

Заимствования в 110-словном списке Сводеша

Илья Егоров (ШАГИ РАНХиГС)

Fragestellung

В основу концепта базисной лексики положены следующие принципы:

- Независимость от типа культуры и природных условий
- Простота и дискретность концептов
- Устойчивость перед лицом заимствований
- Устойчивость перед лицом внутренней замены

Насколько устойчивость к внутренней замене коррелирует с устойчивостью к заимствованию?

Ранжирование базисной лексики

Ранжирование базисной лексики на основании устойчивости к внутренней замене:

Starostin, Sergei A. 2007. Opredelenie ustojčivosti bazisnoj leksiki [Defining the stability of basic lexicon]. *Trudy po jazykoznaniju [Works on linguistics]*, 827–839. Moscow: Jazyki slavjanskix kul'tur.

Эксперимент воспроизведен в статье Жени Коровиной; обсуждаются детали методологии, не эксплицированные в статье С.А. Старостина; предлагаются небольшие исправления.

Korovina, Evgenia V. 2019. Ranžirovanie bazisnoj leksiki v trudax S.A. Starostina: materialy k ulučšeniju [Method of obtaining a ranked list of the basic vocabulary developed by Sergei A. Starostin]. *Indo-European Linguistics and Classical Philology XXIII*. 546–556. doi:10.30842/ielcp230690152345.

Ранжирование (базисной) лексики на основании устойчивости перед заимствованиям:

Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the World's Languages: A Comparative Handbook*. Walter de Gruyter. doi:10.1515/9783110218442.

