

Действительно ли глаголы сложно заимствовать?

Тимофей Архангельский

Universität Hamburg

timarkh@gmail.com

Исследование было поддержано стипендией Александра фон Гумбольдта

XVI конференция по типологии и грамматике для молодых учёных, СПб, 21-23.11.2019

План

- Считается, что глаголы заимствовать сложнее, чем существительные
- Но в доказательствах есть проблемы
- Наблюдаемая синхронная картина может быть следствием другого явления – живучести глаголов (а может и не быть)
- Хотя более аккуратный взгляд на русские данные не фальсифицирует эту гипотезу

Заимствования и части речи

- [Whitney 1881: 19]: шкала; глаголы заимствовать сложнее, чем существительные (без доказательств)
- [Moravcsik 1975]: глаголы сложнее заимствовать, потому что они требуют специальной адаптации (маленькая смещённая выборка)
 - [Wohlgemuth 2009]: это совсем не всегда так. Может быть, это просто иллюзия?

Заимствования и части речи

- [Haspelmath, Tadmor 2009]: World Loanword Database (41 язык, 1460 значений)
- [Tadmor et al. 2010]: таки да, глаголы сложнее заимствовать!
 - Эмпирическое доказательство: среди существительных средняя доля доказуемых заимствований в выборке WOLD равна $1/3$, а среди глаголов – всего $1/6$

Глаголы: проблема

- **Логическая ошибка!**
- Выводы о диахронии делаются на основе синхронных данных
- Но такое же синхронное распределение могло быть получено иным путём

Альтернативное объяснение

- Допустим, что процесс заимствования можно смоделировать так:
 - возникает необходимость ввести обозначение для нового понятия или заменить старое
 - новая лексема либо заимствуется, либо нет
 - если глаголы сложнее заимствовать, то новые глаголы реже будут заимствованиями, чем новые существительные

Альтернативное объяснение

- Допустим, что глаголы легко заимствовать
 - Например, у всех новых слов вероятность оказаться заимствованием равна 50%
- Но при этом они просто более живучи и реже заменяются / появляются
 - Например, за один век исчезает 30% существительных, но только 15% глаголов
- Допустим, в момент времени t_0 в языке L нет заимствований

Альтернативное объяснение

- Заимствования в t_0
 - N: 0%
 - V: 0%
- Заимствования в $t_0 + 100$
 - N: 15%
 - V: 7,5%
- Заимствования в $t_0 + 200$
 - N: 25,5%
 - V: 13,9%

Альтернативное объяснение

- А вдруг так и есть?
- Сейчас посмотрим:
 - основной корпус НКРЯ (283 млн словоупотреблений, тексты с конца XVII в.)
 - размеченный вручную список заимствований (48 тыс. существительных и глаголов) – используется только во второй части доклада

Часть 1: вероятность исчезновения

Вероятность исчезновения

- Сколько глаголов и сколько существительных в среднем исчезает из текстов через t лет?
- Ответ явно разный для лексем разной частотности, поэтому считать нужно отдельно для каждого частотного ранга r
 - Частотный ранг 1 у самой частотной лексемы, 2 – у второй по частотности и т. д.

Вероятность исчезновения

- Например, какова вероятность для глаголов ранга 1000 исчезнуть за 50 лет?
- Делим корпус на десятилетия: 1750-1759, 1755-1764, 1760-1769, ..., 2000-2009
 - на самом деле каждое десятилетие тоже делится на части, чтобы обеспечить сравнимый размер частей

