

XIV чтения памяти С.А. Старостина  
25–26 марта 2019, НИУ ВШЭ, Москва

М.Е. Васильев  
*mvhumanity@gmail.com*

## **Статистический маятник для «глотточасов»: математический подход к решению проблемы**

## Постулаты М. Сводеша

### Постулат 3.

Доля слов из ОС [основного списка], которые остаются неизменными <...> за определённый интервал времени, является постоянной величиной и зависит только от длины этого интервала, а не от того, когда этот интервал был выбран или слова какого языка рассматриваются.

### Постулат 4.

Все слова, составляющие ОС, могут с равной вероятностью остаться или исчезнуть из языка за определённый промежуток времени.

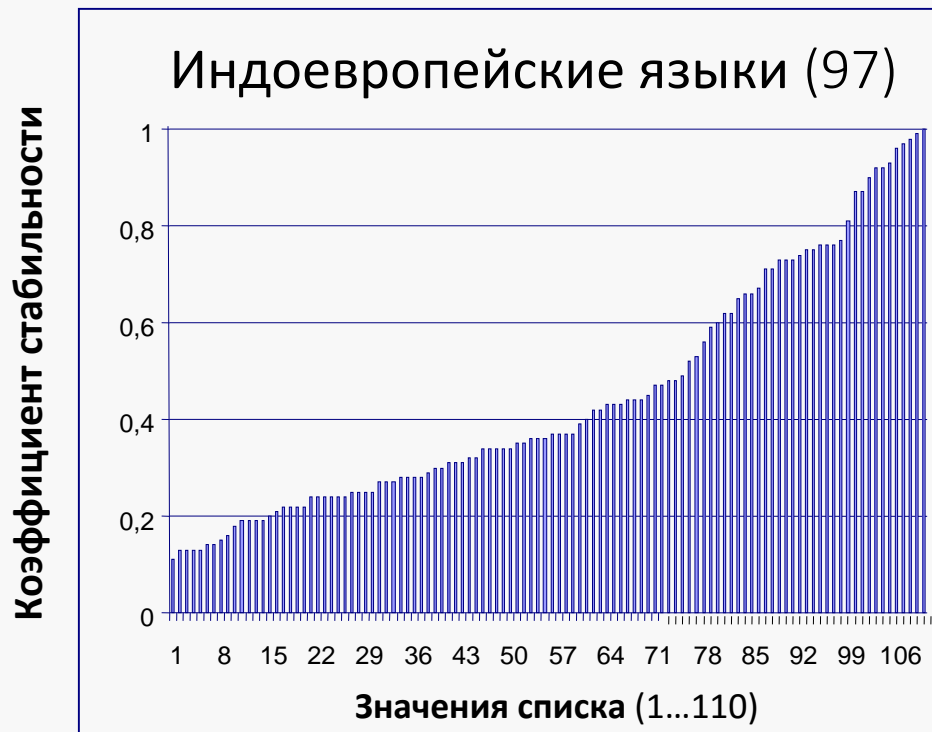
[Арапов, Херц 1974: Математические методы в исторической лингвистике]

## Определение показателя скорости замен $\lambda$ для некоторых языков по историческим данным $t$ и $N(t)$

Язык	$t$	$N(t)$	$\lambda$
Китайский	2,6	0,77	0,10
Английский	1,3	0,88	0,10
Французский	1,5	0,90	0,07
Японский	1,2	0,93	0,06
Испанский	1,5	0,91	0,06
Немецкий	1,2	0,94	0,05

