

**ЦВЕТООБОЗНАЧЕНИЯ
В РУССКИХ РАССКАЗАХ XX—XXI ВЕКОВ:
КОРПУСНОЕ ИССЛЕДОВАНИЕ**

Т. Г. Скребцова¹, А. О. Гребенников^{1, 2}, Н. М. Марусенко¹

¹ Санкт-Петербургский государственный университет,
Россия, 199034, Санкт-Петербург, Университетская наб., 11

² Университет ИТМО
Россия, 197101, Санкт-Петербург, Кронверкский просп., 49 лит. А
Поступила в редакцию 12.01.2024 г.
Принята к публикации 10.11.2024 г.
doi: 10.5922/2225-5346-2025-1-4

Представлены результаты сравнительного анализа частоты употребления прилагательных цвета в русских рассказах начала XX и XXI веков. Исследование базируется на выборке из Корпуса русского рассказа 1900–1930, частотных словарях рассказов отдельных авторов (И. А. Бунина, А. И. Куприна, А. П. Чехова и Л. Н. Андреева), а также материалах Национального корпуса русского языка (НКРЯ). В качестве параметров сравнения используются статистические метрики: относительная частота встречаемости (*irm*) и ранговое распределение. Объектом исследования являются наиболее частотные прилагательные цвета. Сопоставление осуществляется в нескольких направлениях, связанных с творческой индивидуальностью автора, влиянием исторического контекста и спецификой формы художественного произведения.

В результате исследования обнаружено, что индивидуально-авторская вариативность в использовании цветообозначений достаточно высока по обоим статистическим параметрам. Эти различия, однако, полностью нивелируются в выборке из Корпуса русского рассказа, включающей произведения нескольких сотен писателей. Драматичные события начала XX века не сказались на внутригрупповом ранговом распределении прилагательных, но привели к существенному росту параметра *irm*. Предложено объяснение этому феномену, основанное на результатах предшествующего исследования. Анализ современных произведений базируется на выборке из НКРЯ, включающей рассказы, написанные в 2000–2023 годах. При сравнении выборок выявлено, что наиболее частотные прилагательные цвета спустя столетие сохранили свои позиции во внутригрупповом ранговом распределении. Параметр *irm* за редким исключением демонстрирует существенное снижение. В связи с этим выдвинута гипотеза, аналогичная той, что ранее объясняла рост *irm* в 1900–1930 годах. Дополнительное сопоставление, предпринятое на материале рассказов и поэзии И. А. Бунина, свидетельствует о влиянии формы художественного произведения на статистику использования прилагательных цвета.

Ключевые слова: цветообозначения, русский рассказ, лингвостатистика, Корпус русского рассказа 1900–1930, частотные словари, ранговое распределение

1. Введение

Цветообозначения, или цветоименования, — это названия цветов и оттенков. В каждом языке это особая семантическая группа, отличающаяся от подобных групп в других языках с количественной и каче-



ственной точек зрения. Исследования, проводившиеся лингвистами и антропологами в 1960–1970-е годы, позволили выявить степень универсальности / вариативности в различных системах цветообозначений. Было обнаружено, что, несмотря на широкое разнообразие цветовых категорий, выделяемых в разных языках, можно говорить о едином инвентаре «базовых» цветообозначений, а также сформулировать правила, касающиеся их иерархии как в сопоставительном, так и диахроническом аспекте (Berlin, Kay, 1969).

В отечественном языкознании интерес к цветообозначениям проявляется не только в лингвокультурологии и этнолингвистике, но и в исследованиях, связанных с историей возникновения тех или иных названий цвета (Бахилина, 1975), экспериментальным изучением их семантики (Фрумкина, 1984), лексикографическим описанием (Харченко, 2009), прикладным аспектом использования (Василевич и др., 2005), коннотациями прилагательных цвета (Чернейко, Ли, 2019) и пр. Разнообразие теоретических подходов и научных интересов наглядно демонстрируют материалы круглого стола, посвященного цветообозначениям (Проблемы цвета..., 2004).

В настоящем исследовании цветоименования рассматриваются с точки зрения их функционирования в художественной речи (в качестве материала преимущественно используются русские рассказы начала XX и XXI века – подробнее см. ниже). Опираясь на статистические методы обработки материалов электронных корпусов текстов, мы анализируем сравнительную частоту употребления тех или иных названий цвета у различных авторов и в разные эпохи.

Объектом рассмотрения являются восемь прилагательных, обозначающих пять цветов спектра (*красный, желтый, зеленый, голубой, синий*) и три ахроматических цвета (*белый, черный, серый*). Они были отобраны по принципу частотности: в нашем материале данные цветообозначения встречаются чаще прочих. Примечательно, что все они относятся к базовым цветам (единственным исключением, возможно, является *голубой*, чей статус вызывает вопросы, ср. (Перехвальская, 2015, с. 157). Для проверки тех или иных гипотез в статье также используются дополнительные данные по другим цветообозначениям.