Старостин 2007

we, two, I, eye, thou, who, fire, tongue, stone, name

hand, what, die, heart, drink, dog, louse, moon, nail, blood

one, tooth, new, dry, liver, eat, tail, this, hair, water

nose, not, mouth, full, ear, that, bird, bone, sun, smoke

stand, tree, ashes, give, rain, star, fish, neck, breast, leaf

come, kill, foot, sit, root, thin, horn, fly, hear, skin

long, worm, meat, road, know, saltm say, egg, seed, knee

black, head, sleep, burn, earth, year, feather, swim, white, bite

fat, man, person, all, snake, night, see, heavy, go, warm

red, cold, woman, round, near, yellow, lie, green, cloud, far

big, bark, sand, short, good, many, mountain, wind, belly, small

Коровина 2019

I, thou, eye, we, name, tongue, nail, tooth

dog, die, blood, fire, full, thin, eat, what, heart, louse

liver, who, water, hand, salt, star, smoke, ear, give, stone

new, bone, nose, moon, stand, sun, not, hair, fish, year

tail, this, drink, rain, horn, tree, come, long, foot, snake

hear, warm, kill, swim, wind, dry, far, heavy, leaf, black

road, bird, earth, worm, skin, know, egg, fly, knee, head

night, seed, meat, that, sleep, sit, fat, breast, person, feather

white, ashes, short, say, root, near, red, mouth, bite, see

neck, man, woman, cloud, go, burn, lie, green, good, mountain

yellow, all, bark, cold, round, belly, small, sand, big, many

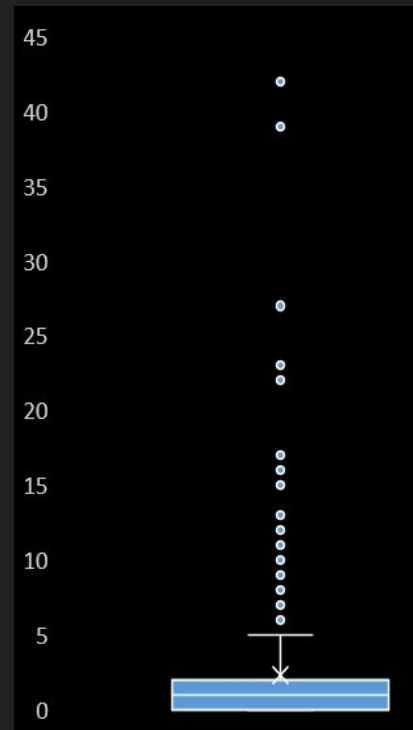
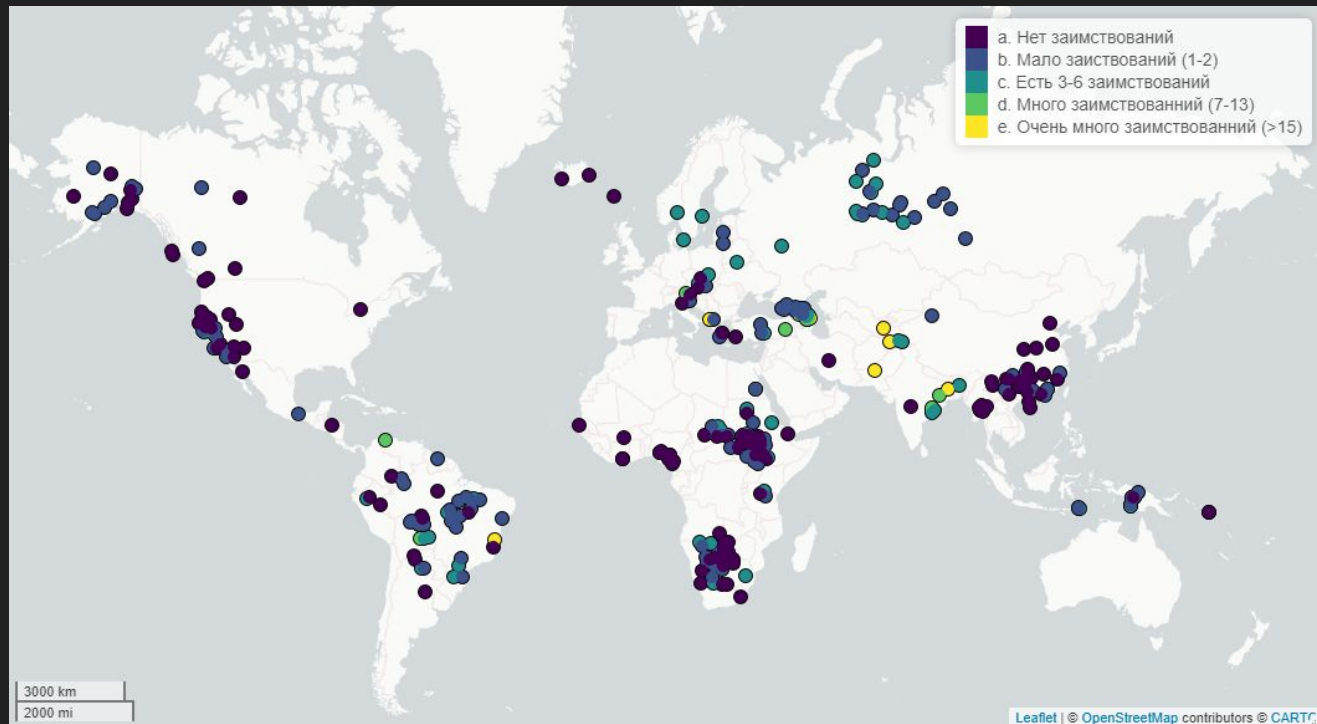
Данные

Выбрано 427 языков из Global Lexicostatistical Database, объединенные в 106 подгрупп. Применяемый стандарт аннотации предполагает обязательное выделение заимствований.