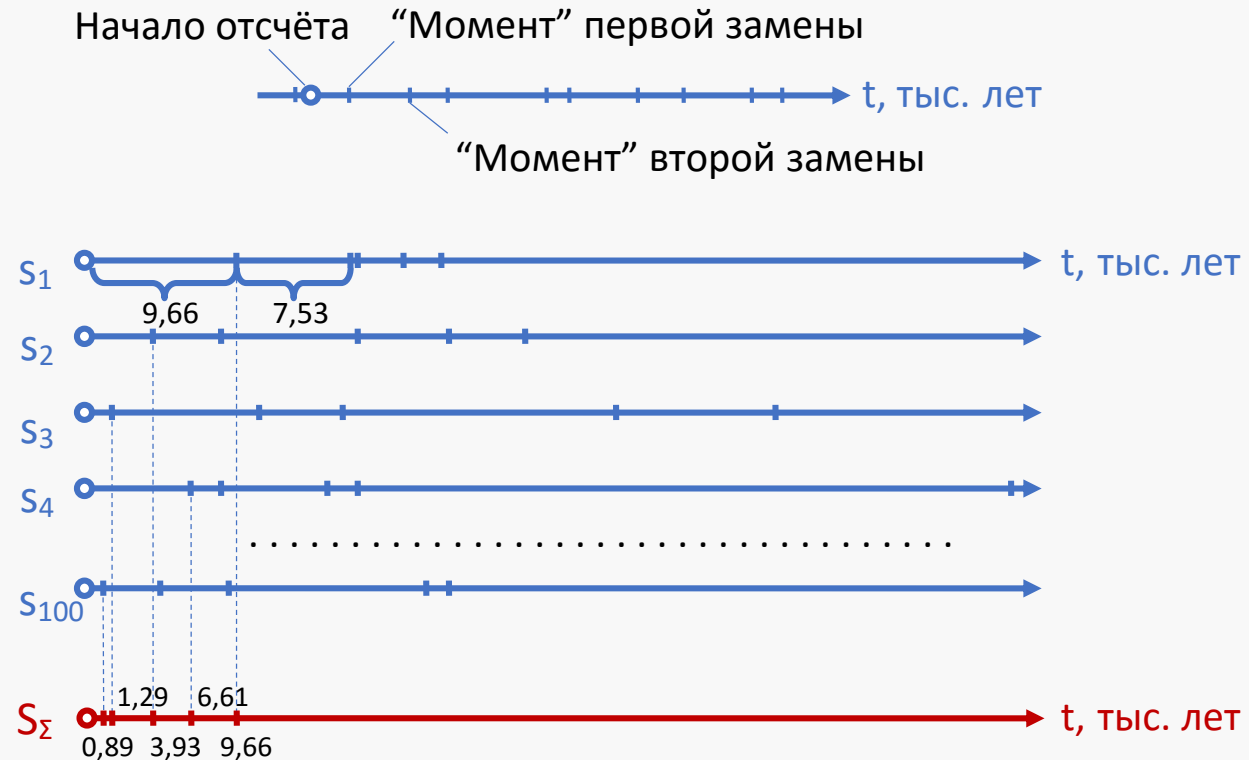
[S. Starosin 2000: Comparative-historical linguistics and lexicostatistics]

## Распределение коэффициентов стабильности различных значений основного списка



[С.А. Старостин 2007: Определение устойчивости базисной лексики]

# Моделирование процесса замен как потока случайных событий



Результаты численного моделирования элементарных потоков  $s_i$  и общего потока  $S_\Sigma$  первых замен лексических значений ОС



Иллюстрация общего потока первых замен (фрагмент)