Вероятность исчезновения

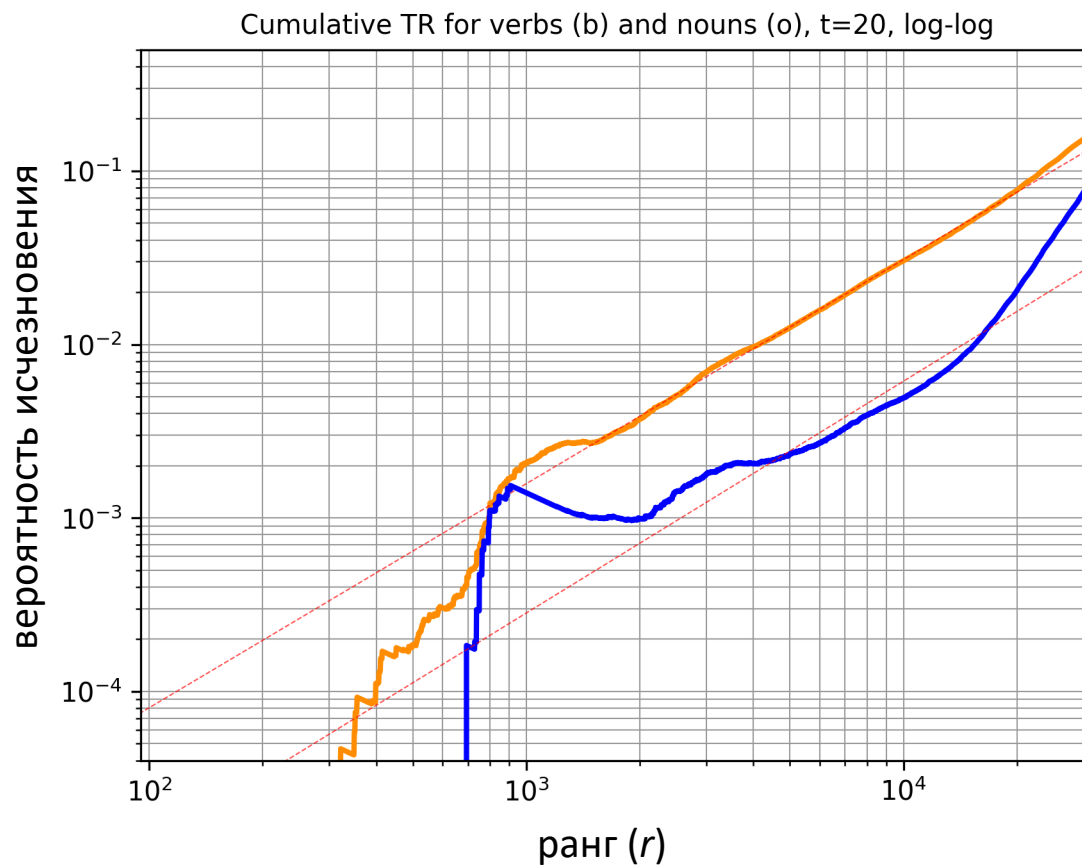
- Смотрим, какой глагол имел ранг 1000 в десятилетии 1750. Присутствует ли он в десятилетии 1800?
- Если исчез, запоминаем это
- Повторяем операцию для других пар (1755-1805, 1760-1810 и т. д.)
- Делим число исчезновений на число рассмотренных пар – получается усреднённая вероятность исчезновения

