

«Множественное пейоративное»: семантика и сочетаемость (по данным веб-корпуса)

О.В. Блинова

Санкт-Петербургский государственный университет

o.blinova@spbu.ru

Аннотация

Статья посвящена формам мн. ч. существительных с отрицательной окраской референтов типа ‘всякие пугачёвы и киркоровы’. Обсуждается технология нахождения форм множественного пейоративного в веб-корпусе (RuTenTen). Приводятся результаты анализа сочетаемости таких форм по данным корпуса.

Во-первых, выделяются компоненты именных групп, возглавляемых формами множественного пейоративного (‘все’, ‘все эти’, ‘всякие’, ‘всякие там’, ‘разные’ и др.). Во-вторых, обсуждаются компоненты именных групп с числительными, счётными существительными, групповыми существительными и наречиями типа ‘пятьсот сидоровых’, ‘десяток сидоровых’, ‘банда сидоровых’, ‘много сидоровых’. В-третьих, рассматриваются конструкции, способные поддерживать выражение смыслов «множественность» и «неуникальность» (элективные конструкции, конструкции с ‘типа’, ‘вроде’).

Результаты исследования могут быть использованы для анализа тональности текстов на русском языке, точнее, для извлечения мнений о людях.

Ключевые слова: русский язык, веб-корпус, RuTenTen, параметры поиска, семантика, частные значения числа, имена собственные, пейоративное множественное, сочетаемость, кванторные слова, отрицательное мнение о людях, анализ тональности.

1. Что такое «множественное пейоративное»?

Основное значение формы числа в русском — количественное. Грамматисты выделяют ряд непрямых (частных, вторичных) значений числа. Одним из таких значений является т. наз. «пейоративное множественное» число существительных.

На то, что формы множественного (Pl) могут быть использованы «вместо» форм единственного, обратил внимание ещё А.А. Потехня, указавший на

существование «множественного гиперболического», которое способно передавать значение «несправедливого пристрастия» [1, с. 3–4]. Со времён Потебни о разнообразных непрямах употребления PI написано довольно много, см., например, [2-8].

Представляется, что наиболее полное обзорное описание непрямы значения PI получили у Н.В. Перцова, который говорит, в частности, о «множественном гиперболическом» типа *Ты что это клумбы-PI топчешь?* (речь идёт об одном предмете) и «множественном пейоративного отчуждения», или «отрицательном» типа *Я верчусь как проклятая, а ты по театрам-PI ходишь* [9]. В качестве отдельного под-значения «множественного пейоративного отчуждения» у Н.В. Перцова приведён пример «пейоративного мн. числа» *Там выступают всякие пугачевы-PI и киркоровы-PI*.

Именно формы типа *пугачевы-PI, киркоровы-PI*, образованные от имён собственных, называющие лиц, обнаруживающие «ярко изобразительный характер множественного числа», с «ясно ощущаемой отрицательной окраской референтов словоформ мн. числа» [9, с. 181] и станут предметом рассмотрения в этой статье.

Отмечался «промежуточный» характер PI от имён лиц. Указано, что такие формы находятся «между» именами собственными и именами нарицательными [10]. Особенности семантики форм множественного пейоративного (PI.Pej), возникающие, согласно Перцову, в результате действия коннотации «ограниченное количество ~ большое количество», должны сказываться на сочетаемости, делая возможными комбинации с разнообразными кванторными словами.

Целью настоящей статьи является корпусное исследование состава именных групп, в которые входит форма PI.Pej. Представляется, что такое рассмотрение позволит выделить стабильные, часто встречающиеся лексические и другие средства, поддерживающие коннотацию «ограниченное количество ~ большое количество» и значение неуникальности референтов именных групп с формами типа *киркоровы-PI*.

2. Поиск форм PI.Pej в веб-корпусе

Рассмотрение форм множественного пейоративного (PI.Pej) выполнено по результатам поиска в веб-корпусе RuTenTen из семейства TenTen [11] объёмом 14,5 млрд. слов.

Были использованы параметры поиска, предложенные для форм типа *чубайсы* в спецификации «MULTEXTEast Morphosyntactic Specifications, Version 4» [12], а именно запросы с тегом «с» — «common» на месте тега грамматического рода. Результаты поиска по запросу [tt_tag="Нрспну"] с указанием условного «общего» рода и падежа (в данном случае — именительного) хорошо соответствовали задачам рассмотрения форм PI.Pej; такой поиск был повторён для всех падежных форм. Поиск происходил по подкорпусу «Russian_domain.ru».

