Phonetic and Lexical Contact Phenomena in the Basic Vocabulary of Turkic Languages in the North Caucasus

Zukhra Aibazova
Institute of Linguistics RAS (Russia)

Zarema Alzhanbekova
Institute of Linguistics RAS (Russia)

According to M. Swadesh's concept, basic vocabulary is the most resistant to borrowing, and its modifications indicate close linguistic contact. This study focuses on contact phenomena in the basic vocabulary of Kumyk, Karachay-Balkar and Nogai. The research is based on fieldwork conducted between 2018 and 2024.

1. Lexical Borrowings

Borrowings in the basic vocabulary from North Caucasian languages fall into two tjpes:
- complete displacement of native words
- frequent synonyms coexisting with native lexemes.

The distribution of borrowings varies across languages and dialects. In the Ak-Nogai dialect of Nogai, borrowings from Caucasian languages were not detected. Kumyk and Karachay-Balkar demonstrate both interlingual differences and internal dialectal variation:

grass: Kum. *ot* (Turkic) vs. KBal. *kɨrdɨk* (Ossetian)
sand: KBal., Nog. *qum* (Turkic), Kum. *qajɨr*, *xum* (Turkic) vs. Balk. dialect *üzmez* (Ossetian) vs. Kum. dialect *čabar/čavar* (Avar)
flower: Kum., Balk. dialect *gül* (Persian), Kum. dialect *čeček* (Turkic), Nog. *šešekej* (Turkic) vs. KBal. *gokka* (Circassian)
egg: Nog., Kum. *jɨmɨrtqa/jumutrqa*, KBal. *ǯumurtxa* (Turkic) vs. KBal. dialect *gakkɨ* (Circassian) and others.

2. Phonetic Contact Phenomena

In the Cherek dialect of the Karachay-Balkar language, labialized velar consonants /k^{w}/ and /g^{w}/ have been documented. Previously, these were automatically attributed to combinatory processes related to the labialization of broad vowels in non-first syllables following labial syllables (cf. *tögerek* → *tögörek*/*tögörök*). Such processes are typical for some Turkic languages, including the Aknogai dialect of Nogai. However, in the Cherek dialect, labialization of vowels combines with labialized velars, which are atypical for Turkic but characteristic of Circassian languages:

• Cherek: *k^{w}ol* 'lake' (standard *köl*)
• Cherek: *kog^{w}ot* 'fruit' (standard *köget*)
• Cherek: *öltürg^{w}endi* 'he killed' (standard *öltürgendi*)
• Chegem: *örg^{w}e* 'upward' (isolated case)
• Chegem: *tögörek kögöt* 'round fruit' (standard *tögerek köget*)
• Nogai: *ölgon mu?* 'Is he dead?', *tögörök* 'round' and others.

Conclusions:

The absence of Caucasian borrowings in the basic vocabulary of the Ak-Nogai dialect indicates a superficial interaction with languages of this region.

Despite close genetic relatedness within the Western Kipchak subgroup, Kumyk and Karachay-Balkar demonstrate different patterns of lexical borrowing, indicating that these changes occurred after the languages separated.

The dialects of the Karachay-Balkar language show a higher frequency of borrowings from Caucasian languages (and Ossetian), indicating more intensive contacts. The labialized velars in the Cherek dialect can be considered a phonetic marker further pointing to this intensity.

The Multiple Partitive Construction in Ossetic and Beyond

Danil Alekseev
Lomonosov MSU, IL RAS

Introduction.[1] The term "multiple partitive construction" (MPC) was first used by Haspelmath (1997) in reference to a special use of the bare interrogative, illustrated by the Russian example below:

(1) *...da v tajg-e i pere-merz-l-i:* ***kto*** *s golod-u,* ***kto*** *s natug-i...*
CONJ in taiga-LOC.SG and PV-freeze-PST-PL who from hunger-PART who from strain-GEN
'{They have all} frozen over in the taiga: someone because of the hunger, others out of strain...' (RNC)

The bare interrogatives introduce multiple parallel coordinate clauses which denote "a subset ... of a larger, contextually given set" (Haspelmath 1997, p. 177).

After Haspelmath, the MPC has received little treatment in the following literature (with the notable exception of Huggard 2015 discussed below). In this paper I explore the properties of the MPC in Ossetic, an Iranian language spoken in the Caucasus, and show how its interaction with other elements of the clause gives rise to a violation of the general principles of syntactic organization in Ossetic.

The MPC in Ossetic. The existence of the MPC in Ossetic has first been noted by Byzova (2022) in an unpublished manuscript:

(2) ***či*** *ag štaw-ə* ***či*** *art ænzar-ə* ***či*** *šug šætt-ə*
who cauldron hang-PRS.3SG who fire start-PRS.3SG who firewood chop-PRS.3SG
'Some hang the cauldron, some start the fire, others chop firewood.' (ONC, Byzova 2022, p. 21)

For Byzova the MPC is of interest since it is one of the few contexts where the bare interrogative is licenced as an indefinite. In most other cases the interrogative root can only function as an indefinite when accompanied by additional markers, such as *-dɜr* (e.g. *čidɜr* 'someone'), stemming from the additive particle, or *iš-* (*išči* 'anyone').

The most striking syntactic feature of the MPC in Ossetic is that interrogatives must be located in clause-initial position (3a), while the canonical interrogative position is preverbal (Lyutikova & Tatevosov 2009; Erschler 2012; Belyaev 2023). Thus attempts to situate the wh-pronoun preverbally lead to an interrogative interpretation (3b).

(3) a. *a-sə qæw-ə* ***či*** *iron u* ***či*** *dəguron* ***či****=ta* *k'ʷədajrag*
this-DET village-IN who Iron be.PRS.3SG who Digor who=CONTR Kudar
'In this village some are Iron, some are Digor, others are Kudar.'

b. *a-sə qæw-ə iron* ***či*** *u* ***či*** *dəguron* ***či****=ta* *k'ʷədajrag*
this-DET village-IN Iron who be.PRS.3SG who Digor who=CONTR Kudar
'*In this village some are Iron, some are Digor, others are Kudar. / OKIn this village, who is Iron, who is Digor, and who is Kudar?'

The bare interrogative used in one of the conjuncts of the MPC can serve as an antecedent to other referring expressions, e.g. a second-position clitic in the next conjunct (4). The anaphor must be singular, even though the speakers in their judgments indicate that in the example described more than one person might be building the house. A plural clitic in (4) is nonetheless infelicitous.

(4) *qæw-ə* ***či*** *hæzar araž-ə* *či****=jən/*šən*** *æxxʷəš kæn-ə či ulæf-ə*
village-IN who house build-PRS.3SG who=3SG.DAT/3PL.DAT help do-PRS.3SG who rest-PRS.3SG
'In the village some are building a house, others are helping them, others are resting.'

Conversely, plural interrogatives, which do exist in Ossetic, are also disallowed in an MPC environment:

(5) **a-sə qæw-ə* ***či-tæ*** *iræ-ttæ u* ***či-tæ*** *dəguræ-ttæ* ***či-tæ****=ta* *k'ʷədajræg-tæ*
this-DET village-IN who-PL Iron-PL be.PRS.3SG who-PL Digor-PL who-PL=CONTR Kudar-PL
'In this village some are Iron, some are Digor, others are Kudar.' (=3a)

Any material that follows the initial interrogative cannot take scope over the whole construction, e.g. the locative noun *qæwə* 'in the village' in (6) only modifies the first clause involved in the MPC.

[1] In interlinear glossing I follow the Leipzig Glossing Rules. I abbreviate the sources used as follows: RNC — Russian National Corpora, ONC — Ossetic National Corpora. Examples without an indicated source were collected during fieldwork in Vladikavkaz in October 2024. For each examples at least three speakers were consulted. I am grateful to all the native speakers of Ossetic who shared their knowledge of the language with me.

(6) *či* ***qæw-ə*** *xæzar araž-ə* *či=jən* *æxx^{w}əš kæn-ə* *či* *ulæf-ə*
who village-IN house build-PRS.3SG who=3SG.DAT help do-PRS.3SG who rest-PRS.3SG
'*In the village some are building a house, others are helping them, others are resting. / ^{OK}Some are building a house in the village, others are helping them, others are resting.' (speaker commentary: "These people might as well be urban dwellers.")

Lastly, the MPC is also embeddable under the so-called "floating subordinators" which are able to occupy a high C position or the preverbal position. The subordinator in this case should precede the whole construction[2]:

(7) a. *aftæ qær kæn-əns* ***səma*** *či baraban-t-æj či gitaræ-jæ či xætæl-æj sæğd-ə*
so noise do-PRS.3PL as_if who drum-PL-ABL who guitar-ABL who trumpet-ABL play-PRS.3SG

b. **aftæ qær kæn-əns či* ***səma*** *baraban-t-æj či gitaræ-jæ či xætæl-æj sæğd-ə*
so noise do-PRS.3PL who as_if drum-PL-ABL who guitar-ABL who trumpet-ABL play-PRS.3SG
'It is so loud as if someone is playing the drums, someone else — the guitar, and someone else the trumpet.' {a=b}

Discussion. The left peripheral position of the interrogatives in an MPC context should come across as a rather unexpected property. Ossetic is a discourse-configurational language with a well-structured preverbal area, which includes narrow foci, interrogatives, and negative pronouns / sentential negation (Lyutikova & Tatevosov 2009; Erschler 2012; Belyaev 2023). Among these elements narrow foci (Belyaev 2023, p. 47) and negative pronouns[3] can also occupy the post-verbal position. Interrogatives normally enjoy no other position, and it is only in the case of the MPC that they are clause-initial, and a substantial amount of material can appear between the predicate and the interrogative.

A close parallel can be found in the ancient Indo-European Anatolian language Hittite, where a similar configurational system is present, see *i.a.* (Lyutikova & Sideltsev 2021; Sideltsev 2014). In particular, indefinite pronouns, including bare ones, usually appear in the immediately preverbal position (Sideltsev 2014, p. 88). However, the MPC is also present in Hittite, with the interrogatives taking the clause-initial position, argued by Huggard (2015) to be derived through movement of the interrogatives to ContrastiveFocusP.

While an explanation of this sort might be deficient for Hittite (Andrei Sideltsev, p.c.), the contransitive nature of the interrogatives involved in the formation of the MPC (although likely a topical one, rather than a focal) is confirmed in Ossetic by examples such as (3), where the introduction of the contrastive clitic *=ta* increases the acceptability of the example for certain speakers.

In the Caucasus the MPC is also present in Georgian (< Kartvelian), as indicated by Haspelmath (1997, p. 177), see (8). Like Ossetic, Georgian otherwise utilizes preverbal clause-internal interrogatives (Harris 1981, p. 16).

(8) ***vin*** *pul-s eloda* ***vin*** *c'eril-s* ***vin*** *gazet-eb-s*
who money-DAT wait-PRS.3SG who letter-DAT who newspaper-PL-DAT
'Some are waiting for money, some for a letter, some for newspapers.'

Conclusion. In this paper I investigated the properties of the multiple partitive construction in the Iranian language Ossetic, showing its distinct violation of the principles of clausal organization in the language. A similar picture can be found in other languages that have a specially dedicated preverbal area, like Hittite and Georgian. Considering that these languages are diverse from genetic and temporal perspectives but show close parallels to one another, further studies are required to determine the source of this commonality.

Belyaev, Oleg (2023). "Phrase structure and configurationality in Ossetic." In: Proceedings of the LFG'22 Conference. **Byzova, Anna A.** (2022). The system of indefinite pronouns in Ossetic (a corpus study). Ms. **Erschler, David** (2012). "From preverbal focus to preverbal "left periphery": The Ossetic clause architecture in areal and diachronic perspective." In: Lingua 122.6, pp. 673–699. **Harris, Alice C.** (1981). Georgian Syntax. Cambidge: Cambridge University Press. **Haspelmath, Martin** (1997). Indefinite Pronouns. Oxford: Clarendon Press. **Huggard, Mattyas G. C.** (2015). "Wh-words in Hittite: A Study in Syntax-Semantics and Syntax-Phonology Interfaces." PhD thesis. UCLA. **Lyutikova, Ekaterina & Sideltsev, Andrei** (2021). "Deriving preverbal position in a verb-final language: the case of Hittite." In: Glossa 6.1. **Lyutikova, Ekaterina & Tatevosov, Sergei** (2009). "The clause internal left edge: Exploring the preverbal position in Ossetian." In: International Conference on Iranian Linguistics. Vol. 3, pp. 11–13. **Sideltsev, Andrei** (2014). "Verb Adjacent Focus in Hittite." In: Vestnik MGGU 3, pp. 86–97.

[2]Repeating the subordinator after **each interrogative** does not improve acceptability judgments.

[3]Post-verbal negative pronouns licence double negation, normally illicit in Ossetic (Oleg Belyaev, p.c.). This seems to confirm the analysis of Lyutikova & Tatevosov (2009) that treats the complementary distribution of preverbal negative pronouns and clausal negation as an instance of the "multiply filled Neg filter".

Expression of spatial and non-spatial meanings in Rutul dialects

Anastasia Alekseeva, Ksenia Dunaeva
HSE University (Moscow, Russia)

One of the salient features of East Caucasian languages is the presence of a great amount of spatial cases, which express various meanings: ‘inside’, ‘on top of’, ‘moving out of’, ‘moving from the top of’, etc. [Comrie & Polinsky 1998]. Spatial cases are remarkable not only for the diversity of spatial meanings they encode but also for their property to express non-spatial meanings, usually specific arguments, e. g. unintentional agent, temporary recipient, etc. [Bokarëv 1948; Testelec 2019]. For example, in Rutul (< Lezgic < East Caucasian) Apud Elative case expresses a spatial meaning ‘moving from X’ (1a), however it can also encode unintentional agent, as in (1b) [Maxmudova 2001: 68].

(1a) *mɨrad laχ-a-da: lu-q-uzu-r=i*
Murad hearth-OBL-APUD.EL PV-RE-1.stand.PFV-CVB=COP
‘Murad moved away from the hearth.’ [Maxmudova 2001: 68]

(1b) *za-da: ɢab jaχ-r=i*
I.OBL-APUD.EL dish 1.break.PFV-CVB=COP
‘I broke a plate (accidentally).’ [Maxmudova 2001: 68]

In contrast to other East Caucasian languages that form spatial cases by combining morphemes positioned in two separate morphological slots, one for localization and one for orientation, e.g. Agul [Alekseev & Sulejmanov 2001], Rutul spatial system is mostly fusional and can be described as bipartite with respect to the number of values of orientation category it distinguishes between (essive/lative vs elative). Rutul dialects have quite divergent systems of spatial cases as well. For instance, in Mukhad dialect (Rutul variety) different case affixes are used to express the meanings ‘under’ and ‘behind’ [Maxmudova 2001: 65], whereas in Ikhrek these meanings are not morphologically differentiated [Ibragimov 1978: 204-205] (Table 1). Accordingly, there is also some variation in the encoding of non-spatial meanings, similarly to (1b), across Rutul dialects.

Table 1. Forms used to encode meanings ‘behind’ and ‘under’ in Mukhad and Ikhrek dialects.

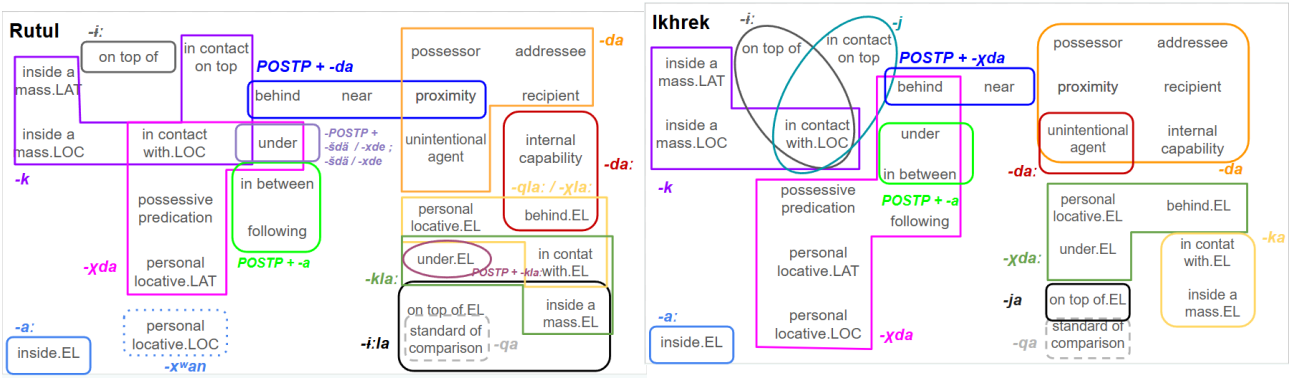
		essive ‘location at’ / lative ‘moving towards’	**elative ‘moving from’**
Mukhad dialect	**‘behind’**	*-χda*	*-χla:*
	‘under’	*-xde*	*-kla:*
Ikhrek dialect	**‘behind’ / ‘under’**	*-χda*	*-χda:*

In this research we examine the expression of some spatial and non-spatial meanings in twelve Rutul idioms: Mukhad dialect (Rutul, Kufa, Kiche varieties), Ikhrek dialect (Ikhrek variety), Myukhrek (Myukhrek and Dzhilikhur), Shinaz dialect (Shinaz varietry), Borch-Khnov (Khnov variety) and “mixed” dialects [Ibragimov 1978: 13-14] (Kina, Luchek, Amsar, Kala). Non-spatial meanings include possessive predication, temporary recipient, addressee, capability subject, to cite a few. The data was elicited using a questionnaire that included contexts for all the meanings [Alekseeva et al. 2025]. For different Rutul idioms we created semantic maps (following the approach of Haspelmath [2003]) which includes all spatial and non-spatial meanings examined. A similar research was conducted by D. Forker [2010] for Tsezic languages, although the diversity of encoding of spatial meanings was not accounted for in her study. Our study also differs in that we consider the forms (i. e. the affixes, like *-χda* in Table 1) by which the meanings are expressed, whereas D. Forker analyzed which paradigmatic labels (e. g. APUD, IN.EL, etc.) are used for which

non-spatial meanings. Additionally, in our study we considered not only affixes, but postpositions marked by spatial cases as well.

The obtained results show that non-spatial meanings tend to be encoded more similarly than purely spatial ones. Consider the following semantic maps (Figure 1) for Rutul and Ikhrek varieties: while there is a variation in the domain of affixes used to express meanings like ‘under’, ‘in contact with’, ‘near’, ‘behind’, etc., the marking of temporary possessor, addressee, temporary recipient by *-da* (Apud Essive), on the one hand, and possessive predication and lative personal locative by *-χda* (Post Essive) on the other remains the same. It suggests that we can reconstruct **-da* as the case marking for temporary recipient, addressee and temporary possessor as early as in Proto-Rutul. While a horizontal transfer of *-da* is theoretically possible, it is highly unlikely, because *-da* is found in these contexts in other Rutul varieties as well. More importantly, it is also found in Khnov, which is the most divergent and geographically isolated variety.

Figure 1. Semantic maps for spatial and non-spatial meanings in Rutul and Ikhrek varieties.



In the talk we will try to explain the asymmetry between the more homogeneous expression of non-spatial meanings and the diverse expression of spatial ones. We will discuss the systems of other Rutul dialects, describe grammaticalization paths for markers with both spatial and non-spatial functions and determine the most and least diachronically stable ones. Additionally, we will consider the cases of asymmetry in coding of the locative and elative values of same localization, cf. ‘under’ in Mukhad (Rutul idiom).

References

Alekseeva, A., N. Beklemishev, M. Daniel, N. Dobrushina, K. Filatov, A. Ivanova, T. Maisak, et al. 2025. *Atlas of Rutul dialects*. Moscow: Linguistic Convergence Laboratory, HSE University.

M. E. Aleskeev, N. D. Sulejmanov. 2001. Agul′skij jazyk [The Agul language]. In *Jazyki mira: Kavkazskie jazyki*, 398–408. Moscow: Academia.

E. A. Bokarëv. 1948. Lokativnye i nelokativnye značenija mestnyx padežej v dagestanskix jazykax [Locative and non-locative meanings of locative cases in Daghestan languages] // Jazyk i myšlenie. XI. Moscow & Leningrad.

Ja. G. Testelec. 2019. Imennye lokativnye formy v dagestanskix jazykax [Nominal locative forms in Daghestan languages]. Manuscript.

B. Comrie, M. Polinsky. 1998. The great Daghestanian case hoax. In Anna Siewierska & Jae Jung Song (eds.), Case, typology and grammar, 95–114. Amsterdam & Philadelphia: John Benjamins.

D. Forker. 2010. Nonlocal uses of local cases in the Tsezic languages. *Linguistics*, 48: 1083–1109.

M. Haspelmath. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In Tomasello, M. (ed.), *The new psychology of language: Cognitive and functional approaches to language structure,* 2: 211–242. Erlbaum.

G. X. Ibragimov. 1978. *Rutul'skij jazyk [The Rutul language]*. Moscow: Nauka.

S. M. Makhmudova. 2001. *Morfologija rutul'skogo jazyka [Morphology of the Rutul language]*. Moscow: Sovetskij pisatel′.

Antonov Anton, Sulaibanov Gasangusen

The quirky syntax of Tsugni ideophones

This talk discusses a typologically very unusual pattern of syntactic integration of ideophones in Tsugni (Dargwa, Nakh-Dagestanian).

Ideophones–also known as "mimetics" in research on Japanese or "expressives" in works dealing with South Asian languages (Lahaussois, Marsault & Treis 2024)–are "marked words that evoke vivid sensory scenes in imitative fashion" (Akita & Dingemanse 2019). Although found (unequally distributed) across the world's languages, they are particularly abundant and well-described in languages of Asia (especially in Japanese and Korean), Africa, and the Americas. Ideophones in Caucasian languages have hardly been the object of any in-depth analyses, and there seem to be only a couple of very recent studies on the subject–a broad overview of the topic (Wier 2023) and a language-specific one (Authier 2024).

It is important to bear in mind that ideophones are not limited to onomatopoeia (such as *cock-a-doodle-doo* or *splash*), but cover a wide range of sensory domains, such as manner of motion (e.g., *kpɔtɔrɔ-kpɔtɔrɔ* 'walking like a tortoise' in Siwu), texture (e.g., *potul potul* 'very soft' in Korean), and psychological states (e.g., *jirijiri* 'fretful' in Japanese) (Akita & Dingemanse 2019).

To the best of our knowledge, crosslinguistically, whenever ideophones are not used on their own as independent utterances or as sentence-initial (or final) elements, their syntactic properties can be described along the lines of Setswana ideophones which Creissels (2001: p. 83) characterizes as "uninflected predicative lexemes that normally occur as the lexical part of compound predicates in which *re* 'say' fulfills the auxiliary function." The auxiliary, or rather light, verb in these compounds is crosslinguistically very often either 'say' or 'do', and one of the two constructions Tsugni ideophones take part in conforms to this universal tendency (see examples 1 and 2), even if the range of light verbs here is wider and also includes 'hit', 'put', 'pull', 'move', 'become' (examples will be given in the talk).

(1) *durħa^{ʕ} qeh w-ik'-ule caj*
boy IDEO M-say.IPFV-PROG.CNV COP:M(3)
'The boy is coughing.'

(2) *durħa^{ʕ} duc' w-ik'-ule caj*
boy IDEO M-say.IPFV-PROG.CNV COP:M(3)
'The boy is running.'

However, Tsugni ideophones also take part in a second construction in which the ideophone appears in the instrumental (=ergative) case. Interestingly enough, the only light verb allowed in this construction is the verb 'carry' (see examples 3 and 4).

(3) *durħa^{ʕ} qeh-li w-iq:-ule caj*
boy IDEO-INSTR M-carry.IPFV-PROG.CNV COP:M(3)
'The boy is coughing.'

(4) *durħa^{ʕ} duc'-li w-iq:-ule caj*
boy IDEO-INSTR M-carry.IPFV-PROG.CNV COP:M(3)
'The boy is running.'

This appears to be formally identical to the antipassive construction of the verb 'carry' (compare examples 5 and 6)

(5) *durħa^{ʕ}-li ʔa^{ʕ}nč'i d-iq:-ule caj/cadi*
boy-ERG clay[NPL] NPL-carry.IPFV-PROG.CNV COP:M(3)/COP:NPL
'The boy is carrying the clay.'

(6) *durħa^{ʕ} ʔa^{ʕ}nčʼi-li w-iq:-ule caj*
boy clay[NPL]-INSTR M-carry.IPFV-PROG.CNV COP:M(3)
'The boy carries clay.' (usually, as part of his job)

In this talk we will present and analyze the syntactic integration of Tsugni ideophones with particular attention to the 'antipassive'-like construction exemplified in examples 3 and 4. As far as we know, this construction has not been reported in other Dargwa languages (Sumbatova & Mutalov 2003, Abdullaev et al. 2014, Sumbatova & Lander 2014, Forker 2019, Daniel, Dobrushina & Ganenkov 2019), nor has there been any mention in the specialized literature of 'inflecting' ideophones (Voeltz & Kilian-Hatz 2001, Akita & Dingemanse 2019, Lahaussois, Marsault & Treis 2024), which means that we are dealing with a typological *rarum*.

Based both on our extensive corpus and some extra elicitation work, we will try to show how it is different from the construction with an uninflected ideophone and a light verb as well as from the antipassive construction itself and suggest a diachronic scenario for its emergence.

We believe our work will contribute not only to the linguistic description of ideophones in Dargwa and Caucasian languages in particular, but also to the typology of the syntactic integration of ideophones in general.

References

Abdullaev, Z. G. et al. 2014. *Современный даргинский язык*. Makhachkala: IYaLI DNC RAN.

Akita, Kimi & Mark Dingemanse. 2019. Ideophones (Mimetics, Expressives). In Mark Aronoff (ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.477.

Authier, Gilles. 2024. Ideophonic verb compounds in Archi. *Linguistic Typology at the Crossroads* 4(1). 325–380.

Creissels, Denis. 2001. Setswana ideophones as uninflected predicative lexemes. In F. K. Erhard Voeltz & Christa Kilian-Hatz (eds.), *Ideophones*, 75–87. John Benjamins.

Daniel, Michael, Nina Dobrushina & Dmitry Ganenkov (eds.). 2019. *The Mehweb language: Essays on phonology, morphology and syntax* (Languages of the Caucasus 1). Berlin: Language Science Press. https://doi.org/10.5281/zenodo.3374730.

Forker, Diana. 2019. *A grammar of Sanzhi Dargwa* (Languages of the Caucasus 2). Berlin: Language Science Press. https://doi.org/10.5281/zenodo.3339225.

Lahaussois, Aimée, Julie Marsault & Yvonne Treis. 2024. Ideophones: honing in on a descriptive and typological concept. *Linguistic Typology at the Crossroads* 4(1). 1–13.

Sumbatova, Nina & Yuri Lander. 2014. *Даргинский говор селения Танты: грамматический очерк, вопросы синтаксиса: монография*. Moscow: Языки славянской культуры.

Sumbatova, Nina & Rasul Mutalov. 2003. *A grammar of Icari Dargwa*. München: LINCOM.

Voeltz, F. K. Erhard & Christa Kilian-Hatz. 2001. *Ideophones*. John Benjamins.

Wier, Thomas R. 2023. Expressive constructions in Georgian and other Caucasian languages. In *Expressivity in European Languages*, 337–358. Cambridge University Press. https://doi.org/10.1017/9781108989084.018.

MULTIPLE EXPONENCE IN THE NORTHWEST CAUCASIAN LANGUAGES

Peter Arkadiev, University of Potsdam

The Northwest Caucasian (NWC) languages are characterised by highly complex morphology, especially in the verbal domain, and can be considered polysynthetic (Lander & Testelets 2017, Arkadiev & Lander 2021, Arkadiev 2023). Their morphology presents a mixture of templatic and layered ordering (Korotkova & Lander 2010, Lander 2016) and generally follows the "one meaning — one form principle". However, the languages of the family also show non-phonological allomorphy and a number of instances of **multiple exponence** (ME), where the same semantic content is expressed by more then one formal element within the word. The latter have not been discussed in the literature so far, and notably are absent from the comprehensive survey in Harris (2017), who only mentions one case from West Circassian and one from Abkhaz. This paper aims at filling this gap by offering a description of a number of instances of ME in various NWC languages from both nominal and verbal domains as well as a discussion of their broader implications. The data come from my own fieldwork, available corpora and published materials.

NWC languages attest cases of at least three out of the four types of ME distinguished by Harris (2017). An example of **periodic** ME (Harris 2017: 55–56) is found in Abkhaz, where the reflexive prefix *čə-* must be preceded by the possessive personal prefixes doubling the features of the ergative agent: ***l-čə-l-š'-wa-jṭ*** **3SG.F.PR**-RFL-**3SG.F.ERG**-kill-IPF-DCL 'she is killing herself' (Chirikba 2003: 38). In the nominal domain, a similar pattern is found in coordinative compounds, e.g. West Circassian ***s-jane–s-jate-xe-r*** **1SG.PR**-mother-**1SG.PR**-father-PL-ABS 'my parents' (cf. Lander 2016: 3513); comparable structures are also attested in Kabardian, Abkhaz and Ubykh.

Most instances of ME in NWC fall into the **reinforcement** type (Harris 2017: 61–64). Thus, in Abaza declarative and polar interrogative verbal forms, negation is expressed twice: by the default negative marker *-mə-*, which occurs as a suffix or as a prefix depending on the tense-aspect form of the verb (Lomtatidze et al. 1989: 111–112), and by the originally emphatic prefix *g'ə-* whose cognates are used as additive markers (Pazov 2019), *jə-**g'**-sə-**m**-dár-ṭ* 3SG.N.ABS-**NEG**-1SG.ERG-**NEG**-know-DCL 'I did not know it'. This pattern is clearly a result of the well-known Jespersen cycle often leading to ME of negation cross-linguistically (Dryer 2013), which in Abaza, however, did not involve all verbal forms, cf. the non-finite *jə-**m**-dár-wa-ta* 3SG.M.ERG-**NEG**-know-IPF-ADV 'he not knowing'. A more peculiar instance of reinforcement ME in Abaza involves the Repetitive (Panova 2019), which is formed by the combination of the prefix *ata-* and the suffix *-χə*: *j-**atá**-qal-**χə**-n* 3PL.ABS-**RE**-go.up-**RE**-PST 'they went up again'. Notably, while the prefix *ata-* normally cannot be used on its own, the suffix *-χə* frequently occurs alone in a broad range of uses belonging to the so-called refactive domain (Stoynova 2013), e.g. *d-ʕa-r-g-**χ**-d* 3SG.H.ABS-CSL-3PL.ERG-lead-**RE**-DCL 'they drove him back'. As argued in Panova (2019), this involves a co-occurrence of a "light" and a "heavy" refactive markers (Wälchli 2006), and a rare situation when the latter is an affix rather than a word.

An interesting case of apparently reinforcement ME is found in the domain of nominal plurality of Abaza and Abkhaz, which have two productive plural suffixes: the general plural *-kʷa* and the human plural *-ĉ(ʷ)a*. With human nouns, the two suffixes often co-occur, cf. Abaza *j-ájš'-**ĉa-kʷa*** 3SG.M-brother-**HPL**-**PL** 'his brothers'. Interestingly, the two suffixes differ in their morphosyntactic scope: whereas the human plural attaches to stems of human nouns, the general plural is a kind of edge inflection that occurs to the right of postnominal adjectival modifiers: [*taba dəw*]-***kʷa*** pan big-**PL** 'big pans'. When the head noun is human, it takes the human plural suffix, and the general plural suffix still occurs at the right edge of the phrase: [*a-bəzŝadərəʕʷ-**ĉa** dəw*]-***kʷa*** DEF-linguist-**HPL** big-**PL** 'the great linguists'. Thus, the ME of plural results from the co-occurrence of stem-based and phrase-based inflectional markers.

A very different instance if ME of plural is found in West Circassian, where in common nouns and demonstratives the plural suffix *-xe* can be followed both by the regular oblique case suffix *-m* and by the special oblique plural suffix *-me*, which can also occur on its own, see Table 1 (cf. a somewhat different treatment of this phenomenon in Harris 2017: 237–238). This case of ME is interesting because it is in apparently free variation with the separative and the cumulative types of exponence.

Table 1. West Circassian case/number paradigm ('girl')

	Singular	Plural
Absolutive	*pŝaŝe-r*	*pŝaŝe-xe-r*
Oblique	*pŝaŝe-m*	*pŝaŝe-xe-m ~ pŝaŝe-**xe-me** ~ pŝaŝe-me*

Ubykh presents a rather unique case of pervasive ME of plurality in the verb, which technically falls under the label of **accidental** ME according to Harris (2017: 64), but is in fact highly systematic (Dumézil &

Esenç 1975: 161–162; Smeets 1997; Fenwick 2011: 135–136). The plurality of the absolutive argument of the verb is expressed by means of the cross-referencing prefixes, suffixes *-a* and *-ne* whose choice is in turn dependent on tense, of the retrospective tense suffixes *-jt̩* SG ~ *-jλ(e)* PL, of the causative prefixes *də-* SG ~ *ʁe-* PL, and root suppletion with a number of verbs. All these means of expression can combine yielding verbal forms with exuberant ME of absolutive plural, e.g. ***a-z-ʁe-dex-á-n*** 3**PL**.ABS-1SG.ERG-CAUS.**PL**-stand.**PL**-**PL**-PRS 'I make them stand up' (Vogt 1963: 112).

Patterns of ME in NWC languages show considerable variation even between closely related members of the same branch of the family (West Circassian vs. Kabardian, Abaza vs. Abkhaz), with Ubykh showing some clearly unique patterns. They involve different morphological features, both inflectional and derivational, and are often optional and/or restricted to particular contexts or subparadigms. While some of them neatly fit into the classes of ME proposed by Harris (2017), others appear to complement and extend this typology in potentially interesting ways.

References

Arkadiev, Peter. 2023. Polysynthesis: Lessons from Northwest Caucasian languages. In: Jenny Audring et al. (eds.), *Comparing Typologies. Online Proceedings of the 13th Mediterranean Morphology Meeting (MMM13), Rhodes (Greece), May 19–22, 2022*. https://pasithee.library.upatras.gr/mmm/article/view/4402

Arkadiev, Peter & Yury Lander. 2021. The Northwest Caucasian languages. In Maria Polinsky (ed.), *The Oxford Handbook of the Languages of the Caucasus*, 369–446. Oxford: Oxford University Press.

Chirikba, Vyacheslav A. 2003. *Abkhaz*. München: LINCOM Europa.

Dryer, Matthew S. 2013. Negative morphemes. In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/112, accessed on 2021-09-04.)

Dumézil, Georges & Tevfik Esenç. 1975. *Le verbe oubykh. Études descriptives et comparatives*. Paris: Klincksieck.

Fenwick, Rohan S. H. 2011. *A Grammar of Ubykh*. München: LINCOM Europa.

Harris, Alice C. 2017. *Multiple Exponence*. Oxford: Oxford University Press.

Korotkova, Natalia & Yury Lander. 2010. Deriving suffix ordering in polysynthesis: Evidence from Adyghe. *Morphology* 20: 299-319.

Lander, Yury A. 2016. Word formation in Adyghe. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds), *Word-Formation. An International Handbook of the Languages of Europe, Vol. 5*, 3508-3526. Berlin, Boston: de Gruyter Mouton.

Lander, Yury A. & Yakov G. Testelets. 2017. Adyghe (Northwest Caucasian). In Michael Fortescue, Marianne Mithun, & Nicholas Evans (eds.), *The Oxford Handbook of Polysynthesis*, 948–970. Oxford: Oxford University Press.

Lomtatidze, Ketevan, Rauf Klychev, & B. George Hewitt. 1989. Abaza. In B. George Hewitt (ed.), *The Indigenous Languages of the Caucasus. Vol. 2. The North West Caucasian Languages*, 91–154. Delmar, NY: Caravan.

Panova, Anastasia B. 2019. Refaktivy v abazinskom jazyke [Refactives in Abaza]. *Acta Linguistica Petropolitana* 15(2): 195–212.

Pazov, Sergej U. 2019. Usilitel'no-podtverditel'naja častica *g'(ə)-* i osobennosti eje upotreblenija v strukture otricatel'nogo glagola v abazinskom jazyke [The emphatic-assertive particle *g'(ə)-* and the peculiarities of its use in the structure of negative verbal forms in Abaza]. In: B. Č. Bižoev (ed.), *Kavkazskaja filologija: istorija i perspektivy. K 90-letiju Muxadina Abubekiroviča Kumaxova. Sbornik naučnyx statej* [Caucasian philology: history and prospects. Papers devoted to the 90th birthday of Mukhadin Kumakhov], 219–225. Nalčik: Institut gumanitarnyx issledovanij.

Smeets, Rieks. 1997. Suffixal marking of plural in Ubykh verb forms. In Sumru A. Özsoy (ed.), *Proceedings of the Conference on Northwest Caucasian Linguistics*, 10–12 October 1994, 37–62.

