

# Дигитализация словаря сочетаемости локальных превербов с суффиксоидами и глагольными корнями [Клычев 1995]

Георгий Алексеевич Мороз

Международная лаборатория языковой конвергенции (НИУ ВШЭ)

«Анатолий Несторович Генко и Кавказ», Черкесск, 25 ноября 2021



# План доклада

Работы Р. Н. Клычева

Процесс дигитализации словаря

Возможное применение полученной базы данных

Заключение

# Работы Р. Н. Клычева, посвященные глагольным аффиксам в абазинском

- [Клычев 1972]: “Глагольные суффиксы направления в абазинском и абхазском языках (по данным всех абазино-абхазских диалектов)”
- [Клычев 1994]: “Локально-превербное образование глаголов абазинского языка”
- [Клычев 1995]: “Словарь сочетаемости локальных превербов с суффиксоидами и глагольными корнями в абазинском языке”
- и другие работы [Клычев 1974, 1983, 1989, 1990].

## Подробнее о [Клычев 1995]

- Словарь представляет собой 4283 глагольные леммы, разделенных на разделы в зависимости от морфем;

**2784. ТшкІара-га-ра** перех. АРС, 363. Занести, завести, заносить, завозить кого, что-л. во двор. ХІахъвгъи нышвгъи тшкІарайгитІ (й-тшкІара-й-г-итІ). Аб. Т. Он (м.) заносит во двор и камень, и глину.

## Подробнее о [Клычев 1995]

- Словарь представляет собой 4283 глагольные леммы, разделенных на разделы в зависимости от морфем;

**2784. ТшкІара-га-ра** перех. АРС, 363. Занести, завести, заносить, завозить кого, что-л. во двор. ХІахъвгъи нышв-гъи тшкІарайгитІ (й-тшкІара-й-г-итІ). Аб. Т. Он (м.) заносит во двор и камень, и глину.

- Каждое словарное гнездо содержит
  - уникальный индекс: 2784;
  - разделенную на морфемы лемму: ТшкІара-га-ра;
  - грамматическую помету: перех.;
  - перевод на русский язык: занести, завести, заносить, завозить кого, что-л. во двор;
  - 1–2 примера употребления с переводами.

# План доклада

Работы Р. Н. Клычева

Процесс дигитализации словаря

Возможное применение полученной базы данных

Заключение

# Процесс дигитализации словаря

- отсканированную версию книги прогнали через пакет tesseract на языке программирования R для применения оптического распознавания символов [Ooms 2021]

**2784. ТшкІара-га-ра перех. АРС, 363. Занести, завести, заносить, завозить кого, что-л. во двор. ХІахъвгъи нышв-гъи тшкІарайгитІ (й-тшкІара-й-г-итІ). Аб. Т. Он (м.) заносит во двор и камень, и глину.**

ТшьТара-га-ра перех. АРС, 363. Занести, завести, заносить, завозить кого, что-л. во двор. ХТахъвъгъи нышв-гъи тшкТарайгит! (й-тшьТара-й-г-ит!). Аб. Т. Он (м.) заносит во двор и камень, и глину.

- полученный текст был преобразован в таблицу, выверен и разделен на поля лемма - грамматическая помета - перевод - пример (Г. Мороз, Е. Кожанова, А. Корженевская, А. Кузнецова, Д. Илларионова, Д. Сосновская, П. Багирова, К. Родионова, З. Керимова);
- полученные леммы разделены на префиксы, корень и суффиксы.

## Результат дигитализации словаря

- В результате получилась таблица со следующими полями:
  - `section_id`: 94
  - `section_name`: ПРЕВЕРБ тшкІара+КОРНИ
  - `pages`: 220
  - `entry_id`: 2784
  - `abaza`: тшкІара-га-ра
  - `category`: перех.
  - `preverb`: тшкІара
  - `root`: га
  - `suffix`:
  - `msd`: ра
  - `root_meaning`: нести
  - `translation`: занести, завести, заносить, завозить кого, что-л. во двор;
  - `examples`: ХІахъвъгы нышвъгы тшкІарайгитІ (й-тшкІара-й-г-итІ). Аб. Т. Он (м.) заносит во двор и камень, и глину.
  - `comment`:
  - `contributor`: Pakiza Bagirova