2. Материал и метод исследования

2.1. Корпус русского рассказа 1900–1930

Сегодня очевидно, что оценку изменений в рамках любой лексической группы следует проводить с опорой на корпусные и статистические методы. С этой целью исследователями СПбГУ и НИУ ВШЭ уже много лет создается представительный Корпус рассказов русских писателей, который включает произведения максимально широкого круга авторов, написанные на русском языке с 1900 по 1930 год и опубликованные в периодических изданиях¹. Художественная проза чутко реа-

¹ URL: <https://russian-short-stories.ru>



гирует на события в социальной и политической жизни, особенно в эпоху бурных исторических перемен. Рассказ, будучи коротким и наиболее распространенным жанром художественной прозы, быстро улавливает изменения в индивидуальном сознании и сознании социума и, таким образом, может быть использован для отслеживания изменений в лексическом составе. На настоящий момент Корпус содержит несколько тысяч единиц.

Историческим центром периода, охватываемого Корпусом, очевидно является Октябрьская революция. Все остальные события и процессы рассматриваются или как преддверие перелома, или как его последствия. В связи с этим Корпус разделен на три хронологически последовательных периода: 1) рассказы начала XX века (1900–1913); 2) рассказы эпохи острых социальных потрясений – Первая мировая война, Февральская и Октябрьская революции и последующая за ними Гражданская война (1914–1922); 3) рассказы послереволюционной эпохи – раннего советского периода (1923–1930).

На материале Корпуса была сформирована тестовая выборка, включающая 310 рассказов (приблизительно по 100 из каждого периода) без каких-либо ограничений по длине, тематике и т. п. Каждый автор может быть представлен не более чем одним произведением за период. Кроме того, необходимо, чтобы рассказ был написан и опубликован в России (или СССР), а не за рубежом. Выборка включает как рассказы признанных мастеров русской литературы, так и (даже в большей степени) произведения авторов, практически неизвестных современному читателю.

При компьютерной обработке тексты рассказов в подкорпусах для каждого периода (а также для выборки в целом) последовательно объединялись, и для этих кумулятивных образований были составлены ранговые частотные словари лексем, которые и служат «полигоном» для нашего исследования:

- 24316 лексем, 376513 словоформ для первого периода;
- 24617 лексем, 303588 словоформ для второго периода;
- 30560 лексем; 383430 словоформ для третьего периода;
- 124081 лексема, 1077970 словоформ для выборки в целом.

2.2. Частотные словари писателей

Кроме того, за прошедшие годы были созданы и изданы частотные словари рассказов выдающихся русских писателей, признанных мастеров жанра (А. П. Чехова, Л. Н. Андреева, А. И. Куприна, И. А. Бунина) (Частотный словарь..., 1999; 2003; 2006; 2011). Эти словари основываются на случайных, но репрезентативных выборках из произведений указанных авторов. Их материалы позволяют многоаспектно исследовать особенности языка писателя, выявлять лексемы, соответствующие ключевым темам его творчества, анализировать наполнение отдельных семантических областей. В настоящем исследовании словари используются для сравнительного анализа цветообозначений.



2.3. Национальный корпус русского языка: рассказы XXI века

Сегодня лингвистам при решении широкого спектра задач предоставляется возможность пользоваться инструментарием Национального корпуса русского языка (НКРЯ). Мы обратились к НКРЯ для того, чтобы получить статистические данные по употреблению прилагательных цвета в XXI веке, а затем сравнить их с соответствующими показателями для начала XX столетия.

С этой целью мы задали аналогичный нашему подкорпус русских рассказов, написанных с 2000 по 2023 год. Он насчитывает 1233 рассказа общим объемом 5 178 237 словоупотреблений; таким образом, его репрезентативность не вызывает сомнений. Как и тестовая выборка из Корпуса русского рассказа, он включает произведения нескольких сотен как известных, так и молодых, совсем недавно заявивших о себе авторов.

Помимо этого материалы НКРЯ были использованы для выявления статистики использования прилагательных цвета в поэзии И. А. Бунина. Это частный аспект настоящего исследования, связанный с влиянием вида художественной литературы на частотное распределение слов.

2.4. Методика

В настоящей статье для количественной оценки употребления цветообозначений мы используем параметр относительной частоты встречаемости слова *ipm* — число употреблений на миллион слов. Такой подход традиционно применяется в условиях, когда объемы выборок существенно различаются.

Вторым параметром, к которому мы прибегаем для наглядности, является ранг. Однако речь идет не о рангах прилагательных цвета в общем ранговом распределении всех лексем, а о ранге того или иного прилагательного относительно других. Иными словами, из общего рангового распределения цветообозначения извлекаются, сравниваются по параметру *ipm* и снабжаются новыми, внутригрупповыми рангами: самому частотному из них присваивается ранг 1, следующему — 2 и т. д. Как правило, в речь идет о рангах от 1 до 8 (в соответствии с объектом нашего исследования), но если привлекаются дополнительные цветообозначения, то и до 13.

Следует подчеркнуть, что из-за невозможности автоматического анализа полисемии каждое цветообозначение рассматривается в качестве единой лексемы, включающей все ее лексико-семантические варианты. Таким образом, хотя на протяжении всей статьи мы говорим о *цвете* (ср. *цветообозначения, цветоименования, прилагательные цвета*), на самом деле в тексте может реализоваться не только основное, «цветовое», значение, но и разнообразные переносные значения, а также окказиональные, индивидуально-авторские употребления. Прилагательное также может подвергаться субстантивации (ср. *красные, белые, зеленые* и т. п.). Подобные формы, однако, автоматически распознаются как прилагательные цвета. Это неизбежное, по крайней мере на сегодняшний день, ограничение.