Stoynova, Natalia M. 2013. *Pokazateli refaktiva* [Refactive markers]. Moscow: AST-Press.

Tabulova, Nurja T. 1976. *Grammatika abazinskogo jazyka. Fonetika i morfologija* [A Grammar of Abaza: Phonetics and Morphology]. Cherkessk: Karačaevo-Čerkesskoe otdelenie Stavropol′skogo knižnogo izdatel′stva.

Vogt, Hans. 1963. *Dictionnaire de la langue oubykh*. Oslo: Universitetsorlaget.

Wälchli, Bernhard. 2006. Typology of heavy and light *again* or the eternal return of the same. *Studies in Language* 30(1): 69–113.

Gilles Authier (EPHE, PSL, Paris)

Michael Daniel (University of Tübingen)

Sergey Say (University of Potsdam)

Argument encoding in East Caucasian: evidence for convergence

While both head-marking and dependent-marking devices employed in argument encoding can be diachronically relatively stable, the lexical extent of argument-encoding patterns is easily susceptible to contact-induced changes (Grossman and Witzlack-Makarevich 2019; Trips 2020). In this study, our objective is to unravel the role of convergent processes in the organization of bivalent valency classes within East Caucasian (EC) languages. This language family serves as an ideal testing ground for our purposes for a number of reasons: i) it is sufficiently large in terms of time depth and the number of languages; ii) these languages showcase intricate valency class systems and iii) their speakers are involved in documented historical patterns of multilingualism (Dobrushina et al. 2017), both within the family and with neighbouring languages which pave the wave for active processes of convergence at all levels of language structure.

The data for this study come from BivalTyp, a typological database of bivalent verbs and their encoding frames (Say ed. 2020-). The database is based on a questionnaire comprising 130 bivalent verbs, such as 'be afraid', 'follow', 'see', 'touch'. The entries in the dataset are annotated for the argument-encoding devices employed for the two arguments, X and Y (e.g., in 'X is afraid of Y'). The entries are further annotated for the "locus of (non-)transitivity", a four-way contrast between transitive pattern (lit. 'X fears Y') and, if not, whether X ('for X is frightening Y'), Y (lit. 'X fears from Y'), or both arguments (lit. 'for X is fear from Y') are encoded by non-core devices.

To compare languages, we employ various distance metrics. One such metric is the simple matching distance based on the lexical distribution of the locus of (non-)transitivity. This metric effectively captures differences in Transitivity Prominence and fundamental types of non-transitivity, but disregards differences between specific argument-encoding devices, such as spatial cases. Another metric captures more fine-grained (dis)similarities in the lexical organisation of the valency classes, based on MI (Mutual Information) between valency class systems, but performs poorly in larger samples of structurally diverse languages. To analyze the resulting distance matrices, we employ Multidimensional Scaling (MDS), Neighbornet and Hierarchical Clustering. Our dataset encompasses 30 EC languages and several further languages of the Caucasus (Kartvelian, Turkic, etc.) that may have been donors of valency patterns.

All structural distances reveal a clear genealogical signal: i) EC languages form a relatively robust cluster amidst the surrounding linguistic landscape, and ii) systematic differences exist between first-level branches within the EC family. However, we also note some areal similarities with neighbouring non-EC languages. A prominent example is Udi: here, various aspects of valency class organisation closely resemble those observed in Turkic languages rather than in genealogically related EC languages. Similarly, certain Dargwa varieties experiencing significant contact with Turkic languages, such as Kaytag and Kadar Dargwa, exhibit an increase in transitivity prominence. Equally significant are clusters that transcend genealogical taxons within the EC family. One case in point is Archi, a Lezgic outlier surrounded by non-Lezgic languages, which showcases deviations from typical Lezgic valency class patterns and shares similarities with languages in its area. In our talk, we will hypothesize about the mechanisms driving these observed convergent processes, such as the evolution of case systems and syntactic calques.

References

Dobrushina, Nina, Daria Staferova, and Alexander Belokon (eds.). 2017. Atlas of Multilingualism in Dagestan Online. Linguistic Convergence Laboratory, HSE. (Available online at https://multidagestan.com)

Grossman, Eitan & Alena Witzlack-Makarevich. 2019. Valency and transitivity in contact: An overview. Journal of Language Contact 12(1). 1–26.

Say, Sergey (ed.). 2020-. BivalTyp: Typological database of bivalent verbs and their encoding frames. (Available online at https://www.bivaltyp.info)

Trips, Carola. 2020. Copying of argument structure: A gap in borrowing scales and a new approach to model contact-induced change. In Bridget Drinka (ed.), Historical Linguistics 2017. Selected papers from the 23rd International Conference on Historical Linguistics, San Antonio, Texas, 31 July – 4 August 2017, 413–434. Amsterdam: John Benjamins. DOI:10.1075/cilt.350.19tri

Deriving Laz causatives

Muhammet Bal & Ömer Demirok
Boğaziçi University

Introduction. Based on novel data from Atina dialect of Laz (Kartvelian), this study explores the morphological causatives with a focus on understanding the structural properties of the embedded constituent and the status of the overt and null causees in causatives of transitives (COTs henceforth) and unergatives respectively. We argue that the causatives of unergatives and COTs are not structured in the same way, with the position of the causee varying across the two structures.

Causatives of transitives. Laz has only one productive causative morpheme for transitives: *-ap*. When embedded under morphological causatives, the causee of such predicates are structured as agents introduced in an active Spec, VoiceP and the whole structure is bi-eventive (as evidenced by 'again' attachment and manner adverb modification). The causee in COTs is an argument despite being optional (contra Nash (2020) for Georgian COTs). Evidence for the causee's agentive role in COTs includes its ability to be modified by agent-oriented adverbs in appropriate contexts, comitatives and instrumentals – reflecting the canonical behavior of agents (Landau, 2010; Legate, 2014) as in (2a). In cases where the causee is null, we argue that the embedded constituent is constructed as an active existential Voice which doesn't project an implicit argument despite exhibiting active syntax as evidenced by binding, agreement, scope of negation, depictive licensing (e.g., as shown in (2b); cf. the causee in COTs licenses a depictive secondary predicate), sluicing and unavailability of oblique adjuncts (Šereikaitė, 2020).

(1) a. Muradi-k biç'i-s urdzeni d-o-3il-ap-u
Murat-ERG boy-DAT grape.NOM AFF-VAL-harvest-CAUS-3SG.PST
'Murat made the boy harvest grape.'

b. Neza-k dişk'a o-ç'it-ap-u
Neza-ERG wood.NOM VAL-chop-CAUS-3SG.PST
'Neza made some or other person person chop wood.'

(2) a. Befa-k bere-s$_i$ taml-epe makasi-te$_i$ dol-o-xorx-ap-u
Befa-ERG child-DAT bush-PL shears-INST PV-VAL-prune-CAUS-3SG.PST
'Befa made the child prune off the bushes with the shears.' *Comitative Attachment*

b. Himu-k t'at'ai t'at'ai ek'na g-o-n3-ap-u
3SG-ERG naked naked door.NOM PV-VAL-open-CAUS-3SG.PST
'S/he$_i$ made some or other person$_j$ open the door naked$_{i/*j}$.' *Depictive Licensing*

Causatives of unergatives. Unergatives allow two types of morphologically distinct causativization patterns: indirect and direct. In the former, they pattern just as in COTs, (2a) where the causee is marked with DAT and the causative morpheme is uniformly *-ap*. However, direct causatives (DCs) of unergatives have three allomorphs, respectively illustrated in (3b-c-d). DCs, moreover, present a challenge regarding the position of the causee. Assuming the long-held view that unergatives are concealed transitives (Hale & Keyser, 1993; Baker & Bobaljik, 2017 a.o.), the causee of a causativizated unergative is always expected to surface in dative just like the causee in COTs (see 1a). This is not borne out as illustrated in (3b-c-d)

(3) a. Tamara-k bozo-s o-nçir-**ap**-u
Tamara-ERG girl-DAT VAL-swim-CAUS-3SG.PST
'Tamara made the girl swim.' *Indirect Causative*

b. İsma-k bozo o-barbal-**ap**-u
İsma-ERG girl.NOM VAL-nag-CAUS-3SG.PST
'İsma made the girl nag/speak nonsense.' *Direct Causative*

c. Ali-k bere o-k'i-**in**-u
Ali-ERG child.NOM VAL-scream-CAUS-3SG.PST
'Ali made the child scream.' *Direct Causative*

d. Neza-k biç'i o-ncir-∅-u
Neza-ERG boy.NOM VAL-sleep-CAUS-3SG.PST
Neza made the boy sleep.' *Direct Causative*

While the causee in indirect causatives behaves like an agent, e.g., it binds the reflexive in (4a), a causee

in DC of an unergative is not able to do so, (4b).

(4) a. Tamara-k bozo-s ti-muşi o-nçir-**ap**-u
Tamara-ERG girl-DAT herself VAL-swim-CAUS-3SG.PST
'Tamara made the girl$_i$ swim herself$_i$.'

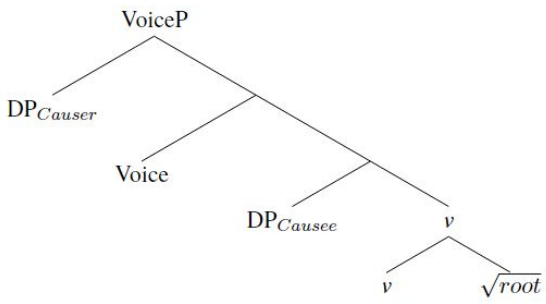
b. İsma-k bozo ti-muşi o-k'i-**in**-u
İsma-ERG girl.NOM ti-muşi VAL-nag-CAUS-3SG.PST
'İsma made the girl$_i$ scream herself$_{*i}$.'

The whole pattern is illustrated in Table 1.

	Direct		**Indirect**	
	Causee	**CAUS**	**Causee**	**CAUS**
Group 1	NOM	*-in*	DAT	*-in-ap*
Group 2	NOM	*-ap*	DAT	*-ap*
Group 3	NOM	∅	DAT	*-ap*

Table 1: Direct and Indirect Causatives of Unergatives

(5)



We argue that, unlike in COTs and ICs of unergatives, the causee in DC of unergatives is not structured as an agent. Instead, unergatives exhibit causative alternation similar to that of unaccusatives where the causee is on par with themes (Neu, 2023). Drawing on evidence from incompatibility with agent-oriented adverbs, compatibility with low applicatives and ability to form reduced relatives, we propose that the causee is merged as an internal argument, distinct from the subject of unergatives in root clauses - corresponding to an overall structure identical to a regular transitive construction illustrated in (5) (See also Legate (2014) for a similar proposal for Acehnese unergatives and Neu (2023) for Turkish and a couple of other unrelated languages).

An insight from variable unaccusativity. Arguing that the causee in DCs of unergatives is merged as a patient-like argument raises an important question: How is the sole argument of the unergative mediated between an agent and a patient-like role? We suggest that the answer lies in what Neu (2023) terms *variable unaccusativity*, whereby roots can exhibit different behaviors in certain contexts. This perspective also raises the possibility that unergatives may behave similarly to unaccusatives in contexts beyond causativization. We propose that this line of inquiry is promising and present two pieces of evidence from the Dative Subject Construction and participles in Laz.

Implications. The study contributes to our understanding of Laz argument structure, supports a Voice-stacking account of causatives (Akkuş, 2020; Nie, 2020) and resolves a challenge to the dependent case theory by demonstrating that in direct causatives of unergatives, the causee is merged as an internal argument, which explains why it fails to receive dependent dative case. Moreover, it brings in support for a recent analysis of DCs by providing additional evidence from Laz.

References. **[1]** Akkuş, F. (2021). (Implicit) argument introduction, voice and causatives (Unpublished doctoral dissertation). University of Pennsylvania. **[2]** Baker M, Bobaljik J. 2017. On inherent and dependent theories of ergative case. In The Oxford Handbook of Ergativity, ed. J Coon, D Massam, LD Travis, pp. 111–34. Oxford, UK: Oxford Univ. Press **[3]** Hale, K., & Keyser, S. J. (1993). On Argument Structure and the Lexical Expression of Syntactic Relations. In K. Hale & S. J. Keyser (Eds.), The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger (pp. 53–109). Cambridge, MA: MIT Press. **[4]** Landau, I. (2010). The explicit syntax of implicit arguments. Linguistic Inquiry, 41(3), 357–388. **[5]** Legate, J. A. (2014). Voice and v: Lessons from Acehnese (Vol. 69). MIT Press. **[6]** Nash, L. (2020). Causees are not agents. In Perspectives on Causation: Selected Papers from the Jerusalem 2017 Workshop (pp. 349–394). **[7]** Neu, E. (2023). Against a low subject analysis of causatives of unergatives. Proceedings of the Linguistic Society of America, 8(1), 5527. **[8]** Nie, Y. (2020). Licensing arguments (Unpublished doctoral dissertation). New York University. **[9]** Pylkkanen, L. (2008). Introducing arguments (Vol. 49). MIT press. **[10]** Šereikaitė, M. (2020). Voice and case phenomena in Lithuanian morphosyntax (Unpublished doctoral dissertation). University

of Pennsylvania.

Background: In the extensive literature on causative constructions, it is standardly assumed that causatives denote a complex event, which consists of two sub-events: the caused sub-event and the causing sub-event (Pylkkänen 2000, 2008, a.m.o.). These two events may or may not be independently represented in the syntax. This approach has been applied to the data from several Nakh-Dagestanian languages, e.g., Tsez (Kulikov 1993), Khwarshi (Sibilev 2023) and Dargwa (Kalyakin 2025). We investigate the event structure of Chechen indirect causatives formed with the suffix /*-iit-*/ by applying the eventhood diagnostics that have been established in the existing literature. It is shown that Chechen causatives pattern with bi-eventive causative constructions. The data was gathered during elicitation with speakers of the Plains dialect of Chechen.

Diagnosing event structure: In the Chechen causative construction formed from a plain transitive base verb, the Causer argument bears ergative case, the Causee argument is marked with allative, and the Theme is absolutive (unmarked). The examples below involve transitive base verbs; the eventhood diagnostics yield the same results for intransitive bases. The first diagnostic we employ is the possibility of modifying each of the sub-events with manner adverbials. In (1), the adverb 'quickly' can modify either the caused or the causing sub-event, suggesting that the two events are represented in the syntax.

(1) naana-s sixa joʕ-ie kiexat d-ieš-iit-ira
mother-ERG quickly daughter-ALL letter(CL4) CL4-read-CAUS-WPST
'The mother quickly made the daughter read the letter'
'The mother made the daughter quickly read the letter'

Secondly, frequency adverbials of the form 'n times' can take scope either over the caused event or over both sub-events (2).

(2) naana-s shozza joʕ-ie bait j-ieš-iit-ira
mother-ERG twice daughter-ALL poem(CL3) CL3-read-CAUS-WPST
'The mother made the daughter read the poem twice' CAUS>2x; 2x>CAUS

Another diagnostic is based on the temporal orientation of the sub-events: the caused and causing sub-events cannot receive independent temporal modification (3).

(3) *sielxana naana-s joʕ-ie ziezag-aš taxana ec-iit-ira
yesterday mother-ERG daughter-ALL flower-PL today buy-CAUS-WPST
Int: 'Yesterday, the mother made the daughter buy flowers today'

It is generally observed that mono-eventive causatives allow negation to only scope over the causing event and bi-eventive causatives allow both scope configurations. In Chechen, sentential negation can take either wide or narrow scope with respect to the causing sub-event (4), although for some speakers, the NEG>CAUS reading is less prominent.

(4) naana-s joʕ-ie kiexat ca d-ieš-iit-ira
mother-ERG daughter-ALL letter(CL4) NEG CL4-read-CAUS-WPST
'The mother did not make/let the daughter read the letter'

To summarize, most of the diagnostics suggest that Chechen causatives involve two syntactically represented events.

Diagnostic	Event structure
Manner adverbials	Bi-eventive
Frequency adverbials	Bi-eventive
Temporal localization	Mono-eventive
Negation	Bi-eventive

Summary of the event structure diagnostics

Analysis and implications: We propose that Chechen indirect causatives pattern with Phase (Voice)-selecting causatives (Pylkkänen 2008; Harley 2017). Firstly, they are highly productive and show no restrictions on the base verb in terms of argument structure (Komen et al. 2020). Secondly, the Causee argument retains subject properties, as it can be associated with agent-oriented modifiers, such as 'on purpose' (5).

(5) as hüttarienna cünga i huma aal-iit-ina
1SG.ERG on.purpose 3SG.ALL DEM thing say-CAUS-PERF
'I made him say this on purpose' CAUS > on purpose; on purpose > CAUS

A prediction of some recent theories of the syntax of morphological causatives (Nie 2022) is that Voice-selecting bi-eventive causatives will allow causative recursion, i.e., the possibility to introduce an unlimited number of Causee arguments. This follows from the assumption that, like any other theta-role, the role of the Causee can only be assigned once per thematic domain. This is not confirmed for Chechen: more than one Causee argument cannot be introduced (6).

(6) Impossibility of causative recursion

a.* da-s k'ant-ie žʕaela id-iit(-iit)-ira
father-ERG son-ALL dog run-CAUS(-CAUS)-WPST
Int.: 'The father made the son make the dog run'

b.* lör-uo neen-ie k'ant-ie molxa mal-iit(-iit)-ira
doctor-ERG mother-ALL son-ALL medicine drink-CAUS(-CAUS)-WPST
Int.: 'The doctor made the mother make the son take the medicine'

Conclusion: We have presented novel data on Chechen morphological causatives. Our findings suggest that the indirect causative is bi-eventive and Voice-selecting (since subject properties are retained). Causative recursion is unavailable. Following the proposal for Dargwa causatives in Kalyakin (2025), we suggest that the Chechen data may also be problematic for Voice-over-Voice approaches to the syntactic structure of morphological causatives.

Acknowledgements: The results of the project "Linguistic and cognitive diversity in formal models, computer tools, and educational resources" (2025-2027), carried out within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University), are presented in this work.

Abbreviations: 1-3 – person; CL1-6 – noun class, order as in Dotton and Wagner (2017); ALL – allative; CAUS – causative; DEM – demonstrative; ERG – ergative; NEG – negation; PERF – perfect; PL – plural; SG – singular; WPST – witnessed past.

References: • Dotton, Z. and Wagner, J. D. (2017). A Grammar of Chechen. • Harley, H. (2017). The "bundling" hypothesis and the disparate functions of little v. *The verbal domain*, 3:3–28. • Kalyakin, I. (2025). Morphological Causative in Kaytag Dargwa and the (Im)possibility of Causative Recursion. Ms. • Komen, E. R., Molochieva, Z., and Nichols, J. (2020). *The Oxford handbook of languages of the Caucasus*, chapter 8, pages 317–365. Oxford University Press, USA. • Kulikov, L. I. (1993). The "second causative": A typological sketch. *Causatives and transitivity*, 23:121–154. • Nie, Y. (2022). Turkish causatives are recursive: A response to Key 2013. *Linguistic Inquiry*, pages 1–14. • Pylkkänen, L. (2000). Representing causatives. In *Semantics and linguistic theory*, pages 132–148. • Pylkkänen, L. (2008). *Introducing arguments*. PhD thesis, MIT. • Sibilev, G. (2023). The Event Structure of the Morphological Causative in Khwarshi Proper. *Typology of Morphosyntactic Parameters*, 6:112–129.

Background: Current approaches to the syntax of derived nominals (Alexiadou 2001; Harley 2009; Borer 2023, a.o.) often state that complex event nominals (Grimshaw 1990) are formed by nominalizing whole verbal structures. The amount of verbal and nominal structure, which determines the verbal and nominal properties a particular nominalization will display, varies within and across languages. This work investigates the verbal and nominal properties of Chechen nominalizations (masdars).

Verbal properties: Chechen complex event nouns retain the argument structure of their source verbs and allow the case marking that is present in their corresponding finite clauses. In (1), the subject of the transitive verb 'read' bears ergative case, while the direct object is absolutive (unmarked). The example also shows that the nominalized verb agrees with the absolutive argument in noun class.

(1) suu-na [aħ bait j-ieša-r] daga-d-ea-ra
1SG-DAT 2SG.ERG poem(CL3) CL3-read-NZ(CL4) heart-CL4-come-WPST
'I remembered you reading the poem'

Secondly, nominalizations allow adverbial modifiers, including manner, aspectual, temporal and some modal adverbs (2). Additionally, the sentential negation morpheme can also be present in masdars (2b).

(2) a. cuo [ħo sixa/minot-eħ/cq'a/hetaħ v-ada-r-iex] laecna d-iic-ina
3SG.ERG 2SG(CL1) quickly/minute-INESS/once/then CL1-run-NZ-LAT about CL4-tell-PERF
'(S)he told about you running away quickly/in a minute/once/then'

b. [iza bilggal ca-qaača-r] xalaxiet-a suu-na
3SG certainly NEG-arrive-NZ upset-PRES 1SG-DAT
'I am upset that he certainly will not come'

Moreover, periphrastic constructions which are formed by a participle and the auxiliary *xila* 'to be' can be nominalized (3). The periphrastic constructions contain the simultaneous converb (3a), the anterior converb (3b) and the future participle (3c).

(3) a. [ħo quza-ħ v-eexa-š xila-r] xazaxiet-a suu-na
2SG(CL1) here-INESS CL1-live-CVB.SIM be-NZ like-PRES 1SG-DAT
'I like you living here'

b. [ħo quza-ħ v-aex-na xila-r] xazaxiet-a suu-na
2SG(CL1) here-INESS CL1-live-CVB.ANT be-NZ like-PRES 1SG-DAT
'I like you having lived here'

c. [ħo quza-ħ v-eexa-r xila-r] xazaxiet-a suu-na
2SG(CL1) here-INESS CL1-live-PTCP.FUT be-NZ like-PRES 1SG-DAT
'I like that you will live here'

Nominal properties: Subjects of nominalizations may be marked with genitive, although some restrictions apply.

(4) suu-na [ħan illi aala-r] xazaxiet-a
1SG-DAT 2SG.GEN epic.song say-NZ like-PRES
'I like your singing of the epic song'

Number marking on nominalizations which retain the arguments of their source clause is highly restricted, although attested in some cases.

Masdars allow adjectival modification, although its acceptability can depend on subject case (genitive vs. ergative/absolutive).

(5) suu-na sielxan-liera [ħan/*ħo sha-ca-qaača-r] daga-d-ea-ra
1SG-DAT yesterday-ADJZ 2SG.GEN/2SG to-NEG-arive-NZ(CL4) heart-CL4-come-WPST
'I remembered you not arriving yesterday'

Masdars are also compatible with the medial demonstartive *i*.

(6) suu-na i [ħo tila-r] daga-d-ea-ra
1SG-DAT DEM 2SG get.lost-NZ(CL4) heart-CL4-come-WPST
'I remembered that time when you got lost'

Finally, nominalizations can serve as complements of postpositions (2a) and bear case markers (7).

(7) naan-na [k'ant dʕa-v-axa-r]-uo xalaxiet-iit-ira
mother-DAT son.CL1 away-CL1-go-NZ-ERG upset-CAUS-WPST
'The fact that the son left caused the mother to get upset'

The verbal and nominal properties of masdars are summarized in the table below.

Structure	Property	Non-Genitive subject	Genitive subject
Verbal	Class agreement	+	+
	Expression of arguments	+	+
	Adverbial modification	+	+
	Negation	+	+
	Constructions with AUX (3)	+	–
Nominal	Adjectival modification	???/–	+
	Compatibility with determiners	+	+
	Case markers and postpositions	+	+

Verbal and nominal properties of Chechen nominalizations by subject case

Account: Our core observation is that masdars which retain a greater amount of verbal structure (e.g., envolve a construction with an auxiliary), exhibit some but not all nominal properties: they disallow genitive subjects and adjectival modifiers, but are compatible with determiners and postpositions. This suggests that the lower nominal projections which host adjectives are absent in their structure, while higher ones, which are associated with determiners, are present. This is in line with the observation that a greater amount of verbal structure in derived nominals correlates with a smaller amount of nominal structure (Alexiadou et al. 2011).

Conclusion: We have investigated the verbal and nominal structure of Chechen compex event nominals and suggested that they exhibit a variable amount of verbal and nominal structure: more verbal structure corresponds to less nominal structure. In our talk, we will present more data and elaborate on the structural correlates of the verbal and nominal properties discussed above.

Acknowledjements: The results of the project "Linguistic and cognitive diversity in formal models, computer tools, and educational resources" (2025-2027), carried out within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University), are presented in this work.

Abbreviations: 1,2,3 – person; CL1-6 – noun class, order as in Dotton and Wagner (2017), ADJZ – adjectivizer; ANT – anterior; CVB – converb; DEM – demonstrative; DAT – dative; ERG – ergative; FUT – future; LAT – lative; NEG – negation; NZ – nominalization; PRES – present; PTCP – participle; SG – singular; SIM – simultaneous; WPST – witnessed past.

References: • Alexiadou, A. (2001). Functional Structure in Nominals: Nominalization and ergativity. *Linguistics Today. John Benjamins Publishing*, 77. • Alexiadou, A., Iordăchioaia, G., and Schäfer, F. (2011). Scaling the variation in Romance and Germanic nominalizations. *The noun phrase in Romance and Germanic*, pages 25–40. • Borer, H. (2023). Argument Structure and Derived Nominals. *The Wiley Blackwell Companion to Morphology*, pages 1–66. • Dotton, Z. and Wagner, J. D. (2017). A Grammar of Chechen. • Grimshaw, J. (1990). *Argument structure.* the MIT Press. • Harley, H. (2009). The morphology of nominalizations and the syntax of vP. In *Quantification, definiteness and nominalization*, pages 320–342. Oxford University Press.

I want ~~Ahmad~~ to eat porridge: the Causative in Melkhi as Switch-Reference*

Gleb Bubnov
HSE University, Moscow
gbubnov@hse.ru

Introduction. Melkhi is a language of the Nakh-Daghestanian family, positioned at the center of the Chechen-Ingush dialect continuum. The data for this presentation was collected during a field expedition in the summer of 2024 in the village of Arshty, Republic of Ingushetia. This presentation explores two functions of the suffix *-iit*: (indirect) causative and **switch-reference** (see Nichols (1985) for similar Chechen data) and provides a **syntactic** analysis that generalizes these two functions that have not yet been proposed.

Data. The causative suffix *-iit* can attach to any verb and conveys a maximally **underspecified** causative meaning. The original subject of a transitive verb is in the allative case.

(1) as moʜmad-ag keʜat yaz d-aʔ-iit-na
1SG.V.ERG Mohammad.V-ALL letter.D write D-do-CAUS-PERF
'I made/asked/allowed Mohammad to write a letter.'

In addition to straightforward causative contexts, *-iit* is used in infinitival clauses as a marker of **switch-reference**, indicating cases where the subject of the embedded clause does not coincide with that of the matrix clause. These are **control** predicate constructions, such as ones with the verb *la* 'want', where the dependent clause is in the infinitive form.[1]

(2) su-n dˤa v-ij v-ax low
1SG.V-DAT COMP V-sleep V-go want.PRES
'I want to go to sleep.'

In cases of non-coreference[2], the causative marker *-iit* attaches to the verb. It is worth noting that the infinitive causative cannot have PRO **controlled** by the subject of the matrix clause, as seen in (3-b).

(3) su-n low Mansur txu-g ca v-aʔ-iit
1SG.V-DAT want.PRES Mansur.V 1EXCL-ALL NEG V-go-CAUS
a. 'I want Mansur not to come to us.'
b. *'I want to make Mansur not come to us.'

Infinitival clauses may also have an **imperative** function. When forms with *-iit* are used (4), a straightforward **imperative** meaning with causation (4-a) or a mere **jussive** meaning (4-b) is possible. The latter one disappears with the **explicit** expression of the subject (5-b).

(4) Aʜmad-ag xudar d-aʔ-iit
Ahmad.V-ALL porridge.D D-eat-CAUS
a. 'Make Ahmad eat the porridge.'
b. 'Let Ahmad eat the porridge.'

(5) aʜ Aʜmad-ag xudar d-aʔ-iit
2SG.V.ERG Ahmad.V-ALL porridge.D D-eat-CAUS

*The results of the project "Linguistic and cognitive diversity in formal models, computer tools, and educational resources" (2025-2027), carried out within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University), are presented in this work.

[1]The dependent clause can also be finite and marked by a subjunctive marker.

[2]In fact, *-iit* can be used even when the subjects of the matrix and embedded clauses coincide, as long as there is no **control**. For example, the sentence *su-n low su-g xudar d-aʔ-iit* is grammatical and means "I want **me** to eat porridge" instead of "I want to eat porridge."

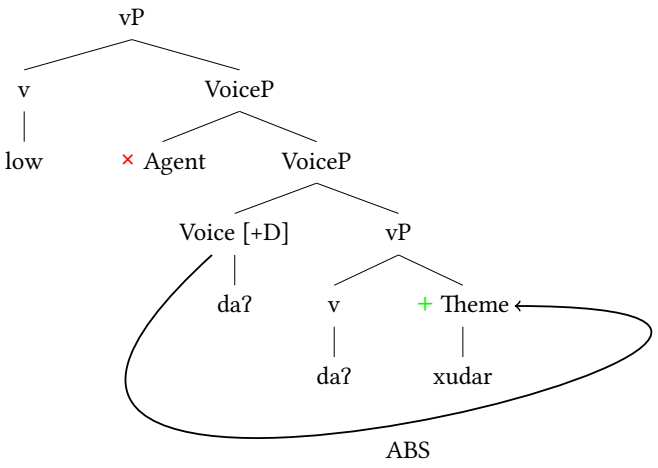
a. ‘(You) make Ahmad eat the porridge.’

b. *‘Let Ahmad eat the porridge.’

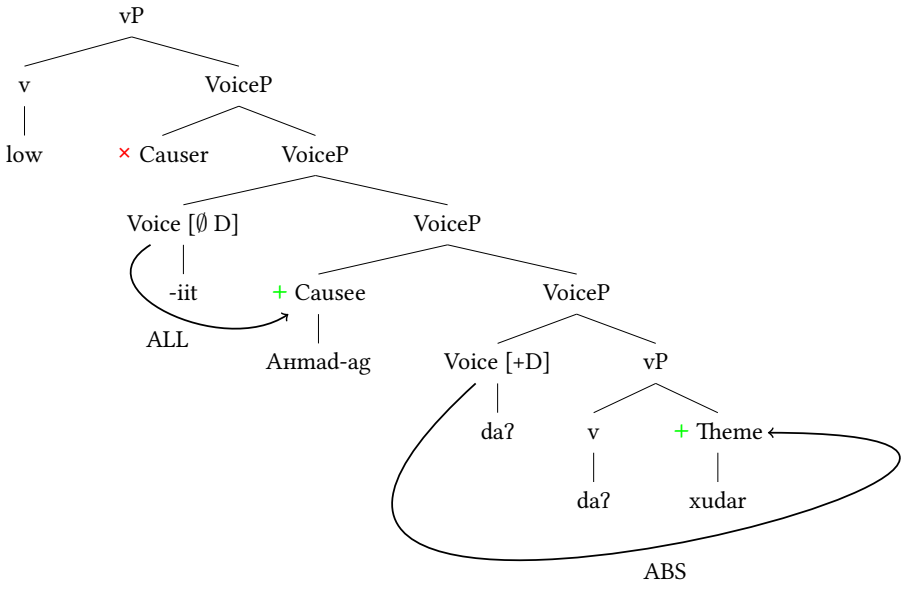
Analysis. I propose that the morpheme *-iit* instantiates *Voice [Ø D]* (Nie, 2020), a voice head that **neither requires nor prohibits** an external argument. Here, *Voice [Ø D]* assigns the allative case to its argument. Following Nie (2020), I also assume that the **highest** *Voice*, that is located under some other heads in the verbal domain, can license **two** arguments, while other *Voice* heads **only one**.

The absence of requirements of *Voice [Ø D]* for a specifier firstly explains the **unspecified** semantics of the causer (1). Secondly, since in non-causative infinitival clauses the highest argument is not licensed (presumably due to the lack of specific projections above *VoiceP*), but is required by the *Voice [+D]* head (6), this leads to subject control in the embedded clause. This explains the use of *-iit* in infinitival clauses in the absence of **control** (7). To account for the **prohibition** of *PRO* when the causative morpheme is present, I suggest that when an argument is optional and cannot be licensed, it never becomes *PRO*, as control is only a **last resort**.

(6) su-n low xudar d-a?
1SG.V-DAT want.PRES porridge D-eat
’I want to eat porridge’

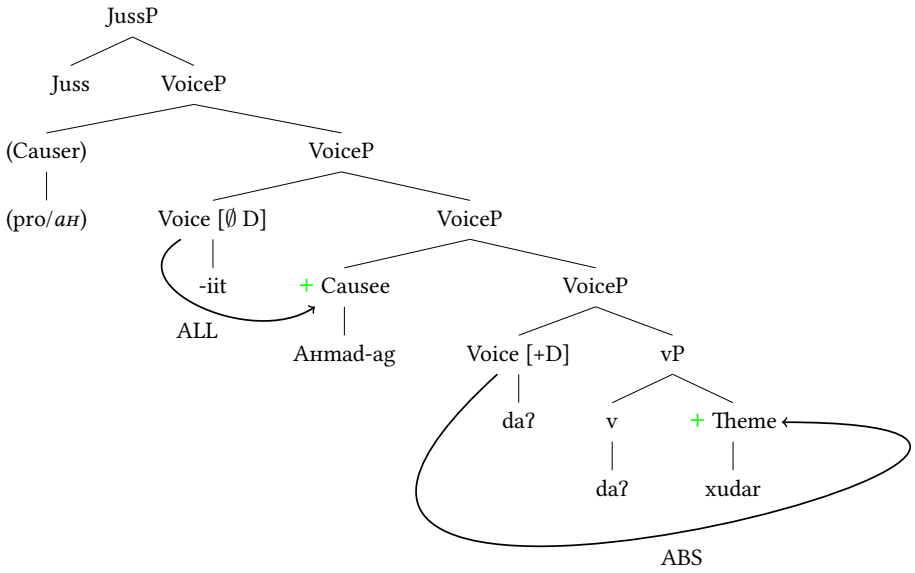


(7) su-n low Ahmad-ag xudar d-a?-iit
1SG.V-DAT want.PRES Ahmad.V-ALL porridge.D D-eat-CAUS
’I want Ahmad to eat porridge’



Concerning the **imperative** function of the causative, I follow Zanuttini et al. (2012) in assuming that imperatives involve a *Jussive* head. This head carries ϕ-features of the addressee and binds the specifier of *VoiceP*. I propose, however, that with an additional *Voice [∅ D]* head licensing a lower external argument and optionally introducing its own, the *Jussive* head's role becomes optional.

(8) Ahmad-ag xudar d-a?-iit
Ahmad.V-ALL porridge.D D-eat-CAUS
'Make Ahmad eat the porridge.'
'Let Ahmad eat the porridge.'



Glossary of Abbreviations. ALL – allative; CAUS – causative; COMP – completive; D – class prefix; DAT – dative; ERG – ergative; EXCL – exclusive; NEG – negation; PERF – perfect; PL – plural; PRES – present; SG – singular; V – class prefix; 1 – first person; 2 – second person.

References. • Nichols, J. (1985). Switch-reference causative. In *Chicago Linguistic Society 21, Part 2: Papers from the Parasession on Causatives and Agentivity*, pages 193–203, Chicago. The University of Chicago Press. • Nie, Y. (2020). *Licensing arguments*. Phd thesis, New York University. • Zanuttini, R., Pak, M., and Portner, P. (2012). A syntactic analysis of interpretive restrictions on imperative, promissive, and exhortative subjects. *Natural Language & Linguistic Theory*, 30(4):1231–1274.

Vladimir de Haldat – Ecole Pratique des Hautes Etudes, PROCLAC UMR 7192

Univerbation and aspect marking in Akhwakh

Akhwakh belongs to the Nakh-Daghestani language family, also known as Northeast Caucasian. In Akhwakh, as in other languages of the western group of this family, aspect is not marked by the verbal root, but by suffixes and periphrastic forms. These analytic forms can morphophonologically merge, which will be the topic of this presentation.

More specifically, the presentation will focus on the study of external sandhi phenomena – which are very flexible, as shown by the provided examples – affecting the verb, i.e. the fusion and univerbation of two or more words, including a verbal lexeme. The forms thus created can be divided into three categories: underlying periphrastic forms with aspectual value, verbal locutions with a light verb and fused forms in which the auxiliary "to be" has univerbated with a noun. Particular attention will be paid to aspect marking. These sandhi periphrases significantly complicate the verbal morphology of Akhwakh, creating new forms and expanding the TAME paradigm.