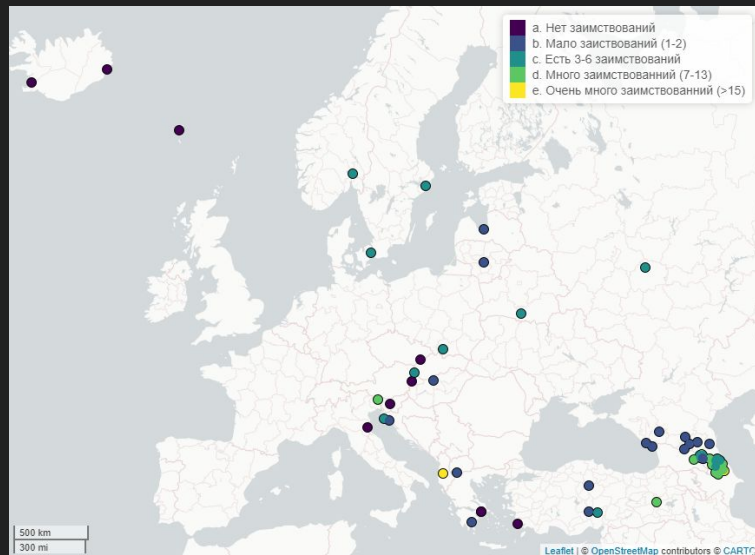
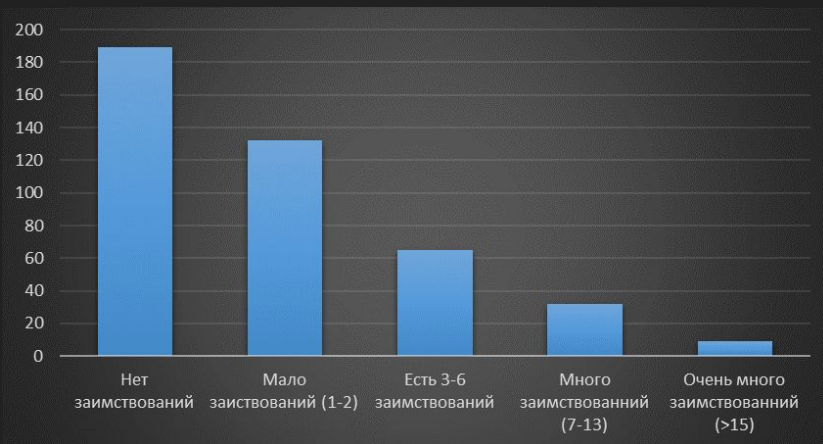
- Исключены списки со слишком большим числом незафиксированных вхождений.
- Исключена очень большая романская база (58 языков).

*Вопрос о корреляции внутренней устойчивости и устойчивости к заимствованиям, конечно, может (и должен) рассматриваться не только на материале 110-словных списков Сводеша.

Результаты: языки



Результаты: языки



- Медианное значение количества заимствований: 1
- Языки с экстремальным количеством заимствований: ягноби (42), парачи (39), брахуи (27), албанский (23), хиналугский (23), будухский (22), махакали (17), алыкский крызский (16), малто (15).
- Выделяются регионы и подгруппы, в которые заимствования распространены больше, чем в других: Кавказ (особенно лезгинские), Центральная Азия, Индия.

Результаты: ранжирование

Ни в одном списке в GLD не обнаружено ни одного заимствования в следующих слотах:

I (1), thou (1), tongue (1), what (2), give (3), hand (3), water (3), who(3), not (4),
drink (5), this (5), kill (6), sit (7), sleep (8), that (8), go (10), cold (11)

В скобках дан номер квантили по ранжированию Жени Коровиной (2019).

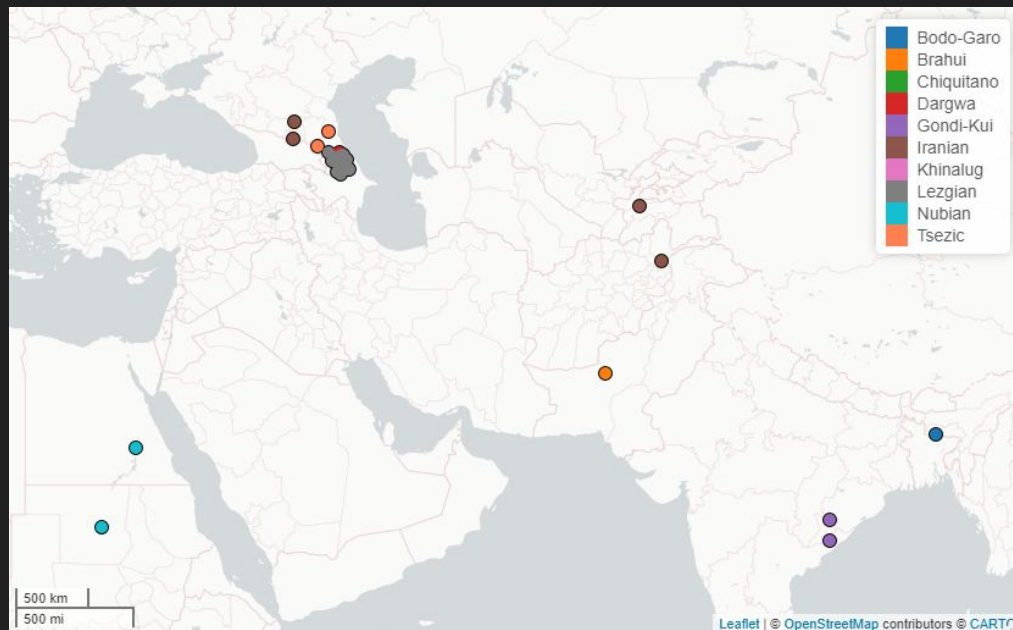
- Среди слов, супер устойчивых к заимствованиям, немного преобладают слова из первой половины списка по устойчивости к внутренней замене. Уже по этим данным видно, что никакой строгой зависимости нет.

