

Основные идеи

1. Чтобы понять, человек строит в голове модель
2. Компьютеру тоже нужна модель, только более "формальная"
3. Для распознавания языка есть модель
4. Модель **Хомского**
5. Модель пригодна для других формальных языков, не только естественных (кран, телевизор, компьютер)

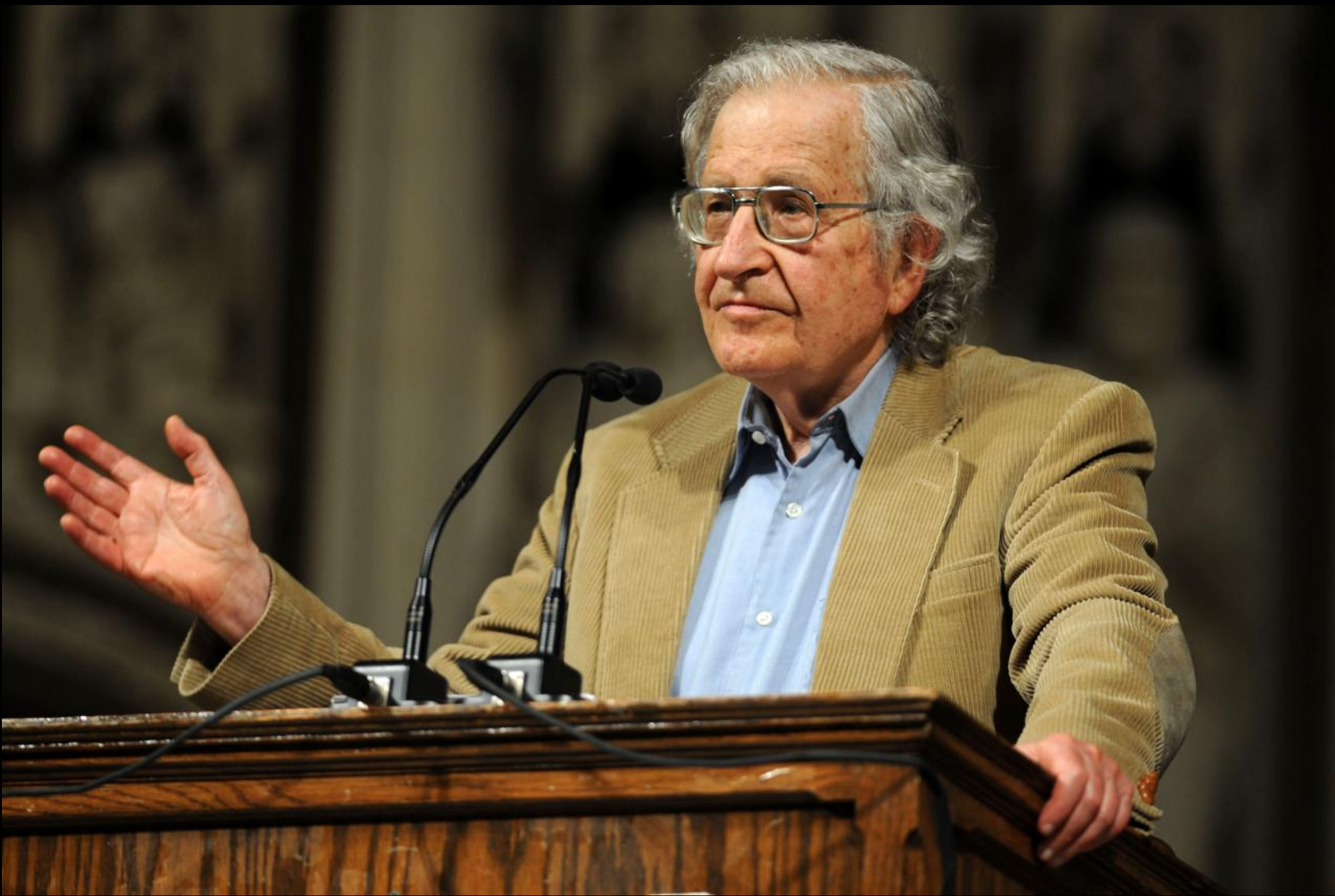
Определение языка

Хомский: язык — множество (конечное или бесконечное) предложений, каждое из которых имеет конечную длину и построено из конечного множества элементов

