соотношение между ними заключает в себе «речевую закономерность», а именно: в пределах языкового набора элементов и отношений между ними речевая закономерность выражается в отношении числа ГД гнезда (вне условий) к количеству высокочастотных единиц как к результату конкурентной борьбы в пределах кластера 2. В частности, для двух гнезд отмечаем рост активности и расширение инвентаря форм. Для гнезда v парадигма увеличивается за счет 4-х словоформ (в итоге с 11 до 15 форм с показателем ipm выше установленного порога): производных причастных (полных форм страдательного залога *увиденное, -ого, -ым*) и деепричастных форм (*увидя*). В гнезде √w устанавливается равновесие между противоборствующими тенденциями, т. к. 5-членная парадигма не увеличивается и не уменьшается в условиях конкуренции гнезд. Для трех других гнезд, наоборот, объем набора элементов убывает в пользу более востребованных в текстах форм. Так, популярность g падает очень значительно: из 7 остается активным только 1. Также заметно падает на 9 единиц популярность форм гнезда √s с 11 до 2. Парадигма √z сокращается с 11 до 1. В результате в матрице обнаруживаем неполноту реализации парадигмы, элементы которой встречаются реже установленной частоты. (В табл. 2 серый цвет фона ячейки матрицы указывает на редукцию ГД в кластере 2 по сравнению с кластером 1.) Например, лакуна в позиции «прошедшее время, ср. р., ед. ч.» появляется после понижения функционального веса глагола *усмотрело*.

Покажем смещение занимаемого места по рангу относительной частоты ipm: а) среди доминант в своем гнезде: с 1 места *усмотреть* на 15, *узрел* → 20, *углядел* → 19, *увидел* = 1, *увидал* → 11; б) вице-доминант: со 2 места *усмотрел* → 18, *узреть* → 21, *углядеть* → 23, *увидеть* = 2, *увидав* →13. Признаки статистической устойчивости сохраняет только ГД лидера – гнезда v, что также подтверждает минимальный перепад числовых значений 1, отмечаемый для *увидят* с 11 на 12 место. То есть словоформа встречается в речи примерно так же часто, как и ожидается по ее доле в гнезде. Она «на своем месте»: не переоценена и не недооценена в реальном

употреблении. В противоположность этому максимальное отклонение фиксируем во всех других гнездах. Так, в ГД √z: перепад числовых значений составил 34 для *узрю* 11 → 45, *узрите* 10 → 44; ГД √s: перепад составил 32 пункта для словоформ *усмотрело* 10 → 42, *усмотрите* 9 → 41; ГД √g: перепад составил 32 для *углядит* 7 → 39; ГД √w: перепад составил 12 для *увидать* 5 → 17, *увидали* 4 → 16. Большой перепад (>10) означает, что реальное употребление сильно отличается от ожидаемого исхода, и словоформа не закрепилась как наиболее удобная и/или стилистически нейтральная и вытеснена синонимами. Подобные крупные отклонения и информантные формы являются наиболее выразительными узлами системы речи. Речевую закономерность можно постулировать так: отношение доли в гнезде к функциональной нагрузке в речи колеблется в среднем между 0 и 34, причем члены гнезд составляют триарную оппозицию, т. к. наибольшие колебания наблюдаются у ГД √z и √s (среднее значение 25–26, размах перепада в широких пределах 19–34 и 14–32 соответственно); наименьшие колебания – у ГД v (средний показатель 0,09, размах 0–1); промежуточное положение с умеренным расхождением между показателями занимают ГД w (средняя величина 11,2, размах 10–12 пунктов).

Выводы

В описании ГД гнезд и моделировании отношений между ними по морфологическим признакам нами выявлена неоднородность данных. Каждое гнездо имеет такой набор словоформ и соотносимых с ними числовых значений, который противопоставляет их в статико-комбинаторном отношении и позволяет группировать их на основе выявленных речевых закономерностей. Матрица встречаемости словоформ наглядно показывает лакуны в структуре, наличие конкурирующих вариантов и обеспечивает моделирование различных преобразований – изменений в системе речи по совокупности парадигматических и частотных корреляций между формами разных гнезд, и на этой основе строить прогноз развития системы языка.