3. Ранговое распределение цветообозначений в рассказах отдельных писателей

Сопоставление частотных списков, представленных в (Частотный словарь..., 1999; 2003; 2006; 2011), обнаруживает довольно большую вариативность в писательском предпочтении цвета (табл. 1). Примечательно, что верхний ранг делят пополам прилагательные *белый* и *черный*. Как будет показано в следующем разделе, прилагательное *белый* в среднем более частотно для рассказов начала XX века. Однако, когда речь идет о творчестве отдельных авторов, его лидерство не является безусловным. Третье место стабильно занимает прилагательное *красный*; четвертое поделено между *серым* и *зеленым*. Забегая вперед, заметим, что в подкорпусе русских рассказов *серый* имеет 4-й ранг на протяжении всех периодов, а также в целом по выборке (табл. 2).

Таблица 1

Ирм и ранговое распределение цветообозначений в рассказах Чехова, Андреева, Куприна и Бунина

| Цвет | А. П. Чехов | | Л. Н. Андреев | | А. И. Куприн | | И. А. Бунин | |
|----------------|----------------|-------|---------------|--------|--------------|--------|-------------|--------|
| | ранг | ирм | ранг | ирм | ранг | ирм | ранг | ирм |
| <i>белый</i> | 1 | 515,0 | 2 | 852,1 | 1 | 1085,8 | 2 | 1800,1 |
| <i>черный</i> | 2 | 348,4 | 1 | 1146,3 | 2 | 1065,0 | 1 | 1994,7 |
| <i>красный</i> | 3 | 343,3 | 3 | 674,6 | 3 | 645,3 | 3 | 1131,2 |
| <i>серый</i> | 5 | 166,6 | 4 | 365,2 | 5 | 239,4 | 4 | 656,8 |
| <i>синий</i> | 7 | 136,3 | 6 | 157,2 | 8 | 225,5 | 7 | 279,7 |
| <i>желтый</i> | 9 ² | 70,7 | 5 | 223,2 | 6/7 | 235,9 | 6 | 437,9 |
| <i>зеленый</i> | 4 | 247,4 | 7 | 131,9 | 4 | 319,2 | 5 | 535,2 |
| <i>голубой</i> | 8 | 90,9 | 8 | 126,8 | 6/7 | 235,9 | 8 | 255,4 |

Таблица 2

Ирм и ранговое распределение цветообозначений в Корпусе русского рассказа 1900 – 1930

| Цвет | 1900 – 1913 | | 1914 – 1922 | | 1923 – 1930 | | 1900 – 1930 | |
|----------------|-------------|-------|-------------|-------|-------------|--------|-------------|-------|
| | ранг | ирм | ранг | ирм | ранг | ирм | ранг | ирм |
| <i>белый</i> | 1 | 934,9 | 1 | 892,7 | 1 | 1006,7 | 1 | 953,6 |
| <i>черный</i> | 2 | 905,7 | 2 | 807,0 | 2 | 978,0 | 2 | 942,5 |
| <i>красный</i> | 3 | 406,4 | 3 | 504,0 | 3 | 751,1 | 3 | 574,2 |
| <i>серый</i> | 4 | 363,9 | 4 | 336,0 | 4 | 508,6 | 4 | 427,6 |
| <i>синий</i> | 5 | 355,9 | 5 | 306,3 | 5 | 461,6 | 5 | 390,5 |
| <i>желтый</i> | 6 | 252,3 | 6 | 204,2 | 6/7 | 313,0 | 6 | 266,2 |
| <i>зеленый</i> | 7/8 | 177,9 | 7 | 174,6 | 6/7 | 313,0 | 7 | 234,7 |
| <i>голубой</i> | 7/8 | 177,9 | 8 | 161,4 | 8 | 247,8 | 8 | 204,1 |

² У Чехова прилагательное *желтый* не входит в круг восьми наиболее частотных цветообозначений (это единственный случай такого рода в рассматриваемом материале). Зато туда вошло прилагательное *рыжий*, занявшее 6-ю позицию.



Важны и ранг прилагательного, и степень его отрыва от ближайшего «соперника». Если обратиться к показателям частотности, нетрудно заметить, что дистанция, разделяющая первую и вторую строчки, весьма значительна у Андреева и Чехова, а у Бунина и Куприна она относительно невелика. Прослеживая показатели частотности по всей таблице, можно также обнаружить, что Чехов в целом гораздо реже использует цветообозначения, чем другие писатели.

Вариативность рангового распределения прилагательных цвета можно считать одним из проявлений творческой индивидуальности писателя, его особенного взгляда на мир. Предпринятое ранее исследование 50 наиболее частотных знаменательных слов в рассказах рассматриваемых авторов продемонстрировало существенные отличия в авторском мировосприятии. Так, было установлено, что творчество Бунина отмечено большим числом лексических единиц, относящихся к миру природы (*ветер, небо, поле, сад, море, лес, солнце, дорога*), в то время как в произведениях Чехова, напротив, основное внимание уделяется миру людей (*жена, муж, должный, нужный, доктор, слово*). Трагическое мироощущение Л. Н. Андреева выражается в высоких рангах слов *смерть, черный, темный, плакать* (Гребенников, Скребцова, 2019).