Решено было искать отдельные падежные формы, поскольку для дальнейшего анализа важна возможность использовать как можно более полные частотные списки словоформ, а TenTen позволяет скачивать списки не длиннее

1000 строк. Частотный же список словоформ, полученный для результатов поиска «Npcp.y» (в меню «Frequency», «Node Forms» > «Save», скачивается список, отсортированный по частоте) недостаточно полно отражает набор форм Pl.Pej.

Таблица 1. Результаты поиска по грамматическим тегам

MSD	запрос	описание	результат
Npcpny	tt_tag [tt_tag="Npcpny"]	Noun Type=proper Gender=common Number=plural Case=nominative Animate=yes	145,310
Npcpgy	tt_tag [tt_tag="Npcpgy"]	Case=genitive	199,509
Npcpdy	tt_tag [tt_tag="Npcpdy"]	Case=dative	1,175
Npcpay	tt_tag [tt_tag="Npcpay"]	Case=accusative	9,882
Npcpiy	tt_tag [tt_tag="Npcpiy"]	Case=instrumental	3,513
Npcply	tt_tag [tt_tag="Npcply"]	Case=locative	623
Npcp.y	tt_tag [tt_tag="Npcp.y"]	Case=any	360,012

В выдаче (особенно в Nom и Gen) присутствует существенное количество контекстов, которые нас не интересуют. Прежде всего, это употребления форм т. наз. множественного ассоциативного типа *Кузнецовы*, а также разнообразные употребления типа *братья Кузнецовы*, *Маша и Петя Кузнецовы*, *семья Кузнецовых*, *род Голицыных*, *Романовых* и др.

Для выделения релевантных результатов было решено рассмотреть частотные списки словоформ, полученные для результатов поиска (см. табл. 1) и сформировать списки написаний со строчной буквы. Выяснилось, что со строчной буквы пишутся в основном общеизвестные фамилии. Прежде всего, это фамилии очень богатых бизнесменов («олигархов»), политиков и чиновников, учёных, террористов.

При дальнейшем поиске по словоформам с сохранением написания по запросам типа «ломоносовы», «эйнштейны» выяснилось, что употребления Pl от фамилий учёных не содержат никаких негативных коннотаций, ср.: *В программе мероприятий — форум талантливой молодежи, в котором примут участие юные ломоносовы из всех уголков страны; Конечно, михайлы ломоносовы из донских сел — большая редкость, Теперь будущие ломоносовы, менделеевы и пушкины станут ежемесячно зарабатывать по 1100 рублей в месяц; Будущие эйнштейны съехали в Новосибирск на Всероссийскую олимпиаду по физике.*

В остальном (за исключением учёных) частотные списки результатов поиска по запросу [tt_tag="Npcp.y"] в части написаний со строчной буквы содержат именно употребления формы Pl.Pej. Сказанное коррелирует с наблюдениями И.И. Ревзина о нейтрализации противопоставления единственности — множественности и «журналистских приёмах» советских газет с унижительными наименованиями типа *зоценки*, *юнгманы* [13, с. 107].

В результате поиска по отдельным словоформам с указанием части речи и сохранением написания типа «сидоровы N.*» было получено 4105 фраз, из которых 3730 фраз содержали употребления Pl.Pej. Контексты были отобраны после скачивания выдачи в формате KWIC (ключевое слово в контексте), ср.

пример: *нам еще долго не видать победы. Все эти «<киркоровы >» и прочие только губят нашу культуру.*

3. Анализ сочетаемости формы Pl.Rej

Согласно [14], имя собственное — это «слово, словосочетание или предложение, которое служит для выделения именуемого им объекта из ряда подобных, индивидуализируя и идентифицируя данный объект». При формах Pl.Rej, по-видимому, должны систематически появляться элементы контекста, указывающие на неединичность, неуникальность объектов, обозначенных с помощью этих форм.

В настоящей статье рассматриваются компоненты именных групп с формами Pl.Rej, учитываются прежде всего элементы левого контекста.

Все реальные фамилии в примерах ниже заменены на *ивановых, петровых, сидоровых*. Формы Pl.Rej в выборке из корпуса чаще входят в длинные сочинённые именные группы типа *все эти ивановы, петровы, сидоровы, кузнецовы, смирновы и другие васильевы* или комитативно-сочинённые именные группы типа *какие-нибудь ивановы с петровыми и прочие сидоровы*. Количество перечисляемых форм Pl.Rej в примерах сохранено (прежде всего — для наглядной демонстрации громоздкости именных групп).

