

Конкуренция русских приименных посессивов в диахронической перспективе

Антон Бузанов, НИУ ВШЭ

27 ноября 2021 г.



HSE
University

План

1. Введение
2. Русские применные посессивы
 1. Данные
 2. Разметка
 3. Результаты
3. Обсуждение и выводы

Введение

- (1) Одна только **спальня принцессина** осталась пощаждённой...
- (2) — Этот сонливый обитатель **фараонова царства** ничего нам не сделает...
- (3) Оказалось, что кто-то похитил волшебный **амулет принцессы**.
- (4) Образно говоря, пасти **стада фараона?**

Обзор

- ▶ Разные способы выражения казалось бы одного значения → разные формальные признаки единиц, в них используемых?
- ▶ Течение времени → изменение свойств конструкции?
- ▶ Разделение конструкций на прототипические и не прототипические? [Van der Auwera и Gast 2010; Lander и Tyshkevich 2015]

Обзор

- ▶ **Гипотеза₀**: стратегии как-то распределены
- ▶ **Гипотеза₁**: притяжательные прилагательные тяготеют к прототипическим контекстам → диахронически более стабильны

- ▶ прототипическое отношение посессивности [Taylor 2019]
 - ▶ буквальное владение
 - ▶ одушевлённый посессор
 - ▶ референтный посессор
 - ▶ активированным в дискурсе посессор
 - ▶ не активированный в дискурсе объект обладания

Посессивные конструкции

Посессивные прилагательные:

- ▶ -ИН
- ▶ -ОВ

vs

- ▶ гЕНИТИВ

Из рассмотрения исключены другие прилагательные:

- ▶ -j (*старуший дом, лисья нора*)
- ▶ -ск (*президентские выборы, пушкинская карта*)

Данные

НКРЯ (www.ruscorpora.ru), снятая омонимия

- ▶ посессивные прилагательные
- ▶ генитивные именные группы

1800—настоящее время

проблема: *стулова ножка vs ножка стула

решение: предварительная разметка посессивных прилагательных

Фильтрация

▶ устойчивые сочетания

(5) — Что же я этому **сукину сыну** сделал? — спросил я запальчиво.

(6) А всё этот **чёртов Василий Петрович!**

▶ топонимы

(7) Вышел на **Куликово поле**

(8) **Ляхово болото** тянется на пять вёрст

Разметка

- ▶ год + период с шагом в 20 лет
- ▶ признаки посессора
 - ▶ собственность/нарицательность
 - ▶ одушевлённость (human, non-human animate, inanimate)
- ▶ признаки объекта обладания
 - ▶ грамматическое отношение
 - ▶ реляционность имени
 - ▶ вес
 - ▶ число

Два этапа разметки

1. полностью ручная разметка (притяжательные прилагательные)
2. полуавтоматическая разметка (генитивные ИГ)
 - ▶ лемматизация с помощью пакета `rumorphy` [Korobov 2015]
 - ▶ разметка грамматических отношений с помощью библиотеки `deeppavlov` [Burtsev и др. 2018]
 - ▶ определение веса объекта обладания

