

Математика синтаксиса:
формализация закономерностей в языковых данных

[24 июля]

Марина Ермолаева

XXIII Летняя Лингвистическая Школа
Переславль-Залесский, 2023

В прошлый раз

- Формализация в лингвистике
- (Формальные) языки и грамматики
- Лексические единицы
- Операция соединения (Merge)



Вспомогательные глаголы в английском

Mary laughs

Вспомогательные глаголы в английском

Mary laughs

Mary will laugh

Вспомогательные глаголы в английском

Mary laughs

Mary will laugh

Mary is laughing

Mary will be laughing

Вспомогательные глаголы в английском

Mary laughs

Mary will laugh

Mary is laughing

Mary will be laughing

Mary has laughed

Вспомогательные глаголы в английском

Mary laughs

Mary will laugh

Mary is laughing

Mary will be laughing

Mary has laughed

Mary has been laughing

Mary will have laughed

Mary will have been laughing

Вспомогательные глаголы в английском

Mary laughs / laughed

Mary will / would laugh

Mary is / was laughing

Mary will / would be laughing

Mary has / had laughed

Mary has / had been laughing

Mary will / would have laughed

Mary will / would have been laughing

Фрагмент грамматики английского языка (I)

Mary :: np

will :: =v np= s

would :: =v np= s

has :: =perf np= s

had :: =perf np= s

have :: =perf v

is :: =g np= s

was :: =g np= s

be :: =g v

been :: =g perf

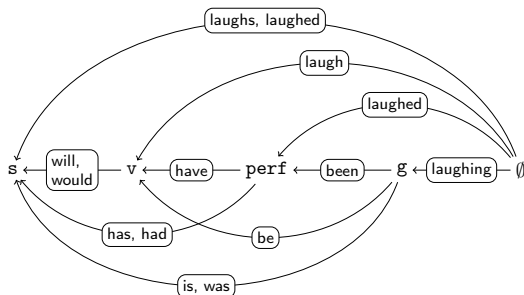
laughs :: np= s

laughed :: np= s

laugh :: v

laughed :: perf

laughing :: g



“Дом, который построил Джек”

This is the maiden all forlorn
That milked the cow with the crumpled
horn
That tossed the dog that worried the cat
That killed the rat that ate the malt
That lay in the house that Jack built.

(английское народное стихотворение)

А это старушка, седая и строгая,
Которая доит корову безрогую,
Лягнувшую старого пса без хвоста,
Который за шиворот треплет кота,
Который пугает и ловит синицу,
Которая часто ворует пшеницу,
Которая в тёмном чулане хранится
В доме,
Который построил Джек.

(перевод С. Маршака)

Относительные предложения в английском

- This is...

Относительные предложения в английском

- This is...
... the mouse

Относительные предложения в английском

- This is...
 - ... the mouse
 - ... **the cat** that **bit** the mouse

Относительные предложения в английском

- This is...
 - ... the mouse
 - ... the cat that bit the mouse
 - ... the dog that caught the cat that bit the mouse

Относительные предложения в английском

- This is...
 - ... the mouse
 - ... the cat that bit the mouse
 - ... the dog that caught the cat that bit the mouse
 - ... the frog that scared the dog that caught the cat that bit the mouse

Относительные предложения в английском

- This is...
 - ... the mouse
 - ... the cat that bit the mouse
 - ... the dog that caught the cat that bit the mouse
 - ... the frog that scared the dog that caught the cat that bit the mouse
 - ... the mouse that the cat bit

Относительные предложения в английском

- This is...

... the mouse

... the cat that bit the mouse

... the dog that caught the cat that bit the mouse

... the frog that scared the dog that caught the cat that bit the mouse

... the mouse that the cat bit

... the mouse that the cat that the dog caught bit

Относительные предложения в английском

- This is...

... the mouse

... the cat that bit the mouse

... the dog that caught the cat that bit the mouse

... the frog that scared the dog that caught the cat that bit the mouse

... the mouse that the cat bit

... the mouse that the cat that the dog caught bit

... the mouse that the cat that the dog that the frog scared caught bit

the frog that scared the dog that caught the cat that bit the mouse



- Последовательные зависимости (serial dependencies)

the frog that scared the dog that caught the cat that bit the mouse



- Последовательные зависимости (serial dependencies)
- Абстрактный аналог:
строки из последовательностей aa и bb в любом количестве
 $\{\epsilon, aa, bb, aaaa, aabb, bbaa, bbbb, aaaaaa, aaaabb, \dots\}$

Язык последовательностей aa и bb

- Язык:

$\{\epsilon, aa, bb, aaaa, aabb, bbaa, bbbb, aaaaaa, aaaabb, \dots\}$

- Грамматика:

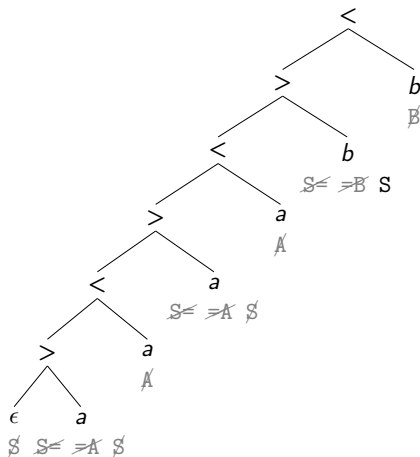
$\epsilon :: S$

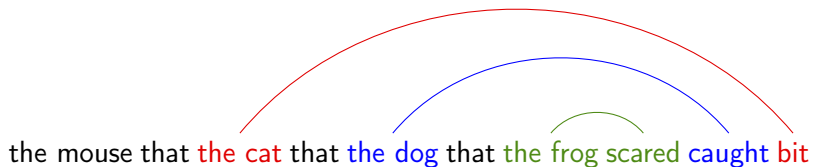
$a :: S = A S$

$a :: A$

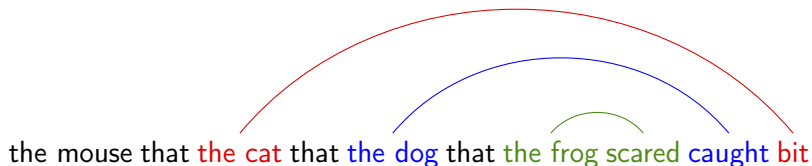
$b :: S = B S$

$b :: B$





- Вложенные зависимости (nested dependencies)



- Вложенные зависимости (nested dependencies)
- Абстрактный аналог:

палиндромы четной длины из a и b

$\{\epsilon, aa, bb, aaaa, abba, baab, bbbb, aaaaaa, aabbaa, abbbba, \dots\}$

Задача: язык палиндромов четной длины

- Язык: $\{\epsilon, aa, bb, aaaa, abba, baab, bbbb, aaaaaa, aabbaa, \dots\}$
- Составьте набор лексических единиц, который:
 - Порождает все палиндромы четной длины из a и b (включая ϵ)
 - Не порождает никаких других строк
- Какова длина получившейся грамматики в символах?
 - Каждое вхождение a или b – 1 символ
 - Пустая строка – 0 символов
 - Каждый разделитель $::$ – 1 символ
 - Каждый синтаксический признак ($S, A, =S, A=, \dots$) – 1 символ