бесконечное множество предложений

Крыса медленно пробиралась краем канавы, прокопанной из фруктового сада в лес. Шла она, - как полагается грызунам, - слегка вприпрыжку. В своем роде она была образцовым и единственным экземпляром, - огромная, желто-бурая, усатая, - потому что жила она в деревенской усадьбе, так хорошо содержимой и охраняемой хозяевами, что из всего прежнего населения разных человеческих "захребетников" гонения выдержала только эта самая крыса да старый-престарый хорек. Добычи всякого рода в усадьбе и кругом нее было сколько угодно, но все-таки это было прескверное место для паразитов, не обладавших такой хитростью и изворотливостью, как крыса и ее старый приятель, хорек.....

конечное множество элементов - слов



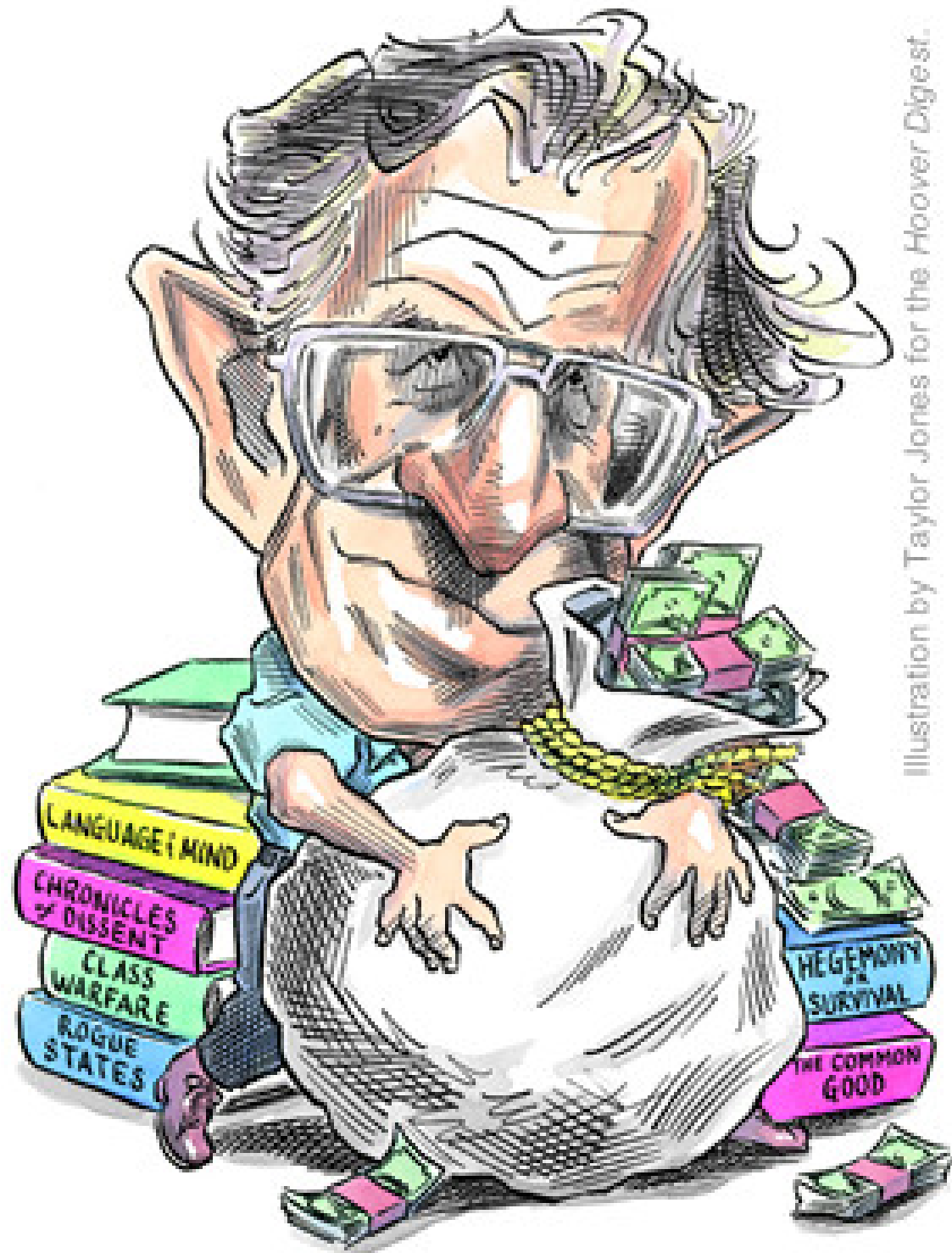


Illustration by Taylor Jones for the Hoover Digest.

Noam Chomsky

Формальный язык

$G = (T, N, S, P)$

T – терминалы

N – нетерминалы

S – корневой символ, один из нетерминалов

P – множество правил (продукций)

Формальный русский язык

$G = (T, N, S, P)$

T – слова

N – синтаксические категории

(существительное, глагол, фраза глагола...)

S – корневой символ, один из нетерминалов

(предложение)

P – множество правил (продукций)

Формальный русский язык

$G = (T, N, S, P)$