Результаты: ранжирование

Дальнейшие подсчеты не могут быть столь же прямолинейны. Часто один семантический слот бывает заполнен заимствованиями во всех или почти во всех языках подгруппы и региона. Если в подгруппе много языков или в одном локальном ареале высокая плотность, это дает сдвиг в данных.

Кажется разумным применить нормализацию следующего типа: подсчитывать количество групп (таксон среднего уровня), в которых данный слот заполнен заимствованием хотя бы в одном языке.

* Принимаются предложения о более математически обоснованной нормализации!



Заимствования в слоте person

Результаты: ранжирование (“нормализованные” данные)

1 we (1), fire (2), ear (3), nose (4), come (5), foot (5), hear (6), know (7), say (9), burn (10), lie (10), belly (11), die (2), eat (2)

2 near (9), tooth (1), head (7), knee (7), see (9)

3 long (5), night (8), name (1), rain (5), big (11), mouth (9), one (1), blood (2), hair (4), stand (4), sun (4), dry (6), bite (9)

4 stone (3), warm (6), horn (5), ashes (9), black (6), heavy (6), full (2), meat (8)

5 neck (10), bird (7), many (11), mountain (10), woman (10), small (11), nail (1), tail (5), smoke (3), far (6)

6 wind (6), two (1), breast (8), earth (7), louse (2), tree (5), road (7), bone (5)

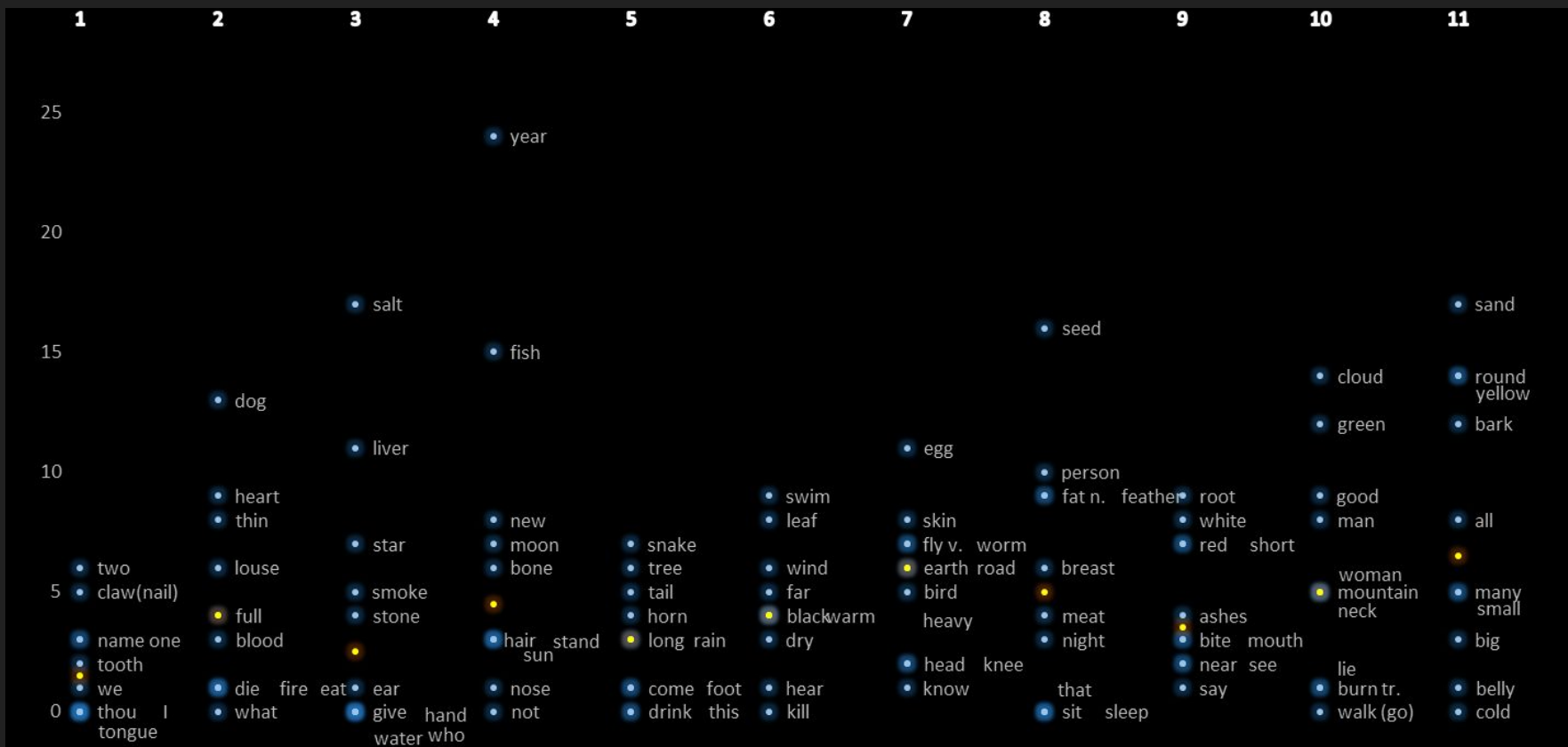
7 moon (4), snake (5), star (3), short (9), fly (7), worm (7), red (7)