4. Ранговое распределение цветообозначений в Корпусе русского рассказа

Обращаясь к частотным словарям, созданным на материале выборки из Корпуса русского рассказа 1900–1930, мы видим практически одинаковое ранговое распределение по всем периодам и выборке в целом (табл. 2). Это может показаться удивительным, если принять во внимание радикальные преобразования, происходившие в стране на протяжении указанных трех десятилетий и существенно повлиявшие на лексику рассказов, в том числе на позиции многих слов в ранговом распределении (Скребцова, 2021; Скребцова, Гребенников, Шерстинова, 2021).

Примечательно, что прилагательное *белый* стабильно занимает первую строчку, хотя его отрыв от прилагательного *черный* относительно невелик во все периоды, а по выборке в целом сводится к минимуму. Пять верхних рангов полностью совпадают: незначительные колебания наблюдаются лишь внизу таблицы. Прилагательное *синий* поднялось до пятого ранга, чего не наблюдалось ни у одного из четырех писателей (см. выше).

На фоне единообразного рангового распределения обращает на себя внимание то, что частотность всех восьми прилагательных в третий период заметно выросла по сравнению с их частотностью в первый период. Это замечание справедливо и в отношении других, менее частотных цветообозначений, занимающих ранги с 9 по 13 (табл. 3).



Таблица 3

**Ирм дополнительных цветообозначений
в Корпусе русского рассказа 1900 – 1930**

| Цвет | 1900 – 1913, ирм | 1914 – 1922, ирм | 1923 – 1930, ирм |
|-------------------|------------------|------------------|------------------|
| <i>рыжий</i> | 85,0 | 92,2 | 229,5 |
| <i>бурый</i> | 21,2 | 52,7 | 54,8 |
| <i>коричневый</i> | 13,3 | 95,5 | 73,0 |
| <i>лиловый</i> | 13,3 | 42,8 | 41,7 |
| <i>оранжевый</i> | 2,7 | 9,9 | 26,1 |

Тенденция столь очевидна, что нуждается в объяснении. Мы предполагаем, что значительное увеличение частотности прилагательных цвета может быть связано с изменениями в тематике рассказов, произошедшими под влиянием исторического контекста. Как было показано ранее, во второй и третий периоды появилось много рассказов, посвященных войне (Первой мировой, Гражданской); кроме того, в третий (советский) период на фоне масштабных аграрных преобразований актуализировались темы, связанные с сельской жизнью, крестьянским бытом (Skrebtsova, 2021). Следовательно, действие повествования чаще происходит на открытом воздухе (в поле, лесу и пр.), на природе с ее многоцветьем. Коричневый, бурый, рыжий – цвет земли или окрас животных; зеленый – цвет травы, растений, листвы; голубой и синий – цвет ясного неба и т. д. (красный тоже встречается в природе, но это также и цвет крови и революционных знамен).

В подтверждение нашей гипотезы заметим, что целый ряд существительных, обозначающих домашних животных и природные объекты, в свою очередь демонстрирует движение вверх в общем ранговом распределении. Наиболее яркие примеры приведены ниже (табл. 4.).

Таблица 4

Ирм некоторых существительных в Корпусе русского рассказа 1900 – 1930

| Существительное | 1900 – 1913, ирм | 1914 – 1922, ирм | 1923 – 1930, ирм |
|-----------------|------------------|------------------|------------------|
| <i>лошадь</i> | 393,1 | 431,5 | 565,9 |
| <i>конь</i> | 37,2 | 128,5 | 331,2 |
| <i>корова</i> | 63,7 | 95,5 | 216,5 |
| <i>собака</i> | 201,9 | 247,0 | 391,2 |
| <i>лес</i> | 58,4 | 59,3 | 75,6 |
| <i>поле</i> | 146,1 | 194,3 | 297,3 |
| <i>куст</i> | 162,0 | 207,5 | 284,3 |
| <i>трава</i> | 201,9 | 266,8 | 294,7 |
| <i>земля</i> | 1035,8 | 872,9 | 1098,0 |
| <i>грязь</i> | 53,1 | 46,1 | 125,2 |



5. Ранговое распределение прилагательных цвета в прозе и поэзии И. А. Бунина

Как было показано выше, общественно-политический контекст, хотя и опосредованно, может влиять на частоту употребления цветообозначений. В этом разделе нас интересует, может ли в качестве значимого фактора выступать форма литературного произведения. На основании частотного словаря рассказов И. А. Бунина, с одной стороны, и собрания его стихотворений в поэтическом корпусе НКРЯ – с другой, можно получить статистические данные по прилагательным цвета (табл. 5).