Вероятность исчезновения

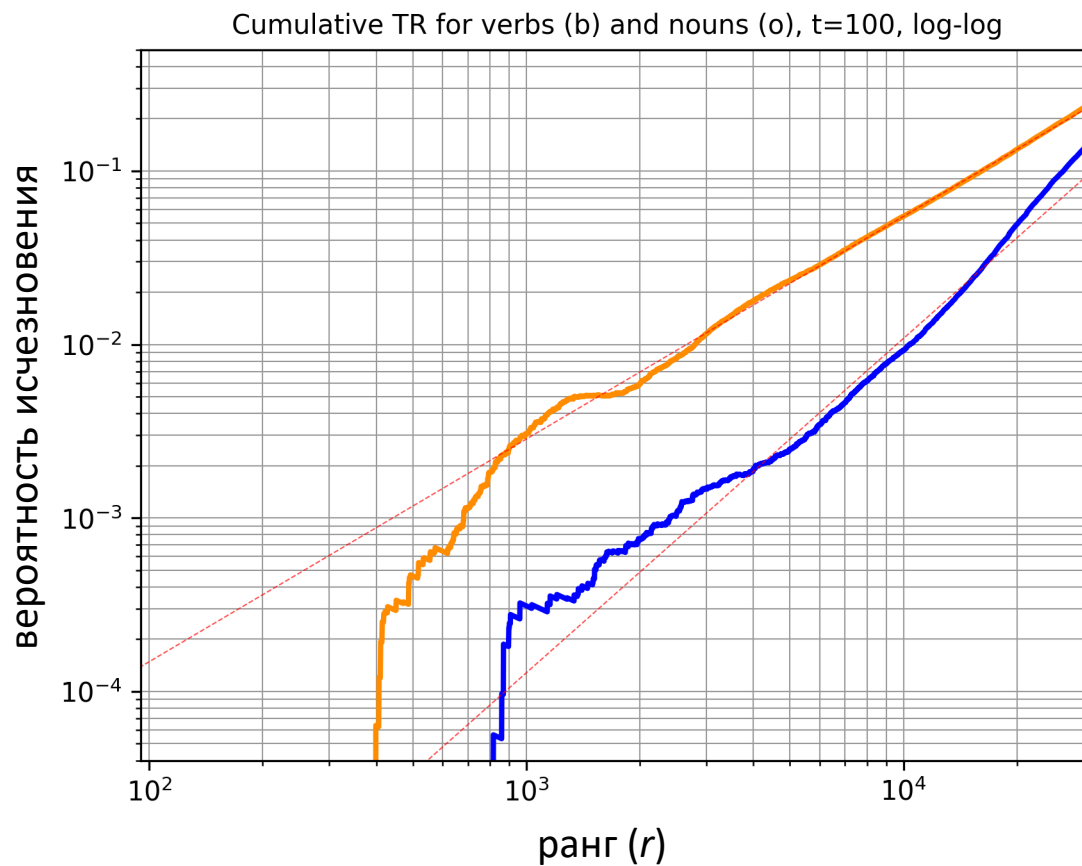
- Рисовать удобнее кумулятивную вероятность:

$TR_c(C, r, t) = P(\text{лексема } L \text{ не встречается в десятилетии } t_0+t \mid L \text{ имеет часть речи } C \text{ и её ранг меньше или равен } r \text{ в десятилетии } t_0)$