Два датасета

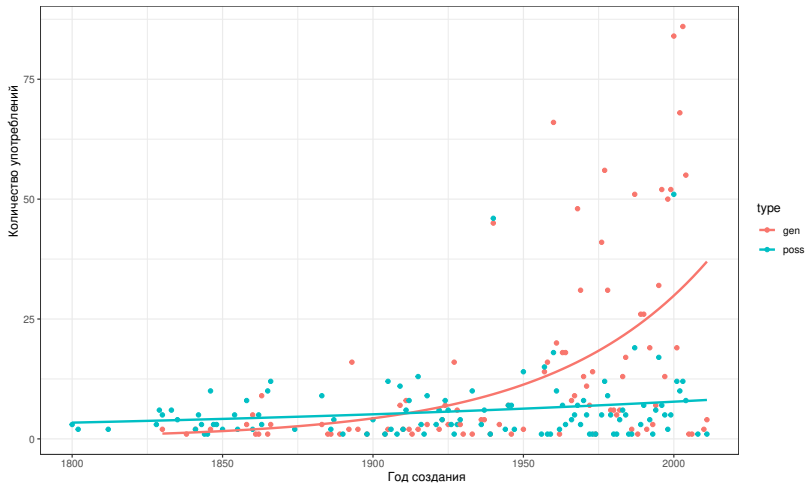
Большой датасет

- ▶ все размеченные данные
- ▶ 627 притяжательных прилагательных
- ▶ 1284 генитивные именные группы

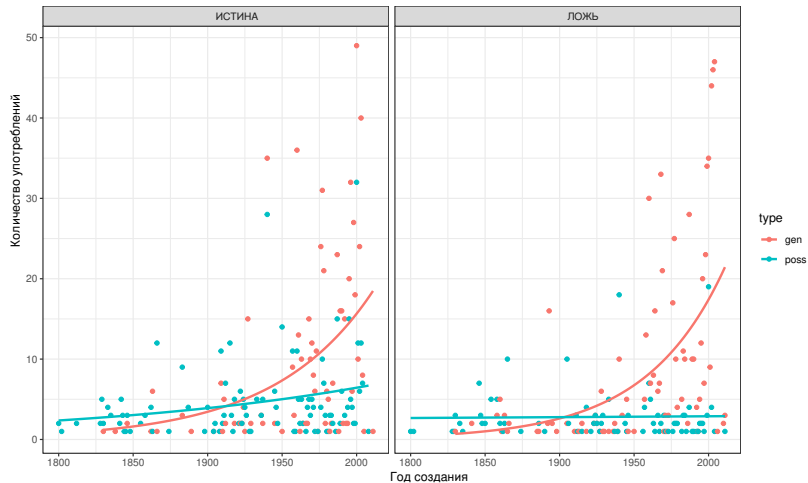
Маленький датасет

- ▶ часть размеченных данных
- ▶ 627 притяжательных прилагательных
- ▶ 504 генитивные ИГ (случайная подвыборка)

