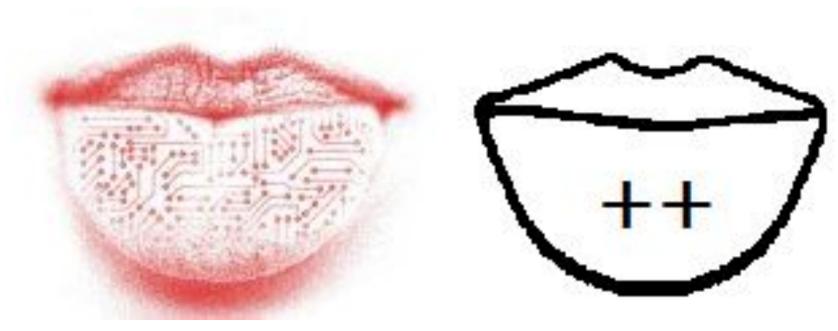
Цифор/Tsifor

байт-код для любого языка из номера основы слова и номера грамматической формы



Главные идеи

```
расчёска-
лакуна
```

блендин

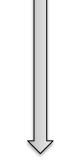
декомпозици я

ЧМИ требует идентификаторы

3483	стойк
3484	стойл
3485	СТОК
3486	стол
3487	столб

на столе ПОД СТОЛОМ к стрлу 3486-24 3486-25 3486-26 бна комо Д





проект «Вофал»

Ком (comb) – расчёска

Бнар – идентификатор основы слова Комод – идентификатор расчёски



Зубцовая классификация языков

```
о 1 2 3 4 5 катализатор предлог префикс трансфикс постфикс послелог расческ
```

изолирующие, аналитические и синтетические языки:

нет зубцов, 1-ые и 5-ые, 2-ые и 4-ые

Лакунная классификация языков

0 1 2 3 4 5 катализатор предлог префикс трансфикс постфикс послелог

01-лакунный (перед 1-м зубцом)

12-лакунный (между 1-м и 2-м, «свой» + «к сыну» = «к своему сыну»: в изолирующих, аналитических и синтетических языках с левым ветвлением, в т.ч. в английском и в русском)

23-лакунный (между 2-м зубцом и основой: в инкорпорирующих с левым ветвлением, например, эскимосских)

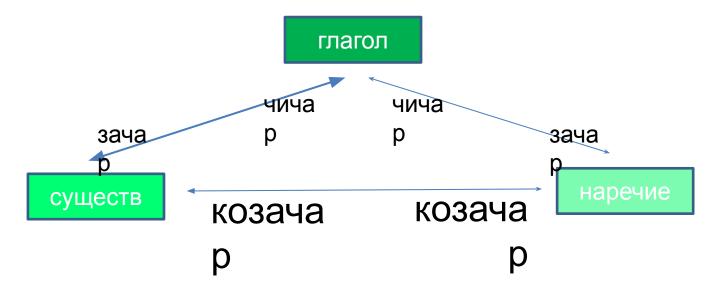
34-лакунный (между основой и 4-м зубцом: в инкорпорирующих с правым ветвлением)

45-лакунный (между 4-м и 5-м: в изолирующих, аналитических и синтетических языках с правым ветвлением) **56-лакунной** (обозначение условно, после 5-го)

Однородные зубцы



Отношения основ



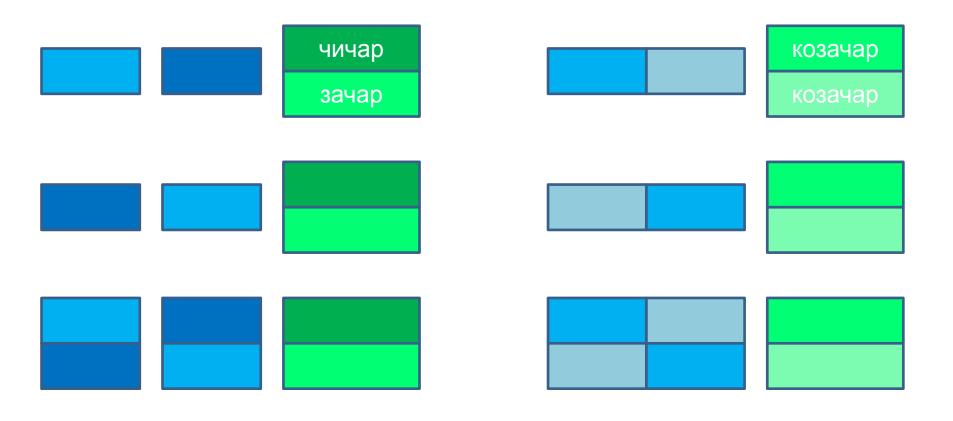
глагольная с существительными > с наречиями,

именная с существительными > с прил. > с числительными

группа прилагательного с существительными («горд братом, чужд разврата, скуп в движениях») > группа прилагательного

C LIONOLIIAGINAIA

Родственные и козачарные зубцы



Чукотский: инкорпорация

га-вэтъат-арма-қорама

(с бодливым сильным



Кавказские: между, мимо,





Немецкий: зубец vs. основа

Er hörte gestern Abend dem Redner eine halbe Stunde lang zu

(Вчера вечером он слушал полчаса выступающего)

Er hat dem Redner eine halbe Stunde lang zugehört

(Он слушал полчаса выступающего)

Hörst du zu? Hast du zugehört?

