Time nouns and temporal constructions in the languages of Dagestan

Timur Maisak
Institute of Linguistics RAS & Linguistic Convergence Laboratory, HSE University

Samira Verhees
Independent researcher

Outline

Introduction - Time nouns and their functions in East Caucasian

Data and methodology

Attested lexemes and their semantics

Distribution of inventories and individual lexemes

Time nouns in temporal constructions

Discussion

Time nouns

East Caucasian languages typically have multiple words for 'time'.

Many languages borrowed terms from Arabic (possibly through Persian), though other sources are also attested (Armenian in Udi, Avar in Andic and Tsezic languages).

Arabisms are widespread but not uniformly distributed; presence of specific lexemes reveals distinct areal patterns.

Time nouns are also used to form temporal constructions.

Time nouns as head of temporal subordinate clauses

(1) Andi

[ƛer-di   w-oqo-b]=**rihi**   hege-šːu-b-o

near-LAT   M-reach-PST.PTCP=time   DEM-M.OBL-INAN1-AFF

haɢo-dːu   b-eč'uχʷa   qala

see-PRF   INAN1-big   palace

'<u>When he came closer</u>, he saw a big palace.'

Time nouns as head of temporal adverbials

(2) Avar

Krym **do-b** **zaman-al-da** *b-uk'-ana*

Crimea DEM-N time-OBL-LOC N-be-AOR

Turk-az-ul *nuχmałi-jal-da*

Turkish-OBL.PL-GEN leadership-OBL-LOC

'Crimea was under Turkish rule <u>at that time</u>.'

Constructions (1) and (2) often employ the same nouns in the same case form, and could reveal additional areal patterns.

Time nouns

In part 1 of this talk we present an inventory of attested time nouns.

We discuss their distribution across the East Caucasian family and other languages in the neighborhood with the aim of revealing patterns of language convergence.

In part 2 we introduce some interesting features of temporal constructions with time nouns.

Data and methodology

Collect time nouns recorded in dictionaries following the methodology of the [Typological Atlas of the Languages of Dagestan](#).

Keyword *time* (or *время* in Russian) and several related notions such as *term*, *period*, *era*, *season*, *age* (as in lifespan).

Annotate lexemes originating from a single source.

Likely borrowing sources were annotated with the help of secondary literature.

Data on temporal constructions come from descriptive grammars and our own previous work with speakers and written texts.

Sample

All East Caucasian
languages
(incl. Caucasian
Albanian);
Azerbaijani, Kumyk,
Nogai; Armenian, Tat;
Georgian

Not included:
Russian, Arabic,
Persian
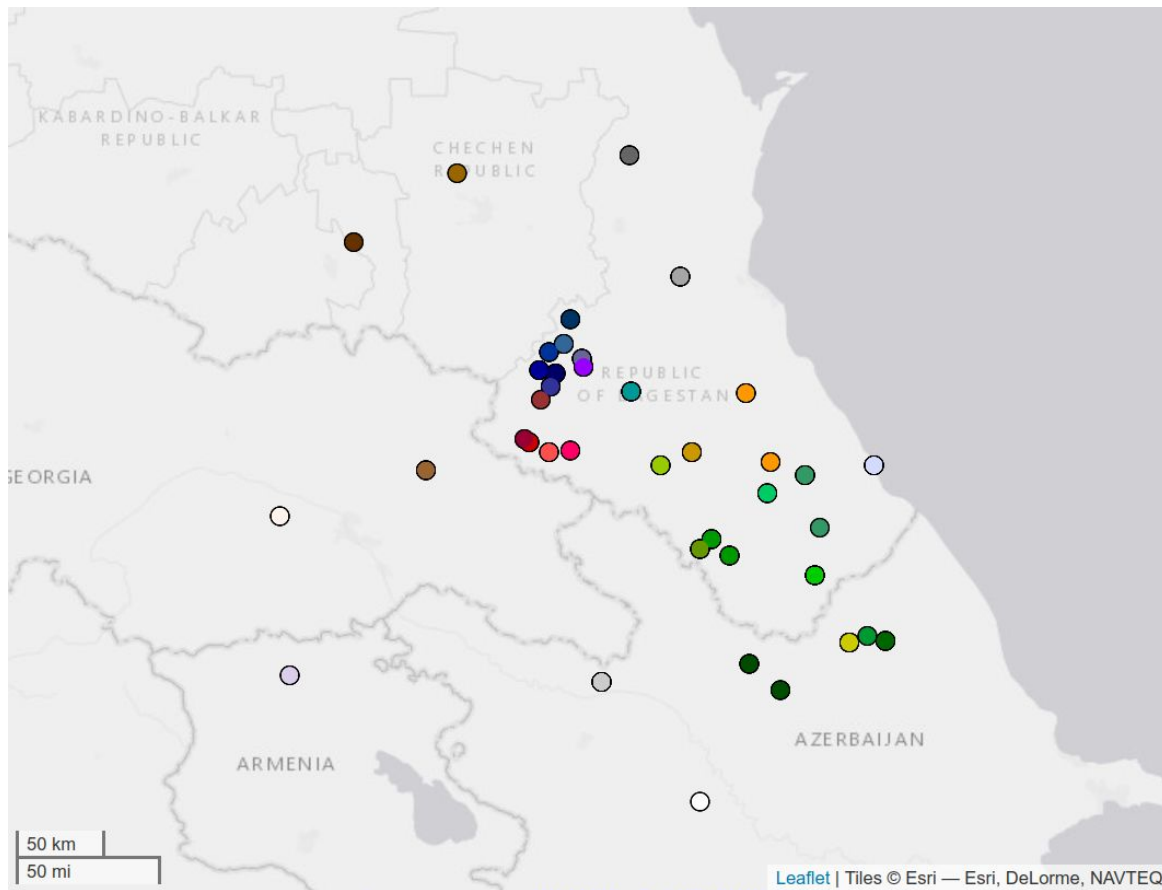
Attested lexemes and their semantics