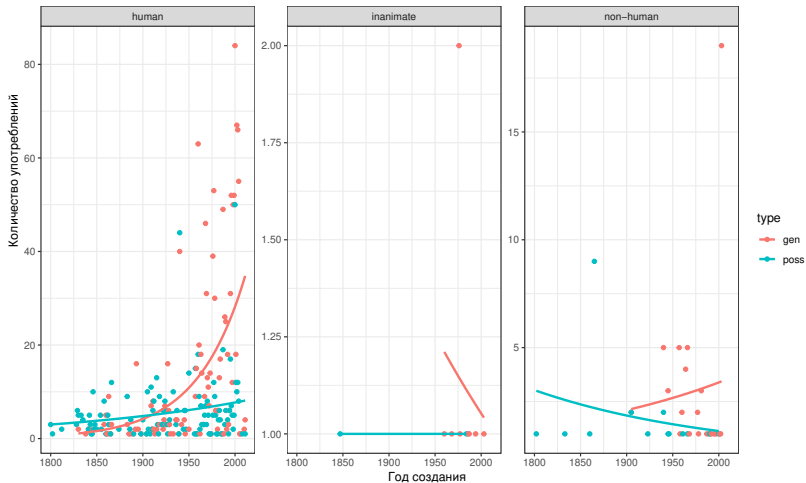
Распределение размеченных примеров



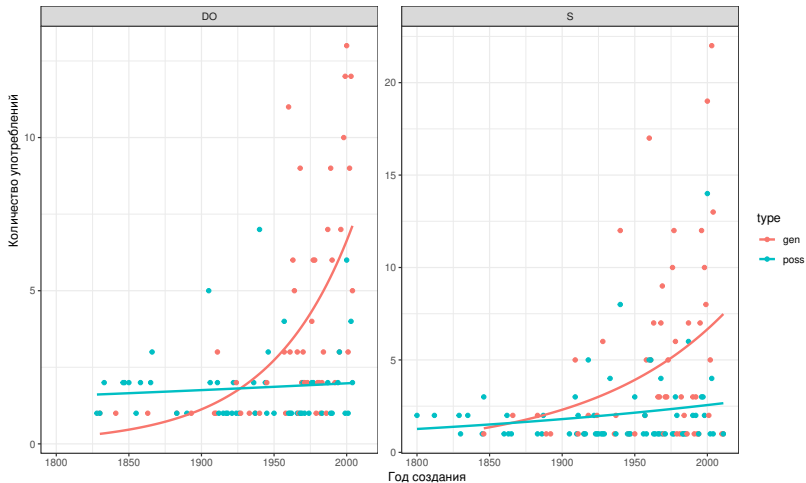
Собственность посессора



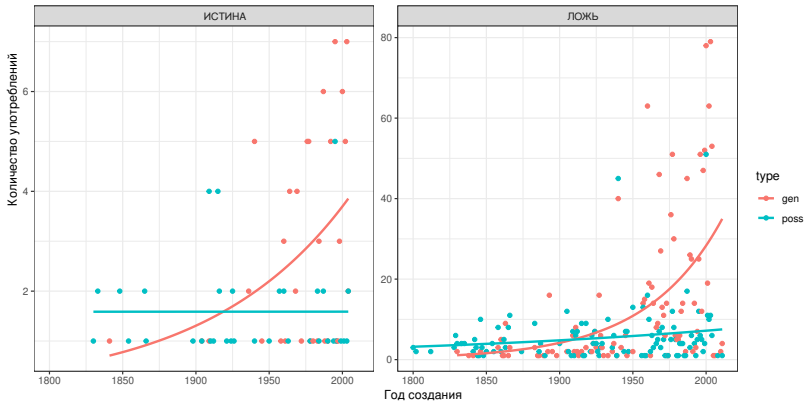
Одушевлённость посессора



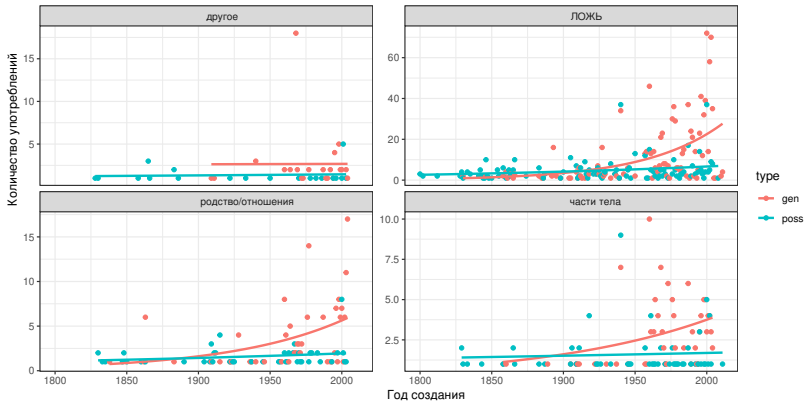
Грамматические отношения



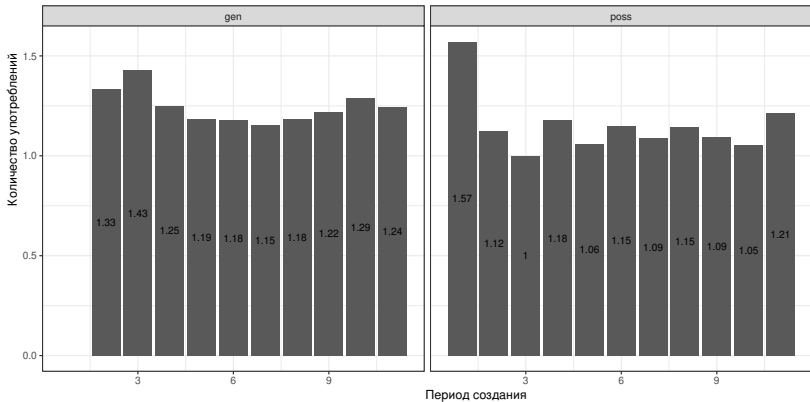
Предикативная позиция



Реляционность объекта обладания



Средний вес объекта обладания по периодам

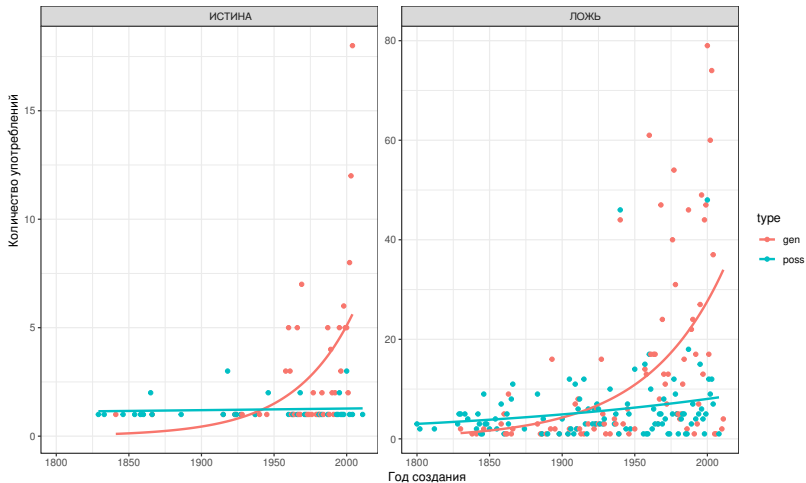


Средний вес объекта обладания. t-тест

Период	p-value
2	0,30029004
3	0,040195489
4	0,316070485
5	0,098178188
6	0,392089033
7	0,166389588
8	0,300904823
9	0,003966831*
10	5,92255E-06*
11	0,293463694