8 man (10), new (4), all (11), skin (7), thin (2), leaf (6), white (9)

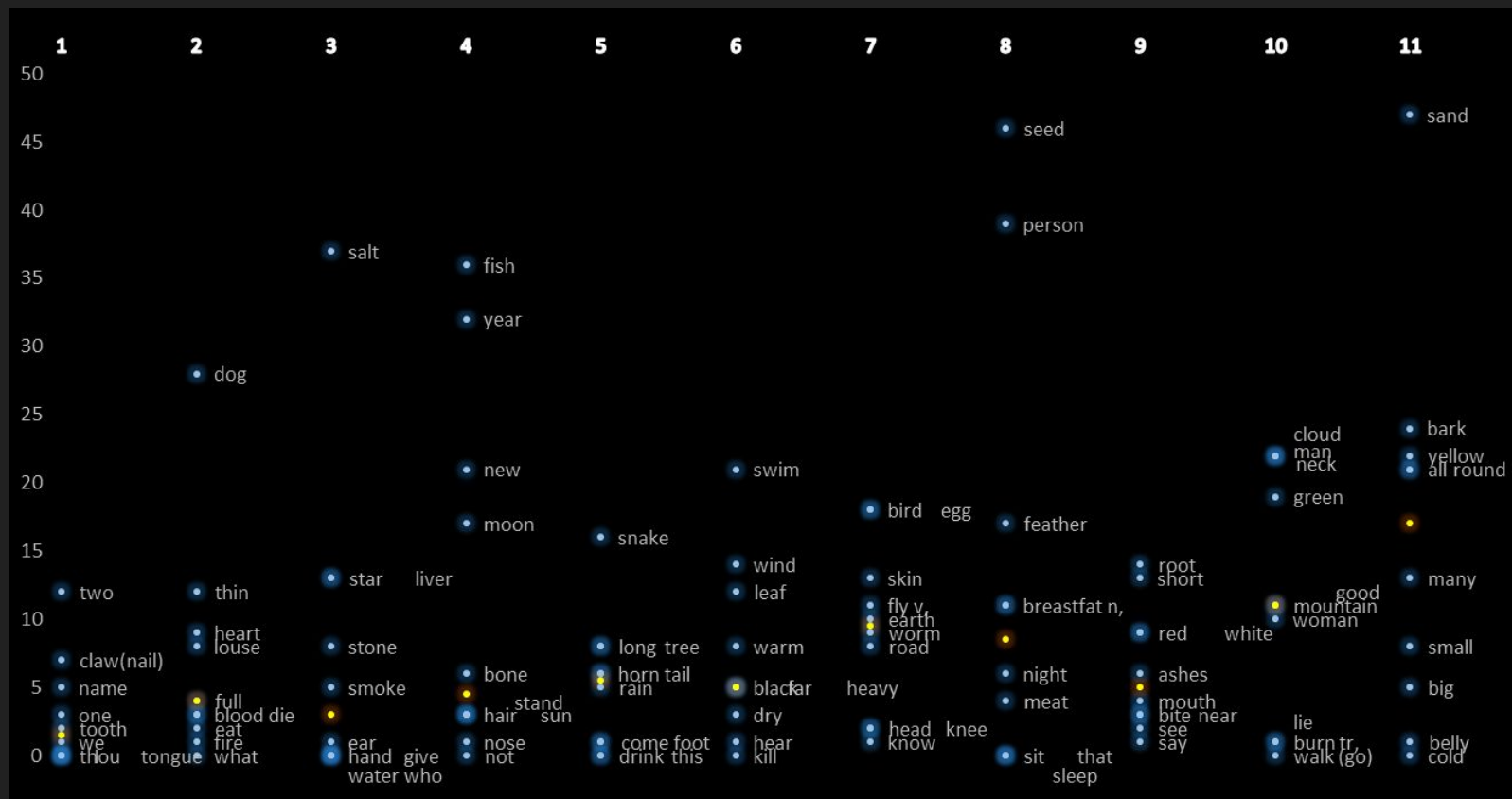
9 swim (6), feather (8), root (9), fat (8), good (10), heart (2)