Time nouns in East Caucasian denote

1) 'time' as an abstract concept
2) a certain timespan
   a) punctual (exact moment when something happened)
   b) durative (denoting a period)

Several more colexified concepts (e.g. 'date', 'age'), depending on the language.

Attested lexemes and their semantics

Dictionary definitions are not very informative because they rely on metalanguage lexemes that cover both types of meanings (e.g. *time* or *время*).

In East Caucasian the abstract meaning seems limited to Arabic borrowings, cf. (3a) with Arabic *zaman* and ungrammatical (3b) with native *meχ* from Avar.

Both can refer to a timespan.

(3a) *di-r*   **zaman** *heč'o*   (3b) **di-r*   **meχ** *heč'o*

    I.OBL-GEN   time   NEG.COP   I.OBL-GEN   time   NEG.COP

    'I don't have time.'   'I don't have time.'

Attested lexemes and their semantics

In Georgian, native დრო *dro* refers to time as an abstract concept, while time spans are referred to with ხანი *xani*, which is likely a borrowing from Nakh languages.

For East Caucasian the two functions are usually not so clearly delimited.

The punctual time measurement ('moment') does not commonly occur as a definition of time nouns in our sample, but whether or not they can denote such a meaning remains to be verified.

Dedicated words for 'moment' are also attested, but it is unclear how common they are.

Colexifications

Colexified concepts occurring more than twice in the sample:

- date
- season
- age (as in lifespan)
- respite, postponement (MUHLAT)
- free time (DEVRAN)

Attested lexemes and their semantics

181 lexemes or tokens in total, 163 of which have 'time' as definition in the dictionary.

Tokens were grouped into TYPES according to the source lexeme.

For example, different phonological adaptations of an Arabic word are in the same set (e.g. *vaχt, vaχt', väχt* etc. > type VAXT).

Cognates are also in a single set.

A type can be limited to one lexeme / token.