* отмечает p-value < 0.01

Распределение множественного объекта обладания



Регрессия: большой датасет

	<i>Dependent variable:</i>
	type
Proper	0.801 (0.609, 0.993) p = 0.000***
Weight	-0.404 (-0.618, -0.189) p = 0.002***
Created	-0.020 (-0.022, -0.018) p = 0.000***
Constant	38.226 (33.925, 42.527) p = 0.000***
Observations	1,911

Note: *p<0.1; **p<0.05; ***p<0.01

Регрессия: маленький датасет

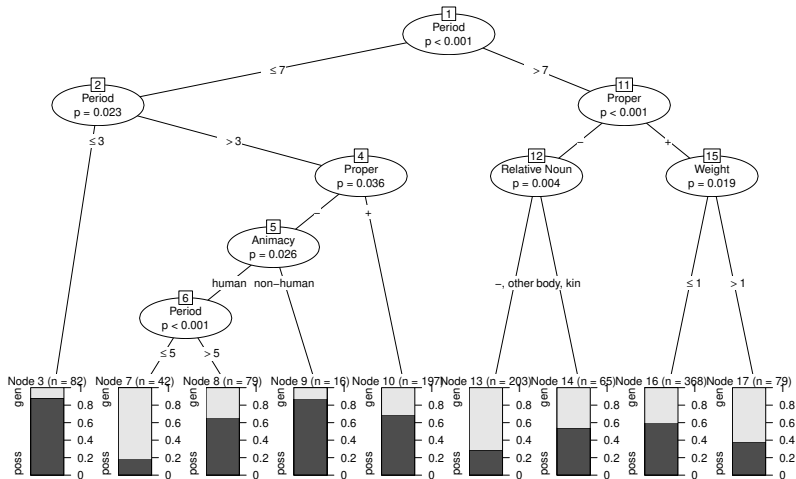
	<i>Dependent variable:</i>
	type
Proper	0.753 (0.534, 0.972) $p = 0.00000^{***}$
Weight	-0.329 (-0.570, -0.087) $p = 0.026^{**}$
Period	-0.163 (-0.207, -0.119) $p = 0.000^{***}$
Constant	1.431 (0.953, 1.909) $p = 0.00000^{***}$
Observations	1,131