Конфликт интересов

Автор декларирует отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Библиографический список

1. Андреев Н.Д. Статистико-комбинаторный метод в теоретическом и прикладном языковедении. Л. : Наука, 1967. 403 с.

2. Борунов А.Б. Разнообразие речи и методы его измерения в тексте (лингвостатистический подход) // Litera. 2017. № 4. С. 81–86.

3. Букаренко С.Г., Медведева Е.В. Глаголы зрительного восприятия видеть и смотреть // Вестник Таганрогского института имени А. П. Чехова. №1. 2022. С. 175–186.

4. Косаренко О.Т., Косаренко С.В. Лингвостатистика словообразовательной системы лексико-семантической группы «Зрительное восприятие» // Вестник Воронежского государственного университета. Серия: Лингвистика и межкультурная коммуникация. 2016. № 4. С. 50–55.

5. Кретов А.А. Основы лексико-семантической прогностики. Воронеж : Воронежский гос. ун-т, 2006. 390 с.

6. Молчанова Л.В. Качественный и количественный аспекты лексико-семантической прогностики. Воронеж : Воронежский гос. ун-т, 2007. 173 с.

7. НКРЯ – Национальный корпус русского языка. URL: http://ruscorpora.ru (дата обращения: 04.06.2026).

8. Петрухина Е.В. Особенности аспектуальной деривации в русском языке (на славянском фоне) // Јужнословенски филолог, издательство Институт за српски језик САНУ Белград, Сербия. 2023. Том 79, № 2, С. 11–29.

9. Савчук С.О., Архангельский Т.А., Бонч-Осмоловская А.А. [и др.] Национальный корпус русского языка 2.0: новые возможности и перспективы развития // Вопросы языкознания. 2024. № 2. С. 7–34.

10. Шереметьева, С. О., Бабина, О. И. Лингвостатистический анализ лексической дихотомии «язык–речь» предметной области // Russian Journal of Linguistics. 2023. Т. 27, № 2. С. 468–499.

References

1. Andreev, N.D. (1967) *Statistiko-kombinatornyi metod v teoreticheskom i prikladnom yazykovedenii* [Statistical-combinatorial method in theoretical and applied linguistics]. Leningrad, Nauka publ. 403 p. (In Russian)

2. Borunov, A.B. (2017) *Raznoobrazie rechi i metody ego izmereniya v tekste (lingvostatisticheskii podkhod)* [Diversity of speech and methods of its measurement in text (linguostatistical approach)]. Litera. (4), 81–86. (In Russian)

3. Bukarenko, S.G., Medvedeva, E.V. (2022) Glagoly zritel'nogo vospriyatiya videt' i smotret' [Verbs of visual perception to see and look]. *Vestnik Taganrogskogo instituta imeni A.P. Chekhova.* (1), 175–186. (In Russian)

4. Kosarenko, O.T., Kosarenko, S.V. (2016) *Lingvostatistika slovoobrazovatel'noj sistemy leksiko-semanticheskoj gruppy "Zritel'noe vospriyatie"* [Linguistic statistics of the word-formation system of the lexical-semantic group "Visual perception"]. Proceedings of Voronezh State University. Series: linguistics and intercultural communication. (4), 50–55. (In Russian)

5. Kretov, A.A. (2006) *Osnovy leksiko-semanticheskoi prognostiki* [Fundamentals of lexical-semantic prognostics]. Voronezh, Voronezh State University publ., 390 p. (In Russian)

6. Molchanova, L.V. (2007) *Kachestvennyi i kolichestvennyi aspekty leksiko-semanticheskoi prognostiki* [Qualitative and quantitative aspects of lexical-semantic prognostics]. Voronezh, Voronezh State University publ., 173 p. (In Russian)

7. NKRYa – *Natsional'nyi korpus russkogo yazyka* [National corpus of the Russian language]. Available from: http://ruscorpora.ru [Accessed 04.06.2026].