$T = \{ \text{мама, раму, мыла} \}$

$N = \{ N, V, S, NP, VP \}$

$S = S$

$P: S \rightarrow NP VP$

$N \rightarrow \text{мама} \mid \text{раму}$

$NP \rightarrow N$

$V \rightarrow \text{мыла}$

$VP \rightarrow V NP$

Порождение предложений

$S \Rightarrow$

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

Порождение предложений

$S \Rightarrow$

$S \rightarrow VP NP$

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

Порождение предложений

S ⇒ NP VP

S → NP VP
NP → N
VP → V NP
N → мама | раму
V → мыла

Порождение предложений

$S \Rightarrow NP VP$

$NP \rightarrow N$

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

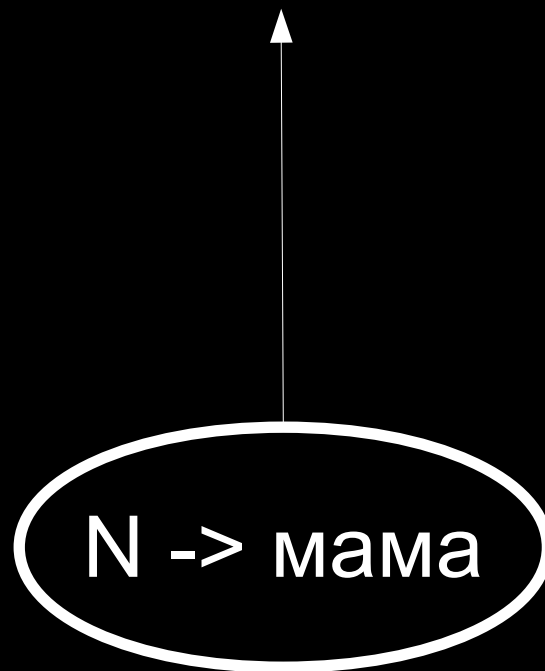
Порождение предложений

S \Rightarrow NP VP \Rightarrow N VP

S \rightarrow NP VP
NP \rightarrow N
VP \rightarrow V NP
N \rightarrow мама | раму
V \rightarrow мыла

Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP$



$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

Порождение предложений

S => NP VP => N VP => мама VP

S → NP VP
NP → N
VP → V NP
N → мама | раму
V → мыла

Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP \Rightarrow \text{мама VP}$

$VP \rightarrow V NP$

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP \Rightarrow \text{мама VP}$