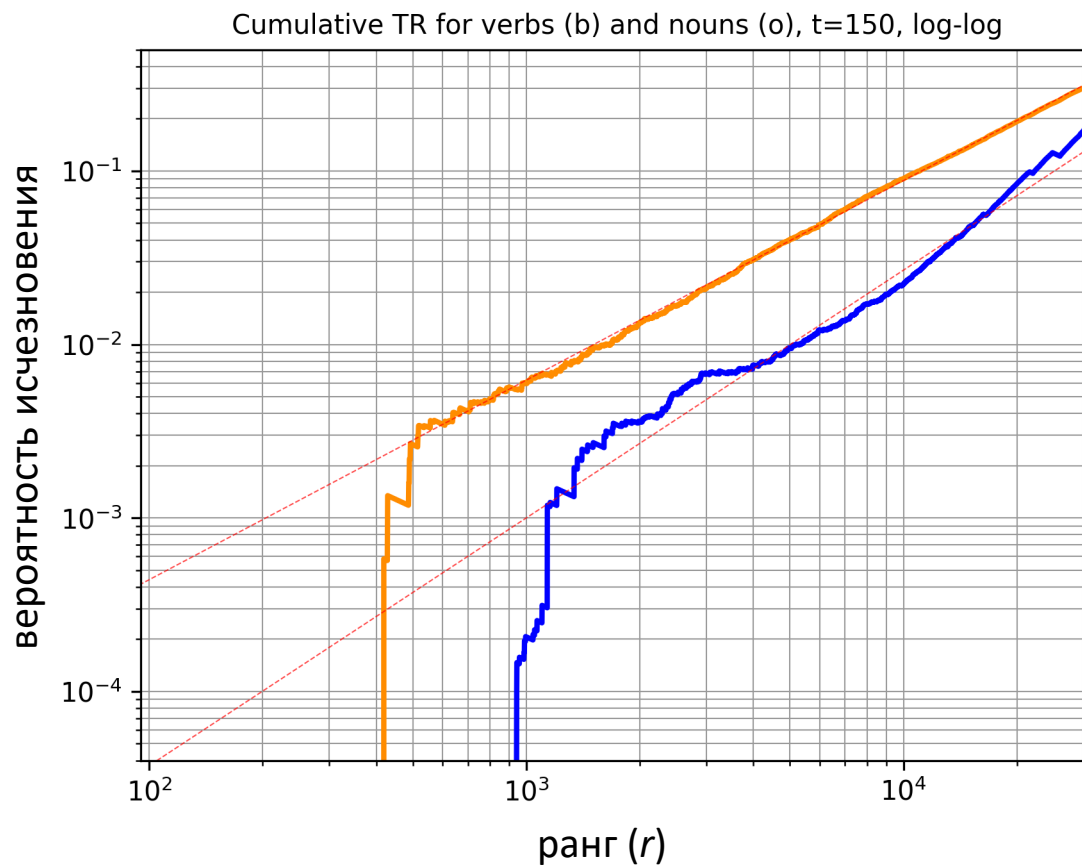
$TR_c(N/V, r, 20)$



$TR_c(N/V, r, 100)$



$TR_c(N/V, r, 150)$



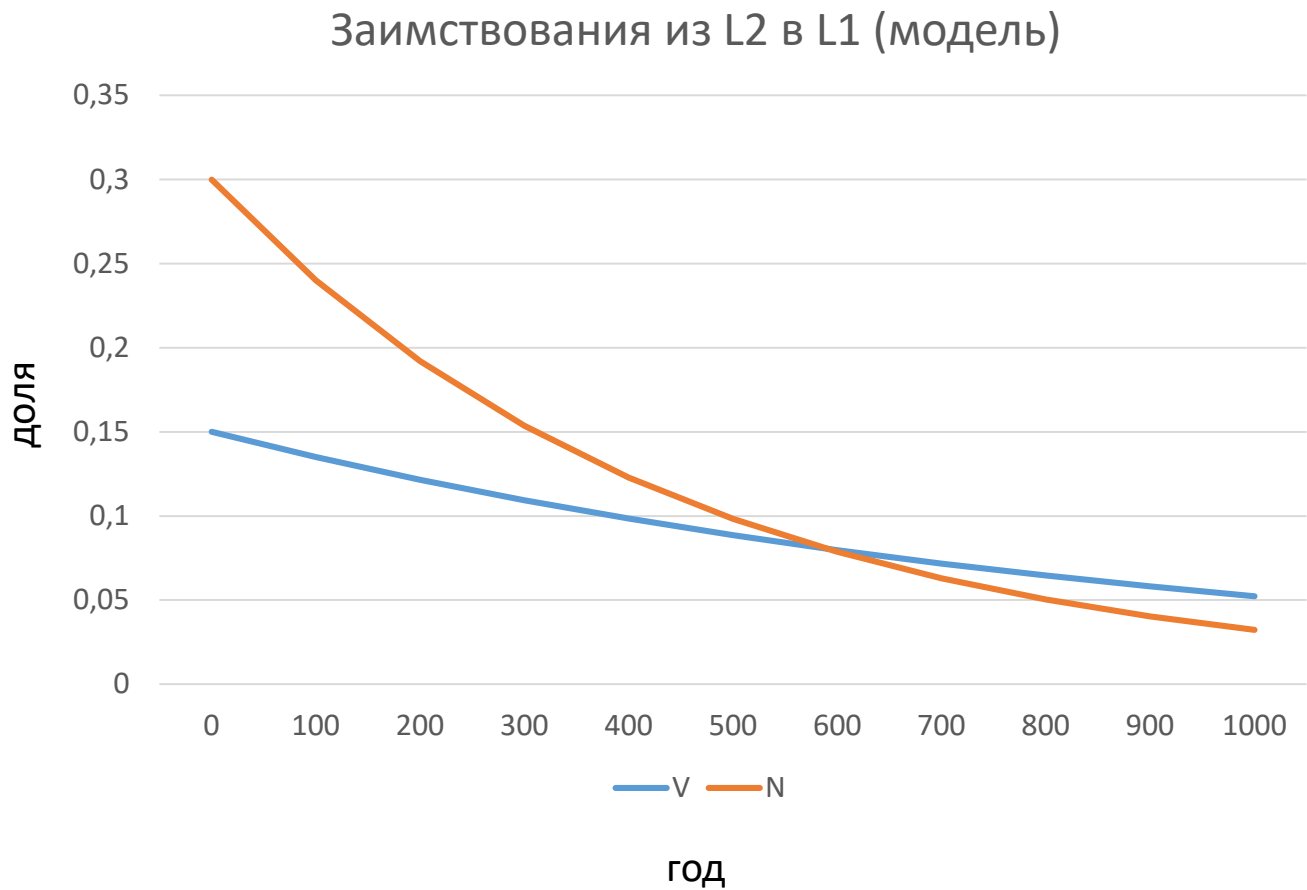
Вероятность исчезновения

- Вывод: да, в русском языке глаголы более живучи, чем существительные
- Гипотеза: глаголы в (почти) любом языке более живучи, чем существительные
- Как бы это проверить на типологической выборке?

Гипотеза

- Допустим, что контакт между L1 и L2 имел место, но давно закончился
- Предсказание: в этом случае доля глагольных заимствований из L2 в L1 зачастую будет **выше**, чем доля именных заимствований
 - Допустим, на момент завершения контактной ситуации 30% существительных и 15% глаголов в L1 были заимствованиями из L2
 - Допустим, что за век исчезает 10% глаголов и 20% существительных

Гипотеза



Гипотеза

- Проверим на данных WOLD
 - Отдельно заимствования из разных L2
 - Контактная ситуация завершилась больше века назад
 - Как минимум 2% выборки в L1 являются заимствованиями из L2
- Всего 22 контактных ситуации

Гипотеза

L1	L2	Период	V _{loan}	N _{loan}
Hausa	Kanuri	15 th – early 20 th century	1.2	4.2
Romanian	South Slavic	before 9 th century – 17 th century	11.2	8.2
Selice Romani	South Slavic	14 th – 17 th (?) century	1.2	2.5
Dutch	Latin	several periods, until 19 th century	3.2	8.4
Dutch	German	until 19 th century	1.7	3.2
British English	French	strong influence in 11 th – 17 th century	22.3	29.5