Целью настоящей статьи является не синтаксический анализ фраз с формами Pl.Rej, представляющий отдельный интерес, и не анализ употребительности таких форм, а первоначальная фиксация некоторых элементов контекста, способных употребляться при именах лиц и поддерживать коннотацию «ограниченное количество ~ большое количество».

3.1. Компоненты именных групп, возглавляемых формой Pl.Rej

В качестве компонентов именных групп типа AN, где N — форма Pl.Rej в любом падеже, а A — атрибут в препозиции к ней, употребляются прежде всего местоимение всеобщности *все*, местоимения и прилагательные, «указывающие на произвольный объект множества» [15] *всякие, всякие там, всяческие*, прилагательные *разные, разнообразные, различные, всевозможные*, а также указательные местоимения, возможности пейоративного употребления которых в русском в литературе упоминались [16], ср. табл. 2.

Перечисленные элементы способны образовывать цепочки из двух, трёх и более компонентов, ср.: *все эти* (174 употребления), *все ваши, всякие эти, всякие разные, всякие прочие, прочие разные, эти самые, те же самые, эти гадкие, свои примитивные, всевозможные бесконечные, всё новые и новые, все и всяческие; все эти ваши, все эти ряженые, какие-нибудь выскочки* и мн. др.

3.2. Компоненты именных групп с кванторными словами типа 'десяток сидоровых'

В именных группах типа NumN, NgroupN.GEN, AdvN.GEN, где N — это форма падежа от Pl типа *сидоровы*, Num — это количественное или собирательное числительное, кванторное существительное типа *сотня, пяток, пара*, Ngroup — это групповое существительное типа *банда, армия, клика*, Adv

— это наречие или наречное выражение типа *много, тьма тьмуца*, употребляются прежде всего компоненты, перечисленные в табл. 3.

Таблица 2. Компоненты именных групп типа AN ‘*всякие сидоровы*’

атрибут (кол-во)	пример употребления
всякие (#114)	если опять не придут всякие ивановы и петровы
все (#43)	даже просто перечислить всех «ивановых» в России не хватит времени
эти (#37)	видели мы этих ивановых и петровых
свои (#33)	вы своих сидоровых с Родиной не равняйте [sic]
разные (#30)	как восторженно выражаются разные ивановы, петровы, сидоровы и другие платные певцы
всякие там (#22)	фамилии всяких там ивановых и петровых и даже сидоровых стали «никами»
всевозможные (#13)	изгнанию всевозможных ивановых собственного разлива
никакие (#11)	и кстати никаких сидоровых я не подерживал
те же (#8)	снова те же ивановы, петровы, сидоровы, кузнецовы
другие (#5)	в отличие от других сидоровых решил вложиться в российский спорт
разнообразные (#4)	вскармливания разнообразных "ивановых" и "петровых"
какие-нибудь (#3)	каких-нибудь петровых или сидоровых
различные (#3)	колоссального роста благосостояния различных ивановых, петровых, сидоровых
очередные (#3)	в руках очередных петровых
всяческие (#2)	хулиганским действиям всяческих ивановых
бесчисленные (#2)	бесчисленные ивановы, петровы, сидоровы, кузнецовы васильевы
ваши (#2)	ваши ивановы - петровы - сидоровы
какие-нибудь там (#1)	воля каких-нибудь там ивановых и петровых
разные там (#1)	допустить до власти разных там ивановых и петровых;
всяческие там (#1)	всяческих там сидоровых
многочисленные (#1)	законодательства, позволяющего многочисленным «ивановым» укрываться от налогов
бесконечные (#1)	бесконечные "ивановы" и "петровы"
всё те же (#1)	да всё те же ивановы, петровы, сидоровы и кузнецовы-васильевы
такие же (#1)	куча таких же сидоровых сосут кровь

В сочетаниях типа NgroupN.GEN встречаются также следующие групповые существительные: *ватага, воровская клика, клан, клика, когорта, команда, узкая группа, шобла, группировка, конгломерат, круг, куча, небольшая когорта, небольшая олигархическая группировка, неприкасаемая каста, нынешняя колода, толпа* и др.