МНОГО 10 - person (8), 11 - liver (3), egg (7), 12 - green (10), bark (11), 13 - dog (2), 14 - round (11), yellow (11), cloud (10), 15 - fish (4), 16 - seed (8), 17 - salt (3), sand (11), 24 -year (4)

Результаты: корреляция (“нормализованные” данные)



Результаты: корреляция (“ненормализованные” данные)



Результаты: корреляция

Среди самых устойчивых к заимствованию слов, немного больше слов из наиболее стабильной половины списка.

Среди самых часто заимствуемых слов, больше слов из менее стабильной половины списка.

И все же сильной корреляции нет.

- Мало (или не) заимствуемые, но внутренне нестабильные слоты: 0 - sleep (8), that (8), go (10), cold (11), 1 - say (9), burn (10), lie (10), belly (11).
- Часто заимствуемые, но внутренне устойчивые: 9 - heart (2), 11 - liver (3), 13 - dog (2), 15 - fish (4), 17 - salt (3), 24 - year (4).

Ареальные паттерны

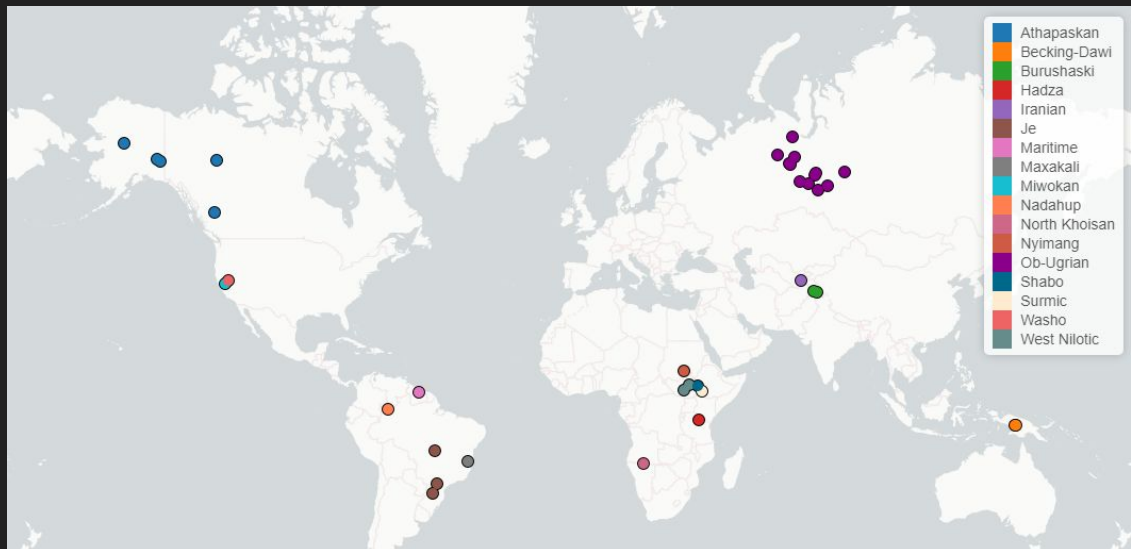
Заимствования в слоте dog

- В некоторых регионах распространяются локальные Wanderwörter.