Aspect can be marked by periphrastic forms, optionally univerbated. Akhwakh thus creates new aspecto-temporal categories that cannot be expressed as such by a specific suffix different from the imperfective and perfective suffixes. In example (1), the form *b-eq'erēk'w-āwudi* consists of the lexical verb *b-eq'-uruƛa* "to know" in the progressive converbial form *-ere* and the auxiliary *b-ik'-uruƛa* in the perfect tense. This combination is used to express the imperfect, and thus constitutes a paradigm augmentation, since synthetic forms cannot express imperfect in Akhwakh. It is also possible to augment the evidential paradigm with the auxiliary *m-ič-unuƛa* "to find (oneself)" in the conditional *m-ič-ala* (2). The auxiliary expresses an event that has not been witnessed. It is therefore an evidential auxiliary univerbated with the short infinitive (marked with *-u* instead of *-uruƛa*), creating a synthetic counterfactual form.

The univerbation of two elements can allow one of them to change part of speech. This is witnessed in Akhwakh with the verb "to come" *b-oɬ̄-uruƛa*, which can, when combined with a noun, form a new denominative verbal lexeme (3). This new verbal lexeme is incorporated into a perfect form. Nonverbal lexemes can also, when merged with verbal lexemes, give rise to preverbal forms (4).

Finally, there is the case of the auxiliary *b-ik'-uruƛa* "to be" merging with a noun (5). In this case, the auxiliary is not univerbated with the auxiliary *b-iχ̄-ada*, but is moved onto a nominal constituent, *beča*, to emphasize it. The form thus obtained is combined with the auxiliary to create a perfect.

Studying merged verb forms is a necessary step in understanding how aspect marking works in Akhwakh. The forms analyzed in this presentation will help clarify the typology of aspect marking in this language.

1)

b-eq'erēkʼʷ-āwudi	<	*b-eq'-ere*	*b-ikʼʷ-awudi*
N-know.CV:PROG+(N)BE-PERF		N-know-CV:PROG	N-be-PERF

hugu-šu-la	*b-eq'erēkʼʷ-āwudi*	*inšu-be*	*koša*	*eq̄-u*
DIST-M-DAT	N-know.CV:PROG+(N)be-PERF	REFL.M-N(GEN)	bad	look-INF
šo-ɬ-ika	*k'eha*	*b-ikʼʷ-e.*		
good-VBLZ-IPF.NEG	eye	N-be-SEQ		

'He knew he had an evil eye and could not look without bringing bad luck.'

2)

b-ikʷeč-ala	<	*b-ik'-u*	*m-ič-ala*
N-be.INF+(N)find-COND		N-be-INF	N-find-COND

t'ẽki	*Zamaq̄o*	*kurak'e*	*kʷīłe*	*b-ikʷeč-ala*
throw.IPF.NEG	Zamako	abricot	wish.MSD	N-be.INF+(N)find-COND
mede-dala	*b-iɬ-u*	*b-ux̄-e*	*b-ik'-uwa*	*ruša.*
2SG.ERG-at.least	N-put-INF	N-fall-SEQ	N-be-PERF?/FUT?	tree

'I won't throw any at you, Zamaqo, if you wanted an apricot all you had to do was topple the tree.'

3)

waranoƛ̄ʷ-ehe	<	*warani b-oƛ̄-ehe*
camel+(N)go-SEQ.N		camel-N-go-SEQ.N

dene	*waranoƛ̄ʷ-ehe*	*gwede.*
1SG.NOM	camel+(N)go-SEQ.N	COP.N

'I turned into a camel.'

4)

ʁadabux̄-alaq'o	<	*ʁadiga*	*b-ux̄-alaq'o*
ground.LAT+N.fall-CV:ANT		ground.LAT	N-fall-CV:ANT

ʁadabux̄-alaq'o	*b-adaƛ-ewudi*	*šari*
ground.LAT+N.fall-CV:ANT	N-laugh-PERF	fox

'The fox laughed his head off.'

5)

bečekʷ-awudi	<	*beča*	*b-ikʷ-awudi*
mountain+(N)be-PERF		mountain	N-be-PERF

bečekʷ-awudi	*Cetãka-s̄ʷ-e*	*b-ix̄-ada.*
mountain+(N)be-PERF	Cetanka-OBL.M-ERG	N-hold-PERF

'Cetanka had taken possession of this mountain.'

References

CREISSELS, Denis, 2010. « Liaison and grammaticalization in Northern Akhvakh », [online] https://disk.yandex.ru/d/fqtPmzQMOPDTEw/Andic/Axvax.

CREISSELS, Denis, 2016. « Univerbation via liaison and the evolution of lexicon and grammar in Northern Akhvakh » [online] http://www.deniscreissels.fr/public/Creissels-liaison_Akhvakh.pdf

Nina Dobrushina
(Laboratoire Dynamique du Langage, CNRS)

Curses are structurally different from blessings

Blessings and curses are a prominent element of traditional discourse in Daghestan. Most Nakh-Daghestanian languages have dedicated verb forms which express blessings and curses by the same morphological marker:

Mehweb Dargwa (Nakh-Daghestanian) (Dobrushina 2019)

(1) blessing

hum-be	ʔaˁχ	d-uh-a-b!
road-pl	good	npl-become:pfv-irr-opt

'May you have a good trip!'

(2) curse

kapul-le	w-ebk'-a-b!
pagan-advz	m-die:pfv-irr-opt

'May he die impious!'

There is some typological evidence that languages may use dedicated forms for curses (Aikhenvald 2020). While in Nakh-Daghestanian languages such cases were not attested, I will show in this paper that a tendency for certain forms and constructions to specialize in expressing curses is indeed observed.

For this study I chose three Nakh-Daghestanian languages which belong to different branches and are not in contact: Avar and Lak (both standard varieties) and Rutul (Mukhad variety). Using bilingual Avar / Lak / Rutul - Russian dictionaries and phraseological dictionaries as well as collections of folklore texts as sources of examples of wishes, I collected between 96 and 238 examples, depending on the language. I coded the examples according to the semantic type of wish (blessing or curse), morphological verb form and certain properties of the constructions.

Avar and Lak were chosen for two reasons. First, both languages possess more than one optative form. Second, due to the fact that several dictionaries were available, the number of examples of wishes was sufficient to make a quantitative study. Rutul was selected as a language for which I could compare the data from the dictionaries with the corpus data.

The study showed that the expressions of curses exhibit similar structural properties in all three languages in the study.

First, in all three languages - Avar, Lak and Rutul - there are forms that are associated primarily with curses: optatives in *-ad* in Avar, *-wu(j)* in Lak and *-dɨ* in Mukhad Rutul, as well as second person imperatives in Lak.

Second, three constructional features were found mainly in curses: the absence of the words for 'God' and 'Allah' (as compared to their frequent use in blessings), the presence of second

person pronouns, and their final position in the utterance. The final position of the second person pronoun proved to be the most robust feature observed in the curses. It was found in all three languages in the study, for example:

Avar (Gimbatov 2006: 456)
(3) dandamaj-gi mun!
swell-OPT you.SG
'May you be swollen!'

In Mukhad Rutul the second person pronoun in the final position was most likely the source of the development of the optative affix *-dɨ*. The existence and origin of the form in *-dɨ* in Rutul, to my knowledge, has never been discussed, although there are eleven examples of this form in the Rutul-Russian dictionary (Alisultanov and Suleimanova 2019), which have been confirmed by the native speakers:

(4) kan **l-ešu-dɨ**
base **PV-1.take.PFV-DɨT**
'May you die!' (Alisultanov and Suleimanova 2019: 178)

In the talk I will argue that the form in *-dɨ* originates from the combination of the participle in *-d* with the second person pronoun *wɨ* (nominative): *l-ešu-d wɨ* (PV-1.take.PFV-ATTR 2SG) → *lešudɨ*.

An important finding of this study is that both morphological and constructional features associated with curses manifest themselves only as tendencies. There are no structures which are found exclusively in curses: even if the data obtained from the dictionaries and corpora does show a strict one-to-one correspondence between a property and expression of ill wishes, language consultants readily reproduce the same structure in blessings in elicitation. This may be one of the reasons why curses are rarely discussed as a special subtype of optative constructions cross-linguistically.

References

Aikhenvald, Alexandra Y. 2020. "Damn your eyes!" (Not really): Imperative imprecatives, and curses as commands. In: edited by N. Nassenstein, A. Storch, eds. Swearing and Cursing, Berlin, Boston: De Gruyter Mouton, pp. 53-78.

Alisultanov, Alisultan S., Sulejmanova, Taisija A. 2019. Russko-rutul′skij slovar′ [russian-rutul dictionary]. Maxačkala: Alef.

Dobrushina, Nina. 2019. Moods of Mehweb. In: Daniel, M., Dobrushina, N. & Ganenkov, D. (Eds.). 2019. The Mehweb language: Essays on phonology, morphology and syntax. Berlin: Language Science Press.

Gimbatov, M. M. 2006. Avarsko-russkij slovar′ [Avar-Russian dictionary]. Makhachkala: DNC RAN.

Abaza Dative and Sociative: Peculiarities and Classification Challenges

Anastasiia Dobrynina, HSE University[1]
aidobrynina12@gmail.com

Key words: Abaza language; applicatives; dative; sociative.

Applicatives in the Abaza language. Abaza is a Northwest Caucasian language with polysynthetic morphology and ergative-absolutive alignment. Non-core arguments can be introduced by postpositions or by special applicative affixes. This paper focuses exclusively on applicative strategy. In Abaza there are numerous locative applicatives and some grammatical applicatives: *benefactive, melefactive, instrumental, comitative, estimative, inadvertitive* (*involuntative*)*, dative, sociative* (see for example (Arkadiev et. al. 2024; O'Herin 2001)). In this paper we consider the peculiarities of dative and sociative applicatives, as described below. Our data come from elicitation sessions with residents of aul Staro-Cuvinsk, L1-speakers of the Abaza language, Ashqarawa dialect, in the year 2024.

What are dative and sociative? In literature two similar applicatives are reported: so-called dative *a-* and sociative *a(j)-* (Arkadiev submitted), whose distribution is lexically determined. The main difference between them lies in their position within the verbal complex: the sociative applicative is the closest prefix to the root, whereas the dative is separated from the root by negation, ergative, and causative markers, if present. We tested this claim for four verbs, and our results partially contradict the classification proposed in previous studies. According to (Arkadiev submitted), the dative occurs with the verb *pšra* ('to look'), and the sociative occurs with the verb *gʷəʁára* ('to hope'). However, we found that the element *a(j)-* is placed after the causative marker in the former verb (1) and before it in the latter verb (2). For the other two verbs, *hʷara* ('to say/tell') and *ŝaʒara/ẑaʒara* ('to talk'), elements *a(j)-* align with the classification of dative and sociative, respectively, as suggested in the literature.

(1) *də-r-d-s-**r-aj**-pš-ṭ*
3SG.H.ABS-3PL.IO-3PL.IO-1SG.ERG-**CAUS-SOC**-look-DCL
'I caused him to look like their'

(2) *sara* *fatima* *d-r-**aj**-gʷ-sə-**r**-gʷəʁaj-ṭ* *awət*
1SG Fatima 3SG.H.ABS-3PL.IO-**DAT**-LOC-1SG.ERG-**CAUS**-hope-DCL DIST.PL
'I caused Fatima to rely on them'

Problems with dative and sociative applicatives. Regardless of their distinction, both differ from other applicatives in all Northwest Caucasian languages (Arkadiev et. al. 2024: 892). In Abaza dative and sociative are peculiar in the following ways. Firstly, the occurrence of these applicatives is lexically determined whereas other applicatives exhibit greater productivity. In Abaza these can occur with verbs *hʷara* 'to say/tell' (3), *ŝaʒara/ẑaʒara* 'to talk' (4), *gʷəʁára* 'to hope' (5), *pšra* 'to look' (conveying the meaning of 'be similar to') (6) and *rejoice* (7), and these are the only verbs that we have observed to take a dative applicative thus far. Secondly, their semantics is peculiar: they either express trivial semantics in the sense that the thematic relation of an indirect object is determined by the stem (3-5, 7) or make non-compositional meaning (6). Finally, both dative and sociative occupy a special position in the verb structure, appearing after other grammatical and locative applicatives (8-9).

(3) *sará* *warά* *jə-**w-á**-s-hʷə-w-j*
1SG 2SG.M 3SG.N.ABS-**2SG.M.IO-DAT**-1SG.ERG-say-IPF-DCL
'I say it to you (male)'.

(4) *fatima* *akhmed* *d-**j-á**-ẑaʒa-w-ṭ*
Fatima Akhmed 3SG.H.ABS-**3SG.M.IO-SOC**-talk-IPF-DCL
'Fatima talks with Akhmed'

(5) *akhmed* *hará* *d-**h-ai**-gʷəʁá-w-ṭ*
Akhmed 1PL 3SG.H.ABS-**1PL.IO-DAT**-hope-IPF-DCL
'Akhmed relies on us'.

(6) *Fatima* *l-án* *d-**l-ái**-pšə-w-ṗ*
Fatima 3SG.F.PR-mother 3SG.H.ABS-**3SG.F.IO-SOC**-look-IPF-DCL
'Fatima is similar to (looks like) her mother'

[1] The results of the project «Linguistic and cognitive diversity in formal models, computer tools, and educational resources» (2025-2027), carried out within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University), are presented in this work.

(7) *w-a-m-**aj**-gʷərʁ'ə-w-š-ta* *j-g'-aʔa-m*
2SG.M.ABS-3SG.N.IO-NEG-**SOC**-rejoice-IPF-FUT-ADV 3PL.ABS-NEG.EMP-be-NEG
'It is impossible not to be happy about it.' (Abazašta 2017, cited in Arkadiev submitted)

(8) **j-s-ĉ-**s-a-ʕa**-r-hʷa-j*
3SG.N.ABS-1SG.IO-MAL-**1SG.IO-DAT-CSL**-3PL.ERG-say-DCL
Intended: 'They said it to me against my will'.

(9) **j-**l-a-s-ĉə**-w-hʷə-w-j*
3SG.N.ABS-**3SG.F.IO-DAT-1SG.IO-MAL**-2SG.M.ERG-say-IPF-DCL
Intended: 'You say it to her against her will'.

Hypothesis. To explain the peculiarities of dative and sociative applicatives, we hypothesize them to be low applicatives. According to (Pylkkänen 2002) low applicatives differ from the high ones in the way that their semantics is unspecified, i.e. "low applied arguments bear no semantic relation to the verb" (Pylkkänen 2002: 19). This difference is derived from the position of the applicatives: structured low (inside the VP) or high (outside the VP). Considering dative and sociative as low applicatives would explain their close position to the verb root and their semantics. To verify this, we need to apply diagnostics for applicative type.

Dative/sociative and depictives. According to (Pylkkänen 2002) only arguments introduced by high applicatives should be available for depictive modification (based on depictive diagnostics, this corresponds to a non-finite imperfective form in Abaza). However, this criterion yields different results depending on the verb. Depictive modification is ungrammatical for arguments introduced by dative or sociative with the verbs *gʷəʁára* 'to hope' and *pšra* 'to look' (10), but grammatical with the verb *ŝaʒara/ẑaʒara* 'to talk' (11). With the verb *hʷara* 'to say/tell' speakers exhibit variation (12).

(10) **rakhmet* *j-án* ***d-gʷərʁ'açʲá-w*** *d-**l-é**-pšə-w-ṗ*
Rakhmet 3SG.M.PR-mother **3SG.H.ABS-happy-IPF** 3SG.H.ABS-**3SG.F.IO-SOC**-look-IPF-DCL
Intended: 'Rakhmet is similar to his happy mother'.

(11) *sara* *s-**l-a-**ĉaẑa-w-n* *a-phʷəspa* ***d-gʷərʁ'açʲa-w***
1SG 1SG.ABS-**3SG.F.IO-SOC**-talk-IPF-PST SP-girl **3SG.H.ABS-rejoice-IPF**
'I talked with the girl happy'.

(12) *aslan* *bara* *arəj* ***b-a**-j-hʷa-jṭ* ***b-gʷərʁ'açʲa-w***
Aslan 2SG.F PROX.SG **2SG.F.IO-DAT**-3SG.M.ERG-say-DCL **2SG.F.ABS-rejoice-IPF**
'Aslan told it to you (f) happy'.

Dative/sociative and inadvertitive. During our fieldwork we observed that verbs under consideration differ in their ability to simultaneously host dative/sociative and inadvertitive applicatives. While this is possible with the verbs *ŝaʒara/ẑaʒara* 'to talk' and *hʷara* 'to say/tell' (13), it is ungrammatical with *pšra* 'to look' (14). For the verb *gʷəʁára* ('to hope'), the acceptability of such constructions varies across speakers (15).

(13) *sará* *aráj* *Akhmed* *j-**sə́-mqa-j-a**-hʷa-jṭ*
1SG 3SG.N Akhmed 3SG.N.ABS-**1SG.IO-INADV-3SG.M.IO-DAT**-say-DCL
'I said it to Akhmed not intentionally'.

(14) **rakhmet* *j-an* *d-**a-mqa**-g'-**l-aj**-pšə-m*
Rakhmet 3SG.M.PR-mother 3SG.H.ABS-**3SG.N.IO-INADV**-NEG.EMP-**3SG.F.IO-SOC**-look-NEG
Intended: 'It is an unintentional case that Rakhmet does not look like his mother'.

(15) %*zalina* *bara* *d-**a-mqa-b-ai**-gʷəʁa-w-ṭ*
Zalina 2SG.F 3SG.H.ABS-**3SG.N.IO-INADV-2SG.F.IO-DAT**-hope-IPF-DCL
Intended: 'Zalina overrelied (relied in vain) on you (female)'.

Conclusion. The verbs that can host dative or sociative markers appear to be inhomogeneous. The test on low applicatives does not provide clear results regarding the status of these applicatives, nor does their distribution with the inadvertitive applicative. This inhomogeneity also cannot be explained in terms of the distinction between dative and sociative. A summary of our findings is presented in Table 1.

Table 1. Properties of verbs with *a(j)-* prefix

	hʷara ‘to say/tell’	*ŝaʒara/ẑaʒara* ‘to talk’	*gʷəʁára* ‘to hope’	*pšra* ‘to look’
position relatively to causative	before	after	before	after
our classification	DAT	SOC	DAT	SOC
depictive modification	%	ok	*	*
combining inadvertitive with dative/sociative	ok	ok	%	*

Further research. Another important test on high and low applicatives by (Pylkkänen 2002) is that only high applicatives can occur with unergatives. We didn’t consider this diagnostic because tests for unergativity and unaccusativity are language-specific and we find it challenging to determine them for Abaza. Another restriction of our research is that we considered only overt dative marker whereas in Abaza there are verbs with recipient participant that is introduced with no applicative (16). Examining non-morphological dative constructions alongside overt dative markers could illuminate differences in applicative behavior and further refine our understanding.

(16) *s-abá* *č’áŝ* *pχá* *j-á-**sə**-j-ta-jṭ*
1SG.PR-father pie warm 3SG.N.ABS-CSL-**1SG.IO**-3SG.M.ERG-give-DCL
‘My father gave me a warm pie’.

Abbreviations

1, 2, 3 — first, second, third person, ABS — absolutive, ADV — adverbial, BEN — benefactive, CAUS — causative, CLS — cislocative, DAT — dative, DCL — declarative, DIST — distal demonstrative, EMP — emphatic, ERG — ergative, F — feminine, FUT — future, H — human, INADV — inadvertitive, IO — indirect object, IPF — imperfective, LOC —locative preverb, M — masculine, MAL — malefactive, N — neutral, NEG — neagtion, PL — plural, PR — possessor, PROX — proximal demonstrative, PST — past, SG — singular, SOC — sociative, SP — specific.

References

Abazašta 2017: «АуагIахъа гIазхчауа абызшва акIвпI» // *Abazašta*, 2017, July 27, p. 2.

Arkadiev (submitted) — Arkadiev, Peter (submitted). Abaza. A grammatical sketch. Submitted to Yuri Koryakov, Yury Lander & Timur Maisak (eds.), The Caucasian Languages. An International Handbook. De Gruyter Mouton.

Arkadiev, Peter, Lander, Yury, Bagirokova, Irina. (2024) Applicative constructions in the Northwest Caucasian languages // Applicative Constructions. A Comparative Handbook, ed. Denis Creissels et Fernando Zúñiga.

O’Herin, Brian. (2001) Abaza applicatives. Language. Vol. 77. No. 3. Pp. 477–493.

Pylkkänen, Liina. (2002). Introducing Arguments. Doctoral Dissertation, MIT.

The Categorical ‘Boundaries’ of the -wa/-o Converb in Abkhaz

Sarah M. Dopierala

Goethe Universität Frankfurt

Abkhaz, Converb, Typological Categories, Linguistic Description

When does a converb stop being a converb and start being something else? One way to address this question is via a case study that examines the (arbitrary) boundaries of the typological category “converb” (see Ross 2021) and the limitations of its descriptive use in a language-specific context. For instance, in Abkhaz (Northwest Caucasian), one of the verb forms that has traditionally been ascribed to the category of “converb” (Rus. *деепричастие*) is the form ending in -wa/-o (*-ya/-o*) which is also morphologically identical to the (dynamic) nonfinite present TAM form of the verb (see, for instance, Aristava 1960; Hewitt 1989; Yanagisawa 2012). The flexibility of the form’s usage – especially when influenced by additional morphological elements such as suffixes – calls into question whether its identity as a “converb” is stable or situational. For instance, consider the following examples which all contain this -wa/-o form in some capacity. Note that only Example 1 is a usage associated with the “converb” category:

1. -wa/-o in an adverbial clause

Атцарақәа цырҷыруа ашәа рҳәоит			
A-ҫara-k°a	ҷ′ərҷ′ər-wa	a-š°a	r-ḣ°-oiṭ
DET-bird-PL	Chirp-PRES.NFIN	DET-song	3PL-say-DYN.FIN
“Chirping, the birds are singing.” (Yanagisawa 2012: 376)			

2. -wa/-o with prefix in a relative (adjectival) clause

Итцәуаа аԥҳәыс дыурҭынчыртә еиԥш убас улацәажәароуп.					
уə-ҫ°wa-wa	a-pḣ°əs	də-w-r-tənč′ə-r-ṭ°	eipš	wəbas	wə-l-a-c°ž°a-r-owṗ
REL-cry-PRES.NFIN	DET-woman	3SG.HUM-2SG.MASC-CAUS-be calm-PURP	like	so	2SG.MASC-3SG.FEM-to-speak-COND-STAT.FIN.PRES
“You should talk to a woman in tears [crying] in such a way as to calm her.” (Hewitt 1987: 53)					

3. -wa/-o with question morphology acting as the main verb of an interrogative sentence

Уара агазеҭ уаԥхьома?		
Wara	agazet	w-apx′-o-ma
2SG.MASC	DET-newpaper	2SG.MASC-read-PRES.NFIN-QUES
“Are you reading the paper?” (Aristava et al. 1968: 120)		

In this small sample set, Example 1 would be the “undisputed” converbal usage of this form in that it is a nonfinite verb form heading an adverbial clause; Example 2 would not be considered a converb as it is modifying a noun (and also possibly because of its special person morphology); and Example 3 would not be considered a converb a.) because of its complex suffixal morphology *-o-ма* and b.) because it is, essentially, acting as finite verb in this instance as it is the only verb in a complete, independent sentence. These examples reveal that while a typological boundary might classify

Example 1 as a converb and, therefore, distinct from the other uses, since this form is clearly identifiable in all three instances as being “versions” of the dynamic nonfinite present, this raises the question: Do the boundaries drawn by the “converb” category meaningfully reflect any natural, language-specific categorizations within Abkhaz itself?

In this case study, I will analyze the morphological and functional contexts in which the -wa/-o form in Abkhaz aligns with or diverges from the typological category ‘converb.’ Specifically, I will investigate whether boundaries arise through certain morphological markers, like special person marking on participles, or through functional shifts where nonfinite verbs take on roles more typical of finite verbs, as seen in interrogative sentences. I will apply these potential language-specific boundaries to other "converb" forms to examine whether such boundaries reveal any clearly observable groups. Finally, I will compare these boundaries with typological expectations to assess whether the typological "converb" category is able to meaningfully identify any potential categorical distinctions in Abkhaz.

References:

1. Аристава, Шота Константинович [Aristava, Shota Konstantinovich]. 1960. Деепричастие в абхазском языке [Converbs in the Abkhaz Language]. Сухуми: Госиздат Абхазии.
2. Аристава, Ш.К., Х. С. Бгажба, М. Циколиа, Л. П. Чкадуа & К. С. Шакрыл [Aristava, Sh.K., X. S. Bgazhba, M. Tsikolia, L. P. Chadua & K. S. Shakryl]. 1968. Грамматика абхазского языка: Фонетика и морфология [Grammar of the Abkhaz Language: Phonetics and Morphology]. Сухуми: Алашара.
3. Hewitt, Brian George. 1987. The Typology of Subordination in Georgian and Abkhaz. Berlin: Walter de Gruyter.
4. Hewitt, Brian George. 1989. Abkhaz. (Croom Helm Descriptive Grammars Series). Routledge.
5. Ross, Daniel. 2021. Pseudocoordination, Serial Verb Constructions and Multi-Verb Predicates: The relationship between form and structure. PhD Dissertation.
6. Yanagisawa, Tamio. 2012. A Grammar of Abkhaz. Hituzi Syobo Publishing.

CROSS-GENERATIONAL ANALYSIS OF LAZ: IMPLICATIONS FOR ENDANGERMENT AND EDUCATION

Ömer Eren, University of Chicago

Goal: The endangered status of Laz, a South Caucasian language spoken in Türkiye (Kutscher 2008, La Croix 2009, Haznedar et al. 2018), remains largely unexamined with systematic linguistic data from younger-generation heritage speakers. To address this gap and identify vulnerable areas in Laz grammar that could inform heritage language education, I conducted a free production task using *The Frog Story* (Mayer 1969) with 73 Laz speakers across age groups, with a particular focus on younger heritage speakers. Statistical analyses of grammatical and lexical variables between heritage and baseline speakers reveal a notable decline in the use of distinctive Laz content words, spatial prefixes, finite subordinate clauses established with the bound complementizer *na-*, and valency-changing operations among heritage speakers. These results underscore the vulnerable areas in Laz grammar and align with patterns observed in other heritage language speakers (Montrul 2016, Polinsky 2018). Additionally, the findings show that the number of distinct Laz words serves as a more accurate measure of linguistic proficiency than speech rate, with a strong correlation between vocabulary inventory and speaker age. This pattern highlights low intergenerational transmission, reinforcing the endangered status of Laz. Such a metric could be effectively applied in governmental heritage Laz programs to assess and classify proficiency levels, enabling tailored educational approaches that optimize learning outcomes. Finally, deviations in forms produced by heritage speakers indicate that, despite reduced input due to language shift to Turkish, Heritage Laz grammar remains systematic because heritage speakers produce similar types of deviant forms despite their heterogeneous linguistic background. The vulnerabilities of Laz grammatical domains appear to follow this hierarchy: *Verbal morphology* > *Case morphology* > *Lexicon* > *Word order*. These insights underscore the need to focus on these specific areas in heritage language education to improve pedagogical effectiveness.

Divergences between Heritage and Baseline Laz: Based on cross-linguistic heritage language research (Montrul 2016, Polinsky 2018), 73 speakers were divided into two groups by age of bilingualism onset: those who were monolingual in Laz until age 7, when they began learning Turkish in school, i.e., baseline group, and those classified as heritage speakers (following Laleko 2010), who were exposed to both Turkish and Laz from birth. Using narrative data from the wordless picture story *Frog, Where Are You?* (Mayer 1969), the speech samples were transcribed and annotated by native Laz speakers. Frequency counts were then conducted for grammatical markers and constructions expected to be vulnerable to change or loss, informed by heritage language research (Polinsky 2018). Statistical analyses then identified key areas of divergence between the two groups, shedding light on the vulnerable aspects of Laz grammar. Results (see Figure 1) show that heritage speakers used significantly fewer of the following features than baseline speakers: distinct Laz content words ($t(71) = 5.85$, $p < 0.001$), valency alternations ($t(71) = 5.28$), spatial prefixes ($t(71) = 6.12$, $p < 0.001$), and complex clauses formed with na- ($t(71) = 3.48$). These differences remained significant when adjusted for length, such as the rate of spatial prefixes and valency alternations per finite verb produced. As for code-mixing practices, the statistical analyses ($t(71) = -2.13$, $p = .036$) indicate a significantly higher frequency of code-mixing patterns among heritage speakers (M = 1.80 > 1.28). To control for the potential influence of speech length on code-mixing frequency, I also calculated a relative measure by dividing the total number of code-mixed utterances by the total number of sentences produced (rate of code-mixing). The statistical analysis of this measure ($U = 397$, $z = -2.918$, $p = .004$) revealed a statistically significant difference, with the heritage group (Mdn = .2425 > .1526) exhibiting a higher rate of code-mixing than baseline speakers. These findings align with heritage language trends, where speakers typically exhibit a reduced lexicon, prefer simpler

	Baseline Group (N=40)				Heritage Group (N=33)			
Variables	**M**	**ss**	**Min**	**Max**	**M**	**ss**	**Min**	**Max**
Laz content word	129.65	63.93	58	330	68.73	36.29	19	61
Valency change-total	19.58	11.91	4	52	8.70	7.80	0	33
SP-Different (out of 27)	14.38	3.821	7	20	8.76	4.00	2	16
NA- total	4.38	4.61	0	23	1.94	3.02	0	10
Rate of valency changing	.1688	.0522	.05	.27	.1192	.0819	0	.33
Rate of spatial prefixes	.76	.13	.31	1.05	.48	.17	.21	1.10
Rate of code-mixing	.04	.06	0	.40	.09	.33	0	.09
Code-mixing total	2.53	2.88	0	12	4.52	4.74	0	18

M: Median, *N*: Number of participants, *ss*= Standard Deviation, *Min*=Minimum, *Max*=Maximum

Figure 1: Descriptive statistics for grammatical variables in baseline and heritage group

clauses over complex structures, and struggle with morphology (Montrul 2016, Polinsky 2018), all of which might render and trigger code-switching as an avoidance strategy. The reduced use of prefixed elements in the Laz verbal complex—especially spatial prefixes and valency-changing markers—illustrates this pattern. Heritage speakers also showed a stronger tendency to use postpositional constructions over synthetic spatial prefixes, likely influenced by Turkish, reinforcing the shift toward increased analyticity seen in heritage languages (Polinsky 2018: 183).

Linguistic Proficiency, Sociolinguistic Factors, and Endangerment of Laz: In heritage language studies, assessing linguistic proficiency helps account for individual variation among speakers. I evaluate two measures for proficiency: (1) the number of distinct Laz words (Laz content words), counting each lexical word only once (excluding repetitions, inflected forms, and grammatical words like determiners and copulas), and (2) speech rate, or the average number of words per minute, calculated by dividing total words (including repetitions) by total minutes spent narrating the story. (Polinsky 2008, Daller et al. 2011, Anstatt 2017). A comparison of two potential measures of linguistic proficiency reveals that the use of Laz content words has stronger correlations with grammatical measures than speech rate. Specifically, correlations between Laz content words and other linguistic measures ranged from moderate (rate of spatial prefixes and rate of code-mixing) to high (NA-total) or very high (SP-Different and total valency change). In contrast, speech rate correlations were consistently weaker across variables, from low (rate of spatial prefixes) to moderate (SP-Different and NA-total). Additionally, significant correlations associated with Laz content words outnumbered those with speech rate: Laz content word usage correlated significantly with seven variables, while speech rate correlated with only five. Both measures showed a similar, low correlation with the sociolinguistic variable of village altitude, though Laz content word usage displayed a slightly stronger correlation ($r = .386$, $p < .01$) than speech rate ($r = .292$, $p < .05$). Laz content word frequency also exhibited stronger correlations not only with grammatical variables but with sociolinguistic factors like age and village altitude. Using Laz content word frequency as a primary measure of linguistic proficiency, I further demonstrate that proficiency declines with participant age, supporting the endangered status of Laz. Overall, the findings suggest that younger speakers and those from urbanized areas, such as towns rather than higher-altitude villages, tend to have lower levels of linguistic proficiency, which qualifies as a practical measure of linguistic proficiency in the absence of proficiency tests at educational institutions. The results also lend empirical support to observations in previous literature that Laz is not being transmitted effectively to younger generations—an argument previously based on self-reported proficiency (Haznedar et al. 2018) or fieldwork observations (Kutscher 2008) but not systematically investigated until now.

Deviant forms & Vulnerability in Laz Grammar: Informed by the results of the statistical analyses reported above, I identify the aspects of Laz grammar vulnerable to change or erosion by analyzing deviant forms (c.f. errors) produced by heritage speakers—forms not typically accepted grammatical by proficient Laz speakers (s). I show that although Heritage Laz grammar shows greater variation due to these deviant forms, it still retains systematicity: heritage speakers of different sociolinguistic backgrounds produce similar deviant forms in predictable ways. These forms are not random; they are consistently produced across different speakers and within individual speakers. To address the variability among heritage speakers, I classify them into three proficiency levels—low, mid, and high—based on lexical proficiency, measured by the number of distinct Laz content words each speaker produces. I find that deviant forms are more prevalent among lower-proficiency speakers, with statistically significant differences in their production of distinct spatial prefixes, complex clauses, and valency alternations compared to mid- and high-proficiency speakers. This analysis reveals the vulnerable areas in Laz grammar when acquired with limited input, highlighting verbal morphology (particularly verbal prefixes) and nominal morphology as the most affected. Heritage speakers often struggle with accurately using prefixes in the verbal template—such as affirmative particles, spatial prefixes, and pre-root vowels—and with temporal-aspectual suffixes (as exemplified in (1)), which show unique allomorphy in the baseline variety. For nominal morphology, they produce deviant forms with spatial, structural, and inherent case markers (e.g., dative, and ergative cases). These findings align with statistical analyses showing lower production of distinct content words, valency alternations, and spatial prefixes in heritage Laz, and mirror patterns in other heritage languages, where aspects like aspectual and case morphology are prone to overregularization or omission. I divided deviant forms into four categories: verbal morphology, case morphology, lexicon, and syntax. Verbal morphology emerges as the most frequently affected area, followed by case morphology (especially in case markers). Additionally, heritage speakers with limited lexical inventories occasionally use content words inappropriately within given contexts, the third most frequent type of deviation. Syntactic deviations, e.g., word order and pro-drop variations, are the least common. Lastly, I argue that the current educational programs in Laz need to be updated based on these findings.

(1) Cur tane kurbağa **(*ko)-**zir-**am**-an. (c.f. the baseline/correct form: zir-um-an)
two piece frog affirmative-find-imprf-3pl
Intended: 'they (=the boy and the dog) find two frogs.'

Question-Answer Coordinations as Answers in South Caucasian

David Erschler and Itai Bassi
Ben-Gurion University of the Negev

In Colloquial Georgian, Svan, and Megrelian, in an answer to a wh-question, the entire question, or, more frequently, only the wh-phrase from it, can be coordinated with the answer. This construction has not been addressed so far in the literature, and appears to be a typological *rarum*. We describe its properties and propose a preliminary analysis for it.

This construction is illustrated in (1) for all the 3 languages. While in Svan and Georgian, the regular co-ordinating conjunction *i/da* is used, in Megrelian, it is *da* 'if' rather than *do* 'and'. We will call this construction a Q-ANSWER.

(1) Upper Bal Svan

a. al di:no:l imʁa igwni?
this girl why cry.PRS.3SG
'Why is this girl crying?'
[(al di:no:l) imʁa (igwni)]=i [legd æri]
this girl why cry.PRS.3SG=& sick is
'She is sick.' (lit. Why is this girl crying and she is sick.)

b. jær-d lale:m dijær?
who.ERG eat.AOR.3SG[1] bread
'Who ate the bread?'
[jær-d (lale:m dijær)]=i [gela-d]
who-ERG eat.AOR.3SG bread=& Gela-ERG
'Gela.' (Lit. 'Who ate the bread and Gela.)

Georgian

c. ra iq'ide? [ra (viq'ide)]=da [p'uri (viq'ide)]
what buy.AOR.2SG? what buy.AOR.1SG=& bread buy.AOR.1SG
'What did you buy?' 'I bought bread.'

Megrelian

d. mu iidi? [mu (ip'idi)]=da [kobali (ip'idi)]
what buy.AOR.2SG? what buy.AOR.1SG=if bread buy.AOR.1SG
'What did you buy?' 'I bought bread.'

Any felicitous answer to a given wh-question may be integrated in a Q-answer.

(2) a. Neg-word, UB Svan
isgu ʦitsw-s iʃæʃd xaq'lu:ni? iʃæʃd=i de:ʃæʃd
your cat-DAT who.BEN fears who.BEN=& nobody.BEN
'Who does your cat fear?' 'No one.'

b. Universal quantifier, UB Svan
ime izge-x mugw-ær? ime=i ʧiæg
where live.PRS.3-PL pigeon-PL where=& everywhere
'Where do pigeons live? 'Everywhere.'

c. 'I don't know', Georgian, https://forum.ge/?f=32&showtopic=34933034&st=375
rat'om gak 8GB RAM-ze 32-iani sist'ema?
why you.have 8GB RAM-on 32-SUFF system?
'Why do you have a 32 bit system on an 8GB RAM?'
rat'om da [ar vitsi ʤma-o araperi]
why & NEG I.know brother-VOC nothing
'(Because) I don't know anything, man.'