$\Rightarrow \text{мама V NP}$

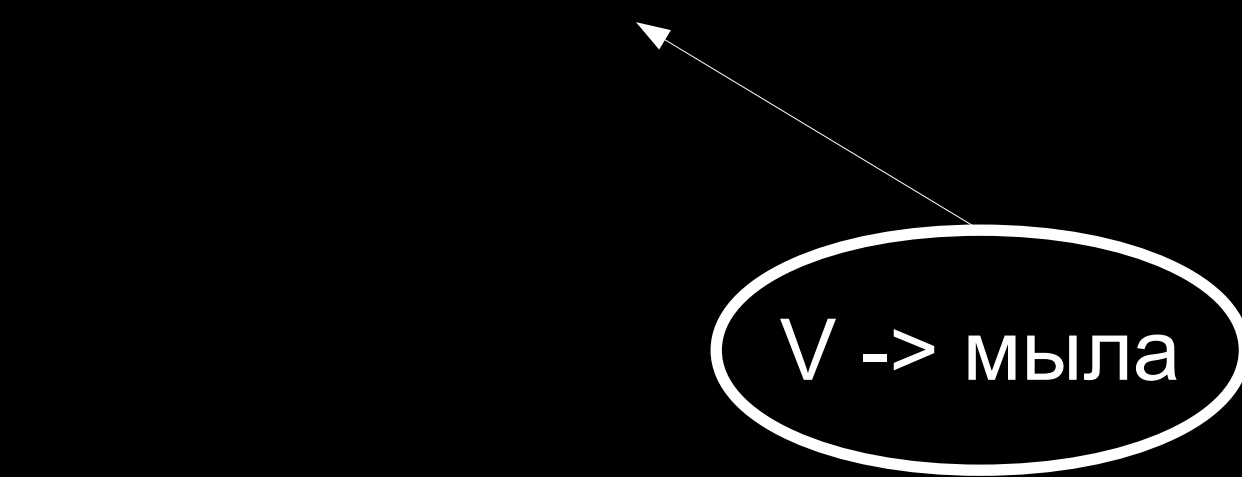
$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP \Rightarrow \text{мама VP}$

$\Rightarrow \text{мама V NP}$

$V \rightarrow \text{мыла}$



$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP \Rightarrow \text{мама} VP$

$\Rightarrow \text{мама} V NP \Rightarrow \text{мама мыла} NP$

$S \rightarrow NP VP$

$NP \rightarrow N$

$VP \rightarrow V NP$

$N \rightarrow \text{мама} \mid \text{раму}$

$V \rightarrow \text{мыла}$

Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP \Rightarrow \text{мама } VP$

$\Rightarrow \text{мама } V NP \Rightarrow \text{мама мыла } NP$

$NP \rightarrow N$

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP \Rightarrow \text{мама } VP$

$\Rightarrow \text{мама } V NP \Rightarrow \text{мама мыла } NP$

$\Rightarrow \text{мама мыла } N$

$S \rightarrow NP VP$

$NP \rightarrow N$

$VP \rightarrow V NP$

$N \rightarrow \text{мама} \mid \text{раму}$

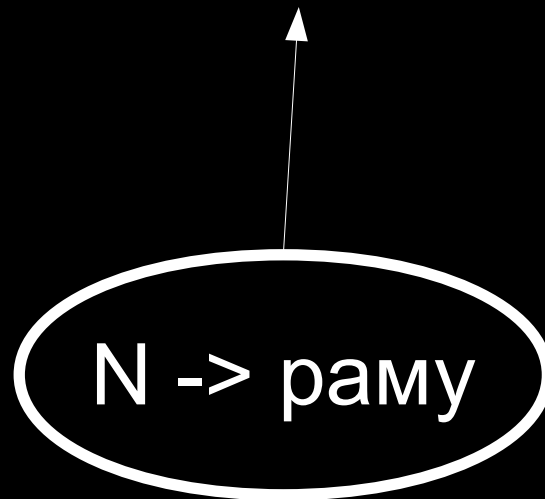
$V \rightarrow \text{мыла}$

Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP \Rightarrow \text{мама } VP$