Таблица 3. Компоненты именных групп типа NumN, NgroupN.GEN, AdvN.GEN

компонент (кол-во)	пример употребления
десяток (#4)	в расстреле десятка сидоровых ЕСТЬ ТОЛК; в руках десятка ивановых и петровых
пара (#2)	пара "удачливых" ивановых
с десяток (#2)	и посадит в тюрьму с десяток кузнецовых
сколько (#2)	а сколько таких «ивановых» у нас в России
столько (#1)	там столько ивановых — от солдата до генерала
два (#1)	двух ивановых и пары сидоровых достаточно
парочка (#1)	расстреляет парочку ивановых
трое (#1)	еще трое иванов ивановых
тысячи (#1)	это не парад из тысяч ивановых и петровых, это-народные массы
двадцать пять (#1)	да его 25 кузнецовых затопчут в секунду
десятки (#1)	на благодатной почве перестройки десятки ивановых, петровых, сидоровых, кузнецовых и васильевых
два и более (#1)	двух и более ивановых наша страна вряд ли выдержит
сотни (#1)	на его месте появится [sic] сотни таких же петровых
сотня (#1)	расстрельная сотня сидоровых знает об этом
триста тысяч (#1)	прибежище 300 тысяч российских «ивановых», которых он фактически и породил
не двести двадцать пять (#1)	не двести двадцать пять петров ивановых
пятьсот (#1)	500 "ивановых" или 100 000 специалистов среднего уровня
двести сорок (#1)	240 петровых нарисовал на одной картине
тысячи и тысячи (#1)	о тысячах и тысячах тогдашних ивановых, петровых и сидоровых
двадцать семь (#1)	если взять капиталы всех 27 наших ивановых и раздать поровну всему населению
пяток (#1)	пяток лесных «сидоровых», и будет очень крупный передел
более сотни (#1)	таких петровых по России более сотни
много (#4)	после 2007 года будет много " кузнецовых"
кучка (#9)	в карманы кучки одуревших от бабла ивановых и петровых
свора (#5)	«чисто» бандитской своре ивановых/петровых
компания (#3)	в компании всех этих ивановых и прочих петровых
сонм (#2)	появление сонма ивановых, петровых, сидоровых, кузнецовых, васильевых, смирных
банда (#2)	Только за банду ивановых печётся!

3.3. Конструкции с формами Pl.Rej

Ещё одним средством выражения смыслов «множественности» и «неуникальности» являются конструкции. Ниже приводятся две наиболее употребительные.

Во-первых, это элективная конструкция с предлогом *из* типа *многие из N.GEN, один из N.GEN, кое-кто из N.GEN, некоторые из N.GEN, никто из N.GEN*, см. примеры: *кое-кого из ивановых и петровых не повесили; современной деятельности кое-кого из сидоровых и кузнецовых, хоть один из*

наших "петровых" вместо яхты или дворца; многие из «сидоровых» начнут пошатривать налево уже не на словах.

Ср. также конструкции с предлогом *среди* типа *не наблюдаю среди ивановых, петровых и сидоровых хотя бы одного профессора Преображенского*.

Во-вторых, это конструкция, выражающая принадлежность к определённом классу (разновидности) объектов с типа N.GEN, *вроде N.GEN, в виде N.GEN* см. примеры: *«великие государственнички» типа ивановых, петровых, сидоровых, кузнецовых, смирных, васильевых; соблазнитель — совратитель типа нынешних сидоровых; вплоть до недоумков типа ивановых; избавится от шушеры типа сидоровых; пиявки, типа ивановых и петровых; пиявок на теле общества вроде разных кузнецовых и смирных; взрастить новую поросль в виде ивановых, петровых, сидоровых и пр.*

Наконец, можно упомянуть употребления оборотов типа *разного рода, всякого рода, всех мастей*, например, такие, как: *всякого рода ивановых, петровых, сидоровых и прочих угнетателей трудящихся России; символы ясны всякого рода сидоровым, кузнецовым, смирным и прочим проводникам и сторонникам капитализма*. Эти выражения в смысловом отношении близки выражениям с местоимениями и прилагательными типа *всякие, разные, разнообразные, различные, всяческие*.