A Q-answer is possible for a multiple wh-question.

(3) UB Svan
jær-d mæj lale:m?
who-ERG what eat.AOR.3SG

[1]Glosses: ALL allative; AOR aorist; BEN benefactive; MOD modal particle; PRV preverb; Q qestion particle; QT quotative particle; VOC vocative; XPV external preverb; & coordinating conjunction.

'Who ate what?'

jær-d	mæj=i	gela-d	dijær	(lale:m)
who-ERG	what=&	Gela-ERG	bread	eat.AOR.3SG

'Gela (ate) bread.'

The verb in the question part of a Q-answer agrees in the person w.r.t the speaker. In Georgian, the preverb will be different in the forms 'X gave Y to me/you' (*mo-*) and 'X gave Y to a 3rd person' (*mi-*). In Q-answer the respective switch occurs (4). This shows that the coordinated question is not an echo-question.

(4) Context: In Gia's presence, somebody asks somebody other than Gia about Gia. Gia answers the question instead of them.

Q:	gia-s	vin	**mi-s**-ʦ-a	dana?
	G-DAT	who(ERG)	PRV-3SG.IO-give-AOR.3SG	knife

'Who gave Gia a knife?'

Gia:	vin	**mo-m**-ʦ-a	dana=da	Gela-m
	who(ERG)	PRV-1SG.IO-give-AOR.3SG	knife=&	Gela-ERG

'Gela.'

In Svan, polar questions allow Q-answers as well (5). This is not the case for Georgian or Megrelian.

(5) UB Svan

gela	esʁri=ma	uʃgul-te?	esʁri=ma=i	adu/ma:ma/esʁri
G.	goes=Q	Ushguli-ALL	go=Q=&	yes/no/goes
'Is Gela going to Ushguli?'			'Yes/no/he's going (=yes).'	

Alternative questions do not allow Q-answers in any of the 3 languages, as illustrated in (6) for Georgian.

(6)

Q:	ʧai	ginda	tu	q'ava?		
	tea	you.want	or.Q	coffee		
	'Would you like tea or coffee?'					
A:	#ʧai	minda	tu	q'ava	da	ʧai
	tea	I.want	or.Q	coffee	&	tea
	'Tea' (intended)					

The construction raises several analytical puzzles.

- Q-answers appear to involve coordination of an interrogative with a declarative.
- Different types of question, viz., wh-questions, altertnative questions, and polar questions behave differently with respect to Q-answers.

To address the first puzzle, we invoke the Performative Hypothesis (Levinson 1983). This hypothesis effectively says that the meaning of an utterance *What did you buy?* is the proposition *I ask you what you bought*, while the meaning of the answer *Bread* is *I tell you that I bought bread*. We hypothesize that South Caucasian languages possess a means to morphosyntactically convert questions and answers into such declaratives, and it is these "derived" declaratives that are coordinated in Q-answers. Furthermore, as Megrelian facts explicitly indicate (1d), Q-answer is actually a conditional. While Svan and Georgian use a coordinating conjunction in Q-answers, the use of coordinators in conditionals is well attested cross-linguistically, see e.g. (Haiman 1983; Ross 2021). Accordingly, we propose that the meaning of a Q-answer can be informally represented as *If you are asking me X, I am answering you Y*.

As for the the 2nd puzzle, we tentatively propose that the ban on Q-answers to alternative questions, and to polar questions in Georgian and Megrelian, stems from prosodic restrictions on coordination rather than from syntactic or semantic constraints.

References

Haiman, John. 1983. Paratactic *if*-clauses. *Journal of Pragmatics* 7(3). 263–281.

Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Ross, Daniel. 2021. Pseudocoordination, Serial Verb Constructions and Multi-Verb Predicates: The Relationship between Form and Structure. University of Illinois Urbana-Champaign Dissertation.

West Circassian bare nouns are full DPs: Evidence from nominal possession

Ksenia Ershova[1] and Nikita Bezrukov[2]
[1]University College London, [2]University of Pennsylvania

The literature on bare nouns and pseudo-noun incorporation (PNI, Massam 2001; Öztürk 2005; 2009) offers a compelling size-based explanation for the non-(co-)occurrence of number and case in bare nouns: they are structurally smaller (nPs) compared to full DPs, lacking both NumP and DP. West Circassian (WC) represents a highly ambiguous case of theoretical importance, where bare nouns exhibit an unusually broad distribution for PNIed nominals.

Recently, Arkadiev and Testelets (2019; hereafter AT2019) argue that WC bare nouns generally align with expectations for PNIed nominals and analyze them as nPs based on their indefinite interpretation, low scope, and number neutrality. However, WC nominals show a list of properties atypical for PNIed elements, such as free word order, full agreement, and the absence of interaction with case marking on other nominals (i.e., no detransitivizing effects).

Bagirokova et al. (2022) characterize bare nouns as an instance of general number (Corbett 2000). We argue that this is represented syntactically: ***nominal arguments are uniformly DPs in WC, but NumP can be omitted, resulting in number neutrality*** (Wiltschko 2008, Kramer 2017). The bare noun pattern results from allomorphy on D. Support for this approach comes from interactions between overt case marking and number in possessed nominals: the absence of NumP leads to number-neutral interpretations and allomorphy on D triggered by Poss. When NumP is present, this allomorphy is disrupted and no number neutrality is observed.

This study breaks down the correlation between overt exponence and syntactic structure. In WC, the omission of number morphology correlates with the absence of NumP. However, the non-exponence of case morphology does not indicate the absence of DP. Correspondingly, this challenges the utility of number neutrality as a diagnostic for DP-level structure.

Bare nominals in West Circassian. In WC, case may be omitted in all syntactic positions, resulting in an indefinite/nonspecific interpretation (1).

(1) a. ʔaze-deʁʷə**-m** w-jə-ʁe-χʷəž'ə-š't
doctor-good**-ERG** 2SG.ABS-3SG.ERG-CAUS-recover-FUT
'The good doctor will cure you.' (AT2019:726)

b. **ʔaze-deʁʷ** jeʁaṣ̂e-m Ø-**j**-e-ʁe-χʷəž'ə
doctor-good always-OBL 3ABS-**3SG.ERG**-CAUS-recover
'S/he is always treated by good doctors. (Bagirokova et al. 2021:288)'

Bare nominals are number neutral and may refer either to a plural, or a singular individual (2) (see also Bagirokova et al. 2022). Overt number morphology must be accompanied with overt case marking (3), which AT2019 connect to the absence of NumP in unmarked nominals, and its obligatory presence in full DPs.

(2) stolə-m txəλ Ø-tje-λ
table-OBL book 3ABS-LOC-lie
'There is a book on the table / there are books on the table.' (AT2019:731)

(3) č̣'ale-xe-r, č̣'ale-xe-m vs. *č̣'ale-xe
boy-PL-ABS boy-PL-OBL boy-PL (*ibid.*)

Based on these properties (and low scope), AT2019 analyze bare nominals as lacking a DP layer, analogous to pseudo-incorporation of arguments in e.g. Turkish (Öztürk 2005).

Wide distribution of bare nominals in West Circassian. Unlike PNI, unmarked nominals need not be adjacent to the predicate, trigger normal ϕ-agreement, and are not limited to internal arguments, as can be see for the ergative agent in (1b). They also do not affect case assignment possibilities (no detransitivization): e.g. the agent in (4) bears ergative case and the bare nominal is not verb-adjacent. Also, while unmarked nominals frequently correlate with

low scope (as argued by AT2019), this is not universally so (4).

(4) **txəλ** ç̣'el-jə-š'**-me** Ø-q-a-hə-ʁ
book boy-LNK-three**-PL.ERG** 3ABS-DIR-3ERG.PL-bring-PST
'The three boys brought a book.' (all three boys are sharing the same book; INDEF > 3)

Based on similar observations, AT2019 propose that bare nominals, despite lacking the DP layer, may appear in all positions associated with DPs, are assigned case and control agreement.

The analogy with PNI is further challenged by the morphosyntactic behavior of possessed nominals, which, in the absence of overt number morphology, are incompatible with case marking (5) and display number neutrality (6; Bagirokova et al. 2022).

(5) **sjə-**nəbʒeʁʷə**(*-m)**
1SG.POSS-friend(*-OBL)
'my friend' (Rogava and Keraševa 1966:70)

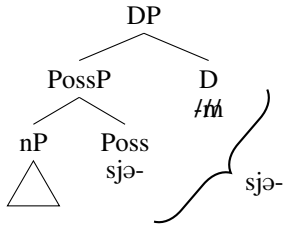
(6) [mə bzəλfəʁe-m **jə-ha**] Ø-z-ʁe-šxe-n-ew
this woman-OBL **3SG.POSS-dog** 3ABS-1SG.ERG-CAUS-eat-MOD-ADV
Ø-je-z-ʁe-ž'a-ʁ
3ABS-DAT-1SG.ERG-CAUS-begin-PST
'I began feeding this woman's dog / dogs.'

Full DPs with null D. We argue that this cluster of properties is best accounted for by abandoning the nP/DP distinction claim. Instead, similarly to Kramer's (2017) analysis of Amharic, ***full DPs may lack NumP***, resulting in number neutrality. The indefinite, number neutral interpretation of bare nominals results from the absence of NumP, combined with an indefinite D, which is spelled out as a null morpheme. Overt case suffixes correspondingly expone definite D, combined with the case feature assigned to the full DP (e.g. ERG in 1a and 4).

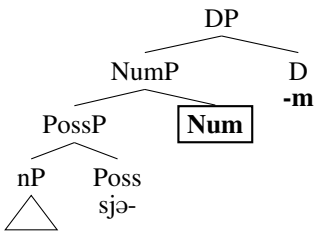
Evidence for this approach comes from the interaction between case exponence and number marking in possessed nominals (5-6): a definite determiner undergoes fusion with the adjacent Poss head, resulting in the absence of overt case morphology: [Poss]-[D] → [Poss,D] (7).

If NumP is present, this fusion is disrupted by the intervening Num head (8): thus, D is spelled out as an overt case suffix in the presence of a plural suffix (9) or numerals (10).

(7)



(8)



(9) **sjə-**nəbʒeʁʷə**-xe-m**
1SG.POSS-friend-PL-OBL
'my friends' (adyghe.web-corpora.net)

(10) jə-**zə-**šolk-ʒene-daxe**-r**
POSS-**one-**silk-dress-beautiful**-ABS**
'one beautiful silk dress of hers'
(Lander 2017:84)

Extension: Other DPs without case. The DP analysis of bare nouns is further supported by personal pronouns and proper names, which are incompatible with overt case (11). Similarly to possessed nominals, D is structurally present, but unpronounced.

(11) **te** **məjeqʷape** mə bzəλfəʁe-m jə-mašjəne-ç̣'e tə-qe-ḳʷa-ʁ
we(ABS) **Maykop(OBL)** this woman-OBL 3SG.POSS-car-INS 1SG.ABS-DIR-go-PST
'We went to Maykop in this woman's car.'

Select references. • Arkadiev & Testelets 2019. *Studies in Language*. • Bagirokova, Lander & Phelan 2022. In *Number in the World's Languages*. • Öztürk 2009. *Lingua*.

Nominal tense in Kartvelian? Some observations on the 'previous-state nouns'

Natalia Evseeva
University of the Basque Country (UPV/EHU)

Keywords: nominal tense, inflection/derivation, temporality, TAM-semantics, Kartvelian

Tense-aspect-mood (TAM) is usually considered a verbal grammatical category; however, in a number of languages it can be encoded in the nominal domain. This phenomenon is known as 'nominal tense' (cf. Nordlinger & Sadler 2004, Tonhauser 2006, Bertinetto 2020, Aikhenvald 2021) and has been described mostly on the basis of the languages of the Americas, whereas few studies address nominal temporal marking in Eurasian languages (Nikolaeva 2015, Gusev 2021). From the very beginning of the research on this topic the difficulty has been highlighted to distinguish inflectional nominal tense markers from derivational affixes (Nordlinger & Sadler 2004: 780). Tense markers are expected to be fully productive and able to attach to any noun regardless of its semantics; in turn, derivational affixes are usually restricted to one or several semantic fields.

Accordingly, a contrast is normally drawn between, on the one hand, such affixes as Tariana *-miki-*, which can be used with a noun like 'eagle' to mean 'the remains of the eagle', 'what used to be the eagle' or Guaraní *-kue*, used with a noun like 'house' to mean 'old house', and, on the other, English *ex-* (?*ex-eagle*, ?*ex-house*). However, it has been shown that nominal temporality markers are also subject to restrictions, which are mostly pragmatically-based (Tohnauser 2006, Bertinetto 2020). Therefore, the boundary between inflection and derivation among nominal tense markers is not as clearly set as assumed in many studies on the topic.

To our knowledge, Kartvelian languages have not been examined through the lens of studies on nominal tense. While all four languages of the family present nominal markers with past temporal semantics (such as the Georgian circumfixes *na- ... -ar-/-al-*, *na- ... -ev-*), they are usually described as derivational affixes that form 'previous-state nouns' (*c'ina vitarebis saxelebi*, Šaniże 1980, Vašakiże 1987). This talk aims to open a discussion on whether one can speak about nominal tense in Kartvelian, as well as define the meanings of the above mentioned morphemes. This ongoing study is mostly based on Georgian material and, to a lesser extent, on Megrelian, Laz, and Svan. The data are drawn from dictionaries (cf. Rayfield 2006, Čikobava 1950-1964, Kajaia 2006), corpora (GDC, GNC, TITUS), as well through elicitation from native speakers.

An analysis of over 600 'previous-state nouns' reveals a high degree of productivity of this morphological device in Kartvelian, as suggested by the fact that the relevant affixes attach to nouns belonging to a wide range of semantic classes. Moreover, the restrictions of compatibility with some types of nominal stems align with those described by Tonhauser (2006) for Guaraní, which supports Bertinetto's (2020) claim that they are pragmatically motivated.

Similarly to other languages with reported nominal tense markers, Kartvelian languages display uses of these affixes with nouns denoting professions and stage-level relations (Georgian *naminist'rali* 'former minister', *nakmrevi* 'former husband'). In turn, no 'previous-state nouns' that convey the meaning of a deceased human or animal have been found. The most common group of nouns with temporality markers refer to a space previously occupied by some entity: Georgian *nasimindari* 'field where maize grew last year', Svan *nalaqwam* 'site of former church'. This is particularly relevant considering that this kind of examples have received little attention in previous studies, cf. Tonhauser (2006: 197-198). Equally remarkable is a subgroup of nouns that share a meaning of a mark left by a brief presence of some entity: Georgian *napexuri* 'footprint', Megrelian *nak'ibira* 'tooth mark'. Finally, it is worth mentioning that in Megrelian a combination of

temporality markers and diminutives results in a deprecatory meaning (*nobergia* 'worthless hoe', *nok'abia* 'old, worthless dress'), cf. Aikhenvald (2021) on Tariana.

Acknowledgments

This research has been supported by the research project *La transversalidad de las categorías gramaticales: un estudio tipológico* PID2021-124769NB-I00, funded by the Spanish Ministry of Science and Innovation / Spanish Research Agency (AEI) /10.13039/501100011033, and the European Regional Development Fund (ERDF "A way of making Europe").

References

Aikhenvald, Aleksandra Yu. 2021. Beyond nominal tense. Temporality, aspect and relevance in Tariana noun phrases. *Studies in Language* 46 (1). 40-75.

Bertinetto, Pier Marco. 2020. On nominal tense. *Linguistic Typology* 24 (2). 311-352.

Čikobava, Arnold, et al. (eds). 1950-64. *kartuli enis ganmart'ebiti leksik'oni* [The explanatory dictionary of the Georgian language]. Tbilisi: SSSR mecnierebata ak'ademiis gamomcemloba.

Gusev V. Yu. 2021. K tipologičeskoj xarakteristike severnosibirskogo substrata [Towards a typological profile of the North Siberian substrate]. *Voprosy Jazykoznanija* 2021 (5). 26–58.

Kajaia, Otar. 2006. *megrul-kartuli leksik'oni* [Megrelian-Georgian dictionary]. Tbilisi: nek'eri.

Nikolaeva, Irina. 2015. On the expression of TAM on nouns: evidence from Tundra Nenets. *Lingua* 166. 99-126.

Nordlinger, Rachel, and Louisa Sadler. 2004. Nominal tense in crosslinguistic perspective. *Language* 80 (4). 776-806.

Rayfield, Donald. 2006. A comprehensive Georgian-English dictionary. London: Garnett Press.

Šaniże, Ak'ak'i. 1980. *kartuli enis gramat'ik'is sapużvlebi* [The foundations of Georgian grammar]. Tbilisi: tbilisis universit'et'is gamomcemloba.

Tonhauser, Judith. 2006. *The Temporal Semantics of Noun Phrases: Evidence from Guaraní*. PhD dissertation. Stanford University.

Vašakiże, Tamar. 1987. c'inavitarebis saxelta c'armoebisatvis kartvelur enebši [Towards the formation of previous-state nouns in Kartvelian languages]. *macne, enisa da lit'erat'uris seria* 1987 (2). 174-178.

GDC. Georgian Dialect Corpus. http://www.corpora.co/#/corpus

GNC. Georgian National Corpus. http://gnc.gov.ge/gnc/

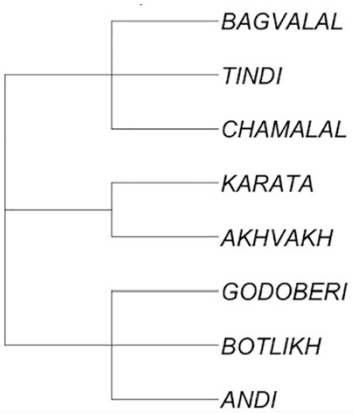
TITUS. http://titus.uni-frankfurt.de

On linkage effects in Andic: a case of Anchiq and its surroundings

Konstantin Filatov
HSE University / ILS RAS

From a comparative-historical perspective, all genealogically related groups of languages can be described either in terms of the tree model (Stammbaumtheorie or cladistic approach) or using the wave model (Wellentheorie). The cladistic approach is more applicable to situations where rapid population separation without much subsequent contact took place. Divergence in this scenario occurs through the separate accumulation of innovations, and separation events are depicted by nodes on a tree. The wave model, on the other hand, yields better results for genealogies where the loss of mutual intelligibility of dialects was gradual and varieties were able to contact without an intermediary language for a long period of time. The second situation, well-known to dialectologists, involves the gradual spread of innovations between large networks of mutually intelligible idioms, resulting in clusters of intersecting isoglosses that are difficult to assign to discrete separation events (François 2014, François & Kalyan 2024). When a dialect continuum loses its mutual intelligibility, it becomes a linkage (Ross 1988).

Among the Nakh-Daghestanian languages the best example of linkage is represented by the Dargwic languages, which are often considered a dialect continuum, despite having diverged approximately 2000 years ago (Sumbatova 2020).



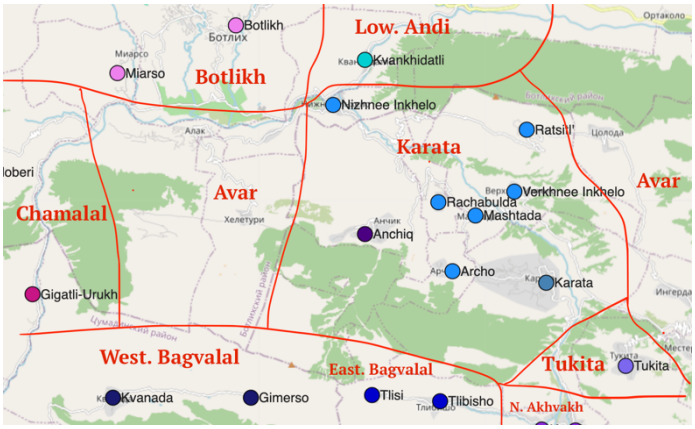


Fig. 1 (left). Alekseev's (1988) classification of Andic. Fig. 2 (right) Anchiq and surrounding idioms

A much less commonly discussed group within the Nakh-Daghestanian family that is believed to exhibit linkage properties is the Andic languages. The most influential classification of Andic by M. Alekseev (1988) groups together Bagvalal, Tindi, Chamalal vs. Karata, Akhvakh vs. Godoberi, Botlikh, Andi (Fig. 1). However, the linkage effects in Andic were noticed already by T. Gudava (1967: 274): "The Andic languages as a whole form a continuous chain of languages and dialects, transitioning rather smoothly into each other." The important property of a linkage is overlapping subgroups, where an idiom can belong to more than one subgroup (François 2014). Such examples of "mixed subgroupiness" are also noted by Gudava (1967: 274) "For example, the Godoberi language resembles a connecting link between the Botlikh and Chamalal languages, having common features with the Gigatli Chamalal, on the one hand, and with the Miarso Botlikh, on the other." Gudava's observation implies the presence of common isoglosses between Godoberi and Gigatli Chamalal that cross-cut Alekseev's subgrouping.

A similar claim was made by Z. Magomedbekova (1971) about the Anchiq dialect of the Karata language: "Some features of Anchiq resemble those of Botlikh; this dialect serves as a connection link between Karata and Botlikh on the one hand, and between Karata and Bagvalal on the other." This talk aims to verify Magomedbekova's claim by investigating common isoglosses between Anchiq and its linguistic surroundings: Botlikh, Bagvalal and other Karata dialects. To demonstrate that this subset of Andic idioms exhibits linkage properties we primarily rely on methodology proposed by François (2014), which involves determining intersecting isoglosses representing common innovations between these languages. We examine (i) isoglosses in regular and lexeme-restricted (irregular) phonetic correspondences, (ii) the distribution of cognate roots in basic and non-basic lexicon, and, finally, (iii) isoglosses in morphological structures.

Some preliminary findings seem to support Magomedbekova's hypothesis. Anchiq and Bagvalal have certain peculiar common isoglosses in different linguistic domains: rhotacism (i. e. phonetic change **d* > *r*, cf. forms of coupla Anchiq *gira*, Tlondoda Bagvalal *ira*, vs. Karata *idʲa*, Botlikh *ida*), direct stem of logophoric

pronoun *ẽ-* (Anchiq, Bagvalal), as opposed to *ži-* in Karata and Botlikh. The most interesting candidate for a common innovation between Anchiq and Bagvalal is the unique system of hearer's sex marking on question particles and interrogative pronouns and the question particles themselves: cf. Kvanada Bagvalal *=išto* (=Q.HEAR.M), *=išta* (=Q.HEAR.F), which regularly corresponds to Anchiq *=sːo* (=Q.HEAR.M), *=sːa* (=Q.HEAR.F). This system is absent in all other Andic varieties except Chamalal, where phonetically different markers are employed.

Between Anchiq and Botlikh there are several isoglosses that are absent in Karata and Bagvalal: e.g. the root for 'big' Anchiq *hič'**uχ***^{w}-, Botlikh *-eč'**uχ***^{w} (cf. Karata *herk'-,* Bagvalal *hĩč'-*), 'to eat' Anchiq *am-*, Botlikh *ʁam-* < **q'am-* (cf. Karata *q'amd-,* Bagvalal *q'an-* < **q'ːam-*) or the prefixal nasal shift (i. e. the change of gender markers *b-, r-* into *m-, n-* under nasalization) as well as some minor lexically restricted irregular sound changes.

The talk will present the full range of data, provide numerical values that will give preliminary estimates of subgroupiness between varieties, and demonstrate that these (and some other) changes are innovative. We will also analyze some common retentions between varieties discussed, as far as the non-introduction of some features may also signal contact in the past (van Gijn & Wahlström 2023: 196). For example the root of the lexical item for 'black' is *beč'at'ir- / beč'et'ir-* in all Karata varieties, while in Anchiq, Kvanada and Tlibisho Bagvalal it is *beč'at'- / beč'et'-*. It is evident that *beč'at'ir-* is a suffixed form of *beč'at'-* , but this innovation was not introduced into the Anchiq variety as a supposed result of Anchiq–Bagvalal contact.

Finally, we will summarize some known evidence for a past contact situation. Speakers of Anchiq and Botlikh have long-standing cultural relations. Despite being politically subordinate to Karata, Anchiq has always been (and still is) more economically dependent on the lowland Botlikh market in accordance with Nichols' model of vertical spread (Nichols 2015). The case of possible Anchiq–Bagvalal contact is much more obscure. Anchiq and three Bagvalal villages (Tlissi, Gimerso and Kvanada) share a common land border; however, this border runs along the high-altitude Alak ridge and no direct road is known between these settlements. The roundabout road is long and passes through Akhvakh- or Chamalal-speaking territories.

References

Alekseev, M. E. 1988. Sravnitel′no-istoričeskaja morfologija avaro-andijskix jazykov [Historical-comparative morphology of Avar-Andic languages]. Moscow: Nauka.

François, A. & S. Kalyan. 2024. Subgrouping: Trees vs. waves. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge Handbook of Historical Linguistics*. 2nd edn. New York: Routledge.

François, A. 2014. Trees, waves and linkages: Models of language diversification. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge Handbook of Historical Linguistics*. 1st edn. New York: Routledge.

Gijn, Rik van & Max Wahlström. 2023. Linguistic areas. In Rik van Gijn, Hanna Ruch, Max Wahlström & Anja Hasse (eds.), *Language contact: Bridging the gap between in- dividual interactions and areal patterns*, 179–219. Berlin: Language Science Press.

Gudava, T. E. 1967. Andijskie jazyki [Andic languages]. In E. A. Bokarëv, K. V. Lomtatidze (eds.), Jazyki narodov SSSR. Iberijsko-kavkazskie jazyki [Languages of the peoples of the USSR. Ibero-Caucasian languages], 272–275. M.: Nauka.

Magomedbekova, Z. M. 1971. Karatinskij jazyk: grammatičeskij analiz, teksty, slovar′ [The Karata language: grammar, texts, dictionary]. Tbilisi: Mecnierebi.

Nichols, Johanna. 2015. Types of spread zones: Open and closed, horizontal and vertical. In Rik De Busser & Randy J. LaPolla (eds.), *Cognitive Linguistic Studies in Cultural Contexts*, vol. 6, 261–286. Amsterdam: John Benjamins Publishing Company.

Ross, Malcolm. 1988. *Proto-Oceanic and the Austronesian languages of Western Melanesia* (Pacific Linguistics). Canberra: Australian National University.

Sumbatova, N. 2020. Dargwa. In Maria Polinsky (ed.), *The Oxford handbook of languages of the Caucasus* (Oxford Handbooks), 147–200. Oxford, United Kingdom ; New York, NY: Oxford University Press.

Grammaticalization of body part terms in the languages of the Caucasus

Diana Forker

Body part terms are frequently sources for grammaticalization processes (e.g., Heine & Kuteva 2002, Brenzinger & Kraska-Szlenk 2014, Zariquiey & Valenzuela 2022). Heine (2014: 17) notices three domains of grammar in which body part terms are particularly frequent: (i) spatial orientation, (ii) reference identity (i.e. reflexivity), and (iii) counting.

In my presentation, I will examine the grammaticalization of body part terms in all three indigenous language families of the Caucasus and explore which patterns are common for which family or subgroup. Numerous Nakh-Daghestanian and Northwest Caucasian languages have (mostly spatial) postpositions, adverbs and preverbs for which body part terms are the source (Arkadiev & Maisak 2018, Nasledskova 2021). This pattern seems to be absent in Kartvelian languages, which, in contrast, have reflexive pronouns that originate from the words for 'head' (e.g. Amiridze 2006).

With respect to semantics, I will explore which body part terms are sources for which topological relations and directional meanings and plot the patterns on a semantic map in order to investigate implicational relations between the meanings. I will show that in addition to spatial orientation and reflexivity, also instances of temporal meaning and even causal meaning are attested among the postpositions and adverbials that have developed from body part terms. Concerning morphosyntax I will analyze the morphological and syntactic properties of the constructions (e.g. case marking) and explore their use in three case studies by means of corpora for the standard languages Adyghe, Avar and possibly Lezgian. Finally, I will come up with a systematic categorization of the respective constructions along a grammaticalization cline, explain the grammaticalization paths by means of metaphorical extension and metonymy (Rubba 1994, Matsumuto 1999) and *embodiment* (e.g. Brenzinger & Kraska-Szlenk 2014) and review the results from a cross-linguistic perspective.

Data for this presentation comes from grammars, dictionaries and field work.

Examples of grammaticalization processes

(1) body part > postposition/adverb/preverb with spatial and temporal meaning in Nakh-Daghestanian and Northwest Caucasian languages

- **Lezgian** (Nakh-Daghestanian, Haspelmath 1993: 206, 208, 215-216): *wil(-i)* 'eye' > *wilik* 'in front; before', *wilikdi* 'forward, to the front', *wilikaj* 'from the front', *wilík(daj)* 'before', *wilíkamaz* 'beforehand', etc.,
- **Hinuq** (Nakh-Daghestanian, Forker 2013: 387-388): *ƛ'ʷoq'o* 'forehead' > *ƛ'oq'ar* 'in front'
- **Abaza** (Northwest Caucasian, Arkadiev 2020): *wacạ* 'intestines' > *-wacạ* 'inside'; *cạqʷa* 'tail' > *-cạqʷa* 'after (that), then'; *š'apʲə́* 'foot' > *-š'apʲə́* 'under',
- **Adyghe** (Northwest Caucasian, Rogava & Kerasheva 1966: 93-94): *ʔʷə* 'mouth, aperture' > *ʔʷə-* 'front part, in front, near, beside', *ḳʷeҫ* 'internal organs, interior' > *ḳʷeҫ* 'inside', *ҫ'əb* 'back' > *ҫ'əb* 'behind'

(2) body part > reflexive: 'head' > reflexive in all four Kartvelian languages

- **Georgian** *tav-i*, **Megrelian** and **Laz** *dud-*, **Svan** *txwim-* (Testelets 2021: 504, Amiridze 2006)

References

Amiridze, Nino. 2006. *Reflexivization strategies in Georgian* (Unpublished doctoral dissertation). University of Utrecht.

Arkadiev, Peter M. 2020. *Abaza sketch grammar*. Moscow, RGGU.

Arkadiev, Peter M., and Timur Maisak. 2018. Grammaticalization in the North Caucasian languages. In Heiko Narogg & Bernd Heine (eds.) *Grammaticalization from a typological perspective*, 116-145. Oxford: Oxford University Press.

Brenzinger, Matthias, and Iwona Kraska-Szlenk (eds.) 2014. The body in language: Comparative studies of linguistic embodiment. Leiden: Brill

Forker, Diana. 2013. *A grammar of Hinuq*. Berlin: De Gruyter.

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

Heine, Bernd. 2014. The body in language: Observations from grammaticalization. In Brenzinger, Matthias, and Iwona Kraska-Szlenk (eds.) *The body in language: Comparative studies of linguistic embodiment*, 11-32. Leiden: Brill.

Heine, Bernd and Tania Kuteva 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Matsumoto, Yo. 1999. The Extension of Body-part Nouns to Object-part Nouns and Adpositions: Shape, Function, and Location in Lexical-semantic Change and in Grammaticalization. Manuscript. Meiji Gakuin University

Nasledskova, P.L. 2021. Istočniki grammatikalizacii otymenniyx poslelogov v naxsko-dagestanskix jazykax. *Acta Linguistica Petropolitana*. 2021. Vol. 17.1. P. 215–252. DOI 10.30842/alp23065737171215252

Rogava Georgij V., and Zajnab I. Kerasheva. 1966. *Grammatika adygejskogo jazyka* [A Grammar of West Circassian]. Krasnodar: Krasnodarskoe knižnoe izdatel'stvo.

Rubba, Jo. 1994. Grammaticalization as semantic change. In William Pagliuca (ed.) *Perspectives on Grammaticalization*, 81-101. Amsterdam: John Benjamins.

Testelets, Yakov G. 2021. Kartvelian (South Caucasian) languages. In Maria Polinsky (ed.), *The Oxford Handbook of Languages of the Caucasus*, 491–528. Oxford: Oxford University Press.

Zariquiey, Roberto and Pilar M. Valenzuela (eds.) 2022. The Grammar of Body-Part Expressions: A View from the Americas. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198852476.001.0001

Syntactic agreement in Chechen and Ingush is semantic in Tsova-Tush (Nakh/East-C)

Caleb Gordon (U. Victoria), Theis Greve (U. Edinburgh),
Jérémy Pasquereau (LLING, CNRS & U. Nantes), Jesse Wichers Schreur (U. Leiden)

Background. Tsova-Tush (Nakh/East Caucasian) has a rich system of verbal inflection, including many TAM forms and various overlapping systems of argument indexing. Though description of this language has improved, its aspect system is still underdescribed, especially its interface with argument and event plurality. In Tsova-Tush, 15 verbs come in pairs described as singular/plural stems reflecting the number of the S/O argument (Desheriev 1953, p. 150; Holisky and Gagua 1994, p. 178; Hauk & Harris 202?: 22), eg. sg *d-ollar* - pl *d-oxk'ar* 'put down'; sg *qollar* - pl *qoxk'ar* 'hang'; sg *xaʔar* - pl *xabžar* 'sit down'; sg *qosar* - pl *qasar* throw, shoot' (*-ar* is a Verbal Noun suffix, *d-* is a gender prefix). Sources remain elusive as to the exact nature of this alternation in Tsova-Tush. Our study seeks to answer the following questions: 1) What kind of number are these verbs sensitive to? Is it agreement with the formal morphosyntactic number specification on the S/O argument as in related Chechen and Ingush? Or it it a distinction reflecting a more semantic notion of number? 2) Is this distinction homogeneous across the Nakh branch given that Chechen and Ingush make the same morphological distinction (Nichols (2011, p. 313) for Ingush, Nichols (2007, p. 1173)). We report our findings so far based on published data and the results of a preliminary study we conducted in Zemo Alvani, Georgia with speakers of Tsova-Tush.

Verbs are sensitive to semantic number Morphosemantic number and semantic number distinctions are aligned in sg count nouns and pl count nouns. They are however misaligned in (i) mass nouns (cumulative and homogeneous reference, but morphosyntactically singular in Tsova-Tush), (ii) numeral noun phrases (which pick out pluralities if numeral > 1, but are morphosyntactically singular in Tsova-Tush), and (iii) pluralia tantum nouns, whose semantic number depends on context even though they have plural morphosyntax. If sg/pl verb pairs in Tsova-Tush involve agreement with the morphosyntactic number on S/O, we expect V and argument number mismatches to be disallowed (formal agreement hypothesis). If on the other hand, sg/pl verb pairs reflect a sensitivity to semantic number, we expect mismatches to be allowed as long as the S/O argument expresses congruent sg or pl semantic number (semantic compatibility hypothesis). As far as count nouns are concerned, sg and pl respectively control Vsg and Vpl, indeed mismatches were found to be unacceptable (2).

(1) a. *lejll-nas* **du-i / do^n*
drive.SG.PFV-AOR.1SG horse-PL / horse
I drove the *horses / horse.'

b. *lejxk'-nas* **do^n / du-i*
drive.PL.PFV-AOR.1SG horse / horse-PL
'I drove the *horse / horses.'

This is consistent with both hypotheses as to what verbal number distinctions may be marking. In what follows, we present three generalizations that favor the semantic compatibility hypothesis.

Generalization 1 Crucially, we found that mass nouns cooccur with Vpl even as they are morphosyntactically sg.[1] Numeral phrases are likewise acceptable (in fact preferred) with Vpl, even though they are morphosyntactically sg (quantified noun cannot take plural morphology).

(2) a. *qejs-nas* *ħejr bat'anmak*
throw.PL.PFV-1SG.AOR flour on_ground
'I threw flour on the ground.'

b. *d-ixk'-nas* *t'q'a tesl*
D-put.PL.PFV-1SG.AOR twenty seed(D)
'I put down twenty seeds.'

This is predicted by the semantic compatibility hypothesis, whereas the formal agreement hypothesis predicts such combinations should be unacceptable since both mass nouns and numeral phrases are morphosyntactically singular in Tsova-Tush, but mass nouns are semantically plural (cumulative reference) and numeral phrases above one describe a plurality of atoms.

Generalization 2 Tsova-Tush (like other East-Caucasian languages) has gender (aka class) agreement with the S/O nominative argument. When mass nouns and numeral phrases agree in gender, they control sg agreement (3 and 4).

[1] Here, D and B refer to two of five Tsova-Tush genders.

(3) *k'ac'k'o^{n} qexk'-b-i-en brinž b-ar.*
little boil-**b.sg**-TR-PTCP.PST rice(B) **b.sg**-be.pst
'There was a bit of cooked rice.'