British English	Latin	mostly until 17 th century	5.9	9.7
British English	Old Norse	9 th – 12 th century	4.8	2.7
Kildin Saami	Proto-Fennic	until ~6 th century	4.7	3.7
Kildin Saami	North Germanic	~11 th – 16 th century	1.3	4.9
Kildin Saami	Karelian	until 20 th century	1.0	4.0
Kildin Saami	Proto-Northwest-IE	until 2 nd millennium BC	3.3	2.5
Archi	Arabic	14 th – 19 th (?) century	3.0	6.7
Yakut	Mongolic	12 th /13 th – 15 th /16 th centuries	11.4	10.3
Thai	Mon-Khmer	until 18 th century	4.2	4.4
White Hmong	Old Chinese	1 st millennium BC	2.7	2.0
White Hmong	Middle Chinese	until 16 th century	7.5	4.0
White Hmong	Modern Chinese	16 th – 19 th centuries	7.3	6.9
Malagasy	Malay	mostly before 8 th century	6.0	5.9
Saramaccan	Suriname Portuguese	17 th – 18 th (?) century	30.9	12.7
Kali'na	Spanish	17 th – 18 th (?) century	0.0	3.7
Mapudungun	Quechua	15 th – 16 th (?) century	1.6	3.8

Гипотеза

- В 10 случаях из 22 так и есть
- Среднее соотношение $V_{\text{loan}} / N_{\text{loan}}$ равно 0,91 (среднее по всему WOLD – 0,45)
- Чем древнее заимствования, тем больше вероятность, что заимствованных глаголов будет больше
- Это подтверждает гипотезу

И что теперь?

- Может быть, живучесть глаголов – независимое явление, из-за которого синхронно среди глаголов меньше заимствований
- Может быть, наоборот: существительные ровно потому чаще исчезают, что они чаще заменяются на заимствования

И что теперь?