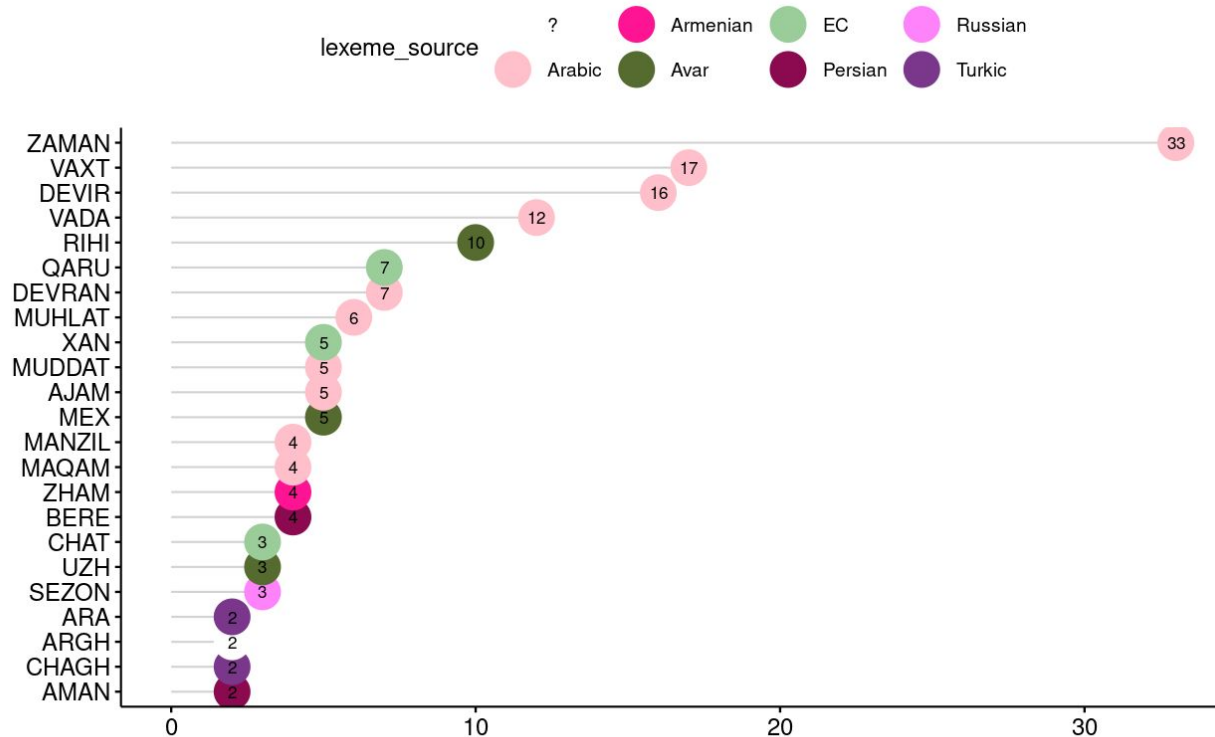# Distribution of lexemes and inventories

181 lexemes in total

- mean/median of 4 nouns per idiom (min 2, max 12)

43 types

- 10 types of native lexemes; 10 from Arabic; 4 Turkic; 3 Avar; 3 Persian; 2 Iranian (Tat); 1 Armenian; 1 Georgian; 1 Russian
- 7 unknown; 1 mixed

Tokens per type (> 1) with source

Distribution of inventories

Similarity between languages based on the number of shared lexemes in a pairwise comparison (Jaccard distance).

0 distance = same inventory.

In case of smaller inventories, it is easier to be similar, e.g. 2 languages with 2 lexemes each, both of which are of the same origin (cognate or borrowing) = 0 distance (Chechen and Ingush).

If one language has 3 lexemes, another 2, and they have 1 in common, distance = 0.75 (Chechen and Georgian).