(4) *ši saldat čuh v-iš-v-ano kox-e.*
two soldier inside **m.sg**-lie_down-**m.sg**-NW.PST hut-OBL[ESS]
'Two soldiers (apparently) lay inside the hut.' (Wicher-Shreur 2024: 139)

Given that mass nouns and numeral phrases control singular gender agreement, if number marking on the stem involved agreement, we would expect Vsg to co-occur with them in S/O position contrary to what we found. This is explained if gender agreement is morphosyntactic, but verb stems are chosen according to a more semantic number distinction.

Argument 3 Finally, while mass nouns and numeral phrases are preferred with Vpl, they are accepted with Vsg but only under a collective construal of the S/O argument. For instance, the numeral phrase wiht Vsg in (5) is acceptable only if the seeds are conceptualized as being together in a pouch.

(5) *d-ill-nas t'q'a tesl st'olmak*
D-put.SG.PFV-AOR.1SG twenty seed(D) on_table
'I put 20 seeds on the table.
SC: Not acceptable, unless the seeds were in a pouch

That Vsg is able to coerce a collective construal of mass and numeral phrases, which otherwise typically occur with Vpl, suggest that the Vsg/Vpl alternation is not governed by shallow morphosyntactic agreement but by a notion of plurality that is semantically interpretable.

Examples with pluralia tantum turned out not to be conclusive, mostly because of properties of pluralia tantum nouns in Tsova-Tush. Though these data are inconclusive with respect to our two hypotheses, we will discuss them in the talk in the interest of transparency.

Nakh comparison. Chechen and Ingush exhibit very similar verbal systems. According to Nichols, Ingush has 23 known sg/pl verb pairs and Chechen has about 20 such pairs. Interestingly, in both Ingush and Chechen, these are explicitly described as "inflectional number agreement in he verb" (e.g. Nichols 2007, p. 1173 for Chechen, Nichols (2011, p. 313) for Ingush). Indeed, Ingush examples show that a numeral phrase is incompatible with pl agreement and must agree according to its morphosyntactic value, which is singular (6).

(6) *ɨž pxi sag ʕa-xeʔira /*ʕa-xeʔišar.*
DEM.PL five person down-sit.PST /down-sit:PL.PST
'Those five people sat down.' Nichols (2011, p. 313)

Conclusion and outlook Tsova-tush Vsg/Vpl pairs reflect the semantic number value of their S/O argument: plural count nouns, mass nouns, and numeral phrases typically co-occur with Vpl, unless they are construed as one unit (eg. if they are packaged) in which case Vsg is possible. This state of affairs differs from what similar verb pairs mark in related Chechen and Ingush where morphosyntactic agreement is involved. Modelling the mechanism involved in Vsg or Vpl selection in Tsova-Tush (e.g. pluractionality, suppletion), as well as understanding how the discrepancy between Chechen and Ingush on the one side and Tsova-Tush on the other emerged are our next goals.

References

Desheriev, Yunus Desherievich (1953). *Bacbijskij jazyk. Fonetika, morfologija, sintaksis, leksika [The Batsbi language. Phonetics, morphology, syntax, lexicon].* Moscow, Leningrad: Akademija Nauk SSSR, Institut Jazykoznanija.

Holisky, Dee Ann and Rusudan Gagua (1994). "Tsova-Tush (Batsbi)". In: *The North East Caucasian languages, Part 2.* Ed. by Rieks Smeets. Vol. 4. The indigenous languages of the Caucasus. Delmar: Caravan Books, pp. 148–212.

Nichols, Johanna (2007). "Chechen morphology (with notes on Ingush)". In: *Morphologies of Asia and Africa.* Ed. by Alan S. Kaye. University Park, USA: Penn State University Press, pp. 1188–1207. ISBN: 9781575065663.

— (2011). *Ingush Grammar.* Berkeley, Los Angeles, London: University of California Press.

The TAM-System of Kist Chechen

A first look based on Dahl's (1985) questionnaire

Maximilian Grübsch, University of Vienna

SHORT INTRODUCTION

The proposed presentation outlines the system of major tense, aspect and mode (TAM) forms of the Kist dialect of the Chechen language, utilizing material from the questionnaire of Dahl (1985). The objective is to describe the usage of major TAM-categories in the terms of Dahl (1985). The data was collected via elicitation during a field trip to the Pankisi Gorge in 2024 and backed up with material from monologues that were recorded during two field trips in 2022. The presentation includes the previously poorly described periphrastic TAM-forms of the dialect and indicates promising topics for future in-depths studies.

DESCRIPTION

Kist Chechen, spoken in the Pankisi Gorge in Northern Georgia by ethnic Chechens called Kists, who settled there no later than mid-19th century, provides the basis for a case study that is going to potentially advance the research in verbal semantics in the Caucasian context. While interference with the Georgian language has left its mark on Kist Chechen (Aliroev, 1962: 65-67) due to close cultural and political ties (Sanikidze 2007), the variety is still being transmitted uninterruptedly and proficiency remains high amongst all generations (Forker and Botkoveli, 2024; Pareulidze, p.c.), which makes inquiries into its grammar valuable for the Caucasus Studies in general. One of the key issues in Kist (and indeed Vainakh) linguistic characterisation is the description of periphrastic TAM-forms, as Standard Chechen is known to have up to 49 of these forms (Nichols and Molochieva, 2018). While for the latter variety there is a monograph about verbal semantics that also includes periphrastic constructions (Molochieva, 2010), the three fundamental descriptions of Kist Chechen (Aliroev 1962; Arsachanov 1969; Pareulidze, 2010) focus rather on the morphology of the synthetic verb forms. It is necessary to investigate their semantics, as well, as Kist Chechen is no less complex than other Vainakh varieties.

One of the available methods facilitating an investigation into the nature of the Kist Chechen TAM-forms is the typological questionnaire of Dahl (1985). This resource, while not allowing for a detailed investigation, and likely to elicit some word-by-word translations (Forker, 2006: 7-9), provides an immediate overview of the categories that bear high functional significance (major TAM-categories), putting them into cross-lingual context, which makes it useful for the purposes of the current study.

The data from the questionnaire is compared to the data collected during field trips by the author of this research in 2022 and 2024. The earlier elicitation shortcomings are mediated by implementing visual stimulae, additional to the spoken ones. This approach is enhanced by a small corpus of monologues.

RESULTS

It is revealed that the major TAM forms of Kist Chechen seem to be the Simple and Durative-Progressive Present tense on the one hand and the Perfective, the Remote and Recent Witnessed Past tenses on the other – employing the traditional terminology, taken from Standard Chechen (Molochieva 2010). A closer look, however, shows that their semantics do not behave quite like the traditional labels suggest.

While the Durative-Progressive Present tense behaves as expected from a progressive tense, the Simple Present tends to occur in habitual-generic contexts – in fact, it could be said that the Kist Chechen Simple Present tense matches Dahls prototypical Habitual-Generic tense quite well. It might thus be possible to state that Kist Chechen Simple Present does not represent a default present tense. This sets it apart from the Remote Witnessed Past, which fulfills the role of a default past tense quite well, as it is indeed neutral towards evidentiality, it can be used for unwitnessed situations, and situations that went on for some time. The Durative Progressive Past is much more restricted than the Recent Witnessed Past (matching Dahls prototype of a Past-Imperfective quite well) and occurs more seldomly than its present counterpart. The Imperfective Past is tied to the explicit mentioning of duration; however, it can be used to express past habits, too.

While evidentiality in the form of witnessing past events plays a role in the selection of past tense forms, the relationship is not straightforward, as witnessed and unwitnessed events can be assigned any past tense form – the interaction between specific morphological tenses and categories of evidentiality seems to be one of preference (e.g. perfective tends to denote unwitnessed events, but can be employed for situations resulting from witnessed events, cf. 1). An interesting finding of the study are addressee-dative pronouns used in evidentiality-like expressions as described for Chechen and Ingush in Nichols and Molochieva (2018), cf. (1). In this example, the speaker informs the hearer, who cannot see the situation himself, that he, the speaker, has opened a door.

(1) *aas* *nawaresh* *hwa-j-ill-ina* *hwun*
PRON1SG.ERG door:PL(J) PVB-J-open-PF PRON2SG.DAT
"I opened the door(s), to your information!"

REFERENCES

Aliroev, I. Ju. (1962). Kistnskij dialekt čečenskogo jazyka. In *Izvestija*, 3 (2), Groznyj: Čečeno-ingušskoe knižnoe izdatel'stvo, pp. 3-88.

Arsachanov, I. A. (1969). Kistnskij dialekt. In *Čečenskaja dialektologija*. Groznyj: Čečeno-ingušskoe knižnoe izdatel'stvo, pp. 155-171.

Dahl, Ö. (1985). *Tense and aspect systems*. Oxford: Blackwell.

Forker, D. (2006). *Tense and Aspect in Bezhta*. Unpublished Master thesis from Universität Leipzig.

Forker, D., & Botkoveli, N. (2024). Resilience through language? A case study of three minority communities in Georgia. In *International Journal of Multilingualism*, 1–21. doi: 10.1080/14790718.2024.2328140

Molochieva, Z. (2010). *Tense, Aspect and Mood in Chechen*. PhD thesis from Universität Leipzig.

Molochieva, Z. (2011). Aspect in Chechen. Linguistic Discovery, 9(2), pp. 104–121. doi:10.1349/PS1.1537-0852.A.397

Molochieva, Z., & Nichols, J. (2018). Tense, Aspect, Mood and Evidentiality in Chechen and Ingush. In *Brill's Studies in Language, Cognition and Culture* (Vol. 16). United States: BRILL, pp. 26–48. doi: 10.1163/9789004361805_003

Pareulidze, R. (2010). *Čečnuri enis kist'uri dialek'ti (gramatik'uli analizi, t'ekst'ebi, lek'sik'oni)*. Tbilisi: Tbilisi State University.

Sanikidze, G. (2007). Islamic Resurgence in the Modern Caucasian Region: 'Global' and 'local' Islam in the Pankisi Gorge. In *Slavic Eurasian Studies*, 4(14), pp. 264-280.

Noun-Verb Complex Predicates in Dargwa Languages

Ivan Kalyakin
Institute of Linguistics, RAS / Russian State University for the Humanities
kalyakin.iv@gmail.com

Introduction. Complex predicate (CPr) is a construction consisting of two elements – a light verb (LV) and a non-verbal element (NVE). The latter provides the core conceptual meaning, while the former serves primarily as a 'verbalizer', expressing relatively abstract meaning (causation, initiation, etc.). The example (1) instantiates the CPr *wana* AGR-*arq'*- 'warm up' built up of the NVE *wana* 'warm' and the LV AGR-*arq'*- 'make'.[1]

(1) waba-li χureg **wana** **b-arq'-ib**
mother-ERG food.ABS warm N-make.PFV-AOR
'Mother warmed up the food.'

Besides adjectives, the NVE in Dargwa languages can also be represented by a noun, ideophone, adverb/postposition, numeral, borrowed verb, and some special element which is no longer used outside the construction. Nominal NVE are of particular interest because, as extensive literature on different languages shows (see Folli et al. 2004 for Modern Persian and Gündoğdu 2016 for Kurmanji Kurdish), they sometimes possess both lexical and phrasal properties. Moreover, even though N(oun)-LV CPrs are abundant in Nakh-Daghestanian languages, for instance, in Lezgian (Haspelmath 1993) and Chechen (Komen et al. 2021), there have not been any in-depth investigations of the phenomenon in any of the languages. The present talk attempts to fill this gap by providing a comprehensive account of N-LV CPrs in three Dargwa languages (< Nakh-Daghestanian): Muira, Icari, and Kaytag. The goals are threefold. Firstly, I discuss the morphosyntactic properties of N-LV CPrs in Muira. Then, I show that the properties are also characteristic of CPrs in all Dargwa languages, presenting evidence from Icari and Kaytag, two languages from different branches of the Dargwa language group. Finally, I investigate how the current understanding of argument structure can help us to account for the discussed properties of the CPrs.

Data. All data come from the author's field work. Muira data were collected in the village of Kalkni (Republic of Daghestan, Russia) during two fieldtrips in August 2022 and June 2023; Icari and Kaytag data were collected during fieldtrips to the villages of Ivan-Kutan and Javgat (Republic of Daghestan, Russia) in 2023 and 2024 respectively. All acceptability judgements were elicited from at least three speakers and subsequently confirmed in different sessions.

Morphosyntax. The two elements of a CPr demonstrate a considerable degree of independence. For instance, an NVE and an LV need not be linearly adjacent (2). Moreover, the order of the elements may be reversed and it is possible to elide only one of the elements of the CPr, leaving the other intact.

(2) pat'imat-li χureg **cːe** kuχni-le-b **w-arq'-ib**
PN-ERG food.ABS salt kitchen-LOC-N M-make.PFV-AOR
'Patimat salted the food in the kitchen.'

Another important property of (the majority of) CPrs is the deficient nature of the NVE. That is, nominal NVEs bear no Case and, consequently, cannot serve as a goal for agreement relation – as (3) shows, the NVE *gap* 'praise' is not in absolutive (which is zero-marked), since there is another zero-marked DP *Rasul* with which the verb agrees in gender. Additionally, the element cannot be modified (4), it is not referential (5), and it does not get ergative case under antipassivization.

(3) učitelʲ-li rasul **gap** {**w-arq'-ib** / ***b-arq'-ib**}
teacher-ERG PN.ABS praise M-make.PFV-AOR / N-make.PFV-AOR
'The teacher praised Rasul.'

(4) * darɦa [qaˤrq-il **duc'**] **Ø-ik'-u-r=ri**
child.ABS quick-ATR run M-say.IPFV-PROG-CVB=PST
Exp.: 'The child ran fast.'

[1] All the unspecified examples are from Muira. The components of the CPr are in boldface.

(5) * murad [iš **taˤħ**] **Ø-uq-un**
PN.ABS this.2 jump M-move.PFV-AOR
Exp.: ‘Murad made that jump.’

Heterogeneity. There is, however, a small group of nominals which possess certain phrasal features and resemble direct objects. They occupy the internal argument slot and thus prevent any absolutive arguments from occurring with CPrs. Consequently, they receive absolutive case and trigger agreement on the verbs which possess an agreement slot, as in (6). Moreover, such NVEs can be modified (7).

(6) asadulla-li sun-i-la aba-s **q**:**ulluq**: **b-irq'-u-li** sa<b>i
PN-ERG SELF-OBL-GEN mother-DAT care.ABS N-make.IPFV-PROG-CVB <N>COP
‘Asadullah cared about his mother.’

(7) murad-li rasul-li-s [χula-l **kumek**] **b-arq'-ib**
PN-ERG PN-OBL-DAT big-ADJ help.ABS N-make.PFV-AOR
‘Murad provided great help to Rasul.’

Homogeneity. Icari and Kaytag demonstrate exactly the same picture. In these languages the two elements of the CPr are rather independent of each other and may not be linearly adjacent, as in (8). The majority of NVEs are also unable to take modifiers and be referential, cf. (9) and (10).

(8) ICARI DARGWA
učitelʲ-li **gap** rasul **w-aˤrq'-ib**
teacher-ERG praise PN.ABS M-make.PFV-AOR
‘The teacher praised Rasul.’

(9) * ICARI DARGWA
učitelʲ-li rasul [χula **gap**] **w-aˤrq'-ib**
teacher-ERG PN.ABS big praise M-make.PFV-AOR
Exp.: ‘The teacher praised Rasul very much.’

(10) * KAYTAG DARGWA
učitelʲ-li rasul [χʷala **gap**] **k-Ø-ar-iw**
teacher-ERG PN.ABS big praise DOWN-M-make.PFV-AOR
Exp.: ‘The teacher praised Rasul very much.’

Just like in Muira, Icari and Kaytag also possess a small class of NVEs that play a dual role, being not only part of a CPr but also an (internal) argument, see (11) with a modified NVE *q:ulluq:* ‘care’.

(11) ICARI DARGWA
musa-li cin-na waba-j [χula **q**:**ulluq**:] **b-irq'-a** ca-b
PN-ERG SELF-GEN mother-DAT big care.ABS N-make.IPFV-PROG COP-N
‘Musa really cares about his mother.’

Analysis Prospects. Given all the facts about morphosyntactic deficiency of the majority of NVEs, I consider an NV to be the realization of a bare root (Marantz 1997) which corresponds to a phonological word. Those NVEs that also function as arguments should be analyzed as full-fledged DPs; the oblique arguments in such cases are introduced by an Appl or some other argument-introducing functional head. Such structures raise the question of whether roots are even necessary in the structure. I will discuss this question and argue that the absence of roots raises no problems for derivation, being sometimes even a desirable possibility.

Conclusions. After examining a range of N-LV CPrs in three rather different Dargwa languages I come to the conclusion that CPrs behave quite uniformly in all Dargwa languages. Within a single language, however, N-LV CPrs do not form a homogenous group; I have shown that (at least) two classes of NVE should be differentiated. Consequently, these two classes require different syntactic analyses.

References: ● Folli, R., Harley, H., and Karimi, S. (2005). Determinants of Event Type in Persian Complex Predicates. *Lingua*, 115(10):1365–1401. ● Gündoğdu S. (2016). Noun-Verb Complex Predicates in Kurmanji Kurdish: A Syntactic Account. In Bellamy, K., Karvovskaya, E., Kohlberger, M., and Saad, G., editors, *Proceedings of ConSOLEXXIII*, pages 279–301. Leiden University Centre for Linguistics. ● Haspelmath, M. (1993). *A Grammar of Lezgian*. Berlin: Mouton de Gruyter. ● Komen, E., Molochieva, Z., Nichols, J. (2021). Chechen and Ingush. In Polinsky, M., editor, The Oxford Handbook of Languages of the Caucasus, pages 317–

367. Oxford University Press, Oxford. ● Marantz, A. (1997). No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In Dimitriadis A., Siegel L., Surek-Clark C., editors, *University of Pennsylvania Working Papers in Linguistics 4*, pages 201–225. University of Pennsylvania, Philadelphia.

Kamarauli Anastasia

DEVELOPMENT OF AN AUTOMATED PROGRAM FOR ANALYSING POLITICAL SPEECHES USING THE GEORGIAN NATIONAL CORPUS API

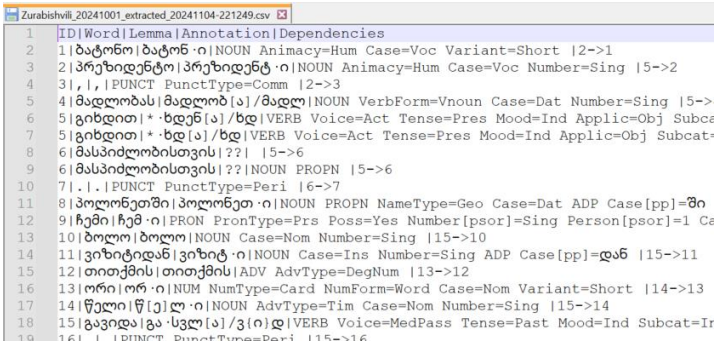
In recent years, the field of political text analysis has witnessed growing interest as researchers aim to explore patterns, sentiments, and rhetorical strategies in political discourse. The role of political language is particularly relevant in young democracies and polarized political discourse as it is the case in Georgia. With advancements in new technologies, particularly in the fields of digital humanities and corpus linguistics, large-scale quantitative analysis has become increasingly accessible. Through the application of these tools, researchers can now analyse vast amounts of linguistic data efficiently, yielding statistically robust insights.

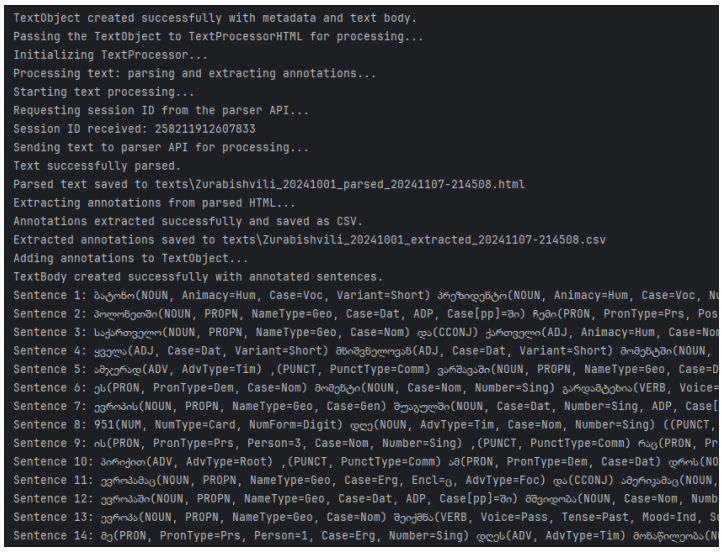
Furthermore, the progress in language model development and artificial intelligence offers unprecedented opportunities for qualitative analysis. By incorporating AI-driven methods, it is possible to delve deeper into the nuances of political language, facilitating a richer, multidimensional analysis. These advancements in technology and methodology enable a dynamic approach to political text analysis, moving beyond static corpora to build custom datasets tailored to specific research needs. However, applying such technologies to low-resource languages like Georgian presents unique challenges. AI models typically rely on extensive datasets for accuracy in tasks like parsing, sentiment analysis, and topic modelling. For low-resource languages, the lack of large, annotated corpora limits model training, often resulting in reduced accuracy.

Additionally, the linguistic features of low-resource languages often differ significantly from those of more widely studied languages, making it difficult for generalized models to perform effectively. A further obstacle is the limited research on certain linguistic aspects of these languages. Unlike well-documented languages, low-resource languages often lack comprehensive analysis of syntactic, semantic, and morphological features. This scarcity complicates the development of high-quality annotations, which are crucial for effective model training. Without reliable annotations, models may struggle to accurately parse and analyse the language, reducing their effectiveness.

This paper presents a system for the automated analysis of political speeches, using the Georgian National Corpus API for parsing and will discuss which aspects are relevant for quantitative political-linguistic analyses as well as the possibilities, difficulties and limitations of the use of AI for quantitative analysis.

The system's modular design includes a GUI for importing speeches, a core parsing module that interacts with the API, and structured XML storage for linguistic metadata. The program allows for the extraction of morphosyntactic annotations and the addition of metadata, enabling comparative analysis of, for example, gender-based or political affiliation-based language use.





The system supports both quantitative and qualitative analysis, storing results in a database and offering visualization options, allowing users to build a local corpus of analysed speeches. The goal is to provide a scalable, flexible solution that accommodates a wide range of research needs, enabling customized corpus creation.

Quantitative analysis focuses on linguistic features such as word frequency, peculiarities at the morphosyntactic level and rhetorical devices, providing statistical insights into speeches. Qualitative analysis examines aspects like sentiment and rhetorical structure, contributing to a comprehensive understanding of language use, themes, and strategies.

In summary, this system automates the analysis of political speeches, combining parsing with advanced analytics to provide a dynamic, adaptable tool for understanding political discourse. Its integration with the Georgian National Corpus API significantly supports political science, linguistics, and data-driven research by enabling nuanced exploration of political rhetoric.

The aim is to provide a program that enables users to analyze political texts efficiently, without requiring knowledge of underlying linguistic or technical details. It is primarily intended for political scientists, sociologists, journalists, and even politicians themselves.

```
Active verb count (Voice=Act): 21
Active verbs - Token IDs: [5, 5, 44, 143, 146, 175, 175, 254, 276, 287, 334, 354, 388, 459, 480, 487, 508, 517, 527, 538, 571]
Active verbs - Sentence IDs: [1, 4, 8, 9, 12, 14, 17, 19, 20, 24, 25, 26, 27]
Adjective count (ADJ): 79
Adjectives - Token IDs: [19, 22, 29, 32, 33, 38, 41, 46, 63, 76, 89, 90, 111, 114, 124, 148, 161, 165, 200, 213, 224, 234, 240, 24
Adjectives - Sentence IDs: [3, 4, 5, 7, 8, 9, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28]
Adverb count (ADV): 47
Adverbs - Token IDs: [12, 24, 62, 79, 84, 101, 102, 109, 131, 132, 142, 145, 152, 166, 167, 172, 177, 192, 193, 212, 226, 243, 252
Adverbs - Sentence IDs: [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 23, 24, 25, 26, 27]
Conjunction count (CONJ): 59
Conjunctions - Token IDs: [18, 26, 36, 54, 60, 66, 73, 87, 92, 104, 123, 135, 140, 144, 171, 182, 198, 203, 209, 219, 222, 228, 24
Conjunctions - Sentence IDs: [3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27]
Imperative count (Imp): 1
Imperatives - Token IDs: [25]
Imperatives - Sentence IDs: [3]
Count of conditional forms (Mood=Cnd): 1
Conditional forms - Token IDs: [160]
Conditional forms - Sentence IDs: [9]
Count of indicative forms (Mood=Ind): 61
Indicative forms - Token IDs: [5, 5, 15, 27, 44, 57, 71, 85, 94, 98, 116, 133, 143, 146, 153, 159, 168, 170, 175, 175, 181, 194, 2
Indicative forms - Sentence IDs: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27]
Number of optative forms (Mood=Opt): 6
```

```
<?xml version='1.0' encoding='utf-8'?>
<TextObject>
    <MetaData>
        <Author>
            <surname>Zurabishvili</surname>
            <middlename />
            <name>Salome</name>
            <dateOfBirth>1952-03-18</dateOfBirth>
            <age>72</age>
            <partyPosition>parteilos</partyPosition>
            <occupation>Präsidentin</occupation>
            <confession>orthodox</confession>
            <ethnicity>georgisch</ethnicity>
            <origin>Paris, Frankreich</origin>
        </Author>
        <textType>Rede</textType>
        <Location>Tbilisi</Location>
        <date>2024-10-01 00:00:00</date>
        <source>https://president.ge/index.php?m=206&amp;appeals_id=416</source>
    </MetaData>
    <TextBody>
        <text>ბატონო პრეზიდენტო, მადლობას გიხდით მასპინძლობისთვის.

პოლონეთში ჩემი ბოლო ვიზიტიდან თითქმის ორი წელი გავიდა. საქართველო და ქართველი ხალხი თავის ევროპულ გზაზე ყოველთვის გრძნობდა და გ

ამჯერად, ვარშავაში ჩემი ვიზიტი მაშინ ტარდება, როცა ჩვენ ყველა გარდამტეხ დროსა და ეპოქაში ვიმყოფებით. ეს მომენტი გარდამტეხია მსოფლიოსთვ
```

```
ბატონ პრეზიდენტთან ძალიან დიდი იმედი გამოვთქვი, რომ საქართველო არამხოლოდ დაუბრუნდება ამ გზას,

კიდევ ერთხელ, მინდა დიდი მადლობა გადავუხადო ბატონ პრეზიდენტს ამ ვიზიტისთვის, ფორუმში ჩემი მო

დიდი მადლობა!</text>
        <sentences>
            <sentence id="1">
                <tokens>
                    <token id="1">
                        <word>ბატონო</word>
                        <lemma>ბატონ·ი</lemma>
                        <annotation>NOUN,Animacy=Hum,Case=Voc,Variant=Short</annotation>
                        <dependency>2</dependency>
                    </token>
                    <token id="2">
                        <word>პრეზიდენტო</word>
                        <lemma>პრეზიდენტ·ი</lemma>
                        <annotation>NOUN,Animacy=Hum,Case=Voc,Number=Sing</annotation>
                        <dependency>5</dependency>
                    </token>
                    <token id="3">
                        <word>,</word>
                        <lemma>,</lemma>
                        <annotation>PUNCT,PunctType=Comm</annotation>
                        <dependency>2</dependency>
                    </token>
```

```
                    <annotation>PUNCT,PunctType=Excl</annotation>
                    <dependency>617</dependency>
                </token>
            </tokens>
        </sentence>
    </sentences>
</TextBody>
<Evaluation>
    <quantitativeEvaluation active="21" adjectives="79" adverbs="47" conjunctions="59" imperative="1" mood.Cnd="1" mood.Ind="61" mood.Opt="6" mood.Pot="3" nouns="189"
    passive="0" pron.me="0" pron_chemi="0" pron_chven="0" pron_chveni="0" pron_tkven="0" pron_tkveni="0" stopwords="0" tempus_fut="12" tempus_past="19" tempus_perf="2"
    tempus_pres="28" tempus_pqp="1"
    type="{'პოლოზეთ·ი', 'პრეზიდენტ·ი', 'რომ', '26', 'და·წყებ[ა]/წყ', 'აღ·ნიშნ[ა]/ნიშნ', 'იმედოვნებ[ა]/იმედოვნ', 'ეტაპ·ი', 'მაშინ', 'აქამდე', 'თავისუფ[ა]ლ·ი', '-',
    'შე·დგომ[ა]/დგ', 'ყველა', 'გამო·წვევ[ა]/წვ', 'პატივ·ი', 'სოლიდარობ[ა]', 'მჭიდრო', 'არცერთ·ი', 'მადლობ[ა]/მადლ', 'სამართლიან·ი', 'კიდევ', 'მოსახლეობ[ა]', 'ლიდინ·ი/ლოდ',
    'კავშირ·ი', 'ფოფხ[ა]/არ', 'ომ·ი', 'ბრძოლ[ა]/ბრძ', 'ფორუმ·ი', 'გარდა·ტყდომ[ა]/ტეხ', 'ორ·ი', '*·ხდ[ა]/ხდ', 'დრო', 'დღიან·ი', 'მსჯელობ[ა]', 'ისტორიულ·ი', 'და',
    'მნიშვნელოვ[ა]ნ·ი', 'ჩამო·სვლ[ა]/სვლ', 'გა·გრძელებ[ა]/გრძელ', 'გამო·ჩენ[ა]/ჩნ', 'არჩევ[ა]ნ·ი', 'გა·ტან[ა]/ტან', 'გზ[ა]', 'და·რღვევ[ა]/რღვ', 'ქართველ·ი', 'ყოფნ[ა]/ყ[ოვ]',
    'ხალხ·ი', 'სად', 'პრაქტიკულ·ი', 'პირიქით', 'მი·ღებ[ა]/ღ', 'წინააღმდეგ', 'საჭირო', 'რაება', 'და·რწმუნებულ·ი/რწმუნ', 'რუსეთ·ი', 'თითქმის', 'მონაწილეობ[ა]/მონაწილე', 'მოკლე[ქ]',
    'შე·რყევ[ა]/რყ[ევ]', 'კამპანი[ა]', ',', 'მსოფლიო', 'სასარგებლო', '(', 'ჩვენ·ი', 'ძალიან', 'რაც', 'რომ[ე]ლ·ი', 'რეფერენდუმ·ი', 'ადგილ·ი', 'ამჯერად', 'სუვერენობ[ა]', '3',
    'მო·ხერხებ[ა]/ხერხ', 'მყოფნ[ა]/მყოფ', 'უკვე', 'გადა·დგმ[ა]/დგ', 'შტ·ი', 'შე·ძლებ[ა]/ძ[ე]ლ', 'მის·ი', 'ერთ-ერთ·ი', 'როგორ·ი', 'ყოველთვის', 'ამერიკ[ა]', 'კომპონენტ·ი',
    'უსაფრთხოებ[ა]', 'ახ[ა]ლ·ი', 'სწორედ', '??', 'პარტნიორ·ი', 'უსასტიკეს·ი', 'არამედ', 'რუსულ·ი', 'წარსულ·ი', 'იანვ[ა]რ·ი', 'გამო·თქმ[ა]/თქვ', 'ნდომ[ა]/ნდ', 'ქონ[ა]/ქვ',
    'უკრაინელ·ი', 'იქ', 'მშვიდობიან·ი', '(ოდნ[ა]/ც', 'რეგიონ·ი', 'მი·მდინარეობ[ა]/მდინარე', 'გამო·ცდილ·ი/ცდ', 'წარ·სულ·ი/ს', 'გაწევრიანებ[ა]', 'ევროკავშირ·ი', 'ბოლო', 'ვარშავ[ა]',
    'ვითარებ[ა]/ვითარ', 'ა·რჩევ[ა]/რჩ', 'ერთხელ', 'პერიოდ·ი', 'დამკვირვებ[ე]ლ·ი', 'მო·მავალ·ი/ვ[ა]ლ', 'რ[ა]', 'დიდ·ი', 'გონებ[ა]/გონ', 'ყოვ[ე]ლ·ი', 'ბატონ·ი', 'მე', ')', 'ყოფნ[ა]/ქნ',
    'თქმ[ა]/თქვ', 'ურყევ·ი/რყ', '*·ხდებ[ა]/ხდ', 'კარ·ი', 'ქცეყნ[ა]ნ[ა]', 'კანონ·ი', 'ის', 'დღემდე', 'მხარდაჭერ[ა]', '951', '*·ცდომ[ა]/ცდ', 'მი·ყვან[ა]/ყვან', 'ევროპ[ა]', 'ევროპელ·ი',
    'მშვიდობ[ა]', 'როგორც', 'თავმჯდომარეობ[ა]/თავმჯდომარე', 'შეგიძლ·ი', 'შე·ჩერებულ·ი/ჩერ', 'უფრო', 'ასევე', 'გადა·ხდ[ა]/ხად', 'მხ[ე]', '*·გრძელებ[ა]/გრძელ', 'ევრაზ[ე]რ·ი',
    'მო·ლაპარაკებ[ა]/ლაპარაკ', 'გვერდით', 'შე·ქმნ[ა]/ქმ[ე]ნ', 'თუ', 'მომჭირნ·ი', 'ნაბიჯ·ი', 'არსებულ·ი/არსებ', 'სასიცოცხლო', 'გადა·მწყვეტ·ი/წყვეტ', 'მთავ[ა]რ·ი', 'გრძნობ[ა]/გრძნ',
    'ჩემ·ი', 'სწრაფ·ი', 'შე·ძლებ[ა]/ძლ', 'სტაბილურობ[ა]', 'აგრესი[ა]', 'მო·ხერხ[ა]/ხერხ', 'ოჭტომბ[ე]რ·ი', 'გა·ვლ[ა]/არ', 'უწდა', 'ეპოქ[ა]', 'სუვერენიტეტ·ი', '*·ტარებ[ა]/ტარ',
    'გა·ხსნ[ა]/ხსნ', 'ჩა·ტარებ[ა]/ტარ', 'მიმართ', 'როცა', 'მაგრამ', 'საკუთ[ა]რ·ი', 'აღმო·ჩენ[ა]/ჩნ', 'ხელისუფლებ[ა]', '!', 'ვიზიტ·ი', 'თავისუფლებ[ა]/თავისუფლ', 'დღ[ე]', 'მო·ხდომ[ა]/ხდ',
    'გარდა·მტეხ·ი/ტეხ', 'წ[ე]ლ·ი', 'უკრაინ[ა]', 'ვინ', 'არჩევნებ·ი', 'მხარდამჭერ·ი', 'გა·სვლ[ა]/ვ[ი]დ', 'დამოუკიდებ[ე]ლ·ი', 'დღეს', 'მთლიანად', 'დასავლეთ·ი', 'ყოფნ[ა]/ყოფ', 'არ',
    'მიმართებ[ა]', 'გა·კეთებ[ა]/კეთ', 'კარგ·ი', 'გაყინლ·ი', 'არამხოლოდ', '*·სოხ[ა]/სვ', 'მეგობ[ა]რ·ი', 'ჭირვებ[ა]/ჭირ', 'იმედ·ი', 'ევროპულ·ი', 'ივნის·ი', 'მიმდინარ[ე]/დენ',
    'დამოუკიდებლობ[ა]', 'გადაწყვეტილებ[ა]', 'ორიენტირებულ·ი', 'ვერ', 'მითუმეტეს', 'ჩვენ', '*·მზადებ[ა]/მზად', 'მხოლოდ', 'მეზობ[ე]ლ·ი', 'შორის', 'მორალურ·ი', '.', 'თავის·ი',
    'ეს', 'და·ბრუნებ[ა]/ბრუნ', 'საქართველო', 'გარემო', 'მომავ[ა]ლ·ი'}" verbs="69" vocative="2"/>
        <activeVerbs_token_list>5,5,44,143,146,175,175,254,276,287,334,354,388,459,480,487,508,517,527,538,571</activeVerbs_token_list>
        <activeVerbs_sentence_list>1,4,8,9,12,14,17,19,20,24,25,26,27</activeVerbs_sentence_list>
    </quantitativeEvaluation>
</Evaluation>
</TextObject>
```

Endoclisis in the Northern Talyshi verb

Steven Kaye, Surrey Morphology Group

This talk will argue for the existence of the striking phenomenon of *endoclisis* (a.k.a. *mesoclisis*) – where clitics are hosted inside morphological words – in the verb system of Northern Talyshi, an Iranian (< Indo-European) language of southern Azerbaijan and northwestern Iran. Although it is known to be unusual in cross-linguistic terms, endoclisis has already been identified in languages that are closely associated with Northern Talyshi both genealogically (Sorani Kurdish, Iraq / Iran; e.g. Walther 2012) and areally (Udi < Nakh-Dagestanian, Azerbaijan / Georgia; e.g. Harris 2002), and I suggest that its presence in Northern Talyshi has gone under the radar largely because of a long-standing misapprehension in the descriptive literature about the morphological makeup of the forms where it is found. Clearing up this misapprehension allows us to recognize the Northern Talyshi phenomenon as part of this broader Iranian and Caucasian context, as well as adding another entry to what is still a rather short list of endoclitic formations found in the world's languages.