$\Rightarrow \text{мама } V NP \Rightarrow \text{мама мыла } NP$

$\Rightarrow \text{мама мыла } N$



$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

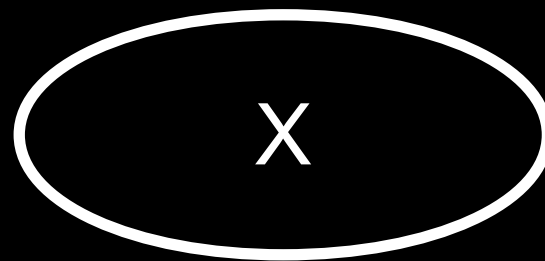
Порождение предложений

$S \Rightarrow NP VP \Rightarrow N VP \Rightarrow \text{мама } VP$

$\Rightarrow \text{мама } V NP \Rightarrow \text{мама мыла } NP$

$\Rightarrow \text{мама мыла } N$

$\Rightarrow \text{мама мыла раму}$



$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

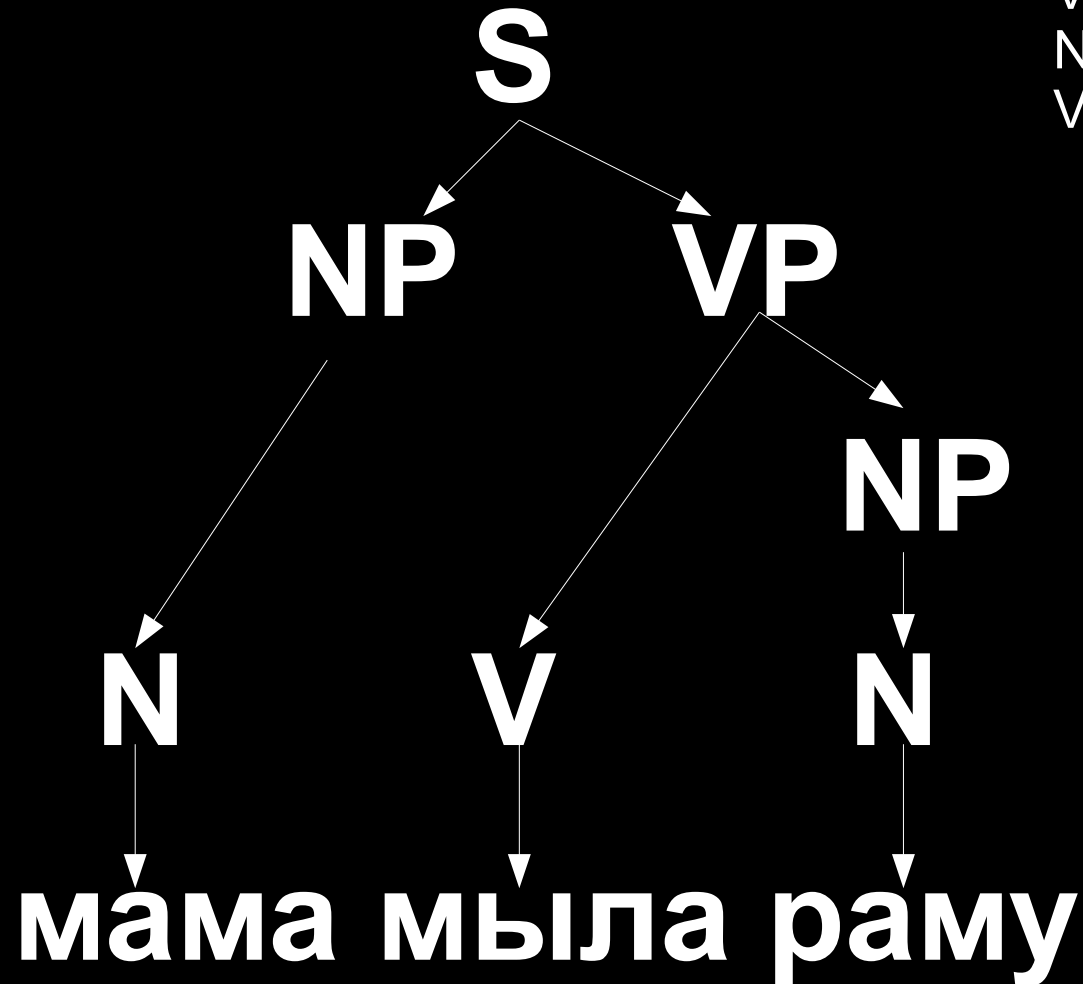
Для чего все это?