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

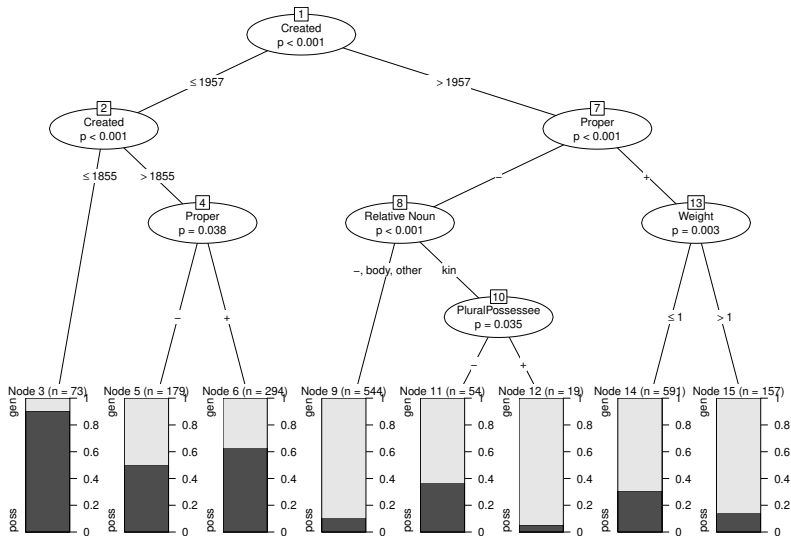
Критерий согласия Пирсона и тест Фишера

Переменная	p-value
Собственность	3.581E-13
Вес	0.0004998
Период	0.0004998
Одушевлённость	0.2669
Множественное обладаемое	0.0266
Предикативность	0.4265
Реляционность имени	0.1282

Решающее дерево: маленький датасет



Решающее дерево: большой датасет



Выводы

Выбор типа посессивной конструкции зависит от следующих факторов:

1. период — с течением времени количество посессивных прилагательных уменьшается
2. собственность посессора — для собственных видим большее число притяжательных прилагательных
3. вес — более лёгкие обладаемые более склонны к притяжательным прилагательным, чем более тяжёлые
4. реляционность имени — для не имён родства посессив используется значительно реже, чем для имён родства
5. число обладаемого — гораздо чаще посессивное прилагательное появляется с единичным посессором

Литература I

- [1] Mikhail Burtsev и др. “Deeppavlov: Open-source library for dialogue systems”. В: *Proceedings of ACL 2018, System Demonstrations*. 2018, с. 122–127.
- [2] Mikhail Korobov. “Morphological Analyzer and Generator for Russian and Ukrainian Languages”. English. В: *Analysis of Images, Social Networks and Texts*. Под ред. Mikhail Yu. Khachay и др. Т. 542. Communications in Computer and Information Science. Springer International Publishing, 2015, с. 320–332. ISBN: 978-3-319-26122-5. DOI: 10.1007/978-3-319-26123-2_31. URL: http://dx.doi.org/10.1007/978-3-319-26123-2_31.
- [3] Yury Lander и Natalia Tyshkevich. “True, liminal and fake prototypes in syntactic typology”. В: *Типология морфосинтаксических параметров*. 2015, с. 185–199.

Литература II

- [4] John R Taylor. "Prototype effects in grammar". В: *Cognitive Linguistics-Key Topics* (2019), с. 127.
- [5] Johan Van der Auwera и Volker Gast. "Categories and prototypes". В: *The oxford handbook of linguistic typology*. 2010.