Because dictionary data can be rather noisy and asymmetrical, these distance measures are preliminary.
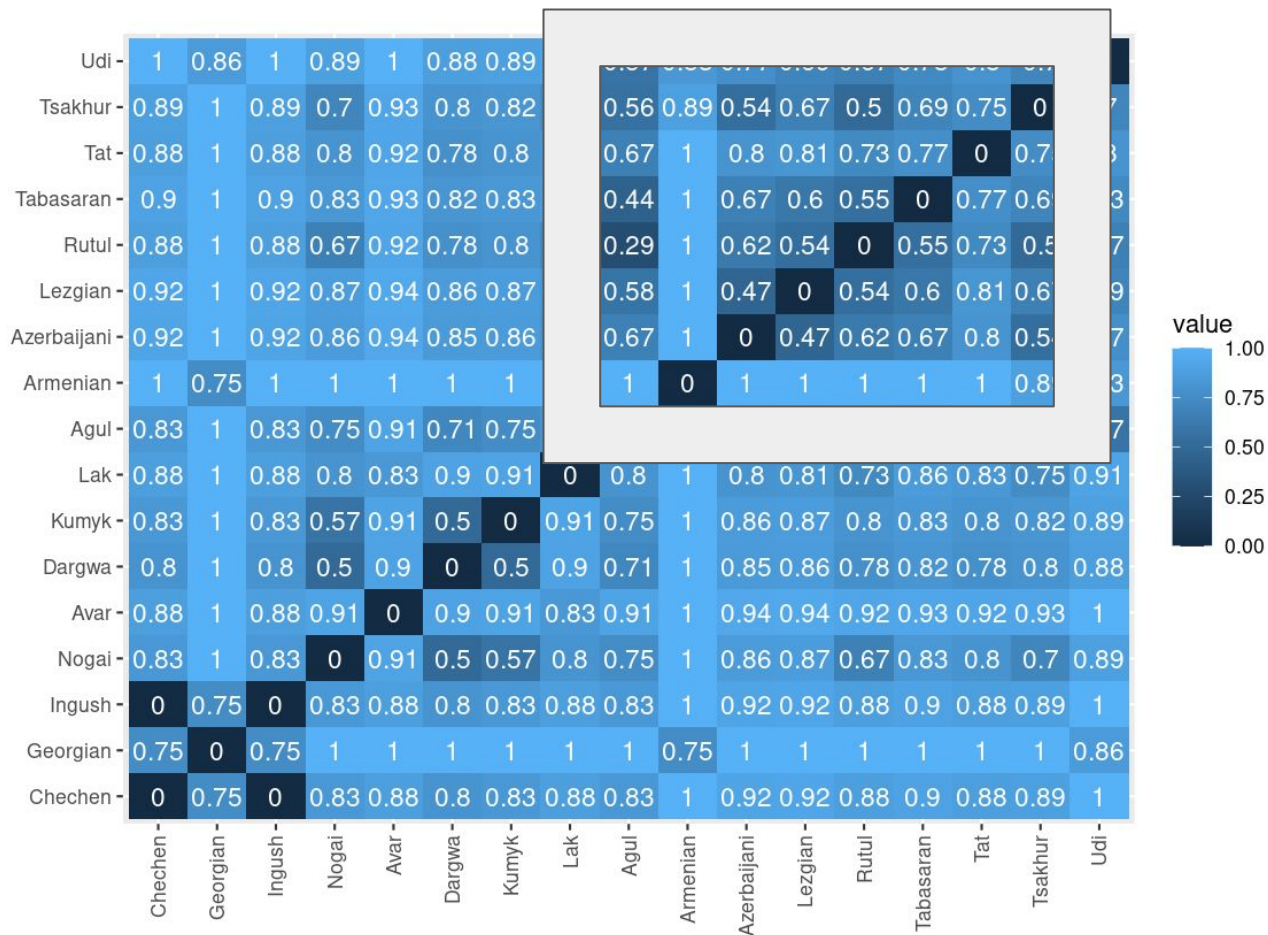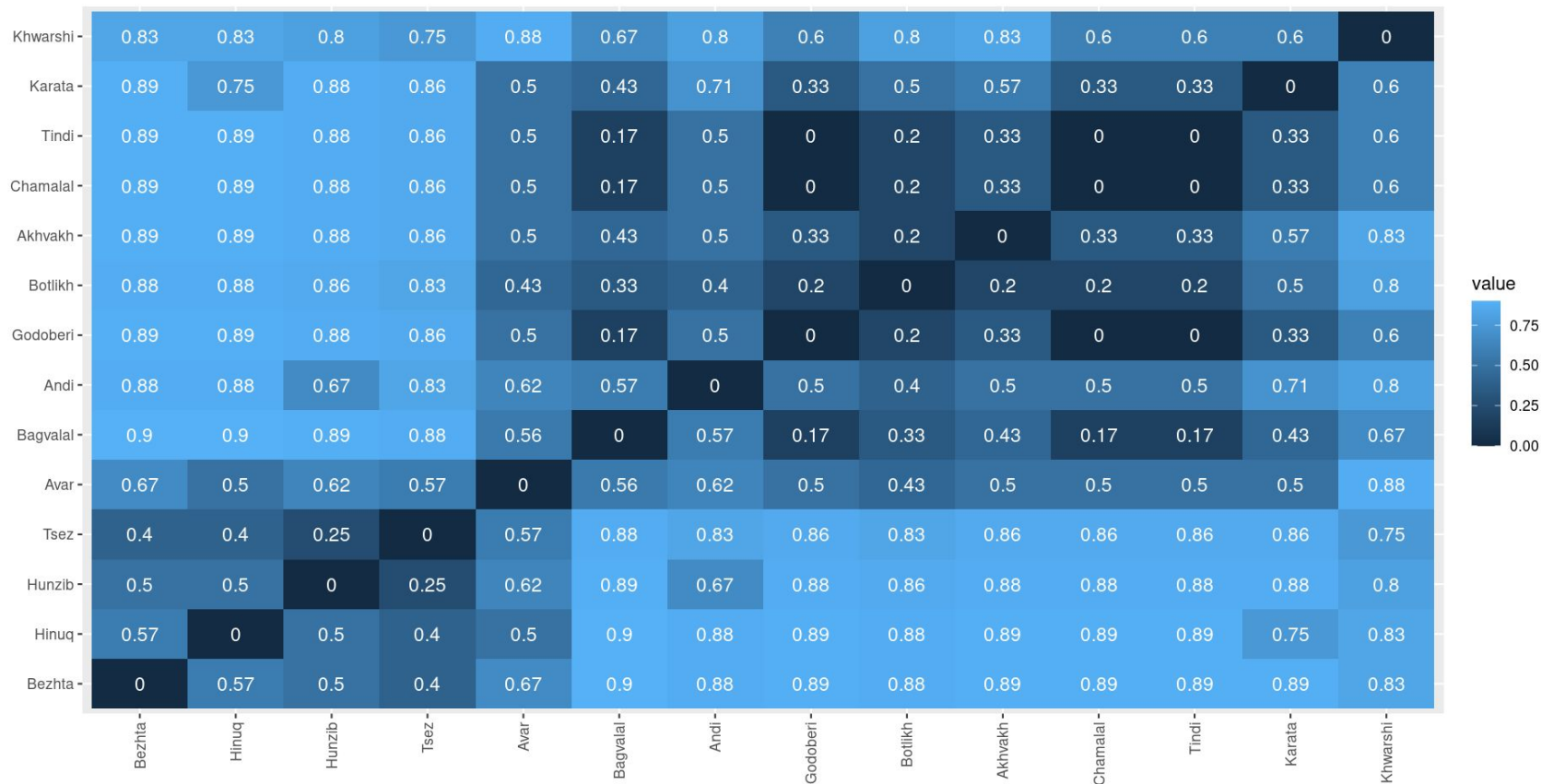
Distribution of lexemes