# План доклада

Работы Р. Н. Клычева

Процесс дигитализации словаря

Возможное применение полученной базы данных

Заключение

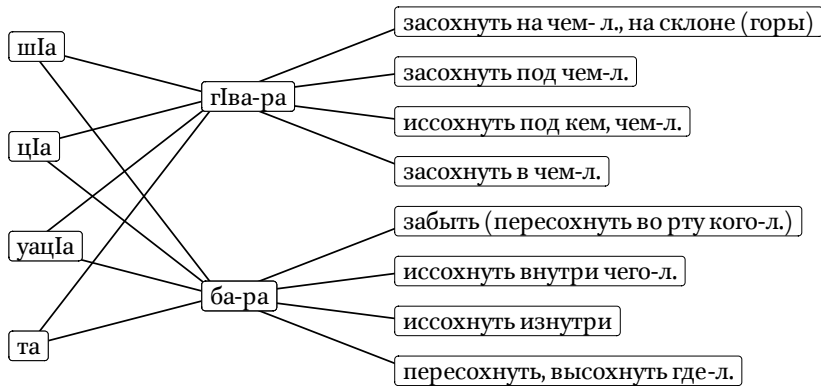
# Возможное применение полученной базы данных

- расширение словарей
  - поморфемный поиск предоставляет дополнительные возможности для исследования
  - это в свою очередь улучшает разные прикладные применения, например, системы проверки правописания, предиктивное правописание
- пополнение корпуса абазинского языка  
[http://lingconlab.ru/spoken\\_abaza/](http://lingconlab.ru/spoken_abaza/) [Panova et al. 2021]
- дальнейшие исследования сочетаемости морфем

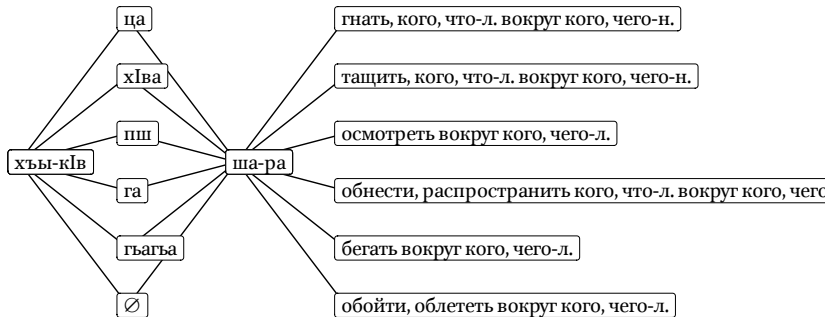
# Поморфемный поиск: копать



# Поморфемный поиск: сохнуть<sub>1</sub>, сохнуть<sub>2</sub>



## Поморфемный поиск: *xъы-кIв*



См. также [Клычев 1974]: “К структуре глагольной основы *кIв* + корень + *ш(a)* в абазинском языке”

# План доклада

Работы Р. Н. Клычева

Процесс дигитализации словаря

Возможное применение полученной базы данных

Заключение

# Планы на будущее

- дальнейшее выверка базы данных
  - опечатки
  - морфемное деление
- завершение аннотации корней и морфем
- публикация онлайн и преобразование в формат словаря

- дигитализация — это важный процесс преобразования опубликованных данных в электронный формат
- результаты дигитализации можно использовать
  - для создания новых инструментов
  - для проверки старых гипотез
  - для постановки новых вопросов



Спасибо за внимание!

## References

- Ooms, J. (2021). *tesseract: Open Source OCR Engine*. R package version 4.1.2.
- Panova, A., Sorokina, A., Arkadiev, P., and Sokur, E. (2021). Spoken corpus of abaza. Moscow: School of Linguistics, HSE University; Linguistic Convergence Laboratory, HSE University. (Available online at: [http://lingconlab.ru/spoken\\_abaza/](http://lingconlab.ru/spoken_abaza/), accessed on 23.11.2021.).
- Клычев, Р. Н. (1972). *Глагольные суффиксы направления в абазинском и абхазском языках (по данным всех абазино-абхазских диалектов)*. Черкесск: Ставропольское книжное издательство Карачаево-Черкесское отделение.
- Клычев, Р. Н. (1974). К структуре глагольной основы кІв + корень + ш(а) в абазинском языке. *Вопросы абазинской и адыгской филологии*, 2:3–6.

- Клычев, Р. Н. (1983). К структуре абазино-абхазских отглагольных имен с превербными. In *Система превербов и послеслогов в иберийско-кавказских языках*, pages 112–117. Ставрополь.
- Клычев, Р. Н. (1989). К семантике и структуре абазинско-абхазских отглагольных имен с превербными. In *Отглагольные образования в иберийско-кавказских языках*, pages 58–61. Ставрополь.
- Клычев, Р. Н. (1990). Из глагольного словообразования в абазинском языке. *Ежегодник иберийско-кавказского языкознания*, 17:47–58.
- Клычев, Р. Н. (1994). *Локально-превербное образование глаголов абазинского языка*. Черкесск: ТОО "Аджьпа".

Клычев, Р. Н. (1995). *Словарь сочетаемости локальных превербов с суффиксоидами и глагольными корнями в абазинском языке*. Черкесск: Карачаево-черкесское книжное издательство.