Заимствования в слоте salt

- Для некоторых регионов слово является культурным.



110-Swadesh vs Leipzig-Jakarta

0 I (1), thou (1), tongue (1), what (2), give (3), hand (3), water (3), who(3), not (4), drink (5), this (5), kill (6), sit (7), sleep (8), that (8), go (10), cold (11)

1 we (1), fire (2), ear (3), nose (4), come (5), foot (5), hear (6), know (7), say (9), burn (10), lie (10), belly (11), die (2), eat (2)

2 near (9), tooth (1), head (7), knee (7), see (9)

3 long (5), night (8), name (1), rain (5), big (11), mouth (9), one (1), blood (2), hair (4), stand (4), sun (4), dry (6), bite (9)

4 stone (3), warm (6), horn (5), ashes (9), black (6), heavy (6), full (2), meat (8)

5 neck (10), bird (7), many (11), mountain (10), woman (10), small (11), nail (1), tail (5), smoke (3), far (6)

6 wind (6), two (1), breast (8), earth (7), louse (2), tree (5), road (7), bone (5)

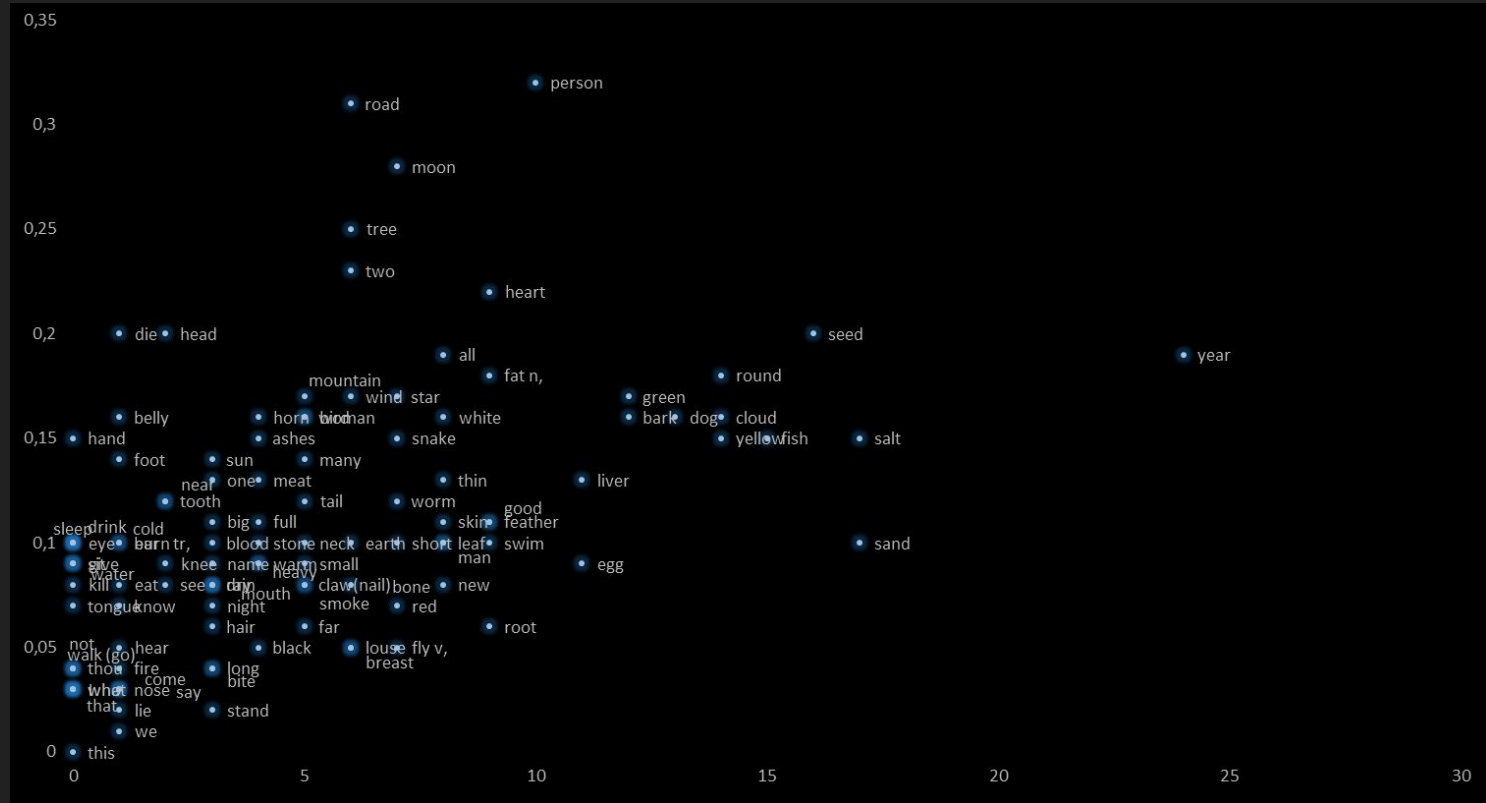
7 moon (4), snake (5), star (3), short (9), fly (7), worm (7), red (7)