Таблица 5

Ирм и ранговое распределение цветообозначений в рассказах и стихотворениях И. А. Бунина

| Цвет | Рассказы | | Стихи | |
|----------------|----------|--------|-------|--------|
| | ранг | ирм | ранг | ирм |
| <i>белый</i> | 2 | 1800,1 | 1 | 2503,9 |
| <i>черный</i> | 1 | 1994,7 | 2 | 1930,8 |
| <i>красный</i> | 3 | 1131,2 | 5 | 921,8 |
| <i>серый</i> | 4 | 656,8 | 7 | 585,5 |
| <i>синий</i> | 7 | 279,7 | 3 | 1382,7 |
| <i>желтый</i> | 6 | 437,9 | 8 | 535,6 |
| <i>зеленый</i> | 5 | 535,2 | 4 | 1295,5 |
| <i>голубой</i> | 8 | 255,4 | 6 | 859,5 |

К сожалению, ввиду отсутствия материала у нас нет возможности распространить сравнение на других авторов, оставивших солидное наследие в обоих видах художественной литературы, – к примеру, В. Набокова, М. Кузмина, А. Белого. Поэтому то, что представлено ниже, вынужденно носит характер кейс-стади.

Приведенные данные демонстрируют значительные расхождения по обоим параметрам. Первые ранги по-прежнему распределены между прилагательными *белый* и *черный*, хотя и в обратной последовательности; у прочих прилагательных ранги сильно отличаются. Обращает на себя внимание значительно более высокая частотность прилагательных *синий*, *зеленый* и *голубой* в поэзии Бунина по сравнению с его рассказами.

Причина различий, по-видимому, кроется в разнице форм художественной речи. Очевидно, что на выбор слова в поэзии сильно влияют фонетические факторы (рифма, ритм, аллитерация, слоговая структура слова и пр.); в прозе их действие по меньшей мере не столь заметно. Можно также предположить, что в целом поэзия характеризуется употреблением слов в фигуральных значениях и индивидуально-авторских окружениях. Напротив, рассказ как нарративный жанр, с присущими ему краткостью и связью с актуальными событиями, скорее описывает конкретную действительность, чем ее художественно пере-



осмысляет и обобщает. Соответственно, в рассказах слова будут чаще выступать в своих прямых, конкретно-предметных значениях. Отсюда следует, что «средний баланс» между прямыми и переносными значениями, а также конвенциональными и новаторскими употреблениями в стихах и рассказах будет различаться. В поэзии он будет склоняться в пользу фигуральных и индивидуально-авторских употреблений (ср.: Масевич, Захаров 2019, с. 47). А поскольку возможности для расширения и сдвига значений являются особенными, специфичными для каждого слова, этот баланс нарушается индивидуально, по-своему у каждого из рассматриваемых прилагательных. Впрочем, эта гипотеза нуждается в дальнейшей проверке.

Вообще, как свидетельствуют множественные статистические сопоставления, предпринятые на материале русской поэзии от XVIII до XXI века (Марголис, 2022), авторская вариативность в использовании цветообозначений в поэзии весьма высока. В частности, в творчестве поэтов Серебряного века не всегда доминирует *белый* или *черный* цвет: так у В. Маяковского наиболее частотным является *красный*, а у С. Есенина — *синий*. В то же время среднее ранговое распределение цветообозначений в поэзии начала XX века схоже с тем, что представлено в таблице 2; в порядке убывания частоты следуют *белый* – *черный* – *красный* – *синий* (прилагательное *серый* в цитируемой монографии не учитывалось, поэтому его позиция в иерархии неизвестна) (Там же, с. 192–195).

6. Ранговое распределение цветообозначений в подкорпусе НКРЯ (2000–2023)

В данном разделе мы обращаемся к современному нам материалу – рассказам, написанным в начале XXI века (табл. 6). Для наглядности в таблицу также включены данные по ранговому распределению и относительной частоте прилагательных цвета из таблицы 2. Таким образом, можно сравнить статистические характеристики совокупностей текстов, разделенных почти столетием.

Таблица 6

Ирм и ранговое распределение цветообозначений в рассказах начала XX и начала XXI века

| Цвет | Рассказы 1900–1930 | | Рассказы 2000–2023 | |
|----------------|--------------------|-------|--------------------|-------|
| | ранг | ирм | ранг | ирм |
| <i>белый</i> | 1 | 953,6 | 1 | 628,0 |
| <i>черный</i> | 2 | 942,5 | 2 | 601,8 |
| <i>красный</i> | 3 | 574,2 | 3 | 342,2 |
| <i>серый</i> | 4 | 427,6 | 4 | 228,5 |
| <i>синий</i> | 5 | 390,5 | 6 | 177,3 |
| <i>желтый</i> | 6 | 266,2 | 7 | 172,8 |
| <i>зеленый</i> | 7 | 234,7 | 5 | 230,0 |
| <i>голубой</i> | 8 | 204,1 | 8 | 123,4 |



Нетрудно заметить, что перераспределение рангов прилагательных имеет место в нижней половине таблицы, в то время как иерархия *белый – черный – красный – серый* остается неизменной. Интересным выглядит сопоставление цветообозначений по параметру относительной частоты употребления. Можно видеть, что почти все они демонстрируют резкое (в полтора-два раза) падение частотности; единственное исключение составляет прилагательное *зеленый*, сохранившее свою относительную частоту употребления и за счет этого поднявшееся в ранге.

Размышляя о причинах этого феномена, можно вновь предположить влияние внешних факторов — на этот раз характерной для последнего столетия тенденции урбанизации. Массовый отток сельского населения в город не может не влиять на тематику рассказов, на их место действия. Впрочем, в данном случае у нас нет специальных исследований, подтверждающих эту гипотезу.