## Исследование свойств потоков случайных событий

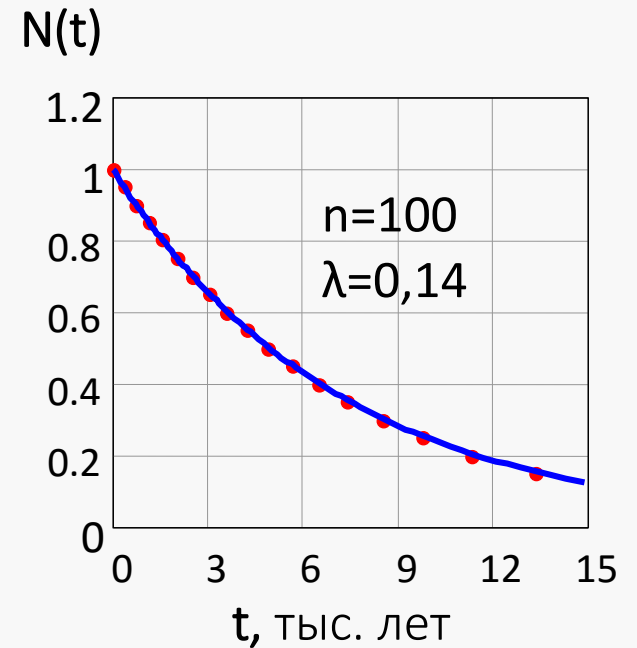
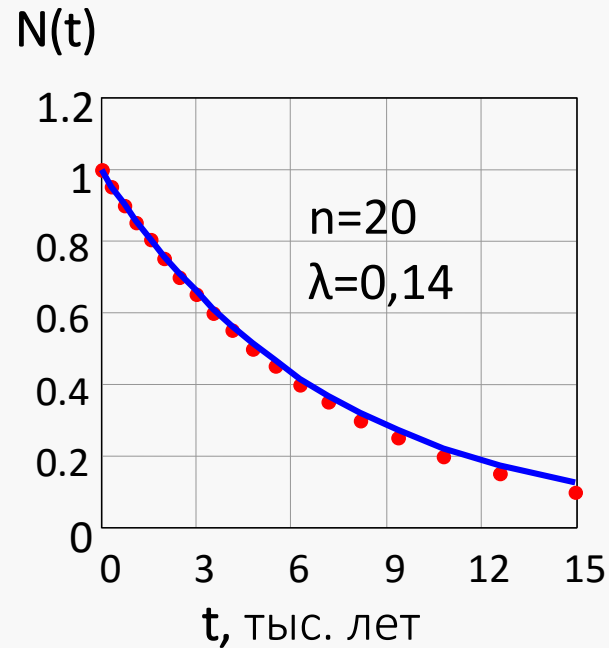
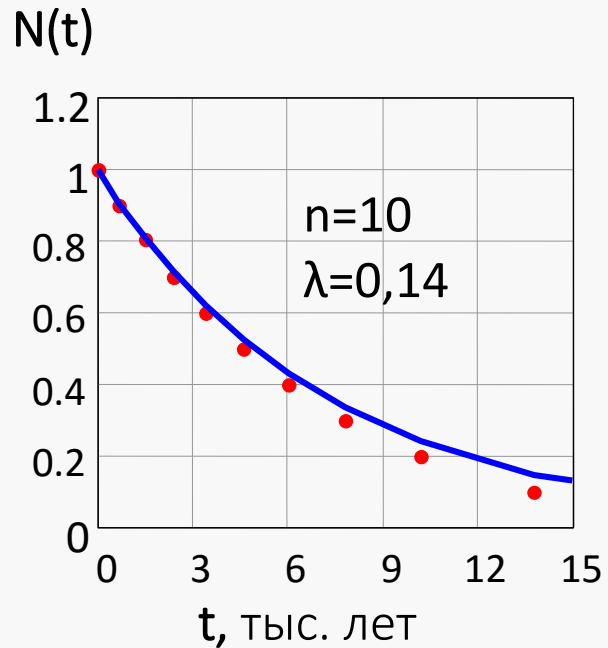