4. Заключение

В настоящей статье показаны некоторые элементы контекста, употребляющиеся вместе с формами Pl.Rej. Эти элементы контекста различны по своему синтаксическому поведению, но способны передавать и поддерживать значения множественности (многочисленности), разнообразия, неуникальности объектов, обозначаемых формами Pl.Rej. Формы множественного пейоративного сочетаются прежде всего с кванторными словами — определительными и другими местоимениями, числительными, счетными существительными, групповыми существительными (наименованиями групп типа *банда, свора*), словами со значением количества типа *много*. Кроме того, они встраиваются в элективные конструкции и конструкции с *типа* и *вроде* (*шушера типа сидоровых*). Наконец, рассматриваемые формы встречаются рядом с оборотами *разного рода, всякого рода, всех мастей*.

В статье рассмотрены элементы левого контекста, в то время как в правом также наблюдаются разнообразные компоненты громоздких сочинённых именных групп, ещё более распространяющие их. Это выражения типа *и прочие ивановы, и другие сидоры, и тому подобные кузнецы, и прочее жульё, и всяческая иная дрянь, а также и пр., и т.д., и т.п., etc, и Ко, и прочая, и компания, сотоварищи* [sic]. Такие выражения передают те же смыслы множественности и неуникальности.

Итак, в статье на выборке фраз с формами множественного пейоративного определены элементы контекста, сопровождающие эти формы. Можно думать, что эти элементы контекста будут способны диагностировать наличие форм Pl от имён лиц с негативной оценкой референтов. Это утверждение будет проверяться в дальнейшем на материале русских корпусов. Результаты работы

могут быть использованы для анализа тональности текстов на русском языке, точнее, для извлечения мнений о людях.

Литература

- [1] Потехина А.А. Значения множественного числа в русском языке. Отдельный оттиск из «Филологических записок». Воронеж, 1888.
- [2] Арбатский Д.Л. Множественное число гиперболическое // Русский язык в школе. 1972. № 5.
- [3] Красильникова Е.В. Имя существительное в русской разговорной речи: функциональный аспект: М., 1990.
- [4] Санников В.З. Русский синтаксис в семантико-прагматическом пространстве. М., 2008.
- [5] Санников В.З. Русский язык в зеркале языковой игры. М., 2002.
- [6] Ляшевская О.Н. Семантика русского числа. М., 2004.
- [7] Князев Ю.П. Грамматическая семантика. Русский язык в типологической перспективе. М., 2007.
- [8] Крысин Л.П. Гипербола в русской разговорной речи // Русское слово, свое и чужое: Исследования по современному русскому языку и социолингвистике. М., 2004. С. 244-261.
- [9] Перцов Н.В. Инварианты в русском словоизменении. М., 2001.
- [10] Пеньковский А. Очерки по русской семантике. М., 2004.
- [11] Russian Web 2011 (ruTenTen1). Corpus description URL: <https://www.sketchengine.co.uk/rutenten-corpus/> (дата обращения 01.04.2017)
- [12] MULTEXT-East Morphosyntactic Specifications, Version 4. 3.10.3. Russian Noun. URL: <https://www.sketchengine.co.uk/russian-tagset/> (дата обращения 01.04.2017).
- [13] Ревзин И.И. Так называемое «немаркированное множественное число» в современном русском языке // Вопросы языкознания. 1969. №3. С. 102-110.
- [14] Лингвистический энциклопедический словарь. М., 2002.
- [15] Левин Ю.И. О семантике местоимений // Проблемы грамматического моделирования. М., 1973.
- [16] Красавина О.Н. Употребление указательной группы в русском повествовательном дискурсе // Вопросы языкознания. 2004. №3. С. 51-68.

«Pejorative plural» in Russian: semantics and co-occurrence patterns (web corpus-based study)

O.V. Blinova

Saint Petersburg State University

The paper discusses plural forms of Russian nouns (in particular, of the surnames) like ‘vsjakiye pugachevy-Pl i kirkorovy-Pl’ (various Pugachevas and Kirkorovs), expressing negative opinion about the referents. Firstly, the ways of searching for such forms of “pejorative plural” (Pl.Pej) in the web corpus (RuTenTen) were considered. Secondly, the co-occurrence patterns of the Pl.Pej forms by corpus data were revealed. The conclusion is that forms of pejorative plural foremost fit together with universal quantifiers including ‘all’, ‘all of these’ etc., and can easily integrate in miscellaneous quantificational expressions, e.g., combinations with numerals, collective nouns, and expressions that include number words like ‘mnogo’ (many). The results of the study could be used for the mining negative opinions about people.

Keywords: Russian, webcorpus, RuTenTen, query parameters, semantics, secondary meanings of plural, proper names, pejorative plural, noun phrase, co-occurrence, quantifiers, sentiment analysis