8 man (10), new (4), all (11), skin (7), thin (2), leaf (6), white (9)

9 swim (6), feather (8), root (9), fat (8), good (10), heart (2)

МНОГО 10 - person (8), 11 - liver (3), egg (7), 12 - green (10), bark (11), 13 - dog (2), 14 - round (11), yellow (11), cloud (10), 15 - fish (4), 16 - seed (8), 17 - salt (3), sand (11), 24 -year (4)

110-Swadesh vs WOLD

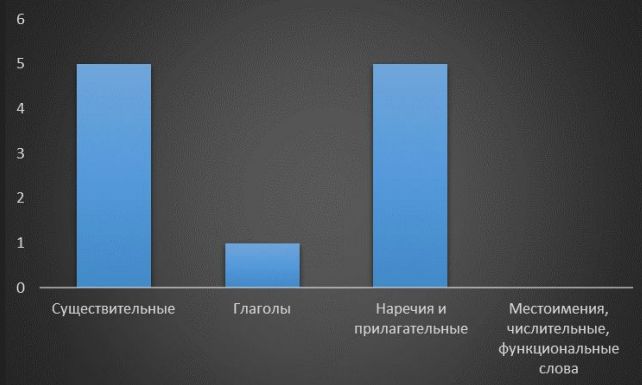


POS и устойчивость

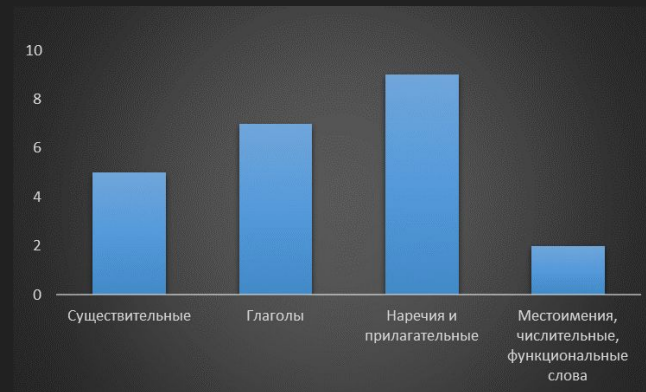
В Haspelmath & Tadmor 2009 делаются следующие обобщения об устойчивости лексики к заимствованию:

- Функциональные слова более устойчивы, чем полнозначные
- Глаголы более устойчивы, чем имена

Относительно устойчивости к внутренней замене, верно лишь первое обобщение.



Медианное количество языков, демонстрирующее заимствования в слотах данной части речи.



Медианная квантиль по ранжированию Коровиной.

Выводы

- Большинство языков демонстрируют не больше двух заимствований в 110-словном списке Сводеша.
- Выделяются некоторые регионы, где заимствования частотнее, чем в других (Кавказ, Центральная Азия, Индия).
- Строгой зависимости между устойчивостью к заимствованию и к внутренней замене нет. Просматривается только незначительная корреляция.
- Отдельные слоты могут быть особенно часто подвержены заполнению заимствованием в определенных ареалах. Это связано с “большей культурностью” данных концептов и с распространением локальных *Wanderwörter*.
- Некоторые очень устойчивые к заимствованию значения из 110-словного списка Сводеша не попали в список Leipzig-Jakarta. При этом borrowed score по данным WOLD хорошо коррелирует с нашим ранжированием. Подтверждаются обобщения относительно связи части речи и подверженностью заимствованиям.

Дякую за увагу