Общение:

Компьютер использует продукции грамматики,
порождает правильные предложения языка,
"говорит"

Дерево разбора

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$



Для чего все это?

Проверка синтаксиса:

Построенное дерево разбора – гарантия что предложение принадлежит языку

Проверить синтаксис – значит построить дерево разбора (восстановить порождение по правилам)

Дерево разбора

$S \rightarrow NP VP$

$NP \rightarrow N$

$VP \rightarrow V NP$

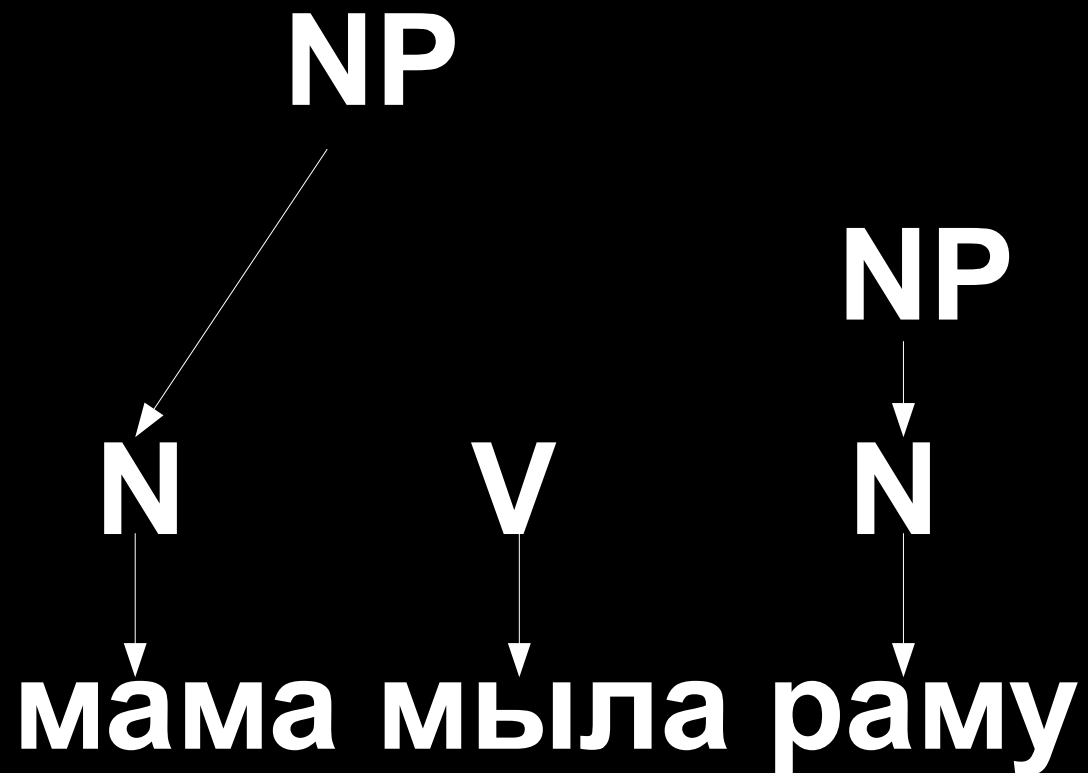
$N \rightarrow \text{мама} \mid \text{раму}$