**А.Я. Хинчин**  
(1894 – 1959)



**Е.С. Вентцель**  
(1907 – 2002)

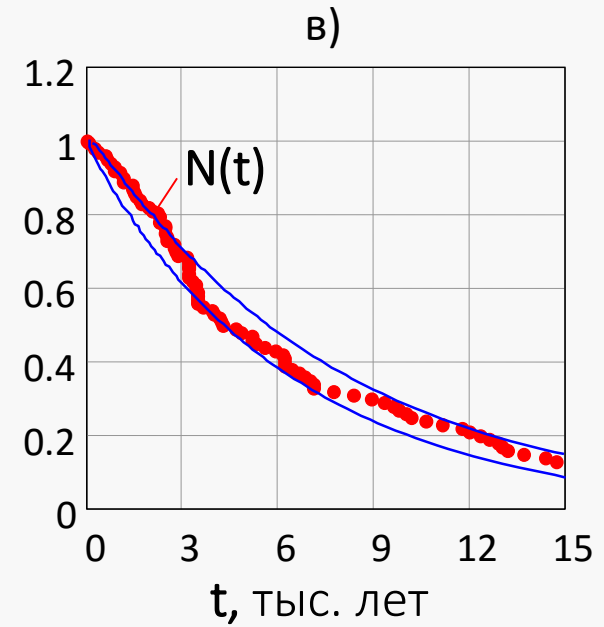
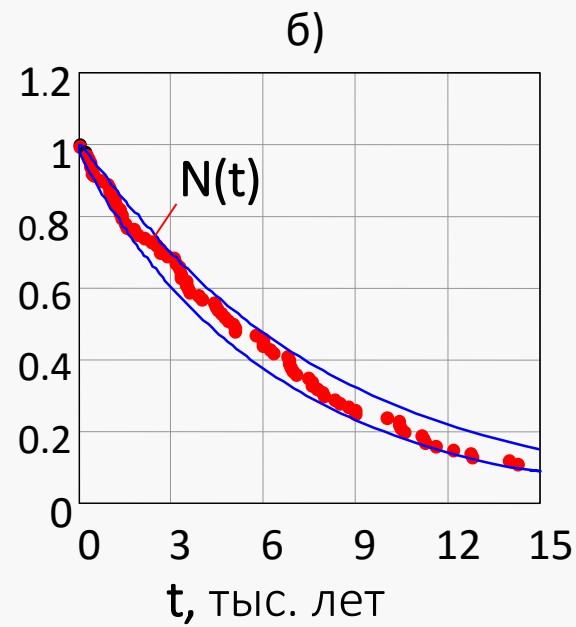
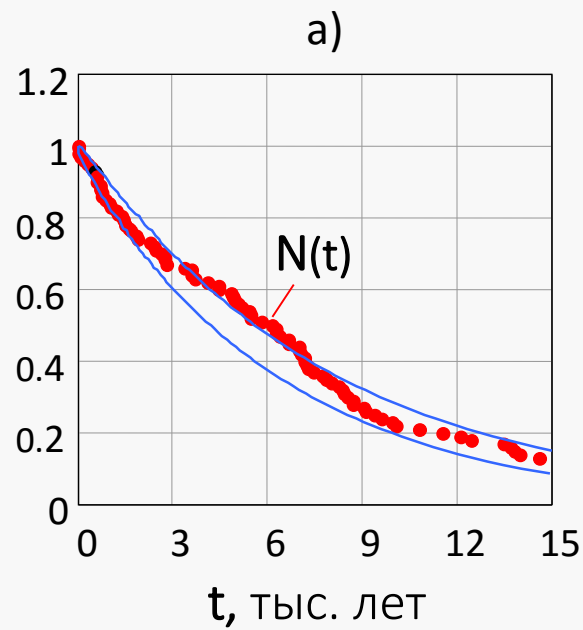
# Сравнение процессов замен $N(t)$ , смоделированных для различного числа слагаемых $n$ в суммарном потоке $S$