ZAMAN / VAXT

A form of Arabic زَمَن *zaman* is found in almost all languages of the sample, including Judeo-Tat, but not Udi, Tsova-Tush, Georgian and Armenian.

In many cases variation *zaman / zamana* is attested; the latter could point to Persian influence according to Zabitov (2001: 109).

A form of Arabic وقت *waqt* is found among languages of the Lezgic branch (except Archi), all three Turkic languages, (Judeo-)Tat, and both Dargwa varieties in the sample (Standard Dargwa and Kubachi).

In all cases except Udi it exists alongside *zaman*.

ZAMAN / VAXT
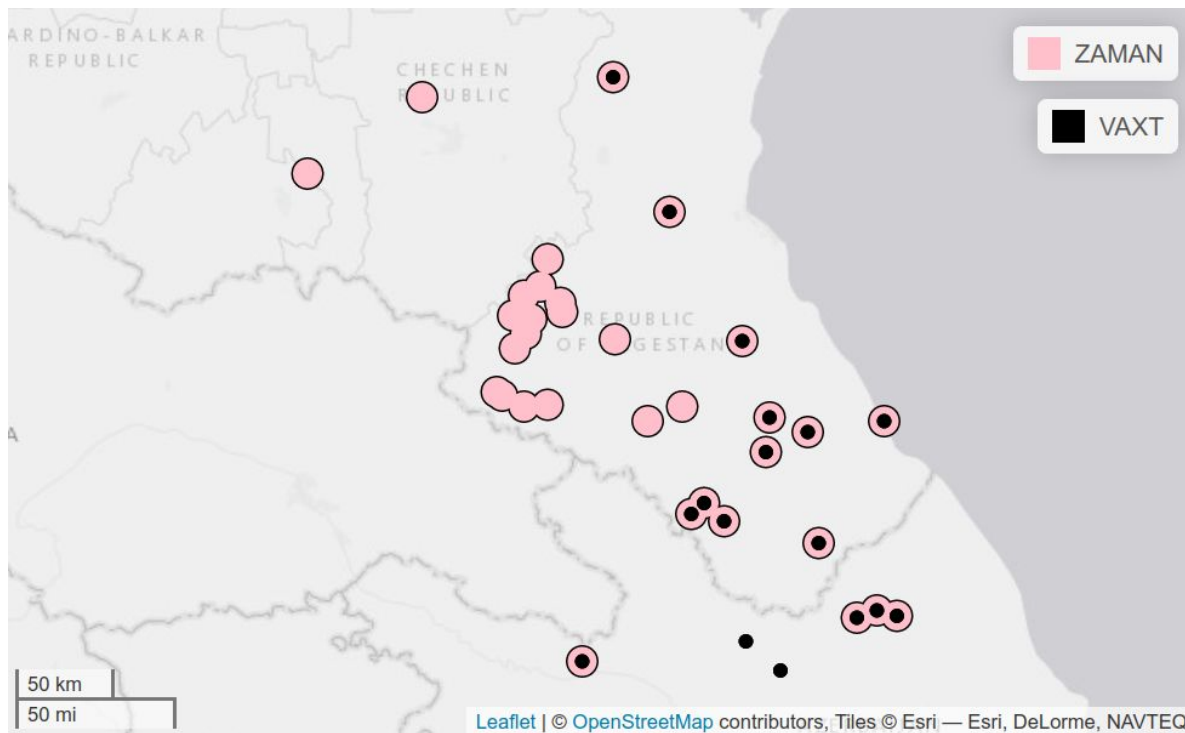
Other Arabic borrowings

Several other words which can have the meaning 'time' were borrowed from Arabic, mostly in languages of the south (Lezgic and Khinalug).