(1) vard**=ışon**-e bə səray
bring**=3PL.CL**-PRET to palace
'They brought [him] to the palace.' (Əboszodə 2004: 76)

An instance of the morphological structure at issue is provided in the first word of example (1), *vardışone* 'they brought'. The unlikely-looking segmentation I offer, in which a preterite marker labelled as a suffix (*-e*) is external to a person marker labelled as a clitic (*=ışon*), captures the typological interest of the forms involved. Drawing on a range of evidence from text collections (Miller 1930; Əboszodə 2004, with selected tales edited as Kaye 2023) and existing descriptions (Miller 1953; Schulze 2000; Stilo 2008, 2018; Paul 2011), the paper will justify this endoclitic analysis, as opposed to one in which both markers are enclitics as has generally been assumed before. It will also situate the endoclitic preterite's behaviour in the broader context of Northern Talyshi's distinctive 'split ergative' morphosyntax; discuss its history; and explore the value and limitations of comparing it with the notorious (and more challenging) endoclitic behaviour found in the verb paradigm of Udi.

The personal clitics involved here, exemplified by 3PL *=ışon* in (1) and descending from oblique pronominal forms already attested in Middle Iranian, are multifunctional in Northern Talyshi: for example, they can encode the possessor (2) and the experiencer argument of *piye* 'want, love' (3). However, as in (1), they are most commonly found as subject markers, as part of a distinctive ergative construction specifically associated with perfective past tenses of transitive verbs (against the backdrop of an otherwise accusative morphosyntax). In this role they can coindex an overt subject argument, which appears in the oblique case; this is seen in (4), which shows a verb form in the perfect tense. Crucially, in any of these uses, the clitic can be hosted either leftwards of the verb, often by a noun phrase as in (2) and (3); or by the verbal complex itself, as in (1) and (4).

(2) qılə=y dışmen**=ım** hest=e
CLF=one enemy**=1SG.CL** EXIST=COP.3SG
'I have an enemy.' (Əboszodə 2004: 172)

(3) çiç**=ı** pi-yeydə?
what**=2SG.CL** want-PROG
'What do you want?' (Əboszodə 2004: 172)

(4) əy kard-ə**=ş**=e bı-zın
3SG.OBL do-PTCP**=3SG.CL**=COP.3SG MOD-know.IMP.2SG
'Know that he has done [it].' (Əboszodə 2004: 172)

As seen in (4), the perfect formation is essentially periphrastic, comprising the general participle in *-ə* and a 3SG copula: the latter is a clitic in all its uses, and can itself float leftwards as a focus marker in some dialects (Stilo 2008). Thus the fact that person marking can surface in between these two components of the perfect is not especially surprising. However, previous analyses have tended to assume (implicitly or explicitly) that the same analysis can also be applied to preterites as in (1), identifying their final segment *e* with the 3SG copula too (e.g. Miller 1953: 168; Schulze 2000; Stilo 2018, 2023). I show that this is not the case, and that the perfect and the preterite feature distinct final morphemes with different properties:

Copular auxiliary *=e*	**Preterite marker *-e***
Can be negated directly > *ni* [NEG.COP.3SG]	Cannot be negated directly
Has a past tense counterpart *be*	Has no past tense counterpart
Effaced immediately after *ə*	Retained (as *-y*) immediately after *ə*
Reliably pronounced as *e* post-consonantally	Pronunciation *ı* also attested
Can float leftwards (Lerik and Astara varieties)	Cannot float leftwards

The upshot of these observations is that the person marker in a preterite form such as *votışe* 'said' (5a) has indeed been inserted between the verb stem and a suffix which would immediately follow it otherwise (5b), confirming the presence of endoclisis here.

(5) a. ibrahim-i vot=ış-e
PN-OBL say=3SG.CL-PRET
'Ibrahim said' (Əboszodə 2004: 40)

b. ibrahim-i bə=ştə həsuyə=ş vot-e
PN-OBL to=REFL.POSS father_in_law=3SG.CL say-PRET
'Ibrahim said to his father-in-law' (Əboszodə 2004: 45)

References

Əboszodə, Fərmoni F. (2004). *Nəğılon* [Tales]. Baku: MBM.

Harris, Alice C. (2002). *Endoclitics and the origins of Udi morphosyntax.* Oxford: Oxford University Press.

Kaye, Steven (2023). *Corpus IMMOCAL : le talyshi* [12 glossed texts with English translation]. URL: https://proclac.cnrs.fr/projets/projets-turcaucase/immocal/talyshi/

Miller, Boris V. (1930). *Talyšskie teksty* [Talyshi texts]. Moscow: Naučno-issledovatel'skij institut nacional'nyx i ètničeskix kul'tur narodov vostoka SSSR.

Miller, Boris V. (1953). *Talyšskij jazyk* [The Talyshi language]. Moscow: Akademija Nauk.

Paul, Daniel (2011). *A comparative dialectal description of Iranian Taleshi*. University of Manchester.

Schulze, Wolfgang (2000). *Northern Talysh*. Munich: Lincom Europa.

Stilo, Donald L. (2008). Two sets of mobile verbal person agreement markers in the Northern Talyshi language. In S. Karimi, V. Samiian & D. Stilo (eds.), *Aspects of Iranian linguistics: Papers in honor of Mohammad Reza Bateni*. Newcastle: Cambridge Scholars Press, 363–390.

Stilo, Donald L. (2018). The Caspian region and south Azerbaijan: Caspian and Tatic. In: G. Haig & G. Khan (eds.), *The languages and linguistics of Western Asia*. Berlin: De Gruyter Mouton, 659–824.

Stilo, Donald L. (2023). A personal narrative in Northern Talyshi with grammatical commentaries and notes on the areal distribution of grammatical features. *Tehlikedeki diller dergisi* [Journal of endangered languages] 13.22: 137–183.

Walther, Géraldine (2012). Fitting into morphological structure: accounting for Sorani Kurdish endoclitics. In A. Ralli, G. Booij, S. Scalise & A Karasimos (eds.), *Proceedings of the 8th Mediterranean Morphology Meeting*, 299–321.

Ideophones in Bezhta

Zaira Khalilova
Institute of Linguistics of the Russian Academy of Sciences, Moscow

The paper is the first report on phonology, morphology, syntax and semantics of ideophones in Bezhta (East Caucasian). The Bezhta class of ideophones, including onomatopoeia, is large and productive, covering various sematic fields. Two main semantic fields are natural and non-natural sounds. Natural sounds are sounds related to humans, e.g. *hik*' –hiccups, *hic-hac* – sneezing, *q'ur* – rumbling in stomach, *χor-χor/χur-χur* – by running nose, *wiš* – stifle, *p'är-p'är* – chatter, *piš-piš* – whisper, *łep'*- whimper, *qis-qas* – chewing, etc.; and to animals: *wic'-wic'* – mouse squeak, *ha^{n}w-ha^{n}w* – bog's barking, *wič-wič* - twitter, *qurq-qurq* – cooing of pigeons, *ʔüʔüʔü* – crowing, *šš* - snake hiss, *кäʔ* -caw, ect. Non-natural sounds are sounds of striking and knocking, e.g. motion sounds: *dip'-dap'* – sound of falling, thud, *qoc-qoc* – sound of heels, *lix-lax* – wiggling, slow rocking, slight movement, *xuc* – collision, *кip'-кap'*- fuss, mess around, *кir-кir* – jolt, rumble-tumble in a car, and others: *žir-žar* – tinkling, *č'iq'-č'iq'* – sound of cutting with scissors, *piq'-poq'* – sound of boiling corn porridge, *ƛ'iq'-ƛ'aq'* – crackling, *ƛ'äq'* – shot, bang, *č'iq'* – click, *diq-däq* – knocking, *q'öp'* – knock, etc. Bezhta definitely implies sound symbolism in the formation of onomatopoeia. For example, final [r] imitates resonant and loud noise: *кir-кir* – jolt, rumble-tumble in a car, *ƛ'är-ƛ'är* – motor of a motorcycle, *žir-žar* – tinkling, *čar* – splashing with a loud noise. Low-intensity sounds are indicated with final [p'], which often imitates muffled and hollow sound, e.g. *q'ip'-q'op'* – stomp, *кip'-кap'*- fuss, mess around, *lip'-lap'* – sound of jolt, shake, *q'öp'* – (dull) knock. Final vowel reflects durable actions, like swing, laughing, birds singing, instrumental melodies, e.g. *hu-hu* –owl hooting, *qi-qi* – laughter, *ku-ku* – swing. Reduplication is also associated with sound symbolism (Hinton et. al. 1995). In Bezhta reduplicated stems are associated with repetitive actions, e.g. *diq-däq* – knocking, whereas non-reduplicated stems are associated with short sounds and momentary actions, e.g. *dap'* – thud.

Ideophones are often used in complex verbs combining with light verbs. Most ideophones denoting non-natural sounds are combined with intransitive 'become' and transitive 'do', e.g. *čuč jaqal* 'to become weak' and *čuč jowal* 'make weak' or with intransitive 'go' and transitive 'send', e.g. *dac je^{n}ƛ'al* 'to lie down' and *dac je^{n}jal* 'to make lie down', *xuc je^{n}ƛ'al* 'to collide' and *xuc je^{n}jal* 'to push'.

Some onomatopoetic verbs are formed by the incorporation of an onomatopoetic element into the verb *iƛ*-, which synchronically means 'call' in Bezhta but has the more general meaning 'say' in some other Tsezic languages (Comrie et al. 2015: 542). Onomatopoetic verbs based on incorporation are unergative and intransitive. Unergatives verbs, for example, *кäʔƛ*- 'caw', *hicƛ*- 'sneeze', *öhƛ*- 'cough', *hik'ƛ*- 'hiccup', *hahƛ*- 'yawn', have the single S_A argument in the Ergative, which is always animate, e.g. *kibba hahƛojo* 'girl.ERG yawn.PST'. Intransitive onomatopoetic verbs have the single S argument in the Absolutive, which is inanimate, e.g. *jeł cacaƛoj* 'oil(ABS) splash.PST' [Wier 2023].

Morphologically, the ideophone class is defective, as it lacks many categories typical of other word classes. Most ideophones denoting unnatural sounds have both nominal and verbal features. Nominal features include gender and, sometimes, number, for example, *xuc-bo* 'collision-PL'; the only verbal feature is the ability to form a masdar from many ideophones. For example, *xuc-ni* 'colliding', *däq-ni* 'knocking'.

Based on comparative analysis of closely related languages, this work provides further evidence for sound symbolism in Tsezic languages. The work covers the main topics in phonology, morphology, and syntax of ideophones, including various degree of integration in a clause. The Bezhta children's lexicon will also be presented.

References

Comrie, B., Khalilov, M. & Khalilova, Z. 2015. *A Grammar of Bezhta: Phonology, Morphology, and Word formation*. Leipzig/Makhachkala: ALEF.

Hinton, L., J. Nichols, and J. J. Ohala. 1995. Introduction. In Hinton, Leanne, Johanna Nichols, & John Ohala (eds.), *Sound symbolism*, 1-12 Cambridge: Cambridge UP.
Wier T. 2023. Expressive constructions in Georgian and other Caucasian languages. In: Williams JP, ed. *Expressivity in European Languages*. Cambridge University Press, 337-358.

Noun phrase conjunction in Jalqan and Mitahi

Irina Khomchenkova
Lomonosov Moscow State University
V. V. Vinogradov Russian Language Institute of the Russian Academy of Sciences
Institute of Linguistics of the Russian Academy of Sciences

I describe noun phrase conjunction in Jalqan and Mitahi (Southwestern Iranian) – an underdescribed language spoken by the Muslim Tats of Dagestan, – using the data collected by elicitation in Dzhalgan, Nizhnij Dzhalgan and Mitagi-Kazmalyar villages in 2023 and 2024. Lexical cognacy data suggest that Jalqan and Mitahi are two dialects of one language (Koryakov 2022), but there are also significant grammatical differences between the two.

NP-conjunction is a cover term for constructions found in sentences that describe a single event with several individuals having the same semantic role: *Amir and Fuad left* (coordinating strategy) or *Amir left with Fuad* (comitative strategy). In the coordinating construction two NPs have identical syntactic functions and form a single constituent, in contrast to the comitative construction (Stassen 2000: 4, 21). This topic is poorly described for other Caucasian Tat languages as well; only basic facts are found in the grammars of Juhuri (Authier 2012: 59–60), Shirvan Tat (Suleymanov 2020: 244), Apsheron Tat (Mammadova 2017: 54–55) and North Azerbaijani Tat idioms (Gryunberg 1963: 51).

In Jalqan, NP conjunction is marked by the postpositive additive particle *=iš* (1a) and the comitative-instrumental marker *be* … *=(r)oz* (2a). In Mitahi the postpositive additive particle *=(e)ni* is used (1b), as well as the comitative-instrumental marker (*be*) … *=beyke* (2b). In some idiolects there is also the prepositive bisyndetic marker *hemi* (*hemi* X *hemi* Y).

(1) a. amir=**iš** marina=**iš** raft-o-yut
Amir=ADD Marina=ADD go$_2$-PRF-3PL
b. amir=**eni** marina=**ni** raxt-e-nd
Amir=ADD Marina=ADD go$_2$-PRF-3PL
'Amir and Marina left.'

(2) a. amir raft-e-y **be** fuad=**oz**
Amir go$_2$-PRF-3SG LOC Fuad=COM
b. amir raxt-e-s (**be**) fuad **beyke**
Amir go$_2$-PRF-3SG LOC Fuad COM
'Amir left with Fuad.'

When connecting NPs, *=iš* and *=ni* mark both conjuncts (1). Cross-linguistically, such constructions can involve both emphatic (such as *both X and Y*) and neutral coordination (Haspelmath 2007: 16). In Jalqan and Mitahi, this marker can be used with symmetrical predicates (3), which supports the neutral interpretation of such coordinating structures. Non-emphatic bisyndetic coordination is widespread in this region (see Stilo 2004: 317 on Western Iranian languages, van der Berg 2004: 215 on Nakh-Daghestanian languages).

(3) a. amir=**iš** fuad=**iš** daʕwo mi-st-en-ut
Amir=ADD Fuad=ADD fight EVT-do$_2$-IPFV-3PL
b. amir=**eni** fuad=**eni** daʕwo mi-st-an-d
Amir=ADD Fuad=ADD fight EVT-do$_2$-IPFV-3PL
'Amir and Fuad are fighting [with each other].'

Haspelmath (2004: 10–12) argues that conjunctions can only cover contiguous regions on the implicative sequence NP — AP — VP — S. In Jalqan and Mitahi, apart from NPs, the particles *=iš* / *=(e)ni* coordinate VPs and clauses, but not adjectives, see example (4) from Jalqan; in this case, asyndetic coordination is preferable. This seems to contradict the above hierarchy. However, it seems that the additive particle simply cannot attach to nominal dependents of this type.

(4) *amir ispohin=iš aqil-lü=iš gede=yu
Amir beautiful=ADD intelligence-PROP=ADD boy=3SG
'Amir is a handsome and smart boy.'

The Jalqan comitative-instrumental marker *be … =oz* is used both in the comitative construction proper (with singular verbal agreement, (2a)), and in its coordinating modification with plural verb agreement (5). The coordinating construction with this marker is limited to the subject function, cf. (6) with the indirect object where only *=iš* is grammatical. The similar situation is found in Mitahi, which I will discuss in the talk.

(5) amir be fuad=oz raft-o-yut
Amir LOC Fuad=COM go$_2$-PRF-3PL
'Amir and Fuad (lit. with Fuad) left.'

(6) me der-am xardeni be amir=iš be marina=iš /
1SG give-PRF.1SG food LOC Amir=ADD LOC Marina=ADD
*be amir be marina=roz
LOC Amir LOC Marina=COM
'I gave food to Amir and Marina.'

Interestingly, the system of NP-conjunction markers is different across Tat varieties. The particle *=iš* is also attested in Shirvan Tat (Suleymanov 2020: 244), Juhuri (Authier 2012: 59) and North Azerbaijani Tat idioms (Gryunberg 1963: 51). In Apsheron Tat the clitic *=(ə)m* is used instead (Mammadova 2017: 53). Apart from the additive particles, in some varieties there are also coordinating conjunctions, for instance, *ve* (from Arabic) and *ne* (from Lezgic languages) in Juhuri (Authier 2012: 60), *və* in Apsheron Tat (Mammadova 2017: 54). As for comitative(-instrumental) markers, in Juuri there is the marker *e … =(r)evoz* (Authier 2012: 109), and in North Azerbaijani Tat idioms – *bæ … -(r)az* (Gryunberg 1963: 30), which are similar to the Jalqan one. In Shirvan Tat the preposition *vo*/*ve* (Suleymanov 2020: 221) is used, and in Apsheron Tat, the preposition *boş* (Mammadova 2017: 107).

In the talk, I will describe the properties of the particles *=iš* and *=ni* in more detail, particularly, its functions beyond NP conjunction: additive and scalar focus, indefinite pronouns, concessives. These functions are typical of additive markers cross-linguistically (Forker 2016), but there are some differences between Jalqan and Mitahi. I will also analyze the properties of comitative constructions in these two languages based on Arkhipov (2009), paying particular attention at the coordination of pronouns. Finally, I will discuss the marker *hemi*, which shows great variability across speakers.

Abbreviations

$_2$ – the second verbal stem; 1, 3 – 1st, 3rd person; ADD – additive particle; COM – comitative; EVT – "eventive" marker (Authier 2012; Suleymanov 2020); IPFV – imperfective; LOC – locative; PL – plural; PRF – perfect; PROP – proprietive; SG – singular.

References

Arkhipov A. V. *Tipologija komitativnyx konstrukcij* [The typology of comitative constructions]. Moscow: Znak, 2009.

Authier G. *Grammaire juhuri, ou judéo-tat, langue iranienne des Juifs du Caucase de l'est*. Wiesbaden: Ludwig Reichert, 2012.

van der Berg H. Coordinating constructions in Daghestanian languages. *Coordinating constructions*. Haspelmath M. (ed.). Amsterdam: John Benjamins Publishing, 2004. Pp. 197–226.

Forker D. Toward a typology for additive markers. *Lingua*, 2016. Vol. 180. Pp. 69–100.

Gryunberg A. L. *Jazyk severoazerbajdžanskix tatov* [The language of North Azerbaijani Tats]. Leningrad: Publishing House of the USSR Academy of Sciences, 1963.

Haspelmath M. Coordinating constructions: an overview. *Coordinating constructions*. Haspelmath M. (ed.). Amsterdam: John Benjamins Publishing, 2004. Pp. 3–39.

Haspelmath M. Coordination. *Language typology and syntactic description. Volume II: Complex constructions*. Shopen T. (ed.). Cambridge: Cambridge University Press, 2007. Pp. 1–52.

Koryakov 2022 — Koryakov Yu. B. Derbent-area Tatic: sociolinguistics and affiliation. Presentation at *HSE ConLab seminar*, 04.10.2022.

Mammadova N. Eléments de description et documentation du tat de l'Apshéron, langue iranienne d'Azerbaïdjan. Linguistique. PhD thesis. Paris: Université Sorbonne Paris Cité, 2017

Stassen L. AND-languages and WITH-languages. *Linguistic Typology*, 2000. 4 (1). 2000. Pp. 1–54.

Stilo D. Coordination in three Western Iranian languages: Vafsi, Persian and Gilaki. *Coordinating constructions*. Haspelmath M. (ed.). Amsterdam: John Benjamins Publishing, 2004. Pp. 269–330.

Suleymanov M. *A grammar of Şirvan Tat*. Wiesbaden: Reichert, 2020.

* The research was supported by Russian Science Foundation, project No. 24-18-00199 at Institute of Linguistics of Russian Academy of Sciences.

Iamitives and semantically related markers in Northwest Caucasian

Evgeniya Klyagina (Institute of Linguistics of the Russian Academy of Sciences)
& Anastasia Panova (Stockholm University)

Iamitive is a term that has been established for grammaticalized 'already' markers in the languages of Papunesia and South-East Asia (Olsson 2013, Dahl & Wälchli 2016; François Forth.). In this talk, we want to contribute to this topic by discussing iamitive markers in the languages spoken outside of the mentioned areas. Specifically, we will focus on iamitive markers in Northwest Caucasian languages.

The Northwest Caucasian family includes Abkhaz and Abaza, constituting the Abkhaz-Abaza branch, and Kabaradian and Adyghe, constituting the Circassian branch. The category of iamitive is not mentioned in grammatical descriptions of Northwest Caucasian languages. However, all four languages have markers which at least partially behave as iamitives.

The marker showing the highest degree of resemblance to iamitive markers described by Ollson (2013) and Dahl & Wälchli (2016) is the suffix *-χ'a* in Abaza. It is compatible with stative verbs (1), regularly appears in natural development contexts (2) and also some contexts typical for perfect, cf. experiential perfect in (3).

(1) *a-pajš'* *ck'a* ***-χ'a-****ṗ*
DEF-room clean-**IAM**-NPST
'The room is already clean'.

(2) [A: Your brother is very handsome!]
B: *mamaw,* *awəj* *d-b-q-r-aštəl.*
no DIST 3SG.H.ABS-2SG.F.IO-LOC-CAUS-forget(IMP)
awəj *d-ʕa-j-g-****χ'a****-ṭ*
DIST 3SG.H.ABS-CSL-3SG.M.ERG-bring(AOR)-**IAM**-DCL
'No, forget about him. He is married.'

(3) *ananas* *ḳarazən*
pineapple ever
*j-b-fa-****χ'a****-ma?*
3SG.N.ABS-2SG.F.ERG-eat(AOR)-**IAM**-Q
'Have you ever eaten pineapples?'

In the iamitive-perfect grammatical space, the cognate suffix *-χ'a* in Abkhaz is closer to perfect. It is used in typical perfect contexts (4) and is not compatible with stative verbs (5) but, at the same time, it is forbidden in the context of sudden events, possible for perfect and impossible for iamitive.

(4) *bara* *znəḳər*
2SG.F ever
lezginka *b-ḳ^{w}aša-****χ'a****-z-ma?*
lezginka 2SG.F.ERG-dance-**IAM**-PST.NFIN-Q
'Have you ever danced lezginka?'

(5) **sar wəs-g'ə sə-pšʒa-**χ'a**-w-ṗ!*
1SG so-EMP 1SG.ABS-beautiful-**IAM**-IPF-NPST
#[A: Darling, put on your earrings to look prettier.'] 'B: I am already pretty'.

(6) **sara a-ṭʷəlawɥ=jə-šʷq̇ʷə s-čʷəʒχ'e-jəṭ*
1SG DEF-citizen-3H.IO-book 1SG.ERG-lose-PFCT-DCL
'I've lost my passport!' (What should I do?)

The Circassian 'already'-marker *-xe*, as well as the 'already'-marker *-ç̌'e* mentioned for Standard Kabardian in (Bagov et al. 1970: 133), predominantly appear in past perfective tenses. In Kuban and Besleney dialects of Kabardian the marker *-xe* is attested in present and future contexts but only occasionally (Somin 2012). In Adyghe the marker *-xe* is used in other tenses and moods but in the intensifying function (Rogava, Kerasheva 1966: 299–302). Despite these TAM restrictions, *-xe* and *-ç̌'e* are obviously not temporal but derivational markers.

To compare the frequencies of the discussed markers in texts, we calculated the number of their occurrences in the Gospel of Luke (Table 1). The results correspond well to our previous observations. We consider the Abkhaz *-χ'a* a tense marker, and Table 1 also shows that it is the most grammaticalized one. The derivational markers *-xe* and *-ç̌'e* in Circassian show expectedly low frequency. Interestingly, the frequency of the iamitive *-χ'a* in Abaza is lower than of the perfect but higher than the frequency of derivational 'already'-markers.

Table 1. The frequency of iamitives and semantically related markers in the Gospel of Luke.

marker	Abaza *-χ'a*	Abkhaz *-χ'a*	Adyghe *-xe*	Kabardian *-ç̌'e*
frequency	11	44	3	3

Thus, the Northwest Caucasian languages present a good example of the intragenealogical variation in the degree of grammaticalization of markers within the 'already'-iamitive-perfect space.

Abbreviations

1, 2, 3 — 1st, 2rd, 3rd person; ABS — absolutive; AOR — Aorist; CAUS - causative; CSL — cislocative; DCL — declarative; DEF — definite article; DIST — distal demonstrative; EMP — emphatic; ERG — ergative; F — feminine; H — human; IMP — imperative; IO — indirect object; IAM — iamitive; IPF — imperfective; LOC — locative; M — masculine; N — non-human; NFIN — non-finite; NPST — nonpast; Q — interrogative; PFCT — perfect; PST — past; SG — singular.

References

P. M. Bagov, B. Kh. Balkarov, T. Kh. Kuasheva, M. A. Kumakhov, G. V. Rogava (eds.). *Grammatika kabardino-cherkesskogo literaturnogo yazyka. Fonetika i morfologiya* [The Grammar of Kabardian. Phonetics and Morphology]. Moscow: Nauka, 1970.

Ö. Dahl, B. Wälchli. Perfects and iamitives: Two gram types in one grammatical space. *Letras de Hoje*. 2016. Vol. 51. Pt. 3. P. 325–348.

A. François. Negation in Dorig. In M. Miestamo & L. Veselinova (eds), in prep. *The Typology of Negation.* Language Science Press.

B. Olsson. *Iamitives: Perfects in Southeast Asia and Beyond*. MA thesis. Stockholm University, 2013.

G. V. Rogava, Z. I. Kerasheva. *Grammatika adygeyskogo yazyka* [Grammar of Adyghe]. Krasnodar, Maikop: Adyghe book publishing house, 1966.

A. A. Somin. *Vremya, vid i modalnost glagola v besleneyevskom dialekte kabardino cherkesskogo yazyka* [TAM system in Besleney Dialect of Kabardian Language]. BA thesis. Moscow: RSUH.

Yuri B. Koryakov
Institute of Linguistics RAS; National Research University Higher School of Economics
E-mail: ybkoryakov@gmail.com

Classification of Cautatic languages based on the lexical distances

The talk will consider an attempt to calculate the lexical distances between the idioms of the Caucasian Tat (Cautatic) group, which are spoken by Muslim Tats, Armeno-Tats and Mountain Jews of Azerbaijan and Dagestan. The "Cautatic" designation is used to distinguish them from North-West Iranian *Tatic* languages spoken in the northern Iran. Up to early 20th century the speakers were concentrated in Azerbaijan and Southern Dagestan but by now all Armeno-Tats and most Mountain Jews had left (to Russia, Israel, etc.) the area.

We use standard 100-item Swadesh lists with semantic specifications made by Kassian et al 2010 for calculations. Ten wordlists were compiled, using both the field data from Dagestan (Jalqan/Jaqlu, Mitahi/Mudei, Nyugdi Juhuri) and Armenia (Madrasa/Dprevank) and the data from published dictionaries and grammars (Derbent & Quba Juhuri, Northern Tat, Central Tat, Apsheron Tat & Shirvan Tat). Completely new data from almost extinct Madrasa variety spoken by Armeno-Tats in Dprevank, Armenia were collected in April, 2024 and is used in the work. Two wordlists (Standard Persian and Tajik) were used for external comparison.

Based on the calculations obtained, it is possible to postulate the existence of four to five clusters, the share of cognates between which is 84 to 89% and which, respectively, can be considered separate languages (Koryakov 2017): (Northern) Juhuri (Judeo-Tat), Mitahi-Jalqan, (Proper or Eastern) Tat (including Apsheron), Shirvan Tat and possibly Madrasa Armeno-Tat. The cognate share with modern Persian was significantly lower - circa 79% for all Cautatic lects.

The Muslim Tat dialects of Derbent area in Daghestan (Jalqan and Mitahi), the data of which were not previously used in scientific consideration, turn out to be a separate language with the distinct identity. The last two elderly speakers of Madrasa Armeno-Tat were found in Armenian village Dprevank and it appeared to be a separate variety more closer to Shirvan Tat and less so to Central Tat.

We still lack the data for Southern Judeo-Tat varieties still (previously) spoken in several locations (Oğuz, İsmayıllı, Şamaxı, etc.) west and south of Shirvan Tat area; it can possibly turn out to be a separate language too.

It is interesting to note that Northern Juhuri has the lowest cognate share with modern Farsi. Neighboring Mitahi and Jalqan have a higher cognate share. This apparently refutes the theory that Mountain Jews arrived later from Iran and switched to the local Cautatic variety.

References:
Kassian, Alexei, George Starostin, Anna Dybo & Vasiliy Chernov. 2010. The Swadesh Wordlist: An Attempt at Semantic Specification. *Journal of Language Relationship* 4: 46-89.

Koryakov, Yuri B. 2017. Проблема «язык или диалект» и попытка лексикостатистического подхода [Language vs. dialect: a lexicostatistic approach]. *Voprosy Jazykoznanija* 6: 79-101.

Lander Yury, Bagirokova Irina

Narrow focus markers in West Caucasian?

Cross-linguistically, markers of narrow focus often evolve from copulas, which originally combined with focus phrases in clefts and pseudoclefts (Harris & Campbell 1995; Harris 2001; Kuteva et al. 2019). This process does not seem to be reported for West Caucasian languages. Presumably the absence of apparent focus markers here is due to the fact that these languages still actively use pseudoclefts with a transparent structure: the focus combined with the copula appears as the predicate while the description of the situation comes as a (normally headless) relative clause functioning as the absolutive subject (1) (the focus is given in brackets); see, e.g., Sumbatova 2009; Kindlein 2016.

Temirgoi West Circassian

(1) *txəλxer* *zestəš'tə-r* [*sšəpχʷəxe-r* *arə*]
the.books whom.I.will.give-ABS my.sisters-ABS COP

'I'll give the books to MY SISTERS.' (Lit., 'Who I will give the books to are my sisters.')

In this talk, we argue that West Caucasian languages nonetheless display the process of developing monoclausal narrow focus constructions which may lead to the rise of focus markers out of the copulas.

We consider several varieties of West Circassian, Kabardian, Abkhaz and Abaza and compare two constructions, namely the externally headed pseudocleft (EHPC) as in (1) and the internally headed pseudocleft (IHPC), where the focused phrase together with the copula appears to be linearly embedded into the apparent relative clause, as in (2).

Temirgoi West Circassian

(2) *txəλxer* [*sšəpχʷəxe-r* *arə*] *zestəš'tə-r*
the.books my.sisters-ABS COP whom.I.will.give-ABS

'I'll give the books to MY SISTERS.' (Lit., 'Who I will give the books to are my sisters.')

Both kinds of pseudoclefts show some connectivity effects (like those described in Higgins 1979; Iatridou & Varlokosta 1997; Sharvit 1999) and hence probably already have some monoclausal features. However, EHPCs and IHPCs may differ in some properties:

- the scope of operators on the lexical predicate (formally, the head of the relative clause) – in some varieties evidential, epistemic and attitude operators on the lexical predicate may have scope over the focus primarily in IHPCs,
- tense marking on the lexical predicate and on the copula – sometimes the lexical predicate in IHPCs but not in EHPCs requires absolute temporal reference, while the lexical predicate in EHPCs may have temporal reference determined relatively to the tense of the copula,
- the possibility of non-absolutive case marking on the focus – IHPCs allow the focus marked with a role assigned within the relative clause more easily than EHPCs (though some Kabardian varieties perhaps allow this even in EHPCs),

- the capacity of the predicate of the apparent relative clause to combine with an external head nominal – normally IHPCs but not EHPCs allow the predicate of the apparent relative clause to combine with an external head nominal.

We suggest, then, that different kinds of apparent pseudoclefts develop into monoclausal structures to different extent (in particular, IHPCs often have more monoclausal properties) and this is observed to a different extent in different West Caucasian varieties, sometimes within a single language. The copula in clauses with more monoclausal properties can be probably interpreted as a narrow focus marker similar to copulas in East Caucasian languages (Kazenin 2002; Harris 2001; Forker 2020). The resulting structures, however, are still problematic for many syntactic theories, as they continue to simultaneously show monoclausal and biclausal properties and their development cannot be described as simple reanalysis.

References

Forker, Diana. 2020. Information Structure in Languages of the Caucasus. In Polinsky, Maria (ed.), *The Oxford Handbook of Languages of the Caucasus*, 973–1002. Oxford University Press. (doi:10.1093/oxfordhb/9780190690694.013.28)

Harris, Alice C. 2001. Focus and universal principles governing simplification of cleft structures. In Faarlund, Jan Terje (ed.), *Grammatical relations in change* (Studies in Language Companion Series 56), 159–170. Amsterdam: John Benjamins. (doi:10.1075/slcs.56.07har) (https://benjamins.com/catalog/slcs.56.07har)

Harris, Alice C. & Campbell, Lyle. 1995. *Historical Syntax in Cross-Linguistic Perspective*. 1st edn. Cambridge University Press. (doi:10.1017/CBO9780511620553)

Higgins, Francis Roger. 1979. *The Pseudo-cleft Construction in English*. New York, London: Garland.

Iatridou, Sabine & Varlokosta, Spyridola. 1997. Pseudoclefts crosslinguistically. *Natural Language Semantics* 6(1). 3–28. (doi:10.1023/A:1008272704531)

Kazenin, Konstantin I. 2002. Focus in Daghestanian and word order typology. *Linguistic Typology* 6(3). 289–316. (doi:10.1515/lity.2003.002)

Kindlein, David. 2016. Aspects of information structure in written Kabardian narrative. University of Gloucestershire, ms. (Unpublished manuscript.)

Kuteva, Tania & Heine, Bernd & Hong, Bo & Long, Haiping & Rhee, Seongha & Narrog, Heiko. 2019. *World Lexicon of Grammaticalization. 2nd ed.* New York, NY: Cambridge University Press.

Sharvit, Yael. 1999. Connectivity in specificational sentences. *Natural Language Semantics* 7(3). 299–339. (doi:10.1023/A:1008390623435)

Sumbatova, Nina. 2009. Kommunikativnaja struktura adygejskogo predloženija: perspektuva i fokus. In Testelets, Yakov G. (ed.), *Aspekty polisintetizma: očerki po grammatike adygejskogo jazyka*, 559–611. Moscow: RGGU.

Lin Chia-Wei

Title: Pattern replication and Semitic syntax as diagnostic features for identifying medieval Arabic-Georgian translations

Abstract:

Old Georgian Christian literature is significantly influenced by Greek and other Christian literary traditions in the neighboring regions, including Syriac, Armenian, and Arabic. Prior to the late 19th century, it was often uncritically assumed that the Old Georgian Christian texts were translated directly from Greek. Nonetheless, recent studies show that the picture is in fact much more complicated. On the one hand, the bibliographical survey of PATARIDZE (2013) points out that around a dozen of medieval Georgian historiographies and hagiographies, previously assumed to be translated directly from Greek, have now been proved to be translated via an Arabic intermediary. On the other hand, ALEKSIDZE (2021) demonstrates that two texts translated from Arabic into Georgian (*Balavariani* and *Abukura*) served as the source of the Greek version. PATARIDZE (2013: 48) notes that "[T]he philological criteria for identifying translations from Semitic languages into Georgian have never been well define."

The present study aims to bridge this gap described by PATARIDZE by providing a checklist of clearly defined linguistic features to determine if an Old Georgian text of unknown *Vorlage* is translated from Arabic. Based on the corpus listed in PATARIDZE (2013) and ALEKSIDZE (2021), the paper first summarizes the linguistic patterns that have already been proposed in the previous research to distinguish translations mediated through Arabic from those directly translated from Greek, including (a) mistranslations, (b) polysemy of Arabic words, (c) phonological structure of proper names, (d) "Semitic *da*".

(a) Similarity between certain Arabic graphemes (e.g. ج *ǧ*/ ح *ḥ*/ خ *ḫ*, ب b/ ي *y*/ ن n/ ت *t*) can lead to misreading by the Georgian translator. For instance, in the context where 'taste' is expected, the Georgian translation reads *ċamis ġopa* 'wink'. This likely indicates that the translator misread Arabic لمج *lamağa* 'to taste' (or its derivatives) as لمح *lamaḥa* 'to glance' (PEETERS 1921: 185).

(b) Arabic polysemy sometimes cannot be conveyed accurately in Georgian. For instance, Arabic *burǧ* 'tower; constellation' is translated into Georgian as *ķošķi* 'tower' (PEETERS 1921: 284), but the intended meaning 'constellation' cannot be conveyed in the Georgian translation.

(c) Proper names align more closely with their Arabic forms than with their Greek counterparts. For instance, in the story of *Balavariani*, the Georgian name იოდასაფ *Iodasap* resembles Arabic بوذاسف *Būḏāsaf* and its variant يوذاسف *Yūḏāsaf* rather than Greek Ιώασαφ *Iōasaph.*

(d) I use the term "Semitic *da*" to refer to Georgian conjunction *da* 'and' that serves as a literal translation of Arabic *wa* introducing a circumstantial (*ḥāl*) clause, even if in Georgian hypotaxis would be preferred over parataxis, e.g.

