

# UD for Abaza

## Tokenization and Word Segmentation

- All words are converted to lower case.
- All punctuation marks except full stops and question marks are removed from the texts.
- Words are delimited by whitespace characters and punctuation.
- Nominal complexes written as single words, e.g. *артистду* ‘big artist’, are treated as single tokens.

## Morphology

### Tags

Abaza uses all universal POS tags except `PART` and `SYM`.

1. `ADJ` (adjective): nominals which modify another nominal.

xl-ан	рыцхla	л-хлатыр-ла
<i>h-an</i>	<i>rəcha</i>	<i>l-hatər-la</i>
1PL.PR-мать	бедный	3SG.F.PR-услуга-INS
‘благодаря нашей бедной маме’		

2. `ADP` (adposition): all Abaza postpositions.

аллахl	й-пны	xl-аргваны-та
<i>allah</i>	<i>j-pnə</i>	<i>h-argʷanə-ta</i>
аллах	3SG.M.IO-возле	1PL.ABS-близкий-ADV
‘чтобы мы приблизились к Аллаху’		

3. `ADV` (adverb): verbal modifiers which specify manner, time or place.

а-рахв	хвит-та	дара-дара	ты-цл-и-тl	щымта-ла
<i>a-raxʷ</i>	<i>xʷəjt-ta</i>	<i>dara-dara</i>	<i>tə-ç-əj-t</i>	<i>šʷəmta-la</i>
DEF-скот	свободный-ADV	3PL-3PL	LOC.ELAT-идти-PRS-DCL	утро-INS
‘Коровы сами выходят по утрам’				

4. `AUX` (auxiliary): the verb *čʷpa-* ‘do’ with borrowings, copula (not sure).

а-намыца-ква	бомбить	гла-р-чпа-н
<i>a-naməsa-kʷa</i>	<i>bombitʷ</i>	<i>ɣa-r-čʷpa-n</i>
DEF-немец-PL	бомбить	CSL-3PL.ERG-делать-PST
‘(Ростов) немцы бомбили’		

5. `CONJ` (coordinating conjunction): *jgʷəj* ‘and’, *ja* ‘or’, etc.

я	тшчlвара-kl	я	уарад-kl
<i>ja</i>	<i>ččara-k</i>	<i>ja</i>	<i>warad-k</i>
или	куплет-indf	или	песня-indf
‘или куплет или песню’			

6. `DET` (determiner): demonstrative pronouns and personal pronouns used as possessors (do we really need this category?).

- a. 

ани	а-дзы	й-цл-на-хlва-ква-з
<i>anəj</i>	<i>a-ɬə</i>	<i>j-ç-na-hʷa-kʷa-z</i>
MED	DEF-вода	3PL.ABS-LOC-3SG.N.ERG-вытягивать-PL-PST.NFIN
‘эта вода высосала из нас всё’		

- б. 

сара	с-ан	д-адыгъа-б
<i>sara</i>	<i>s-an</i>	<i>d-adəgʷa-b</i>
1SG	1SG.PR-мама	3SG.H.ABS-черкез-NPST
‘моя мама черкешенка’		

7. `INTJ` (interjection): *šʷta*, maybe something else.

щта тш-а-д-хl-кlылы-н

*š'ta č-a-d-h-qlə-lə-n*

да RFL-DAT-LOC-1PL.ERG-собираться-PST

‘Мы вместе собрались’

8. NOUN (noun): nominals functioning as syntactic arguments.

й-жвпlа-ха-уа-мца а-жвумп-ква

*j-žpa-xa-wa-mca a-žwmp-k'a*

3PL.ABS-густой-INC-IPF-CVB DEF-ягода-PL

‘Гуще стало ягод’

9. NUM (numeral): words denoting numbers.

а-сквшы зкыи жвшви гlважвы жвба асхъан

*a-sk'šə zk'əj žšəj ɣ'ažə žba asqan*

DEF-год тысяча девятьсот двадцать девять во.время

‘(она родилась) в 1929-ом году’

10. PRON (pronoun): personal pronouns (used as arguments), demonstrative pronouns in the function of personal pronouns.

сара ачlвы-йа йы-с-чп-у-шы-з

*sara ač'wə-ja jə-s-č'p-əw-šə-z*

1SG что-QN REL.ABS-1SG.ERG-делать-IPF-FUT-PST.NFIN

‘Что мне было делать?’

11. PROP (proper noun): personal names, names of places, etc.

абдулгамид й-аща-хlба

*abdulgamid j-aš'a-hba*

Абдулгамид 3SG.M.PR-брат-старший

‘старший брат Абдулгамида’

12. PUNCT (punctuation): full stops, question marks.

13. SCONJ (subordinating conjunction): *h'a* when used as a complementizer.

жена тащи галушки хlва ауи й-хlвА-д-йа

*žena taš'i galuški h'a awəj j-h'a-d-ja*

жена(R) тащи(R) галушки(R) QUOT DIST.SG 3SG.M.ERG-говорить-DCL-PTCL

‘[жена, тащи галушки (по-русски)] - он говорит’

14. VERB (verb): all wordforms with verbal morphology (see the list of verbal affixes in the next section).

апхъа а-заман уаɣlа.хъа-та а-дунай йы-квы-з

*apqa a-zaman waɣa.qa-ta a-dəwnaj jə-k'ə-z*

раньше DEF-время народ-ADV DEF-мир REL.ABS-быть\_на-PST.NFIN

ш-абга-з й-азкlкы-та зджьара й-бзаз-у-н

‘в то время люди, которые находились на свете, собравшись, жили где-то’

*š-abga-z*

*j-azḳḳə-ta*

*zž'ara*

*j-bzaz-əw-n*

REL.MNR-целый-PST.NFIN

ЗPL.ABS-собираться-ADV

где\_то

ЗPL.ABS-жить-IPF-PST

‘в то время люди, которые находились на свете, собравшись, жили где-то’

15. x (other).

## Features

•

## Syntax

•

## Treebanks

There are N ([../treebanks/LCODE-comparison.html](#)) LANGUAGE UD treebanks:

- LANGUAGE-A ([../treebanks/LCODE\\_a/index.html](#))
- LANGUAGE-B ([../treebanks/LCODE\\_b/index.html](#))