These are *vədə* 'time', *dövr* and *dövran* 'epoch, era', *əyyam* 'time(s)', and also more specific *müddət* and *möhlət* 'time period needed to do smth' (Azerbaijani words cited). They are rare and, as a rule, are not mentioned in dictionary examples outside of respective lexical entries.

In East Caucasian, the *vədə*-type word can retain its Arabic phonology, namely the pharyngeal consonants or vowel pharyngealization, cf. Agul *weʕda* / Khinalug *wäʕdä*, or Rutul *waʕhde* / Tsakhur *waʕʔda*.

According to some sources the word ساعة *saʕa* 'hour' can also mean 'time' in an appropriate context, though it is unclear how common this meaning is.

East Caucasian terms

Native East Caucasian lexemes meaning 'time' are attested in the northwestern and central areas.

Several Andic languages preserve a native lexeme *qaru*, which could be cognate with Tsova-Tush *qavl*. Perhaps it was borrowed from Tindi (Andic) into Khwarshi (Tsezic).

Nakh languages have *xan*, which was borrowed into Georgian (see slide 11).

Lak has *č'un,* Lezgian *č'aw,* Archi *č'em* and Udi *č'at*. Not clear if they are related, see also **moč* reconstructed for Caucasian Albanian (Gippert et al. 2008: ii-39, ii-77). Caucasian Albanian also attests *p'ˤi* with no known analogs in modern languages.

Several words from Avar have been borrowed into other languages.

Avar borrowings

Xalidova (2006: 110-111) identified *rih* and its variants as borrowings from Avar into the Andic languages, see also [North Caucasian etymology database](). Signs of contact rather than cognates:
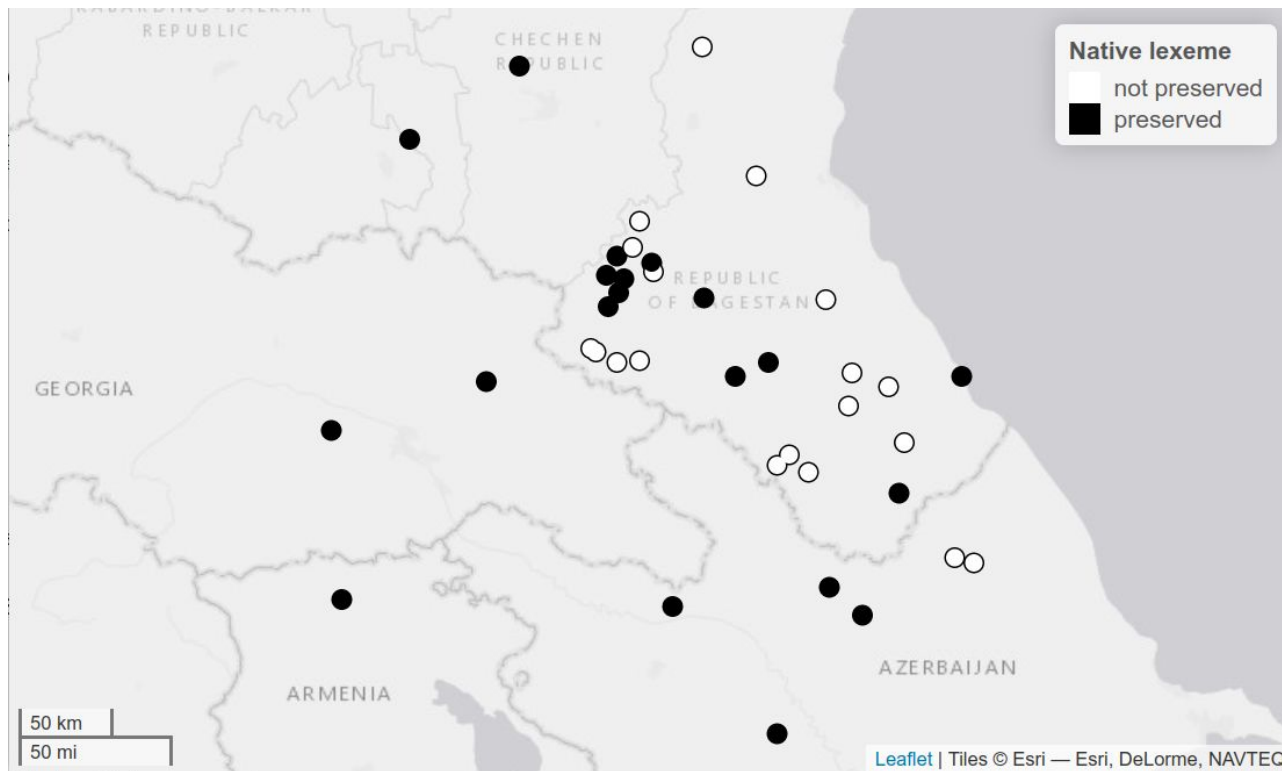
- minimal phonological change (+ a final vowel)
- also present in Archi of the Lezgic branch