Greek:	*ékhōn*	*ónon*		(ed. SCHWARTZ 1939: 227)
	have.ACT.PTCP.PRS	donkey.ACC.SG		
	'having a donkey'			
Georgian:	*da* *mis*	*tana*	*saqedar-i*	(ed. GARITTE 1962: 419)
	and DEM.GEN.SG	with	donkey-NOM	
	'and with him a donkey'			

The Georgian translation seems to reflect the Arabic construction (though unattested here):

**wa-maʿa-hū*	*ḥimārun*
and-with-3SG.M	donkey
'and with him a donkey'	

The second part of the paper proposes further distinctive features of Arabic-Georgian translations found in the selected corpus: (e) morphological calque, (f) phraseological calque, (g) cognate object construction (*figura etymologica*), and (h) Arabic word order.

(e) In some cases, Georgian replicates Arabic morphological formation, e.g. Arabic *ḫāliq* 'creator' (active participle of √ḪLQ 'to create') is translated by an active participle with the *m-X-ul* circumfix:

da-m-bad-eb-ul-i (ed. GVARAMIA 1973: 41)
PVB-PTCP-create-THM-PTCP
'creator'

(f) An idiomatic expression in Arabic is translated literally into Georgian, e.g.

Arabic: *wa-huwa fī l-waqti bnu ṯamānī ʿašri sanatin* (MS Sinai Ar. 395, 126v)
and-he in time.GEN son eight-ten.GEN year.GEN.SG
'and he (was) at that time 18 years old' (lit. 'a son of 18 years')

Georgian: *da iq̇o igi ʒē atrvameṭ-isa ṣli-sa-y* (ed. GARITTE 1962: 411)
and was DEM.SG son eighteen-GEN.SG year-GEN-NOM
'and he was a son of 18 years'

(g) Cognate object construction is a salient feature of Arabic to express any kind of emphasis of the verb (FISCHER 1972: 172), but it is less common in Georgian. The following example illustrates the Georgian translator's attempt to replicate Arabic √*ḤZN* 'to be sad' with Georgian root *ṣux* '(be) sad':

Arabic: *ḥazinū ḥuzn šadīd* (ed. GVARAMIA 1973: 38)
be_sad.3PL.M.PF sadness strong
'They were very sad.' (lit. 'sad (with a) strong sadness')

Georgian: *še-ṣux-n-es ṣuxil-ita did-ita* (ed. GVARAMIA 1973: 38)
PVB-sad-OBJ.PL-3PL.AOR sadness-INS.PL great.INS.PL
'They were sad with a strong sadness.'

(h) Although Old Georgian is known for is relatively free word order, there is an obvious tendency to place the verbal predicate after the subject and object, positioning it at the end of a clause (FÄHNRICH 2012: 348-354). In Arabic-Georgian translations, some sentences closely follow Arabic VSO order.

With the examples illustrated above, the paper aims to establish a list of linguistic features extracted from the selected corpus of medieval Arabic-Georgian translations, which can be applied to Georgian manuscripts whose immediate source language are yet to be determined.

References:

ALEKSIDZE, Nikoloz. 'Translations I: From Other Languages into Greek, Section IV: Georgian'. In *The Oxford Handbook of Byzantine Literature*, edited by Stratis Papaioannou, 219–26. Oxford: Oxford University Press, 2021.

FÄHNRICH, Heinz. *Die georgische Sprache*. Leiden: Brill, 2012.

FISCHER, Wolfdietrich. *Grammatik Des Klassischen Arabisch*. Wiesbaden: Otto Harrassowitz, 1972.

GARITTE, Gérard. 'La Version Géorgienne de La Vie de Saint Cyriaque Par Cyrille de Scythopolis'. *Le Muséon* 75 (1962): 339–440.

GVARAMIA, Rusudan. ამონიოსის 'სინა-რაითის წმიდა მამათა მოსრვის' არაბულ ქართული ვერსიები *Amoniosis 'sina-raitis ṣmida mamata mosrvis' arabul kartuli versiebi*. Tbilisi: Mecniereba, 1973.

PATARIDZE, Tamara. 'Christian Literature Translated from Arabic into Georgian: A Review'. *Annual of Medieval Studies at CEU* 19 (2013): 47–65.

PEETERS, Paul. 'La Version Ibéro-Arménienne de l'autobiographie de Denys l'Aréopagite'. *Analecta Bollandiana* 39 (1921): 277–313.

SCHWARTZ, Eduard, ed. *Kyrillos von Skythopolis*. Leipzig: J.C. Hinrichs Verlag, 1959.

ცეცხლისანი თუ ცეცხლოვანი:
diachrony of the Object Modifier construction in Georgian

Diego Luinetti, *Università degli Studi Guglielmo Marconi (Roma)*

Keywords: Relational Adjectives, *Suffixausnahme*, Parts of Speech Diachrony.

Old Georgian is characterized by the so-called *Suffixausnahme* (Plank 1995): a strategy of agreement consisting in the extension of the case ending of the governing noun to its adnominal genitives, as it is shown in the following examples:

ორნი	ფრთენი	არწივისა	მის	დიდისანი (Ap. Jo. 12, 14)
*or-**n-i***	*prte-**n-i***	*arçiv-is*	*mis*	*did-isa-**n-i***
two-PL-NOM	wing- PL-NOM	eagle-GEN	DET.GEN	great-GEN- PL-NOM

'The two wings of the great eagle'

შვიდნი	ლამპარნი	ცეცხლისანი (Ap. Jo. 4, 5)
*švid-**n-i***	*lampar-**n-i***	*cecxl-isa-**n-i***
seven- PL-NOM	torch-PL-NOM	fire-GEN-PL-NOM

'Seven torches of fire'

Suffixausnahme is found exclusively if the adnominal genitive follows the governing noun, and it can be regarded as an optional strategy. While in Old Georgian *Suffixausnahme* is a common strategy, it progressively gets abandoned and disappears completely in spoken Modern Georgian (Fänrich 1982: §267; Boeder 1995, 1999; Gippert 2022).

In the noun phrases where *Suffixausnahme* takes place, the genitive has a modifying function in respect of the governing noun: however, it does not express a quality, rather it expresses some kind of relation to another entity designed by noun (*e.g.*, the possessor, the material, etc.). Such genitives have thus the same function associated with relational adjectives (Bisetto 2010), insofar as both strategies can be seen as ways of encoding Object Modifiers constructions (in the sense of Croft 2001 and Alfieri 2020). Moreover, both Old Georgian and Modern Georgian actually have relational adjectives too, which can be productively derived from nouns by means of specialized affixes (e.g., *-ian-i*, *-ean-i*, *-ier-i*, *-ul-i*, *-ovan-i*, *-osan-i*, *u-…-o*, etc.), as it is exemplified below (adapted from Fänrich 1982: §70):

მარილი →	მარილიანი	ცეცხლი →	ცეცხლოვანი
maril-i	*maril-**ian-i***	*cecxl-i*	*cecxl-**ovan-i***
salt-NOM	satl-ADJ-NOM	fire-NOM	fire-ADJ-NOM
'salt'	'salty'	'fire'	'burning'

Relational adjectives are thus obtained by means of a word-class-changing derivation (Aikhenvald 2011).

In this paper it is aimed at describing the variation of the encoding of the Object Modifier construction throughout the diachrony of the Georgian language, between the noun and the adjective word-class. Three diachronic segments will be taken into account (Old Georgian: *Ioanes Gamocxadeba*; Middle Georgian: *Vepxisṭqaosani*; Modern Georgian: *Data Tutašxia*): for each segment, a comparable portion

of text will be analyzed, extracting and classifying the first 300 occurrences of Object Modifier constructions. The kind of encoding strategy and the kind of relation expressed by each occurrence will be recorded. Then, the frequency of the *Suffixausnhame* strategy will be compared to the one of the relational adjectives and of the bare genitives in each segment considered.

Preliminary results show that that the portion of functional domain freed by the progressive loss of the *Suffixausnhame* strategy is taken over partly by bare genitives and partly by the expansion of the relational adjective class: in other words, the morphosyntactic encoding of the same functional concept shifts from an almost exclusively analytic nominal strategy to a split between an analytic nominal and a synthetic adjectival strategy. As a consequence, a wider variety of denominal adjectivalizers, with a higher degree of semantic specialization, is found in Modern Georgian, where relational adjectives become a widespread strategy of encoding Object Modifiers.

References

Aikhenvald, Alexandra. 2011. Word-class-changing derivations in typological perspective. In A. Aikhenvald & R.M.W. Dixon. *Language at Large*. Leiden / Boston: Brill. 221-289.

Alfieri, Luca. 2020. The lexicalization of the adjective class as an innovative feature in the Indo-European family. *Poznań Studies in Contemporary Linguistics* 56(3): 379-412.

Bisetto, Antonietta. 2010. Relational adjectives crosslinguistically. *Lingue e Linguaggio* 9(1): 65-85.

Boeder, Winfried. 1999. Attribution in Georgian. In H. van der Berg (ed.). *Studies in Caucasian Linguistics*. Leiden: Research School of Asian, African, and Amerindian Studies.

Boeder, Winfried. 1995. Suffixausnahme in Kartvelian. In F. Plank (hrsg.): 151-215. https://doi.org/10.1093/oso/9780195087758.003.0004

Croft, William. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.

Fänrich, Heinz. 1982. *Grammatik der Altgeorgischen Sprache*. Tbilisi. [translation of Šaniʒe, Aḳaḳi. *Elemenṭaruli çigni ʒveli kartuli enis šesçavlad*]

Gippert, Jost. 2022. Old Georgian "Suffixaufnahme" revisited. *Digital Kartvelology* 1.

Plank, Frans (hrsg.). 1995. *Double Case: Agreement by Suffixaufnahme*. Oxford University Press, New York. ISBN 0-19-508775-5

Scope variability in two classes of Georgian indefinite pronouns

Richard Luo (Yale University) & Natasha Thalluri (Harvard University)

Introduction: It has long been recognized that indefinites come with an existential quantificational meaning (Fodor and Sag, 1982; Heim, 1982; *a.m.o.*). For some languages such as Dutch, *wh*-words can function as both interrogative or existential indefinite pronouns, while in others such as English, existential indefinites are distinct items from *wh*-words. It is also very common cross-linguistically to find indefinites morphologically derived from *wh*-items as in Japanese, Greek, Hungarian, Icelandic, etc. (cf. English *somewhere/somehow*). We investigate this phenomenon in Georgian, which exhibits two different types of indefinites that are derived from *wh*-words. The first series is derived by adding the suffix *-ǧac*, and the second is formed by adding the suffix *-me*. We will show that the *me*-class and *gǎc*-class differ in their licensing requirements and scope taking properties.

(1) *Two classes of Georgian indefinite pronouns*

a. minda **vin-me** v-i-p'ov-o, vin-c lap'arak'-ob-s kartul-ad
need.PRS.1SG **who-IND1** 1-PV-find-OPT, who-REL speak-TS-PRS.3SG Georgian-ADV
'I need to find someone who speaks Georgian.'

b. minda **vi-ǧac** v-i-p'ov-o, vin-c lap'arak'-ob-s kartul-ad
need.PRS.1SG **who-IND2** 1-PV-find-OPT, who-REL speak-TS-PRS.3SG Georgian-ADV
'I need to find someone who speaks Georgian.'

Although the sentences in (1) are more or less synonymous, the two classes do not share identical distributions. *me*-class indefinites cannot be used in simple declaratives (2), while both are acceptable and have an indefinite interpretation inside of polar questions (3).

(2) šota-m icno **vi-ǧac/*vin-me**, vin-c a-t'ar-eb-d-a be-em-ve-s
Shota-ERG recognize **who-IND2/*IND1**, who-REL PV-ride-TS-IMPF-3SG BMW-DAT
'Shota recognized someone who was driving a BMW.'

(3) še-g-i-dzl-i-a mo-m-c-e **romel-ǧac/romeli-me** c'ign-i?
PVB-2-PV-be.able-PERF-SG PVB-1-give-OPT.2SG **which-IND1/IND2** book-NOM
'Could you pass me a book?'

The two classes also yield different interpretations when embedded under an attitude predicate and in the antecedent of conditionals — in both cases, we see that *ǧac*-class indefinites can refer to a specific individual, but *me*-class indefinites do not. (4a) has a reading where Keti asked if someone specific (whom the speaker knows) is coming. Similarly, (4b) could mean that if some person (who the speaker knows) comes to the party, Maia will be happy.

(4) a. keti-m m-k'itx-a, **vi-ǧac** mo-d-i-s tu ara
Keti-ERG 1-ask-AOR.3SG, **who-IND2** PVB-come-PV-PRS.3SG or not
'Keti asked me if someone is coming.'

b. ?tu **vi-ǧac** mo-v-a c'veuleba-ze, maia-s ga-u-xar-d-eb-a
if **who-IND2** PVB-come-FUT.3SG party-on, Maia-DAT PVB-PV-happy-INTR-TS-FUT.3SG
'If someone/at least one person comes to the party, Maia will be happy.'

In contrast, *me*-class indefinites cannot refer to specific individuals; (5a) means Keti asked if anybody is coming, and (5b) has the interpretation that if anyone comes to the party, Maia will be happy. In both cases, the speaker is not referring to any particular person in mind.

(5) a. keti-m m-k'itx-a, **vin-me** mo-d-i-s tu ara
Keti-ERG 1-ask-AOR.3SG, **who-IND1** PVB-come-PV-PRS.3SG or not
'Keti asked me if anyone is coming.'

b. tu **vin-me** mo-v-a c'veuleba-ze, maia-s ga-u-xar-d-eb-a
if **who-IND1** PVB-come-FUT.3SG party-on, Maia-DAT PVB-PV-happy-INTR-TS-FUT.3SG
'If anyone comes to the party, Maia will be happy.'

The Proposal: The difference between the classes of indefinites is a matter of **licensing** — *me*-class indefinites must be bound by an intensional operator, taking obligatory narrow scope, whereas the *ǧac*-class indefinites can take narrow or wide scope. The availability of wide scope for the *ǧac*-class is indicated by the possiblility of referential readings in (4), see Matthewson (1998). Note that (4b) demonstrates the ability for *ǧac*-indefinites to existentially scope out of islands, such as the antecedent of a conditional (cf. Ruys, 1992).

However, such readings are not always available for *ǧac*-class indefinites. In (1b), the *de re* reading in which *viǧac* refers to a specific individual that the speaker is looking for is absent. As in (1), the *ǧac*-class and *me*-class indefinites in (3) are also synonymous, requesting that the hearer pass the first book they find, not a particular book that the speaker has in mind. This restriction on *ǧac*-class indefinites dissuades an alternative analysis for them as *specific indefinites* — they cannot take exceptional scope out of intensional operators.

The *me*-class indefinites, on the other hand, are only restricted to narrow scope readings. While they are not licensed in simple declaratives (2), they do appear in other contexts like the complement of modals (1a), questions (5a), and the antecedent clause of conditionals (5b). In such contexts, they have a meaning similar to NPIs. Crucially, however, the *me*-class indefinites cannot appear in negative sentences — there is a distinct class of negative concord items in Georgian, e.g. *aravin* 'no one', used to express an equivalent meaning (6).

(6) mariam-s **ara-vin** da-u-nax-av-s c'veuleba-ze
Mariam-DAT **NEG-who** PVB-PV-see-TS-PERF.3SG party-on
'Mariam didn't see anyone at the party.'

We claim that *me*-class indefinites are only licensed under intensional environments, and being bound by an intensional operator forces them to take narrow scope. Another piece of supporting evidence is that across the possible *wh*-words in Georgian, there is one *me*-class indefinite which remains **unattested**, despite the existence of its *ǧac* counterpart (7). We argue that this is due to the fact that *why*-adjuncts, e.g. ReasonP, scope above the TP-domain where intensional operators apply (see Shlonsky and Soare, 2011 for a similar view), which would not allow a hypothetical **rat'omme* indefinite pronoun to be bound.

(7)

	ra 'what'	*romeli* 'which'	*vin* 'who'	*sad* 'where'	*rodis* 'when'	*rogor* 'how'	*rat'om* 'why'
ǧac-class	raǧac	romelǧac	viǧac	sadǧac	rodesǧac	rogorǧac	rat'omǧac
me-class	rame	romelime	vinme	sadme	(r)odesme	rogorme	***rat'omme**
MEANING	something	some(thing)	someone	somewhere	sometime/ever	somehow	for some reason

References.

• Fodor and Sag (1982). "Referential and quantificational indefinites". *Linguistics and philosophy* 5 • Heim (1982). *The semantics of definite and indefinite noun phrases* • Matthewson (1998). "On The Interpretation of Wide-scope Indefinites". *Natural Language Semantics* 7.1 • Ruys (1992). *The Scope of Indefinites*. OTS dissertation series • Shlonsky and Soare (2011). "Where's 'Why'?" *Linguistic Inquiry* 42.4

From ‘one’ to many: grammaticalization and lexicalization paths of the numeral ‘one’ in Agul (Nakh-Daghestanian)

Timur Maisak — *Institute of Linguistics RAS, Moscow & HSE University, Moscow*

Solmaz Merdanova — *Moscow State University of Psychology & Education*

Numeral ‘one’ is known to be the source or component of many grammatical markers and expressions: thus, the World Lexicon of Grammaticalization (Kuteva et al. 2019) lists nine recurrent grammaticalization paths with ‘one’ as the source (the targets are (1) ALONE, (2) INDEFINITE, (3) INDEFINITE PRONOUN, (4) ONLY, (5) OTHER, (6) SAME, (7) SINGULATIVE, (8) SOME, (9) TOGETHER). Nakh-Daghestanian are not an exception: ‘one’ is used in the languages of the family with a range of functions. At the same time, dedicated studies of the functions of ‘one’ in Nakh-Daghestanian seem to be lacking, and the present paper is aimed at filling this gap. We address the many functions of ‘one’ in one of the languages of the family, namely Agul (< Lezgic), although we take a broader intragenetic perspective, looking both at functions attested in Agul and closely related languages, as well as those functions that are known to exist in related languages but are absent in Agul.

One domain where ‘one’ plays a major role is that of **indefiniteness**. Although Nakh-Daghestanian languages are usually described as lacking articles, ‘one’ (usually, without phonetic reduction) is used in most of them as a kind of a weakly grammaticalized indefinite article, especially in introductory contexts. For Agul, *sa* ‘one’ has been analyzed as a presentational article by Becker (2018: 148–149). As an article, *sa* precedes a noun phrase, which can be singular or plural (1). Besides, *sa* can be used on its own as an NP head expressing an indefinite referent, e.g. *sa-d adi-ne* [one-SBST come.PFV-AOR] ‘someone came’. Also, *sa* is used, albeit optionally, as a preposed marker of negative indefinite expressions like *(sa) kas=ra* [one person=ADD] ‘no one’, *(sa) fi=ra* [one what=ADD] ‘nothing’, etc.

(1) čʷe χul.a-s sa qːari / qːari-jar šu-ne.
your.PL house-DAT one old_woman old_woman-PL go.PFV-AOR
‘Some (unknown) old woman / old women went to your house.’

Another function related to indefiniteness marking is **approximation**: preposed to a numeral phrase, *sa* expresses approximate number, e.g. *sa ic’u gur* [one ten bowl] ‘about ten bowls’ (cf. a similar use of the indefinite *some* in English: *some ten bowls*). ‘One’ is also found in expressions of approximate quantity like *sa č’uq’*, *sa bic’i*, *sa išan* ‘a little’ and *sa χali* ‘quite a few’.

A reduplicated stem *sa~sa* is used with the meaning ‘some of’ (2). This use may be related to the semantics of distributivity associated with reduplication of cardinal numerals (cf. *sa~sa jaʁluqː* ‘one headscarf each’, e.g. when each woman is given a headscarf as a gift[1]).

(2) sa~sa ħač-ar ut’u-naji.
RDP~one apple-PL rot.PFV-PRF.PST
‘Some of the apples got rotten (i.e. certain apples did and certain did not).’

The **reciprocal** pronoun in Agul is also based on the repetition of ‘one’, although the structure is different: here, each of the two copies of *sa* bears the case of the respective argument. Thus, in (3) the first part of the reciprocal is in the absolutive case (experiencer: who), whereas the second one is in the dative (stimulus: at what), the case encoding being determined by the verb *χutːurfas* ‘look’.

(3) ʕut’a-j, ʕut’a-j, sa-d sa-j.i-s χutːurfa-a...
eat.IPFV-CVB eat.IPFV-CVB one-SBST one-SBST-DAT look.IPFV-PRS
‘While eating, they keep looking at each other.’ (text corpus)

[1] Note that distributive cardinal numerals only combine with the singular form of the noun (e.g., *ʔu~ʔu ħač* ‘two apples each’, *xi~xibu ħač* ‘three apples each’ etc.), which is a general rule for cardinal numerals in Agul.

As a particle, *sa* is used with the **restrictive** meaning ‘only’ (4) and also with a more general **emphatic**/intensifying function (5).

(4) sa bagah za-s waχtː a-a.
one tomorrow I-DAT time IN.be-PRS
‘I have time only tomorrow.’

(5) sa lik’a-j-e gi har jaʁ.a!
one write.IPFV-CVB-COP DEM(ERG) every day(TMR)
‘He writes so much every day!’

Besides looking at the use of *sa* (i.e. the bare form of the numeral ‘one’), we take into account various lexicalized expressions, which seem to include *sa* diachronically. Thus, attributive pronouns *saje* and *sara* (and also *sasra*, which seems to be a fusion of *sa sara*), which all mean ‘other’, are based on ‘one’, which reflects a cross-linguistically common path (Lander & Maisak 2022). In particular, *sara* is historically a lexicalized combination of ‘one’ and the additive enclitic *=ra* ‘also, even’. This derivation is interesting in that *sara* is used not only as an adjective (6), but also as an adverb ‘more, else’ (7). Probably, it is on the basis of the adverb that a pragmatic particle *sara* has developed, which is found, in particular, in mild requests (8).

(6) qu-χa-a uč.i-s sara xir dad.a.
RE-bring.IPFV-PRS self-DAT other wife father(ERG)
‘Then Father brings another wife.’ (text corpus)

(7) han, sara fi qatːq’a?
PTCL more what tell.IPFV
‘Now, what else can I tell?’ (text corpus)

(8) jada pu-ne zun, ma-hata sara zun armi-s...
INTJ say.PFV-AOR I PROH-send.IPFV PTCL I army-DAT
‘Hey, – said I, – please don’t send me to the army...’ (text corpus)

Even given an impressive list of functions associated with *sa* ‘one’ in Agul, it is not the case that Agul possesses all the developments of this numeral attested in the other languages of the family. Thus, in Udi a combination of ‘one’ with the additive clitic (*saal*) can be used as a monosyndetic coordinating device (‘X and Y’), which is not the case of its Agul counterpart (*sara*). In the talk, we will draw parallels between Agul and the related languages, showing both similarities and differences between them.

Abbreviations
(other than those found in the Leipzig Glossing Rules)

AOR – aorist; IN – localization ‘inside’; INTJ – interjection; PTCL – particle; RDP – reduplication; RE – repetitive; SBST – substantivizer; TMR – temporal.

References

Becker L. 2018. *Articles in the world’s languages*. PhD dissertation. University of Leipzig.
Kuteva T., B. Heine, B. Hong, H. Long, H. Narrog, S. Rhee. 2019. *World lexicon of grammaticalization*. 2nd, extensively revised and updated edition. Cambridge: Cambridge University Press.
Lander Y., T. Maisak. 2022. “Other” Strategies in the Eastern Caucasus (Part II): Typology. *Iran and the Caucasus* 26(3), 272–288.

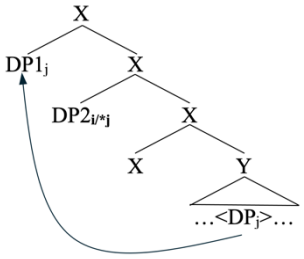
Argument Structure and Binding in Tsova-Tush[1]

Martha McGinnis, Léa Nash, Elmiskhan Kunizheva, Diana Kakishvili, Nino Amiridze, Sascha Alexeyenko

The distribution of reflexive and reciprocal anaphors in Tsova-Tush (Northeast Caucasian; Nakh) is complex, allowing the reversal of c-command asymmetries in some contexts but not others. Internal arguments in a ditransitive clause show asymmetrical Theme>Goal binding, but symmetrical binding of genitive possessors is found for certain other argument structures. We propose an analysis of these observations involving locality and Lethal Ambiguity (McGinnis 2004). A NOM or DAT internal argument merges lower than an ERG external argument, but can undergo A-movement over it, allowing limited new binding possibilities.

Each constituent is labelled with a projection of its head; a phrase's ADDRESS is the label of the constituent it merges with. The copy of an A-moved DP cannot be successfully linked to its antecedent if the moved DP (DP1) shares both a referential index and a syntactic address with another argument (DP2). Such a configuration results in a lethal ambiguity (LA) that causes the derivation to crash at LF (1). Thus, in derivations of this sort, a moved argument (DP1) cannot bind the argument it moves over (DP2)—but can bind an anaphor contained in DP2, since this has a different address.

(1)



This analysis captures a wide range of binding contrasts in Tsova-Tush. For example, a DAT experiencer subject can bind a NOM object (2), regardless of word order (Hauk & Harris 2018:18). The gender affix in (2) indexes the NOM reciprocal, which agrees in gender with its feminine antecedent.

(2) oqarn d-abc' vaša^{n}.
3PL.DAT F.PL-know RECP.NOM
'They (F) know each other.' (DAT>NOM, *NOM>DAT)

However, Hauk & Harris note that such verbs do not allow a NOM object to bind a DAT subject. This follows if the DAT argument merges in the specifier of a high Experiencer applicative phrase (ApplExpP), and the NOM argument can only move over the DAT one by moving into or through a higher spec-ApplExp. This would be required if ApplExpP is a phase, and movement through its specifier is required by phase impenetrability (McGinnis 2004), or because the NOM DP can only move over the intervening DAT DP via spec-ApplExp because of relativized locality. Either way, if NOM and DAT DPs share an index and an address at any stage, an LA results. This analysis correctly rules out a reversal of the binding relations in (2). (The prediction that a NOM object can bind a possessor contained in the DAT subject has yet to be tested.)

Similarly, a higher internal argument can bind a lower one (3). Here the NOM Theme is the higher argument, and can bind a DAT Goal anaphor. The asymmetry is predicted if these structures involve a low DAT Goal that merges below the Theme. It was impossible to elicit examples similar to (3) but with the Goal binding a Theme anaphor: when asked to translate Russian examples involving the Theme binding the Goal or vice versa, speakers invariably translated them with the Theme binding the Goal. One speaker expressed a distinction by replacing DAT *-n* on the Goal with ALL *-gǒ*; another rejected such examples. No speaker produced examples with a Goal binding a NOM anaphor.

(3) diana-s [nodar=a giuirg=a] vašban b-abc'-b-i-e^{n}.[2]
Diana-ERG Nodar.NOM=ADD George.NOM=ADD RECIP.DAT M.PL-know-M.PL-TR-AOR
'Diana introduced Nodar and George to each other.' (NOM>DAT, *DAT>NOM)

This binding asymmetry also obtains when the bound element is a genitive possessor. (4) shows the NOM DO binding a genitive (GEN) possessor contained in the DAT IO. As above, it was impossible to elicit examples similar to (4) but with the IO binding a GEN possessor contained in the DO; speakers invariably translated such examples in Russian into Tsova-Tush examples with the Theme binding a possessor contained in the Goal. These asymmetries suggest that there is no way for an IO to A-move over the DO.

---

[1] We use 'Tsova-Tush' following Wichers Schreur (2024). Some speakers prefer 'Tush.' 'Bats'/'Batsbi' is also used.

[2] Abbreviations: 1SG=first person singular; ADD=additive; ALL=allative; AOR=aorist; B, D=(gender agreement); DAT=dative; ERG=ergative; F=feminine; GEN=genitive; M=masculine; NOM=nominative; OBL=oblique; PL= plural; PFV=perfective; PRES=present; POSS=possessive; RECIP=reciprocal; SF=stem formant; REFL=reflexive; TR=transitive.

This is predicted if they both merge as arguments of a low applicative head (Pylkkänen 2008): any higher head will target the closer DO, with no way to attract IO to a specifier position above DO (McGinnis 2004). If clauses such as (4) are indeed ungrammatical, this constitutes a surprising contrast with parallel examples in French, which allow either the Goal to bind the Theme or vice versa (Boneh & Nash 2012, ex.(22)).

(4) as [levan=a giuirg=a] vašba^{n} nan-i-n b-abc'-b-i-n-as.
1SG.ERG L.NOM=ADD G.NOM=ADD RECIP.GEN mother-PL-DAT M.PL-know-M.PL-TR-AOR-1SG.ERG
'I introduced Levan and George to each other's mothers.' (NOM>DAT, *DAT>NOM)

The situation is different in transitive clauses. An ERG subject can bind either the NOM DO itself (5a) or a GEN possessor contained in it (5b). However, the DO can also bind a GEN possessor contained in the ERG DP (5c), regardless of word order. This suggests that the DO can A-move over the ERG argument. A similar possibility exists in Georgian, but only when the DO precedes the ERG DP (Amiridze 2006:129).

(5) a. [levn-es=a nodr-es=a] vašba^{n} keba-d-b-i-e^{n}.
Levan-ERG=ADD Nodar-ERG=ADD RECIP.NOM praise-SF-M.PL-TR-AOR
'Levan and Nodar praised each other.' (ERG>NOM)

b. [levn-es=a nodr-es=a] [vaša^{n} badr-i] keba(-d)-d-i-e^{n}.
Levan-ERG=ADD Nodar-ERG=ADD RECIP.GEN child-PL.NOM praise-SF-D-TR-AOR
'Levan and Nodar praised each other's children.' (ERG>NOM)

c. [vaša^{n} badr-i-v] keba-d-b-i-e^{n} [levan=a nodar=a].
RECIP.GEN child-PL-ERG praise-SF-M.PL-TR-AOR Levan.NOM=ADD Nodar-NOM=ADD
'Each other's children praised Levan and Nodar.' (NOM>ERG)

An ERG subject can also bind a DAT reflexive anaphor (6a), or a DAT reciprocal DP (6b), regardless of word order. It was impossible (in the sense described above) to elicit examples like (6b) with a DAT DP binding an ERG reciprocal; this might suggest that a DAT DP cannot A-move over an ERG one. Amiridze (2006:118) shows that, likewise, a non-subject cannot bind a subject reciprocal in Georgian.

(6) a. aħ ħui-n ejɬ-n-aħ.
2SG.ERG 2SG.REFL-DAT say.PFV-AOR-2SG.ERG
'You said (it) to yourself.' (ERG>DAT)

b. vašba-n d-aɣup'a(-d)-d-ie^{n} [mara-v pst'uina-v].
RECIP-DAT D-destroy-SF-D-AOR husband-ERG wife-ERG
'The husband and wife destroyed each other.' (ERG>DAT, *DAT>ERG)

However, a DAT DP can bind an ERG reflexive, as in (7), suggesting that a DAT internal argument can indeed A-move over an ERG one. I propose that the contrast between (6b) (which disallows DAT>ERG binding) and (7) (which permits it) follows from LA. Crucially, while a reciprocal is coindexed with its antecedent (as shown by gender agreement in (5a)), a reflexive based on *kortŏ* is not; instead, it is the reflexive possessor (*šer* in (7)) that is coindexed with its antecedent. This coindexation pattern accounts for the absence of gender agreement between antecedent and the anaphor in (8), where the gender suffix (type B) instead reflects the class of the noun *kortŏ* 'head'. On this analysis, (7) is parallel to (5c), where a lower argument A-moves over a higher one and binds a possessive anaphor contained in it. It is crucially distinct from the ill-formed counterpart of (6b), where a lower argument A-moves over a higher one and directly binds it. The morphosyntactically simple form *ħuin* in (6a) is predicted to agree in gender with its antecedent and trigger LA effects, parallel to the reciprocal in (6b). Another possible analysis is that (7) allows a non-agentive reading and (6b) does not; this reading permits ergative anaphors in Georgian (Amiridze 2006:222).

(7) miše-n [šer korti-v] d-i-e^{n}.
Misha-DAT 3SG.REFL.POSS.OBL self-ERG D-do-AOR
'Misha did it to himself.' (DAT>ERG)

(8) [(se^{n}) kort] keba(-d)-b-o-s.
1SG.GEN head praise-SF-B-PRES-1SG.ERG
'I praise myself.' (Hauk & Harris 2018:37) (ERG>NOM)

References

Amiridze, N. 2006. *Reflexivization Strategies in Georgian*. LOT Dissertation Series 127. UIL, Utrecht University.

Boneh, N. & L. Nash. 2012. Core and noncore datives in French. *Variation in Datives*, ed. B. Fernandez & R. Etxepare. Oxford UP.

Hauk, B. & A. C. Harris. 2018. Batsbi. To appear in *The Caucasian languages: An international handbook*, ed Y. Koryakov, Y. Lander, & T. Maisak. HSK Series. De Gruyter Mouton.

McGinnis, M. 2004. Lethal ambiguity. *Linguistic Inquiry* 35:1, 47–95.

Wichers Schreur, J. 2024. Intense language contact in Eastern Georgia: The case of Tsova-Tush. Doctoral dissertation, Johann Wolfgang Goethe-Universität at Frankfurt am Main.

Suprasegmental emphasis in vowels of Arzni Christian Urmi Neo-Aramaic

Maksim Melenchenko (HSE, Moscow) & Varvara Petrova (MSU, Moscow)[1]

This study investigates the effect of suprasegmental emphasis on acoustic properties of vowel phonemes in one of the varieties of Christian Urmi Neo-Aramaic (‹ Semitic). Christian Urmi was traditionally spoken in northwestern Iran, but due to migrations there are now diasporic communities in Iraq, Turkey, Armenia, Azerbaijan and southwestern Russia. In Krymsk (Krasnodarsky Krai, Russia), there are several such communities, the largest of which consists of recent resettlers from Arzni (Armenia). Our consultants for the current work were native speakers of the Arzni Christian Urmi in Krymsk (ACU). The Arzni variety is reported to be different from those of other Armenian villages in some linguistic features [Khan 2016: 28].

As well as other Northeastern Neo-Aramaic languages, Christian Urmi exhibits lexically determined suprasegmental pharyngealization (“emphasis”). G. Khan described this phenomenon for the varieties of Christian Urmi spoken in Iran [Khan 2016: 50–91]. Diachronically, this feature originated in emphatic consonants, which are nowadays lost. The phonological domain of emphasis is a word. According to Khan [2016: 50], in vowels emphasis is manifested as tendency to retraction and lip-rounding. Retraction affects some vowels more than others: in the majority of Christian Urmi varieties, only /a/ and /ə/ are significantly affected [Khan 2013: 116]. The goal of our study was to measure the effect of emphasis on acoustic properties of vowels in the ACU and compare it to the one described for Iranian Christian Urmi by Khan.

Three speakers of the Arzni community were asked to pronounce a list of words, which was compiled using Christian Urmi-Russian dictionary published in [Khan 2016]. The word list contains 32 stimuli, which cover all six vowel phonemes (/aeiouə/) in stressed word-medial, unstressed word-medial and word-final positions in words with and without the emphasis. The positions were selected to match those discussed in [Khan 2016] for the purposes of subsequent comparison. The speakers were instructed to pronounce each word three times. After that, they were asked to pronounce the word in a pre-determined carrier phrase “Say [word] three times” (also thrice, resulting in 2 types of environment × 3 repetitions = 6 tokens per word). The sessions were recorded in a quiet room, using a Zoom H5 recorder and a Shure WH20XLR microphone. The segmentation was performed using Praat [Boersma, Weenink 2024]; the formant values for the vowel midpoint were received through the Fast Track algorithm [Barreda 2021] in Python (*fasttrackpy*). The ANOVA test was used to determine statistically significant differences (Table 1).

Results show that the tendency for emphasized vowels to be retracted in comparison to their non-emphasized counterparts is present in ACU and extends beyond /a/ and /ə/. In contrast to the lack of significant retraction of /eio/ reported for Iranian Christian Urmi [Khan 2016: 76], we found that the F2 values of emphatic realizations in ACU are significantly smaller ($p < 0.01$) than those of plain sounds in both pairs of /i/–/+i/ and /e/–/+e/ in all positions. However, F2 has weaker overall correlation with the contrast of /u/–/+u/ ($p > 0.05$ in the stressed medial position, with almost complete neutralization for one of the speakers). F1 values also contribute to the contrast in emphasis of /i/ in all positions ($p < 0.01$), /e/ and /a/ in non-final positions. For /a/, the nature of contrast differs depending on its accentual marking: stressed plain /a/ is significantly lower and unstressed plain /a/ is significantly higher than their respective emphatic counterparts. This can be interpreted as a stark reduction of the plain /a/ in the unstressed position and establishes F1 as a cue not for emphasis but for the prosodic status of the syllable (Fig. 1). These results present a picture quite different from that of Khan [2013; 2016] for Iranian Christian Urmi. However, because the more careful tracing of vowel environment was implemented in our research, whether these differences are caused by dialectal variation or methodological improvement remains a question for further studies.