- Так или иначе, делать выводы о лёгкости/сложности заимствования на основании синхронных данных не стоит
- Было бы интересно проверить, как обстоит дело в языках, которые редко что-либо заимствуют (китайский)

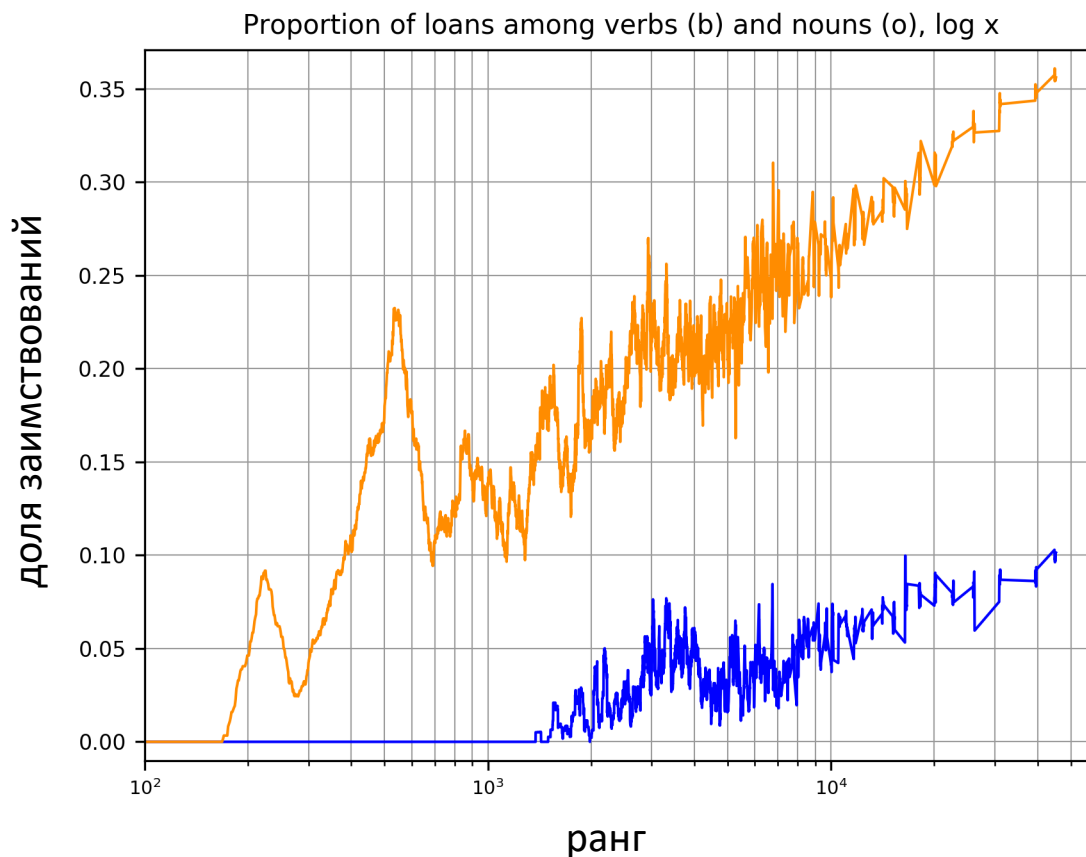
Часть 2: заимствования в русском

Займствования

- Считаются только заимствования, впервые засвидетельствованные не ранее 1750 г.
- Собственные имена не учитывались
- Лексемы с заимствованными корнями и незаимствованной словообразовательной морфологией считались заимствованиями (в отличие от WOLD)