However, Botlikh has *rih* 'age' and *rihu* 'time'; the language typically does not modify borrowings from Avar.

Tsezic languages have borrowed *meχ* from Avar.

ʕuž 'time, season, period' is represented in Hinuq (Tsezic) and Karata (Andic).

Preservation of native lexicon

Other borrowing sources

Iranian loans are represented by *gah* 'time' in Agul and Tabasaran, and *bere* 'time' in Lezgian, Agul, Tabasaran and Rutul. In all these languages, these nouns are highly marginal, although in Agul and Tabasaran *gah* gave rise to a temporal suffix (see below).

Turkic *čaʁ* 'time, age' is mentioned for Tsakhur and also occurs in a few fixed expressions in Lezgian. Another Turkic word *ara* is sometimes mentioned with 'time' as one of its meanings, although its main meaning is more general, 'period, distance (both temporal and spatial)'.

Other borrowing sources

Armenian *žamanak* 'time' was borrowed into the Vartashen dialect of Udi as *žamanak'*, *žam* 'hour, time' appears in Georgian, and in Tsova-Tush (probably through Georgian) as part of the compound *xaⁿ-žam*, which also means 'time'.

Borrowings from Russian are only marginally attested in the time domain, e.g. *sːizon* 'season' in Andi.

Strange correspondences

anta in Bezhta (Tsezic) and *a-amta* in Abkhaz (West Caucasian), cf. Klimov & Khalilov (2003).

Arabic *manzil* 'house, temporary stop for travelers' borrowed into Dargwa and Kubachi (*manzil* 'time, period, distance') and Kumyk and Nogai (*mezgil* 'time, period'), with a similar meaning that is rather different from the source lexeme.

arʁ 'weather, time, age' in Standard Dargwa and *arʁi* 'season' in Bezhta.

Summary

Time noun inventories show a division between the south (Lezgic and Azerbaijani) and the rest, due to larger inventories of Arabic nouns in the southern languages.

ZAMAN is attested throughout the area with a few exceptions.

VAXT is attested in the southern part of the area and few central languages.

Tsezic and Andic form distinct clusters based on which noun they borrowed from Avar.

Borrowings co-exist with native lexicon.

Several interesting patterns of convergence at micro-level (e.g. Tindi to Khwarshi).

Time nouns in temporal constructions

Temporal clauses and converbs

Nouns for 'time' are common grammaticalization sources for temporal markers (see TIME > TEMPORAL in Kuteva et al. 2019: 438), especially for subordinators introducing temporal clauses.

In East Caucasian, quite a few languages possess dedicated temporal **'when'-converbs**. In some of them, the etymology of their (suffixal) markers is

- unknown, as e.g. *-la* in Lezgian or *-tːa* in Archi.

But in some language, it is clear that the suffix of the converb is

- related to a noun 'time' (not necessarily the one which is synchronically the main lexical expression for 'time').

Temporal markers from GAH

In a number of Lezgic languages – Agul (*-guna*, *-gana*, *-guni*), Tabasaran (*-gan*) and some Rutul dialects (*-ga*), the suffixes of 'when'-converbs go back to the now obsolete noun *gah* 'time', a Persian loan. The suffixes are added to participles.

Agul, Kurag village (Magometov 1970)

zun	*kitab*	*χura-**guna**,*	*wun*	*zas*	*mišat*	*m-arq'a !*
[I.ERG	book.ABS	read.IPFV-TEMP]	you(SG).ERG	I-DAT	disturb	PROH-do.IPFV

'When I am reading a book, don't disturb me!'