$V \rightarrow \text{мыла}$

мама мыла раму

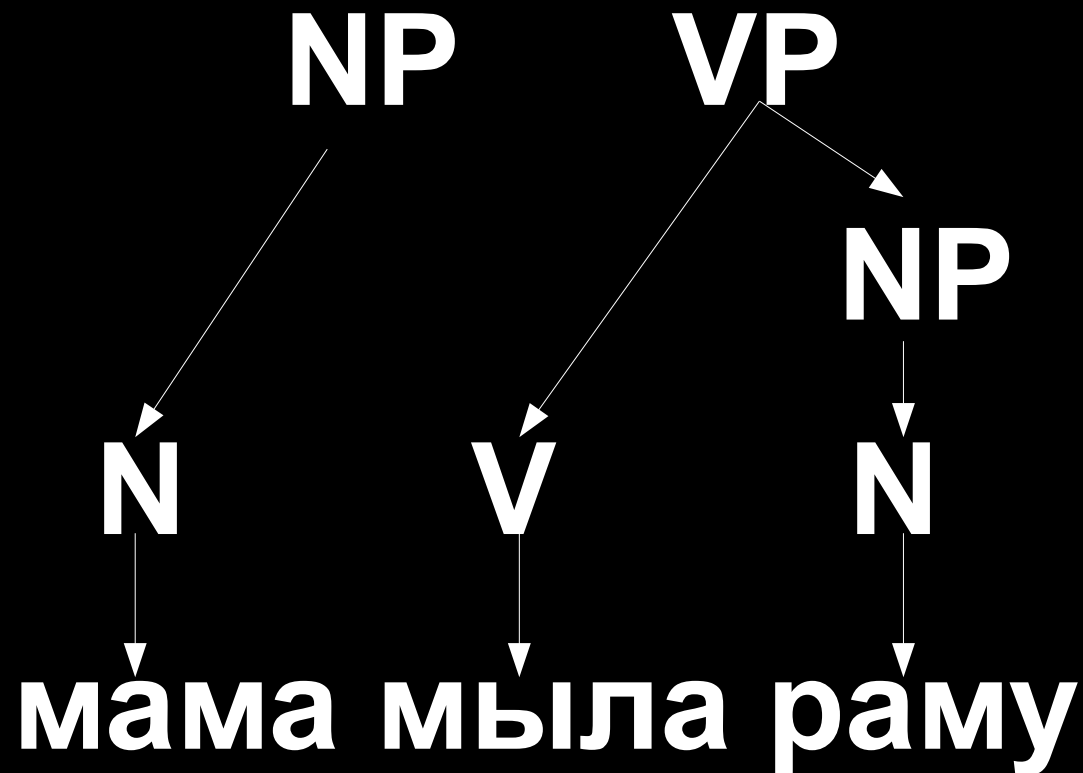
Дерево разбора

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$



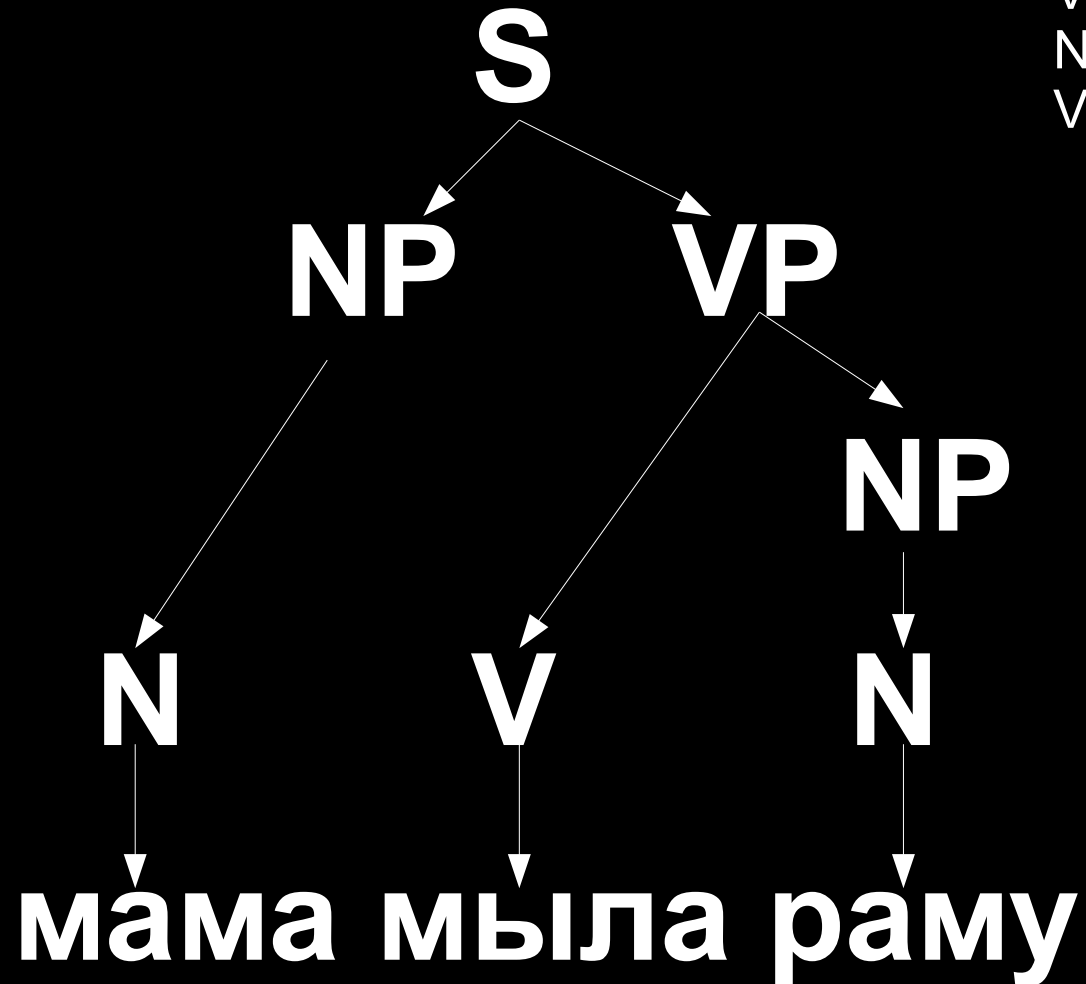
Дерево разбора

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$



Дерево разбора

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$



Дерево разбора

$S \rightarrow NP VP$
 $NP \rightarrow N$
 $VP \rightarrow V NP$
 $N \rightarrow \text{мама} \mid \text{раму}$
 $V \rightarrow \text{мыла}$

мама рыла раму

Контекстно – свободные грамматики

Контекстно-свободные грамматики

Context-free grammars (CFG)

$S \Rightarrow NP VP \Rightarrow N V N \quad \Rightarrow$ мама мыла раму

\Rightarrow ???

Контекстно-свободные грамматики

Context-free grammars (CFG)

$S \Rightarrow NP VP \Rightarrow N V N \quad \Rightarrow$ мама мыла раму

\Rightarrow раму мыла мама

\Rightarrow мама мыла мама

\Rightarrow раму мыла раму

Контекстно-свободные грамматики

Context-free grammars (CFG)

$S \Rightarrow NP VP \Rightarrow N V N \quad \Rightarrow$ мама мыла раму

\Rightarrow раму мыла мама

\Rightarrow мама мыла мама

\Rightarrow раму мыла раму

Контекстно-свободные грамматики

Признак CFG:

в левой части продукций – один нетерминал

S -> NP VP

VP -> V NP

V -> мыла

.....

Контекстно-свободные грамматики

Неадекватность модели:

- склонение существительных**
- СМЫСЛ**

Разрешение неадекватности?