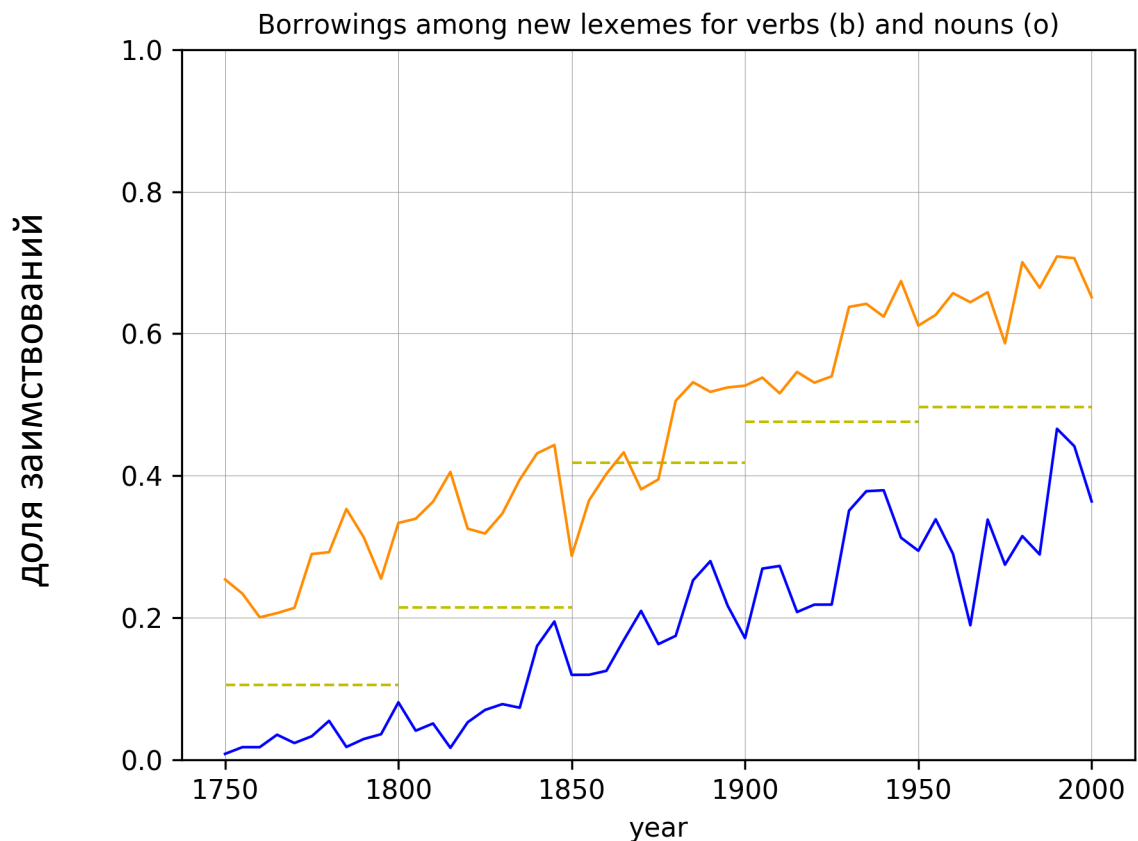
Займствования в русском

- Синхронные данные (тексты 2000-2009):



Заимствования в русском

- Диахронические данные (заимствования среди новых слов):



Заимствования в русском

- Если посмотреть на синхронные данные, может показаться, что глаголы в 4-5 раз сложнее заимствовать
- Если посмотреть на заимствования среди новых слов, то соотношение для последних 150 лет будет ближе к 2
- Тем не менее, среди новых глаголов всё-таки меньше заимствований, чем среди новых существительных

Выводы

- Глаголы типологически более живучи, чем существительные
 - Кроме того, исчезновение лексем подчинено степенному закону
- Это ставит под сомнение доказательства того, что глаголы сложнее заимствовать
 - Это, вероятно, всё-таки правда, но доказать нужно получше – чисто синхронные данные не годятся

Спасибо за внимание!

Литература

- Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the World's Languages. A Comparative Handbook*. Berlin: De Gruyter Mouton.
- Moravcsik, Edith. 1975. Verb borrowing. *Wiener Linguistische Gazette* 8. 3–30.
- Tadmor, Uri, Martin Haspelmath & Bradley Taylor. 2010. Borrowability and the notion of basic vocabulary. *Diachronica* 27(2). 226–246.
- Whitney, William Dwight. 1881. On Mixture in Language. *Transactions of the American Philological Association (1869-1896)*, vol. 12, 5–26. The Johns Hopkins University Press.
- Wohlgemuth, Jan. 2009. *A typology of verbal borrowings* (Trends in Linguistics. Studies and Monographs.). Vol. 211. Berlin: Mouton de Gruyter.

Данные

- Arkhangelskiy, Timofey. 2019. "Replication data for: Verbal borrowability and turnover rates", <https://doi.org/10.18710/JFNESU>, DataverseNO, V1