$$N(t) = e^{-\lambda \cdot t}$$

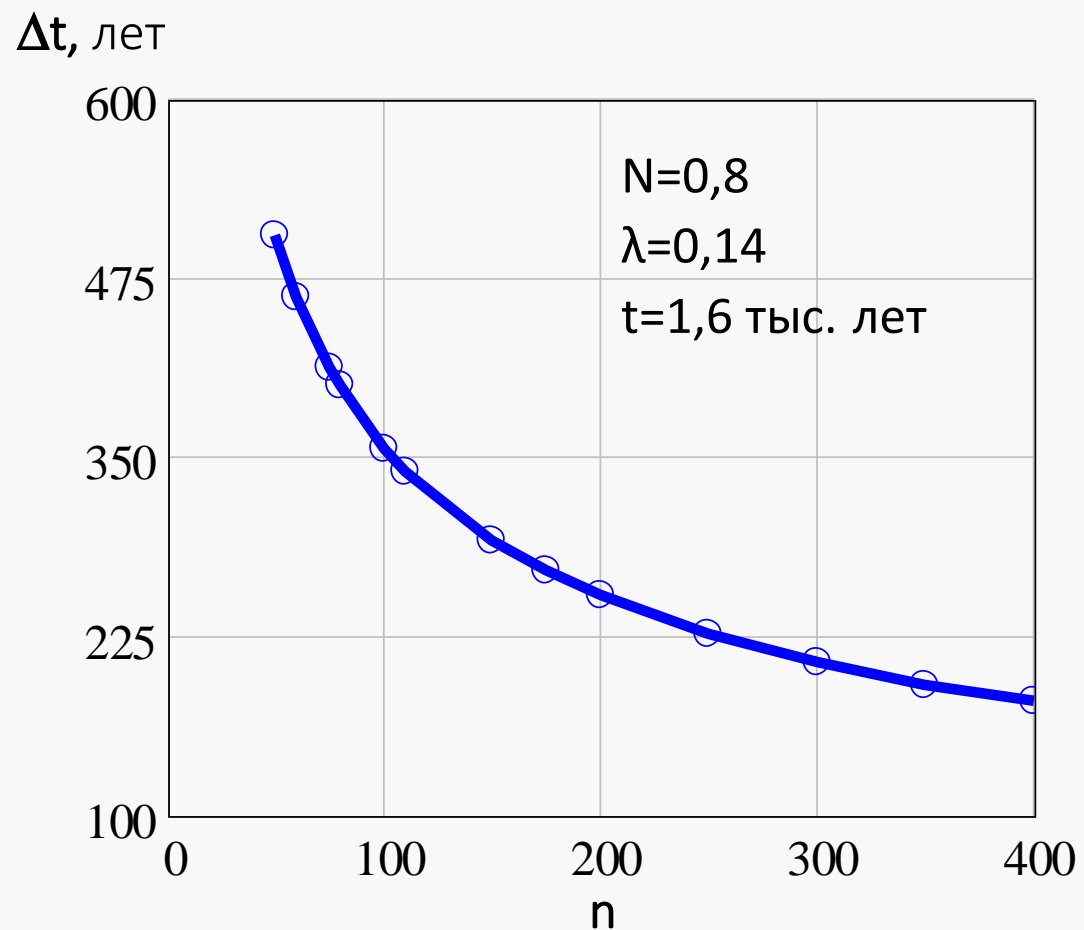
## Доверительные интервалы $\Delta(t)$ для $n=100$ ; $\lambda=0,14$ ; $p=0,7$

$t$ , тыс. лет	0,25	0,5	1	2	3	4	5	6	7	8	9	10	11
$\pm\Delta(t)$ , тыс. лет	0,13	0,19	0,28	0,41	0,52	0,62	0,72	0,82	0,92	1,02	1,13	1,25	1,36
$\Delta(t)t$ , %	54	38	28	20	17	16	14	14	13,1	12,8	12,6	12,5	12,4