Для более широкой проверки обнаруженной тенденции мы, как и ранее, обратились к статистике по менее частотным цветообозначениям (табл. 7). Однако здесь картина получилась не столь единообразной, ср. почти пятикратное падение частотности прилагательного *бурый*, но трехкратный рост *оранжевого*.

Таблица 7

**Ирм и ранговое распределение дополнительных цветообозначений
в рассказах начала XX и начала XXI века**

| Цвет | Рассказы 1900–1930 | | Рассказы 2000–2023 | |
|-------------------|--------------------|--------|--------------------|------|
| | ранг | ирм | ранг | ирм |
| <i>рыжий</i> | 9 | 148,43 | 9 | 96,6 |
| <i>бурый</i> | 10 | 121,52 | 12 | 26,5 |
| <i>коричневый</i> | 11 | 59,37 | 10 | 56,4 |
| <i>лиловый</i> | 12 | 32,47 | 13 | 18,7 |
| <i>оранжевый</i> | 13 | 12,99 | 11 | 37,1 |

Таким образом, сравнивая данные по рассказам начала XX века и начала XXI века, можно констатировать, что ядро описываемой группы (наиболее частотные цветообозначения) демонстрирует схожую динамику употребления в том, что касается рангового распределения и относительной частоты этих цветообозначений. Что касается менее центральных членов группы, заметную роль в статистике их встречаемости, по-видимому, играют индивидуальные факторы, влияющие на жизнь слова в языке.

7. Выводы

В настоящем исследовании были проанализированы частотные характеристики прилагательных цвета в русских рассказах начала XX и XXI веков. В центре внимания были такие параметры, как динамика относительной частоты и внутригрупповое ранговое распределение.



Было обнаружено, что различия в индивидуально-авторском предпочтении цвета могут быть довольно значительными, но они нивелируются при опоре на репрезентативную выборку, включающую произведения сотен писателей. Внутригрупповое распределение рангов прилагательных цвета остается стабильным на протяжении первой трети XX века. При сравнении его с соответствующим распределением в рассказах начала XXI века можно видеть, что четыре наиболее частотных цвета (*белый, черный, красный и серый*) сохраняют свои позиции, а менее частотные демонстрируют некоторую перегруппировку.

Относительная частота встречаемости цветообозначений в тексте оказалась более вариативным параметром: она демонстрирует заметный рост от первого, дореволюционного, периода (1900–1913) к третьему, советскому (1923–1930). Напротив, при сравнении периодов 1900–1930 и 2000–2023 наблюдается, за редким исключением, существенное снижение. В качестве объяснения отмеченных тенденций было выдвинуто предположение о влиянии внешних факторов (войны, революции, отмены частной собственности на землю, коллективизации, урбанизации) на тематику рассказов и, соответственно, на их лексику.

Отдельное пилотное исследование было посвящено сопоставлению рассказов и стихотворений И. А. Бунина с точки зрения использования прилагательных цвета. Оно выявило существенные различия по обоим используемым параметрам. Это дает основания предположить, что форма художественного произведения является значимым фактором, влияющим на частоту употребления цветообозначений в тексте.

Список литературы

- Бахилина Н. Б.* История цветообозначений в русском языке. М., 1975.
- Василевич А. П., Кузнецова С. Н., Мищенко С. С.* Цвет и названия цвета в русском языке. М., 2005.
- Гребенников А. О., Скребцова Т. Г.* Языковая картина мира в русском рассказе начала XX века // Седьмая научно-практическая конференция «Философия и культура информационного общества»: тез. докл. : в 2 ч. СПб, 2019. Ч. 2. С. 141–143.
- Марголис Я.* От Пушкина до поэтов XXI века. Сопоставительный анализ поэтического языка в цифровую эпоху. М., 2022.
- Масевич А. Ц., Захаров В. П.* Частотное поведение прилагательных цвета в русских поэтических текстах // Вестник НГУ. Сер.: Лингвистика и межкультурная коммуникация. 2019. Т. 17, №1. С. 21–48. doi: 10.25205/1818-7935-2019-17-1-21-48.
- Перехвальская Е. В.* Этнолингвистика : учеб. пособие. М., 2016.
- Проблемы цвета в этнолингвистике, истории и психологии : круглый стол / отв. ред. А. П. Василевич.* М., 2004.
- Скребцова Т. Г.* Новые реалии общественно-политической жизни 1920-х гг. и их отражение в русской литературе и лексикографии (на примере сложносокращенных слов) // Политическая лингвистика. 2021. №2. С. 146–154. doi: 10.12345/1999-2629_2021_02_14.
- Скребцова Т. Г., Гребенников А. О., Шерстинова Т. Ю.* Динамика лексического состава русской художественной прозы (на материале частотных словарей кор-



пуca русских рассказов 1900–1930) // Компьютерная лингвистика и интеллектуальные технологии: матер. ежегодной науч. конф. «Диалог». М., 2021. Вып. 20. С. 646–659.

Фрумкина Р. М. Цвет, смысл, сходство. Аспекты психолингвистического анализа. М., 1984.

Харченко В. К. Словарь цвета: реальное, потенциальное, авторское. М., 2009.