8. Petrukhina, E.V. (2023) Osobennosti aspektual'noi derivatsii v russkom yazyke (na slavyanskom fone) [Features of aspectual derivation in Russian (on the Slavic background)]. *Juzhnoslovenski filolog, izdatel'stvo Institut za srpski jezik SANU Belgrad, Serbiya.* 2(79), 11–29. (In Russian)

9. Savchuk, S.O., Arkhangelskiy, T.A., Bonch-Osmolovskaya, A.A. et al. (2024) Natsional'nyi korpus russkogo yazyka 2.0: novye vozmozhnosti i perspektivy razvitiya [Russian National Corpus 2.0: New opportunities and development prospects]. *Voprosy Jazykoznanija.* (2), 7–34. (In Russian)

10. Sheremet'eva, S. O., Babina, O. I. (2023) Lingvostatisticheskii analiz leksicheskoi dikhotomii «yazyk–rech'» predmetnoi oblasti [Linguistic and statistical analysis of the lexical 'Langue-Parole' dichotomy in a restricted domain]. *Russian Journal of Linguistics.* 2 (27), 468–499. (In Russian)

Поступила в редакцию 31.07.2026
Подписана в печать 30.09.2026

Original article
UDC 81`23
DOI 10.47438/2309-7078_2026_3_248.

MODELING THE DISTRIBUTION OF GRAMMATICAL DERIVATIVES OF VERBS "UVIDET', UVIDAT', USMOTRET', UGLYADET', UZRET'" IN THE SPEECH SYSTEM

Oksana T. Kosarenko[1], Sergej V. Kosarenko[2]

Voronezh State Pedagogical University[1, 2]
Voronezh, Russia

[1]*Cand. Philolog. Sci., Associate Professor of the Department of Russian Language, Modern Russian and Foreign Literature, ORCID ID: 0000-0002-5657-7023, e-mail: o.gordey@mail.ru*
[2]*Cand. Philolog. Sci., Associate Professor of the Department of Theory, History and Methodology of Teaching Russian Language and Literature, ORCID ID: 0000-0003-1338-7534, e-mail: serg.v.kosarenko@gmail.com*

Abstract: The article examines the peculiarities of the use of verbs of visual perception "uvidet', uvidat', usmotret', uglyadet', uzret'" and their grammatical derivatives, observed through the speech system. The results of statistical analysis indicate that the distribution of word forms in usage responds to differences and similarities in the semantic organization of synonyms and to the possibility of interchangeability in the same context. Using the material from the national corpus of the Russian language, a probabilistic display of speech patterns dominant in each cluster in a synchronic section of the modern Russian language from 1926 to 2026 is analyzed. This paper identifies, for the first time, speech patterns that form the basis for modeling the morphology of verbs and their derivatives. The study allows us to reconstruct the occurrence matrix of grammatical derivatives belonging to five different word-formation families, taking into account the functional load on high-ranking units. The article demonstrates that the combination of paradigmatic and frequency correlations between forms in different families allows for the recognition of language structure based on the structure of the speech system, as well as the prediction of the system's development.

Keywords: speech system, lexical prognostics, grammatical derivatives, family, paradigm, word form, verb, synonym, relative frequency of use, parametric core.

Cite as: Kosarenko O.T., Kosarenko S. V. Modeling the distribution of grammatical derivatives of verbs "uvidet', uvidat', usmotret', uglyadet', uzret'" in the speech system. *Izvestia Voronezh State Pedagogical University.* (3), 248–252. (In Russ., abstract in Eng.). DOI: 10.47438/2309-7078_2026_3_248.

Received 31.07.2026
Accepted 30.09.2026