In Khinalug, the suffix *-ga* is used to derive temporal adverbs.

a. *tü-**ga***
    that-TEMP
    'then, at that time'

b. *q'ätːi-**ga***
    daytime-TEMP
    'at midday'

Temporal markers from RIH

In Andic languages Andi and Botlikh, the noun *rihi* 'time' grammaticalized into a suffix of temporal converbs, having fused with the participle and partly changed its own form (in an irregular way).

Andi

a. *he-rbihi* < *he-b-**rihi***
    that-TEMP that-IV-TIME
    'then, at that time'

b. *a‹b›ža-rbihi* < *a‹b›ža-b-**rihi***
    ‹IV›begin-TEMP ‹IV›begin-PTCP-TIME
    'when s/he started'

Botlikh

waša *w-ac'a-rudi* < **w-ac'a-**rih-u-di***
son M-reach-TEMP M-reach-time-OBL-ERG
'when (our) son arrived'

Participial clauses with 'time'

The grammaticalization sources for temporal converbs based on 'time' are **relative clauses with 'time' as the head**, i.e. 'in the time [when V happened]'.

Syntactic constructions with 'time' as the head and a preposed participial clause are often mentioned as a regular means of expressing temporal relations. These can be constructions both with very common and with less common time nouns.

Standard Lezgian (Haspelmath 1993)

Zi	*buba.di*	*ajal-ar*	*kaš.a-kdi*		*telef*		*xu-n.i-kaj*
[I:GEN	father(ERG)	child-PL	hunger-SBDIR		perishing		be-MSD-SBEL
ixtilat		*awu-r*	**č'awuz**	*za-z*	*šel*	*q:we-da.*	
conversation		do-AOP]	when	I-DAT	crying	come-FUT	

'When my father tells me how children are starving I have to cry.'

(AOP = aorist participle, *č'awuz* = dative case of *č'aw* 'time')

Temporal conjunctions based on 'time'

Both Caucasian Albanian and modern Udi seem to prefer **finite temporal clauses** introduced by means of **temporal conjunctions** based on 'time'-nouns

CA: *hamočke* 'when' (conjunction) < *hamoč* 'when' (pronoun) < **moč* 'time', subordinator *ķe*

> *ičēebowr cex-al he-boķa-hamočke-oen* …
> *ič=ē-e=bowr   cex-al   he-boķa-hamočke-o=en*
> REFL=GEN-REF=PL:ABS:N all-FOC   hither-lead:PRES-when-he=ERG
> 'When he leads all his own ones out …'

Udi: *hevaχtː (ki), evaχtːe* 'when' (conjunctions based on *vaχtː* 'time', with subordinators *ki* or *te*)

Cf. a similar strategy in modern Azerbaijani: *nə vaxt (ki), nə zaman (ki), havaxt* 'when' (conjunctions based on *vaxt* or *zaman* 'time', optional subordinator *ki*)

Discussion and research perspectives

Time nouns play a role in grammar, becoming sources of temporal suffixes and conjunctions, as well as heading a common type of temporal relative clauses.

Certain areal patterns can be revealed in the use of 'time' in temporal constructions. Thus, the use of 'time' as the head of relative clauses seems to be common to all East Caucasian languages, but only in some of them 'time' grammaticalized in to a suffix:

- *gah* only in a few Lezgic languages
- *rihi* only in a few Andic languages

Only in CA/Udi and Azerbaijani we see finite temporal clauses with temporal conjunctions based on the interrogative 'when' (based on 'time'), which is probably the Iranian influence.

Also, the grammaticalized use of 'time' can show which 'time'-words used to be more common in the past (cf. *gah* in Agul and Tabasaran, *č'aw* in Lezgian).

Probably, more areal and genealogical patterns can be revealed in the future.

Thank you for your attention!

References

Gippert, Jost, Wolfgang Schulze, Zaza Aleksidze & Jean-Pierre Mahé. 2008. *The Caucasian Albanian palimpsests of Mt. Sinai* (2 vols.). Turnhout: Brépols.

Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin: Mouton de Gruyter.

Khalidova, R. S. 2006. *Avarsko-andijskie jazykovye kontakty*. Makhachkala: Dagestanskij naučnyj centr RAN.

Klimov, G. A., & Khalilov, M. S. 2003. *Slovar' kavkazskix jazykov. Sopostavlenie osnovnoj leksiki*. Moscow: Vostočnaja literatura.

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World lexicon of grammaticalization*, 2nd edn. Cambridge: Cambridge University Press.

Magometov, A. A. 1970. *Agul'skij jazyk. Issledovanie i teksty*. Tbilisi: Mecniereba.

Zabitov, S. 2001. *Arabizmy v leksike vostonokavkazskix jazykov*. Moscow: Institut jazykoznanija RAN.