Частотный словарь рассказов А. И. Куприна / авт.-сост. А. О. Гребенников; под ред. Г. Я. Мартыненко. СПб., 2006.

Частотный словарь рассказов А. П. Чехова / авт.-сост. А. О. Гребенников; под ред. Г. Я. Мартыненко. СПб., 1999.

Частотный словарь рассказов И. А. Бунина / авт.-сост. А. О. Гребенников; под ред. Г. Я. Мартыненко. СПб., 2011.

Частотный словарь рассказов Л. Н. Андреева / авт.-сост. А. О. Гребенников; под ред. Г. Я. Мартыненко. СПб., 2003.

Чернейко Л. О., Ли Я. Коннотации непроеизводных прилагательных цвета в аспекте когнитивного анализа их сочетаемости с абстрактными именами существительными // Вестник Московского университета. Сер. 9: Филология. 2019. №5. С. 144–161.

Berlin B., Kay P. Basic color terms: Their universality and evolution. Berkeley, 1969.

Skrebtsova T. Thematic tagging of literary fiction: the case of early 20th century Russian short stories. // Proceedings of the International Conference “Internet and Modern Society” (St. Petersburg, 17–20 June, 2020). CEUR Workshop Proceedings. St. Petersburg, 2021. Vol. 2813. P. 265–276.

Об авторах

Татьяна Георгиевна Скребцова, кандидат филологических наук, доцент, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия.

ORCID ID: 0000-0002-7825-1120

E-mail: t.skrebtsova@spbu.ru

Александр Олегович Гребенников, кандидат филологических наук, доцент, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия; доцент, Университет ИТМО, Санкт-Петербург, Россия.

ORCID ID: 0000-0003-2856-5049

E-mail: a.grebennikov@spbu.ru

Наталья Михайловна Марусенко, кандидат филологических наук, доцент, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия.

ORCID ID: 0000-0002-3347-1373

E-mail: n.marusenko@spbu.ru

Для цитирования:

Скребцова Т. Г., Гребенников А. О., Марусенко Н. М. Цветообозначения в русских рассказах XX–XXI веков: корпусное исследование // Слово.ру: балтийский акцент. 2025. Т. 16, № 1. С. 59–73. doi: 10.5922/2225-5346-2025-1-4.





COLOUR TERMS IN THE RUSSIAN SHORT STORIES
OF THE EARLY XX AND XXI CENTURIES: A CORPUS STUDY

Tatiana G. Skrebtsova¹, Alexander O. Grebennikov^{1, 2}
Natalya M. Marusenko¹

¹ St. Petersburg State University
11 Universitetskaya Nab., St. Petersburg, 199034, Russia

² ITMO University
49 Lit. A, Kronverksky Ave., St. Petersburg, 197101, Russia

Submitted on 12.01.2024

Accepted on 10.11.2024

doi: 10.5922/2225-5346-2025-1-4

The paper presents a comparative analysis of colour terms frequencies in the Russian short stories of the early 20th and 21st centuries. The study draws on the Russian Short Stories Corpus 1900 – 1930, word frequency lists compiled from the stories by particular authors (Ivan Bunin, Alexander Kuprin, Anton Chekhov, Leonid Andreev), and the Russian National Corpus. Two major statistic parameters are used, i. e. relative frequency (ipm) and rank distribution. The study focusses on the most common colour terms. A multifold comparison has revealed differences due to the author's personality, historical context, and literary form. Thus, considerable divergence is shown to exist in the individual writers' preferences for the particular colour terms. However, the differences come down to a unified rank distribution in a sample from the Russian Short Story Corpus 1900 – 1930, containing stories by a few hundred authors. While not affecting the intra-group rankings of colour terms, the turbulent historical events seem to have spurred their ipm index. A hypothesis is put forward to account for this, based on an earlier paper. The analysis of the 21st century material draws on a sample from the Russian National Corpus, containing short stories produced in 2000 – 2023. A comparison between the two samples has shown that the most frequent adjectives have retained their intra-group rankings while the ipm index has largely dropped. Included in the paper is also a comparison between Ivan Bunin's short stories and poetry, exemplifying the impact of literary form on the colour terms' frequencies.

Keywords: colour terms, Russian short stories, linguistic statistics, Russian Short Stories Corpus 1900 – 1930, frequency wordlists, rank distribution

References

Bakhilina, N.B., 1975. *Istoriya tsvetooboznachenii v russkom yazyke* [The history of color terms in the Russian language]. Moscow (in Russ.).

Berlin, B. and Kay, P., 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley.

Frumkina, R.M., 1984. *Tsvet, smysl, skhodstvo. Aspekty psikholingvisticheskogo analiza* [Color, Meaning, Similarity: the Aspects of Psycholinguistic Analysis]. Moscow (in Russ.).

Grebennikov, A.O. and Skrebtsova, T.G., 2019. Linguistic picture of the world in the Russian story of the early XX century. In: *Sed'maya nauchno-prakticheskaya konferentsiya «Filosofiya i kul'tura informatsionnogo obshchestva» : tezisy dokladov* [Proceedings of the 7th Scientific and Practical Conference "Philosophy and Humanities in the Information Society"]. Vol. 2. St. Petersburg, pp. 141 – 143 (in Russ.).