Перечень лакунностей групп

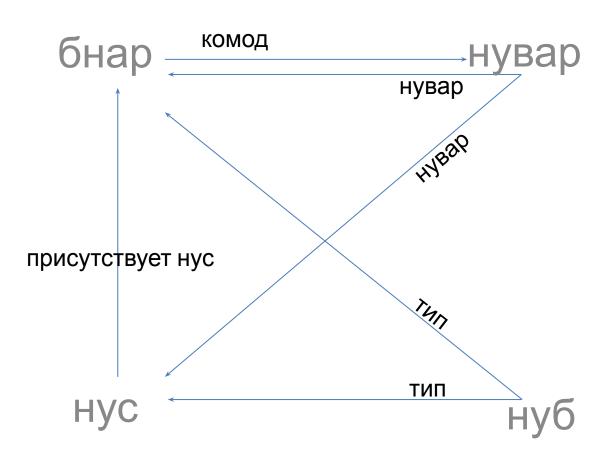
You walk. Do you walk? Сто быков. Быков сто

Группы с несколькими *лакунностями*: в английском вопросительном и повествовательном предложениях, в русском языке в точном и приблизительном выражении количества и т.д.

глагольная с существительными глагольная с наречиями именная с существительными именная с прилагательными именная с числительными группа прилагательного с существительными группа прилагательного с наречиями

Блендинг: граммемы и константы

(«молоко&творительный пять литр» \rightarrow «пятью литрами молока»)



Нувар

Нувар состоит из не более двух нузнов (которые будем отличать от грамматической категории «число», которой числительное в некоторых языках обладает)

1-й зубец («десять человек» в русском, в индонезийском), 2-й зубец (приставки СИ и МЭК вроде «гига» и «гиби»), 5-й зубец («человек десять» в русском, в суахили) и т.д.

Бнар, нус, нувар, нуб

«тройка сотней» → *полковое*

«семи целых трех пятых» → наращивание

«вторые три четверти, пятая тройка сотней, семь целых три пятых»

Классификация грамматических категорий

вогр (неизменяемые и хранящиеся в ОС: спряжение, склонение, именной класс)

лугр (изменяемые и кодируемые в цифоре: время, залог, вид, эвиденциальность, наклонение, адмиратив, краткая/полная форма, отрицание, модальность, число)

ныгр (изменяемые, но некодируемые, а вычисляемые по положению в клаузе: «первый, промежуточный, последний» групповой флексии сущ. в тюркских и осетинском)

Декомпозиция на граммемы

0 1 2 3 4 5 катализатор предлог префикс трансфикс постфикс послелог

0-й в австралийских изменяется по всем глагольным категориям (несет вместо глагола), 1-й в полинезийских – по «собственное-нарицательное», в ирландском – по числам и лицам, 4-й в русском от склонения и бита отрицания («вижу смысл»

- «не вижу смысла», «есть смысл» - «нету смысла»)

Супплетив – основа, тефорвированная не из одного бнара, а из бнара и инида

Пул(pull) – начальная форма расчёски

Комод – идентификатор расчёски

Пулид – идентификатор начальной формы расчёски

(«не» глагола в актанты: «вижу смысл – не вижу смысла», «есть смысл – нету смысла)

малугр смысла)

чилугры залугры
Чичар

чивогры

мавогры

(число, лицо, род актанта в им.п. в гл.: число и лицо в наст. и буд.вр., число и род – в прош.вр.)

зачар

завогры

(падеж, число, род сущ. в прил., прич., притяж. местоим.)

Форма Бэкуса-Наура

«глагол-существительное-прилагательноенаречие»

```
ГЛАБУР ::= гланость #глаг {ЗОНТ} [ГЛАБУР] [Ин]
                           [CH][\Pi H]
                                                  [\mathsf{Mn}][\mathsf{Cn}]
[\Pi\Pi].
3OHT ::= суность #сущ {3OHT} [ГЛАБУР] [Ин]
                           [Объем] [Иллю] [#нус] [Ип] [Сп]
[\Pi\Pi].
 Ип ::= приность #прил {ЗОНТ} [ГЛАБУР] [Ин].
 Сп ::= приность #прил {ЗОНТ} [ГЛАБУР] [Ин].
 Пп ::= приность #прил {ЗОНТ} [ГЛАБУР] [Ин].
  Ин ::= наность #нач [#зач].
  Сн ::= наность #нач [#зач].
  Пн ::= наность #нач [#зач].
   Объем ::= [нуб нувар].
   M_{\text{BB}} = \dots = M_{\text{BB}}
```

Резюме

Дает

единое устройство языков с т.з.

ИНЖенера (vs. философствований Хомски)

перевод нейросетью атомарного

субстрата (vs. "aaaaaaaaaaaaa" → осмысленный перевод)

уменьшение и ускорение

(хранение корпусов, поиск)

описание семантически сегментированных изображений

Докладчик ищет безвозмездную помощь

программистов и лингвистов

Пишите письма