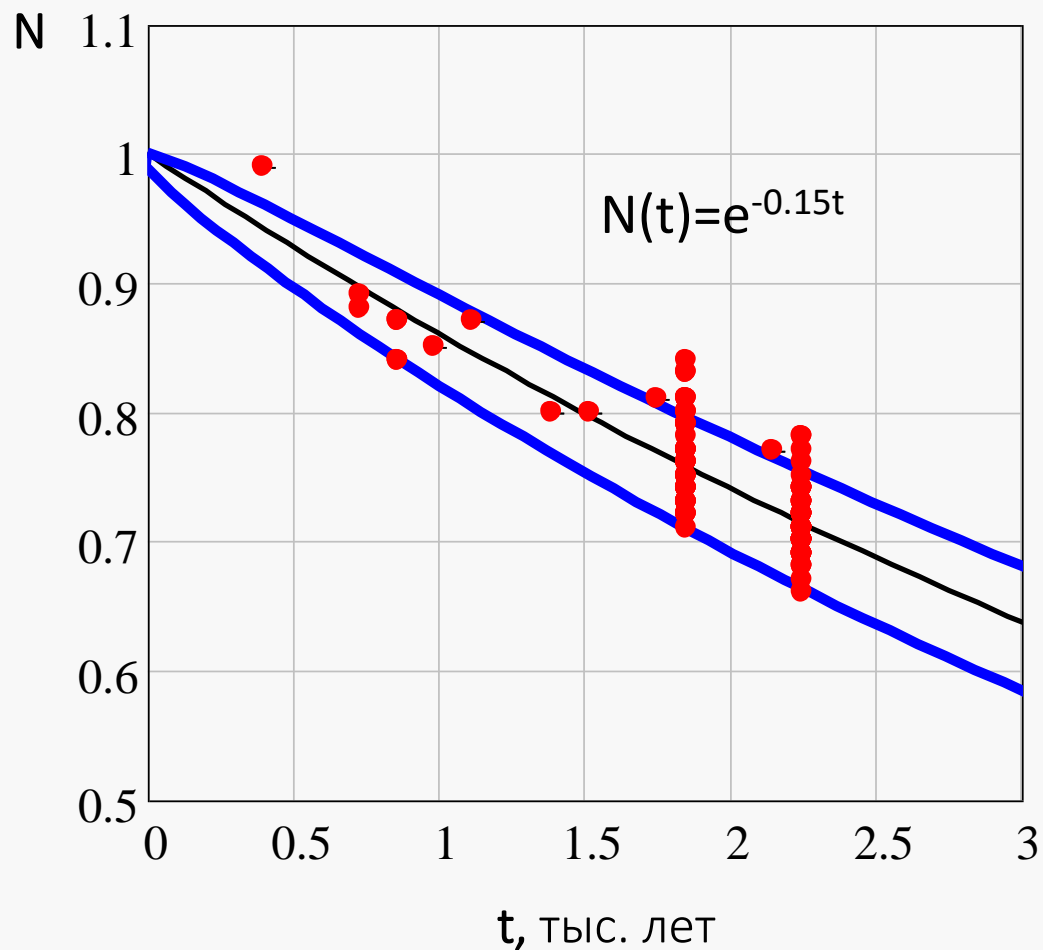


## Зависимость доверительного интервала $\Delta t$ от числа лексических значений $n$ в сравниваемых списках языков



# Процесс замен по фактическим данным для 57 романских языков

из 116 значений 95 (82%) вошли в доверительный интервал ( $p=0,7$ )



## Выводы

1. Незнание единой закономерности в изменениях каждого отдельно взятого лексического значения не является препятствием для проявления и установления такой закономерности для суммы процессов замен этих значений.
2. В условиях непостоянной скорости замен каждого значения основного списка и отличающейся скорости замен разных значений общая доля изменившихся слов во всём списке в среднем остаётся неизменной и может быть описана как случайная величина с известными отклонениями.

Иными словами, понятие «глотточасов» имеет статистическую природу и не нуждается в допущении о постоянстве скорости замен всех значений в любом языке.

3. Полученные выводы снимают существующую проблему теоретического обоснования лексикостатистики и позволяют с заданной вероятностью количественно определить возможности метода в виде доверительных интервалов глоттохронологических датировок.