Grebennikov, A.O., ed., 1999. *Chastotnyi slovar' rasskazov A.P. Chekhova* [Frequency dictionary of short stories by Anton Chekhov]. St. Petersburg (in Russ.).



Grebennikov, A.O., ed., 2003. *Chastotnyi slovar' rasskazov L.N. Andreeva* [Frequency dictionary of short stories by Leonid Andreev]. St. Petersburg (in Russ.).

Grebennikov, A.O., ed., 2006. *Chastotnyi slovar' rasskazov A.I. Kuprina* [Frequency dictionary of short stories by Alexander Kuprin]. St. Petersburg (in Russ.).

Grebennikov, A.O., ed., 2011. *Chastotnyi slovar' rasskazov I.A. Bunina* [Frequency dictionary of short stories by Ivan Bunin]. St. Petersburg (in Russ.).

Kharchenko, V.K., 2009. *Slovar' tsveta: real'noe, potentsial'noe, avtorskoe* [The Dictionary of Colour: Actual, Potential, Author's]. Moscow (in Russ.).

Margolis, Ya., 2022. *Ot Pushkina do poetov XXI veka. Sopostavitel'nyi analiz poeticheskogo yazyka v tsifrovuyu epokhu* [From Pushkin to 21st Century Poets: A Comparative Analysis of Poetic Language in the Digital Era]. Moscow (in Russ.).

Masevich, A.Ts. and Zakharov, V.P., 2019. Frequency Behavior of Color Adjectives in Russian Poetic Texts. *Vestnik NGU. Seriya: Lingvistika i mezhkul'turnaya kommunikatsiya* [NSU Vestnik. Series: Linguistics and Intercultural Communication], 17, pp. 21–48, <https://doi.org/10.25205/1818-7935-2019-17-1-21-48> (in Russ.).

Perekhval'skaya, E.V., 2016. *Etnolingvistika: uchebnoe posobie* [Ethnolinguistics: a Textbook]. Moscow (in Russ.).

Skrebtsova, T., 2021a. Thematic tagging of literary fiction: the case of early 20th century Russian short stories. In: *Proceedings of the International Conference "Internet and Modern Society"*. St. Petersburg, 17–20 June, 2020. Vol. 281, pp. 265–276.

Skrebtsova, T.G., 2021b. New Sociopolitical Realities of the 1920s as Reflected in Russian Literature and Lexicography (on the Basis of Syllabic Acronyms). *Politicheskaya lingvistika* [Political Linguistics], 2, pp. 146–154, https://doi.org/10.12345/1999-2629_2021_02_14 (in Russ.).

Skrebtsova, T.G., Grebennikov, A.O. and Sherstinova, T.Yu., 2021. The dynamics of vocabulary in Russian prose (based on frequency dictionaries of the Corpus of Russian short stories 1900–1930). In: *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: materialy ezhegodnoi nauchnoi konferentsii «Dialog»* [Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"]. Vol. 20, pp. 646–659 (in Russ.).

Vasilevich, A.P., 2004. *Problemy tsveta v etnolingvistike, istorii i psikhologii: kruglyi stol* [Problems of Color in Ethnolinguistics, History and Psychology. Panel Discussion]. Moscow (in Russ.).

Vasilevich, A.P., Kuznetsova, S.N. and Mishchenko, S.S., 2005. *Tsvet i nazvaniya tsveta v russkom yazyke* [Color and color terms in Russian]. Moscow (in Russ.).

Cherneyko, L. and Li, Y., 2019. Connotations of Non-Derivative Color Adjectives in the Aspect of Cognitive Analysis of Their Compatibility with Abstract Nouns. *Vestnik Moskovskogo universiteta. Seriya 9: Filologiya* [Moscow University Bulletin. Series 9. Philology], 5, pp 114–161 (in Russ.).

The authors

Dr Tatiana G. Skrebtsova, Associate Professor, St. Petersburg State University, St. Petersburg, Russia.

ORCID 0000-0002-7825-1120

E-mail: t.skrebtsova@spbu.ru

Dr Alexander O. Grebennikov, Associate Professor, St. Petersburg State University, St. Petersburg, Russia; Associate Professor, ITMO University, St. Petersburg, Russia.

ORCID 0000-0003-2856-5049

E-mail: a.grebennikov@spbu.ru



Dr Natalya M. Marusenko, Associate Professor, St. Petersburg State University, St. Petersburg, Russia.

ORCID: 0000-0002-3347-1373

E-mail: n.marusenko@spbu.ru

To cite this article:

Skrebtsova, T.G., Grebennikov, A.O., Marusenko, N.M., 2025, Colour terms in the Russian short stories of the early XX and XXI centuries: a corpus study, *Slovo.ru: Baltic accent*, Vol. 16, no. 1, pp. 59–73. doi: 10.5922/2225-5346-2025-1-4.



SUBMITTED FOR POSSIBLE OPEN ACCESS PUBLICATION UNDER THE TERMS AND CONDITIONS OF THE CREATIVE COMMONS Attribution-NonCommercial 4.0 International Deed (CC BY-NC 4.0) LICENSE ([HTTPS://CREATIVECOMMONS.ORG/LICENSES/BY-NC/4.0/DEED.RU](https://creativecommons.org/licenses/by-nc/4.0/deed.ru))