[1] The study was implemented in the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University) in 2024.

Figure 1. Vowels /aəeiou/ by position (stressed medial / unstressed medial / final), normalized. The vowel /ə/ does not appear word-finally. All word-final vowels are unstressed, except /e/. The ellipses cover 65% of the data.

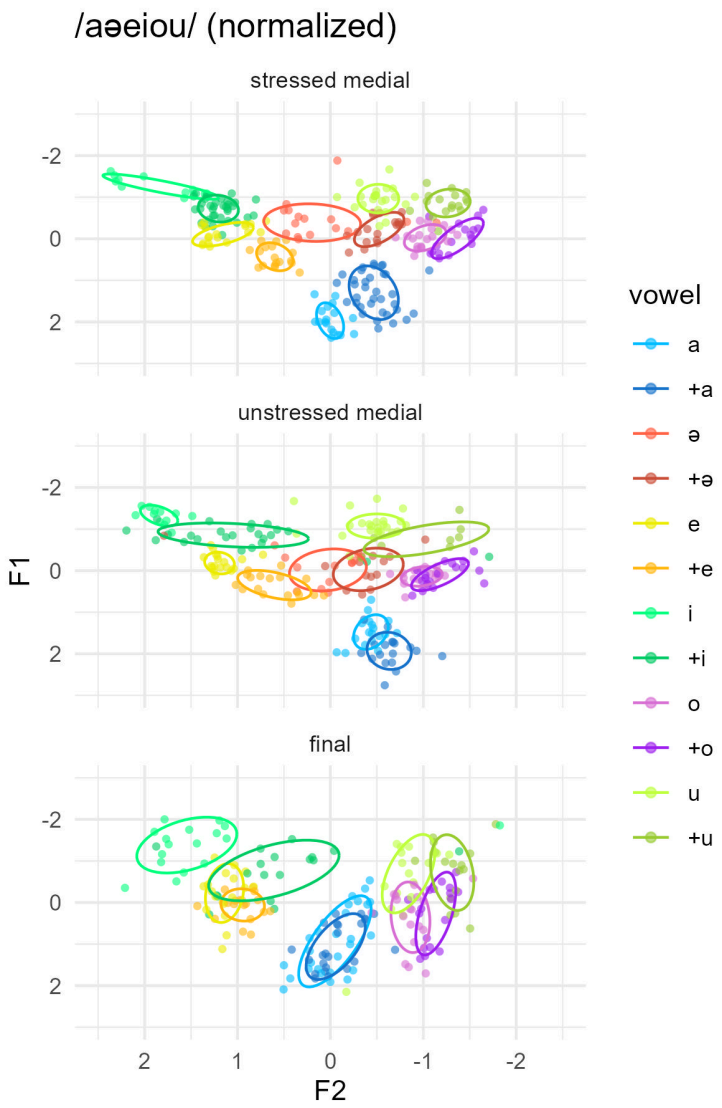


Table 1. Results of the ANOVA test for the association of emphasis with F1 and F2 formants; "ns" stands for 'not significant' ($p > 0.05$); ** stands for $0.001 < p < 0.01$, *** stands for $p < 0.001$.

	stressed medial		unstressed medial		final	
	F1	F2	F1	F2	F1	F2
a	***	***	***	***	ns	ns
ə	ns	***	ns	ns	—	—
e	***	***	***	***	ns	**
i	***	***	***	***	**	***
o	ns	***	ns	**	ns	***
u	ns	ns	**	ns	ns	***

References

Barreda, Santiago (2021). Fast Track: fast (nearly) automatic formant-tracking using Praat. Linguistics Vanguard, 7(1).

Boersma, Paul & Weenink, David (2024). Praat: doing phonetics by computer. Version 6.4.19. http://www.praat.org/

Khan, Geoffrey (2013). Phonological Emphasis in North-Eastern Neo-Aramaic. J. L. Léonard & S. Naïm (eds.) Base Articulatoire Arrière. Backing and Backness (pp. 111–132). Lincom Europa.

Khan, Geoffrey (2016). The Neo-Aramaic Dialect of the Assyrian Christians of Urmi. Brill.

Contact and innovation in the diachrony of East Caucasian numeral systems

Maksim Melenchenko, Linguistic Convergence Lab, HSE (Moscow)[1]

The Caucasus is one of the linguistic areas abundant with vigesimal (base-‘20’) systems, but it also has many languages with decimal (base-‘10’) systems. Both systems are found in the three endemic families of the region: Kartvelian, West Caucasian and East Caucasian. In vigesimal systems of the Caucasus, ‘20’ is used as the multiplicand base and ‘10’ as an additive base (e.g. ‘72’ in such systems is expressed as [3×20] + [10+2]). For proto-East Caucasian the vigesimal system is traditionally believed to be the “original” system (i.e., to have existed in the proto-language) [Alekseev 1985: 68–69], while decimality is usually ascribed to contact influence from other languages (e.g. Turkic) [Chirikba 2008: 58–60]. In the talk, I discuss the distribution of numeral systems in East Caucasian languages as well as several specific features of these systems and their implications for the diachrony. The study is based on an overview of a wide variety of grammars and articles on East Caucasian languages. The presented data challenges the traditional view that the vigesimal system is more ancient than the decimal one and raises the issue of “contact influence” as a universal explanation in studies of numeral systems.

branch	**numeral systems**
Nakh branch	**vigesimal**
Andic branch	mostly **decimal** (exceptions due to contact: Akhvakh, Tukita)
Avar	mostly **vigesimal** (exceptions in several dialects, in many cases due to contact)
Tsezic branch	**both** (**vig**. in West Tsezic and **dec**. in East Tsezic)
Lak	**decimal**
Dargic branch	**decimal**
Lezgic branch	**both** (**vigesimal** in Budukh, Kryz, Lezgian, Udi, southern dialects of Agul and Tabasaran, **decimal** in other lgs)
Khinalug	**vigesimal**

One of the arguments which Alekseev [1985: 68–69] provides in support of the hypothesis that the vigesimal system is more archaic for the Lezgic branch (and the whole East Caucasian family) is that the vigesimal system is “widespread” in the family. However, as the table on the right shows, the number of languages with vigesimal and decimal systems across branches is roughly equivalent.The majority of the branches adhere strictly to one of the two types of systems, but there are exceptions, which are often explained by language contact. For example, Andic languages are predominantly decimal, but Akhvakh and Tukita have vigesimal systems, apparently due to intense influence from Avar. Avar is one of the largest languages of Dagestan, spoken on a vast territory, and it is often mentioned as an example of a language with a consistently vigesimal numeral system. Indeed, this holds true for its standard variety, however, some dialects of Avar have a decimal or “mixed” decimal-vigesimal system. In some cases the divergence from the standard language is also caused by contact with neighboring Andic and Lezgic languages, but in others it is not as clear. For example, in one of the southern varieties (Keleb), which is not geographically close to any other language, the presence of a decimal system is not as clearly explainable by outside influence [Mikailov 1959: 422–423].

The distribution of the two numeral systems in the Tsezic branch corresponds to the genealogical subdivision into West Tsezic (two~three languages with vigesimal systems) and East Tsezic (three languages with decimal systems), with the two groups divided by a mountain range.

[1] The study was implemented in the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University) in 2024.

Curiously, the distribution of numeral systems in these groups is exactly opposite to those of their nearest and most influential neighbors: West Tsezic languages border Andic languages (which have the decimal system), while East Tsezic languages border Avar (which has the vigesimal system). Therefore, this divergence is not easily ascribed to contact influence.

Lezgic languages are the most diverse branch with regard to numeral systems. Roughly a half of the languages of the family have a vigesimal system, and another half decimal. The numeral systems of Agul and Tabasaran show dialectal divergence: northern varieties have decimal systems and southern vigesimal systems. This divergence is most easily explained with contact influence from Lezgian, a large and influential language with a vigesimal system spoken to the south of Agul and Tabasaran. Moreover, the decimal systems of Agul and Tabasaran share several morphological features with Rutul, another Lezgic language, which is not their closest relative or neighbor. This implies that these decimal systems are of shared origin, and in the past, Agul and Tabasaran had decimal systems. Among other Lezgic languages, the vigesimal system is present in southernmost languages (Budukh, Kryz, Udi), which have been otherwise strongly influenced by Azerbaijani. This, again, shows that contact influence (e.g. with Turkic languages) cannot straightforwardly explain the contemporary distribution of the systems [Melenchenko 2024].

These and other facts, which will be presented in the talk, show that the hypothesis of the “originality” of the vigesimal system is not supported by the collected data. I argue that the modern genealogical and areal distribution of the systems does not allow us to make conclusions about the “original” East Caucasian system at this stage of research, and both decimal and vigesimal systems should be treated as viable candidates for this role. Moreover, they show limitations of the contact explanation, which is often evoked in the literature on numerals. It is true that numeral systems are borrowed rather easily. However, the mountainous terrain and the prevalence of a single language family (East Caucasian) make Dagestan a complex area whose history is really hard to pinpoint when it comes to such easily borrowable phenomena. These facts also emphasize the importance of innovation as an alternative explanation for changes in numeral systems. While some scholars even claim that decimal systems cannot spontaneously change into vigesimal systems [Edelman 1999: 232–233], there are cases of such changes, which are apparently not motivated externally (e.g. Danish [Seiler 1990: 202]). In the case of East Caucasian languages, which are spoken in small and highly individualized Dagestani communities, the possibility of spontaneous innovation should not be overlooked.

References

Alekseev 1985 — Alekseev, Mikhail. 1985. *Вопросы сравнительно-исторической грамматики лезгинских языков. Морфология. Синтаксис* [Problems of comparative grammar of Lezgian languages. Morphology. Syntax]. Moscow: Nauka.

Chirikba 2008 — Chirikba, Viacheslav. 2008. The problem of the Caucasian Sprachbund. Pieter Muysken (ed.), *From Linguistic Areas to Areal Linguistics*, 25–93. John Benjamins.

Edelman 1999 — Edelman, Joy. 1999. On the history of non-decimal systems and their elements in numerals of Aryan languages. Jadranka Gvozdanović (ed.), *Numeral types and changes worldwide*, 221—242. Berlin, New York: De Gruyter Mouton.

Melenchenko 2024 — Melenchenko, Maksim. 2024. Numeral systems in Lezgic languages. *Folia Linguistica* (ahead of print).

Mikailov 1959 — Mikailov, Shikhabudin. 1959. *Очерки аварской диалектологии* [Essays on Avar dialectology]. Moscow, Leningrad: Press of the Academy of Sciences of the USSR.

Seiler 1990 — Seiler, Hansjakob. 1990. A dimensional view on numeral systems. *Studies in Typology and Diachrony. Papers Presented to Joseph H. Greenberg on his 75th Birthday*. Amsterdam, Philadelphia: John Benjamins.

N. Mezane KONUK
Université Sorbonne Nouvelle - mezanekonuk@gmail.com
Correlative constructions in Abzakh – a West Circassian dialect

Correlative constructions were first studied in Indo-Aryan languages, where a sentence-initial relative clause has a demonstrative correlate in the main clause as a common variety of the adjoined relative clauses (Lehmann 1986). The example (1), from Hindi (Indo-Aryan), illustrates a canonical correlative construction:

(1) Hindi (Srivastav 1991, 639)

[jo	laRkii	khaRii	hai]$_k$	vo$_k$	lambii	hai
REL	girl	standing	is	**that**	tall	is

"The girl who is standing is tall." (lit. '**which** girl is standing, **that** is tall.)

However, recent typological work has raised the issue of morphosyntactic differences among correlative constructions. Various classification methods for typological classification have been proposed so far, such as relativization-based vs. interrogative-based (Demirok&Öztürk 2020), or DEM-type vs WH-type (Belyaev& Haug 2020). From a semantic point of view, these classification methods have much to commend them; on the other hand, use of these terminologies entails risks of confusion from a morphosyntactic point of view. The present work points out the limitations of these proposals by bringing out language-specific morphosyntactic properties of Abzakh (West Circassian, Northwest Caucasian) correlatives. The data comes from diasporic Abzakh speakers of Yeleme village in Antalya, Turkey.

In view of its use of interrogative pronouns, Abzakh (and Circassian at large) might be considered to have an interrogative-based construction like the correlatives in Kabardian, an East Circassian dialect, which are classified as interrogative-based correlatives by Demirok & Öztürk (2020). However, the morphosyntactic properties of a correlative construction and that of a content question are clearly different. Thus, a more detailed morphosyntactic analysis of correlatives in Circassian is called for.

(2) Content question (elicited)

xɜt	[wə-z-dɜ-ʃxə-ʁɜ-r]
who	2SG.ABS.S-IPM-COM-eat$_{ptcp}$-PERF-ABS

Litt : Who (the one) that you ate with ?
"With whom did you eat?"

(3) Relative clause (elicited)

zɜ-ʃ^{w}ɐ-ʁɜ	wətʂ͡ʔə-r	jɜ-zɜʁ-ɐʁ-ɜp
RES-drink$_{PTCP}$-PERF	herb-ABS	3SG.OBL.R-avail$_{ptcp}$-PERF-NEG

« The pills (litt :herb) that s/he took did not make any benefit (to her/his health)."

(4) Correlative clause (https://doi.org/10.24397/pangloss-0006058#S5)

xɜt	jə-bɜʃ	wərbɐn-əm	nax	bɜ	ʃəʒɜ-w	qɐ-nɜ-m-əj
who	POSS-stick	gap-OBL	more	much	far-ADV	CIS-stay$_{PTCP}$-COND-ADD
ɐ-r		səd	χ^{w}ə-ʃt-əʁɜ ?			
DEM$_{DISTAL}$-ABS		what	be-CERT-PERF			

Litt : if the stick of who stays as more far to the gap too that would be what?
« Which role plays the one whose stick falls farthest from the gap? » (Context: in a game where one explains the rules)

The content questions, as illustrated in example (2), follows the relativization schema (example 3), where the participial form is marked by the resumptive pronoun *z(V)-* which replaces the indexed person marker of the relativized oblique argument. In the correlative construction in example (4), however, the participial form is marked by the conditional marker *mɜ-*, whose vowel is elided for morphophonological reasons, and by the additive marker *əj-*, which is obligatory in the construction. In the main clause, the distal demonstrative pronoun *ɐ,* which is also the 3rd person singular, functions as the anaphoric correlate of the subordinate clause

yet it is not obligatory. When omitted, the construction might yield a universal concessive conditional (UCC) (Haspelmath & König 1998) as illustrated in (5) (see Lander (2016) for a detailed analysis of Kabardian UCCs), which clearly supports the idea that the structure is conditional-based, rather than interrogative-based:

(5) Universal concessive conditional (elicited)

səd	jə-ʃˀɐ-ʁɜ-m-əj	
what	3SG.OBL.AGT-do$_{PTCP}$-PERF-COND-ADD	
qɐlɜ	ʃxʷɜ-m	dɜ-sə-ʃʷə-ʁ-ɜb-ɐ
city	big-OBL	LOC-to_be_sitting-HBLT-PERF-NEG-INTER

"Whatever s/he did, could s/he not stay in the big city ?" (="Couldn't s/he stay in the big city even though she did everything she could?"

The difference between the correlative construction in example (4) and the UCC construction in example (5) is the use or non-use of a demonstrative correlate in the matrix clause. Thus, the resemblance between these two constructions supports that Abzakh correlatives are conditional-based.

As for the syntactic status of correlatives, the tests at hand such as embeddedness, TAM dependency, illocutionary force operators, etc. (Bril 2010; Bril & Rebuschi 2006) yield subordinate clauses, unlike the comparative correlatives like "The more you work, the more you earn", which are considered as cosubordination. The scope of both the interrogative marker and that of the negation marker is restricted to the main clause, it cannot be extended to the correlative clause, which clearly demonstrate an embeddedness as an evidence for subordination: "In these cases of subordinate nexus, the illocutionary force is a feature of the superordinate junct, and the subordinate junct is unaffected by it, but remains semantically backgrounded, but stated, information." (Foley & Van Valin 1984, 240)

Thus, the present analysis discusses the morphosyntactic properties of correlative constructions in Abzakh and brings out an overt conditional-based schema by comparing their morphosyntactic features with those of relative clauses and content questions, and by drawing parallelism between the correlative constructions and UCCs. It also discusses their syntactic status, which clearly points to subordination.

Keywords: Correlative constructions, conditional-based constructions, subordination, Abzakh, Circassian.

References

Belyaev, Oleg, et Dag Haug. 2020. « The Genesis and Typology of Correlatives ». *Language* 96 (4): 874-907. https://doi.org/10.1353/lan.2020.0065.

Bril, Isabelle. 2010. « The syntax and pragmatics of clause linkage and clause hierarchy: Some new perspectives ». *Clause linkage and clause hierarchy: Syntax and pragmatics*, 1-50.

Bril, Isabelle, et Georges Rebuschi. 2006. « Coordination, subordination et co-jonction: faits linguistiques et concepts ». *Faits de langues* 28 (1): 7-20.

Demirok, Ömer, et Balkız Öztürk. 2020. « Correlatives in Languages of the Caucasus ». *The Oxford handbook of the languages of the Caucasus*, 909-42.

Foley, William A., et Robert D. Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge studies in linguistics 38. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

Haspelmath, Martin, et Ekkehard König. 1998. « 9 Concessive conditionals in the languages of Europe ». In *Eurotyp, 3, Adverbial Constructions in the Languages of Europe*, édité par Johan Van Der Auwera, 563-640. De Gruyter Mouton. https://doi.org/10.1515/9783110802610.563.

Lander, Yury. 2016. « Kabardian universal concessive conditionals ». *Rhema. Рема*, nº 2, 44-66.

Lehmann, Christian. 1986. « On the typology of relative clauses ». *Linguistics* 24 (4). https://doi.org/10.1515/ling.1986.24.4.663.

Srivastav, Veneeta. 1991. « The syntax and semantics of correlatives ». *Natural Language & Linguistic Theory* 9 (4): 637-86.

Evaluating linguistic distances in Daghestan and adjacent territories: evidence for contact?

George Moroz (HSE University), Chiara Naccarato (HSE University)

The "Typological Atlas of the Languages of Daghestan" (TALD, see http://lingconlab.ru/dagatlas) is an online resource for aggregating and visualizing grammatical and lexical information on the languages spoken in Daghestan and adjacent territories. The current version of the database includes 54 features (some of which consist of several subfeatures) and 218 idioms. Each feature dataset is accompanied by a chapter describing the rationale behind the choice of values, and providing examples and relevant references.

In this talk we would like to present the preliminary results that we achieved by analyzing the database as a whole with an aim to calculate linguistic distances between some of the idioms included in our database. Since the Atlas is based almost exclusively on data from published grammars and dictionaries, it is not always possible to collect data for each datapoint in the Atlas: information about a certain feature in a specific idiom might be missing in the available literature. Thus, we chose a subsample of more represented features and idioms for the analysis leaving aside more marginal ones. As a result, we came up with a table consisting of 50 idioms (Figure 1) and 84 linguistic features (e.g., Presence of nasal vowels in the phonological inventory). Since we still had some missing values in the table, we replaced them with the most common value for that feature. We applied Anderberg's dissimilarity measure (Anderberg 1973; Chandola et al. 2007), which assigns higher weights to infrequent matches and mismatches. Anderberg's dissimilarity measure provides us with the distance between each of the idioms under analysis. Afterwards we conducted several statistical analyses. First, we applied Multidimensional scaling (MDS) to the distance matrix in order to visualize linguistic distance against geographic distance (Embleton 1993), see Figure 2. Second, we applied hierarchical clustering using the complete linkage method (Sorensen 1948), see Figure 3.

The results of both MDS and hierarchical clustering are interesting. The most uninteresting scenario would be one in which idioms cluster together according to their genetic affiliation, i.e., all idioms belonging to the Nakh branch would end up in one cluster, all idioms belonging to the Lezgic branch in another cluster, and so on. However, although some idioms do cluster together according to their genetic affiliation, some other clusters are more difficult to explain, e.g., Standard Dargwa, Rikvani Andi, Mukhad Rutul and Upper Gakvari Chamalal end up in one cluster, which is rather unexpected. The exceptional behavior of some of these idioms could be explained in terms of contact. For example, Archi clusters together with Lak and not with other Lezgic languages, which is somewhat expected since there is a long history of contact between Archi and Lak, with the former being isolated and quite divergent from the rest of the Lezgic branch, cf. (Giginejšvili 1967: 3-4) and references therein. However, some idioms show up in rather surprising clusters, e.g., Itsari and Sanzhi are closely related Dargic idioms, but they appear to be very distant from one another both in the MDS visualization and clusterization. These seemingly unwarranted results of clustering could be possibly explained as due to some inconsistencies in our data: different features were collected independently by different researchers, which might have led to uneven representations of different idioms. As different as the reasons for such remarkable results may be, the findings of this quantitative analysis are important in that they point us to the spots that are worthy of further investigation. In some cases we find that certain idioms indeed show a typologically unexpected (but not necessarily inexplicable) behavior in some areas of the grammar and/or the lexicon. In other cases we rather observe some flaws in data documentation, collection, and/or interpretation.

References

Anderberg, M. R. (1973). *Cluster analysis for applications*. New York: Academic Press.

Chandola, V., Boriah, S., & Kumar, V. (2007). *Similarity Measures for Categorical Data: A Comparative Study*. https://doi.org/10.1137/1.9781611972788.22.

Embleton, S. (1993). Multidimensional scaling as a dialectometrical technique: Outline of a research project. In *Contributions to Quantitative Linguistics: Proceedings of the First International Conference on Quantitative Linguistics, QUALICO, Trier, 1991* (pp. 267-276). Netherlands: Springer.

Giginejšvili, B. K. (1967). Preface to: Mikailov, K. Š., *Arčinskij jazyk* [Archi]. Makhachkala: IJaLI.

Sorensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biologiske skrifter*, *5*, 1-34.

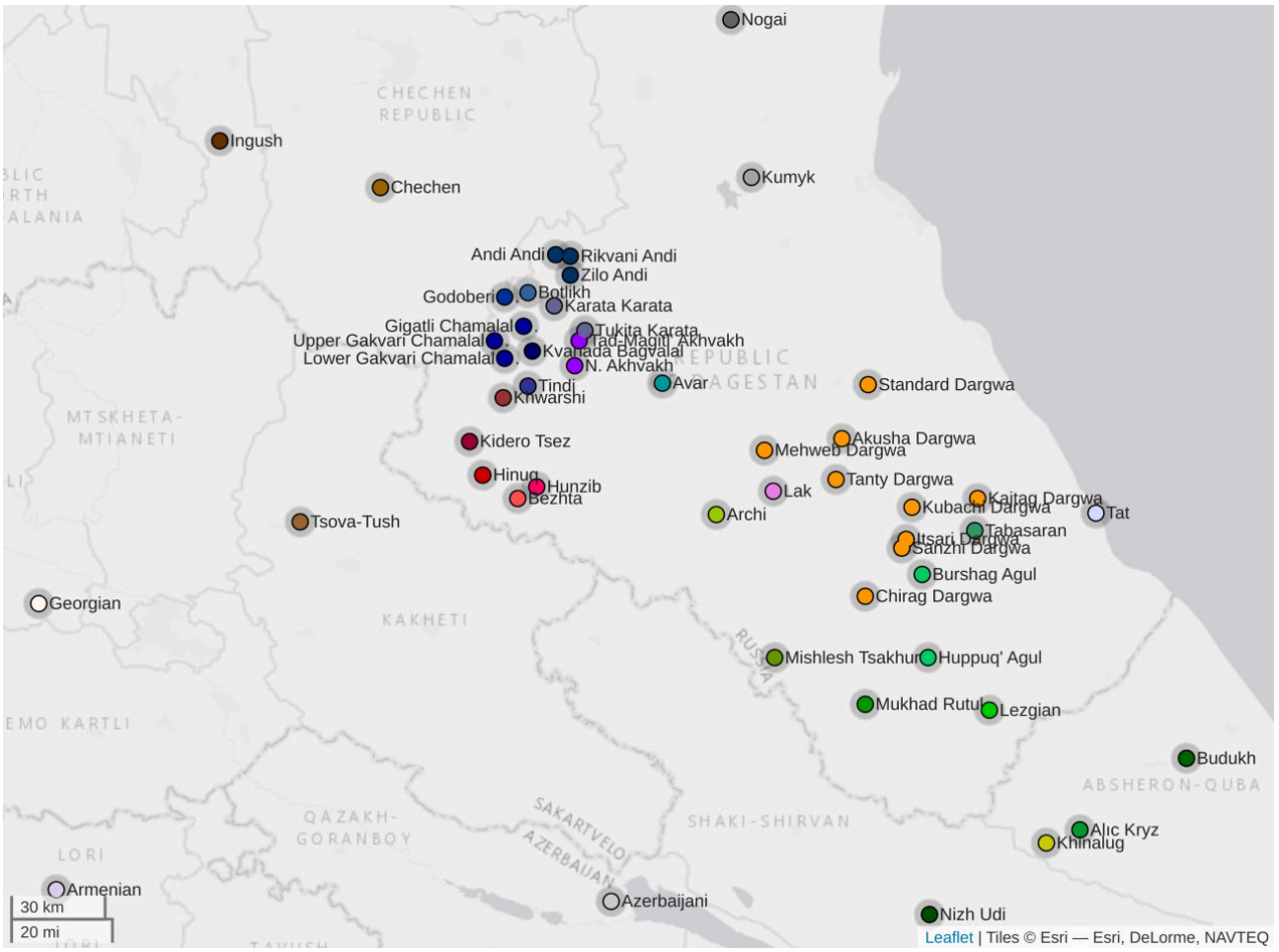


Fig 1 Geographical distribution of the idioms under analysis

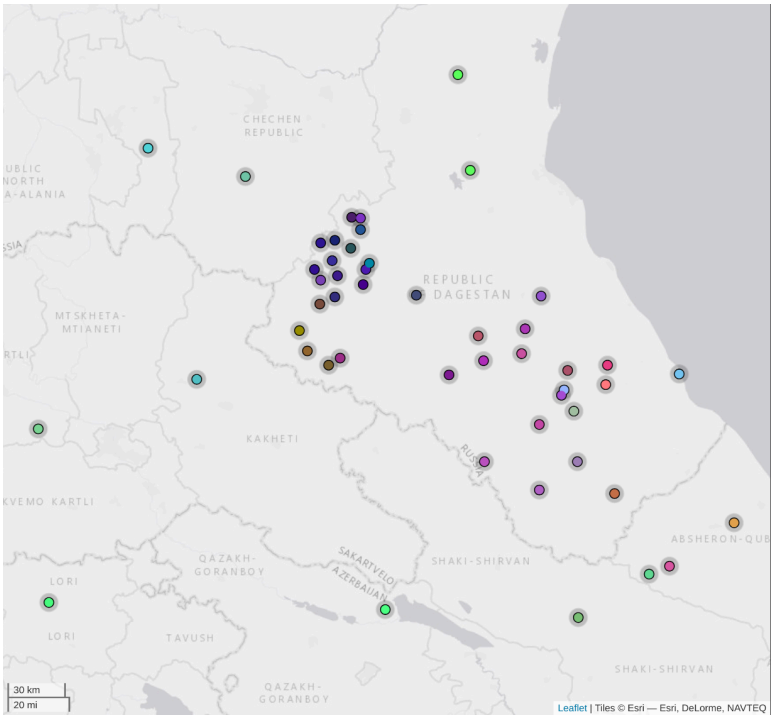


Fig 2 Map representing MDS conversion to RGB

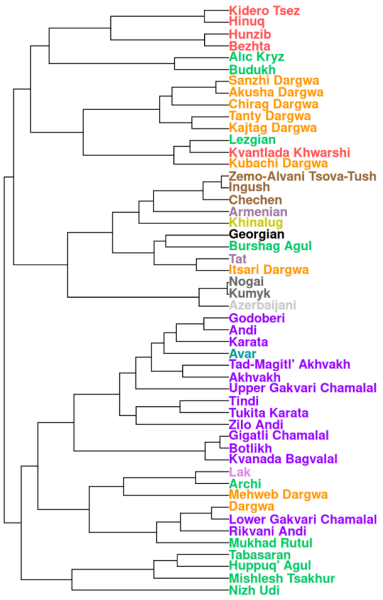


Fig 3 Dendrogram of hierarchical clustering

Mutalov Rasul Osmanovich
Doctor of Philology, Professor,
Chief Scientific Associate,
Institute of Linguistics of the Russian Academy of Sciences.
125009, Russia, Moscow, lane. Bolshoy Kislovsky, 1/1
mutalovr@mail.ru

Sound change and ancient Dargwa migrations

Dargwa languages are a branch of East Caucasian language family spoken in five administrative district in central highland Daghestan (Akusha, Dakhadaev, Kaytag, Levashi and Sergokala) by the overall number of 626 thousand speakers (2021 Census), with quick rate of loss even in remote villages. Classification of lects is often unobvious for at least two reasons. First, Dargwa languages form a dialectal continuum with similarities boosted by adjacency of varieties; second, some of the lects remain under-documented. Linguistic classifications range from three languages in the 1930s (Dargwa, Kaytag and Kubachi) to up to 17 or 18 distinct languages (Koryakov 2006, Forker 2020). Below, we distinguish 6 Dargwa languages, including Northern Dargwa, Mehweb, Tsudakhar-Sirhi, Kajtag, Kubachi and Chirag (Mutalov 2021, see also http://jazykirf.iling-ran.ru/). According to a hypothesis in (Nichols 2023), the proto-Daghestanians migrated from the south of the Caspian Depression and were settling from the east uphill, or westward. The issue of locating the Dargwa homeland proper, however, has not yet been a subject of a special study. Below, we make a suggestion based on observations of sound change specific to different Dargwa varieties.

We suggest that the homeland of Dargwa, or the place where they parted with the proto-Laks, are the headwaters of Ulluchay, which is the southwestern part of today's Dargwa-speaking area. We do not suggest a specific timing of the expansion; instead we hypothesize that it went in several waves. We must also allow for the possibility that some of the waves have been eliminated or decimated by conflicts or natural causes. Empirical evidence comes from the sound change shared by or distinguishing the languages and varieties.

As the **first migration**, we suggest expansion to northeast, where the proto-Dargwa settled in the caves and ravines near the present day Mekegi (Levashi district), providing the source of the modern northern Dargwa varieties. The innovations separating these languages include [d] → [n] (*du* → *nu* 'I'), [d] → [r] (1 and 2 person markers *-da*, *-di* → *-ra*, *-ri*), [ʕ] → [ħ] ('you.sg' *ʕu* → *ħu*); supported by the dative and infinitive markers changes *-j* → *-s*.

Subsequently, two groups splitted away from proto-Mekegi, including Akushi-Huraqi (Russian Урахи, Dargwa Х1урхъи) and Mugi-Muiri. The Akushis settled down in the northern part of Levashi Plateau. The Urahis splitted away from the Akushis and settled in the new territories of what is now easternmost part of the Dargwa speaking area. Akushi is distinguished by palatalization (e.g. [k] → [č], as in Mekegi *kebaʔes* → Akusha *čebaʔes* 'see'). Huraqi is further separated by the subsequent spirantization ([č] → [š], as in Akusha *čebaʔes* → Huraqi *šibaʔis* 'see').

Northern Dargwa lects were additionally forcibly displaced by Tamerlan's invasion in the late 13th century. R. Magomedov argues that the Akushis had to move westward, where they came into contact with the Tsudakhars. The proto-speakers of today's Levashi did not resettle. The proto-Mehwebs, who had lived near Mugi and spoke an early Mekegi variety moved to today's Hunib district, where their language came under very strong influence from Avar and Lak. Mehweb has four noun classes instead of three; and lacks the second person marker; the auxiliary sa<>i is replaced by the existential *le-*.

Several migrations of the proto-Mekegi speakers were directed to the north, northeast and east. In the north, the language of the re-settlers to present day Kadar has been distinguished by the change [q] → [ʁ] ('calf' *qača* →

ʁača). The proto-Mekegis who migrated northeastward founded Gubden, and those who moved eastward founded the Mekegi-speaking village of Murego.

The **second migration** from the homeland in the headwaters of Ulluchay was directed to the south and to the east, where the proto-southern Dargwa splitted into Chirag, Kubachi, Itsari, Kajtag and Sirha. Further migration

southward founded Chirag and its close relative Kunki. While Chirag-Kunki has several innovations, including vowel reduction in verbal stem (Chirag *čakdatur*, Kunki *čekadtur* 'made them sit down'), it also has many local retentions.

Some Dargwas migrated to the southeast, to the basin of Dirbagkatta. This became the territory of Ashty, linguistically close to Kubachi. Some proto-Ashtys migrated to the cliffs to the south from today's Kubachi. Kubachi is considered one of the most divergent Dargwa languages, distinguished by both many retentions and many innovations, including contexts where loss of [r] and [ħ] leads to compensatory lengthening of the preceding vowel; spirantization of [ts] to [s] ('one' *ca* → *sa*), and despirantization ([s] → [t] in *sime* → *time*; [z] → [d] in *waza* → *wada*).

The talk will present additional evidence for the migration westward, the one that followed the valley of Ulluchay and gave rise to Kaytag and Sirha-Tsudqur languages. Additionally, we will address the problem of mixed dialects. As one example, Gapshima is phonetically close to Sirha-Tsudqur, while structurally close to Mekegi. Gapshima (lit. three villages) was formed from the merger of three different villages, one of which probably spoke a variety of Sirha-Tsudakhar, while the other two spoke a variety of Mekegi.

Коряков Ю.Б. Атлас кавказских языков. - М.: Ин-т языкознания РАН, 2006.

Магомедов, Р.М. Даргинцы в дагестанском историческом процессе. Махачкала: Дагестанское книжное издательство. 1999.

Магометов А. А. Кубачинский язык. Исследование и тексты. Тбилиси: Издательство Академии наук Грузинской ССР. 1963.

Муталов Р.О. Мекегинский диалект – основа севернодаргинского языка? // «Язык как он есть» к 60-летию Андрея Александровича Кибрика, издательство ООО «Буки-Веди». М., 2023.

Муталов Р.О. Классификация даргинских языков и диалектов // Социолингвистика, том 6, № 3. М., 2021.

Forker D. A grammar of Sanzhi Dargwa (Languages of the Caucasus 2). Berlin: Language Science Press. 2020.

Grammatical evidence on the East Caucasian (Nakh-Daghestanian) dispersal

Johanna Nichols, University of California, Berkeley and HSE University

Of recent work on the East Caucasian homeland and dispersal, Schulze (1998:169-186, 2018b:371-372) misses the importance of the Caspian corridor and Nichols (2024) overlooks the early importance of the highlands (Amirkhanov 2022). To improve the picture, this paper examines three morphological features of obvious great antiquity:

1. Verbal pre-root elements. These are reserved as a system only in Lak and Nakh. A clear example is *lV- (followed in Lak by gender marker in its medial allomorphic form, Friedman's transcription #):

Lak	la# + qan	'ascend, go up, hang up, ignite, light
	lä#- + qan	'appear, be born',
	li# + qan	'run away'
Ingush	laatt-	'stand, be standing'
	ghott-	'fly away, take off (plane, bird); get up, stand up'
	uott-	'stand up; become, begin, start; drop in (on someone)

Lak has more of these *lV- sets than Ingush, and (as in the examples here) the Lak sets generally have wider semantics. Importantly, the vowel of the Lak morpheme cannot be back rounded (Friedman 2024:44), and to judge from this set and others the same is true of Ingush *lV-. Ingush does not have medial gender marking (or indeed virtually any gender marking other than root-initial). From these facts we can reconstruct a now-affixed pre-root morpheme that can be preposed to a gender-prefixed morpheme in Lak but only to a bare root or stem in Ingush; and it has a non-back-round vowel. The lack of an associated gender marker in Ingush may suggest something about the relative chronology in the two languages of gender prefixation and affixation of the pre-root element, or it may tell us something about the resolution of -CC- sequences if those arose with univerbation of the pre-root element and the root.

2. CV- ~ VC- stem-initial alternations in a few lexemes, preserved in Lak, Nakh, and Khinalug:

Lak		barz	'moon'	oblique	zurul
Chechen	pronouns	so	'1sg'	erg.	as
		ħo	'2sg'		aħ
Khinalug	pronouns	zə	'1sg'	dat.	ás(ər)
		və	'2sg'	dat.	úx(ər)(Schulze)

This appears to be a shared archaism of these branches, suggesting peripheral location.

3. Bantu-type gender prefixes on nouns. The following nouns all belong to the gender class marked by /b-/ or a similar labial. All but 'sun' in Lezgi begin with a labial consonant (underlined), while in Lezgi the consonant /r/ that is medial in most of the others is initial. (More work is needed on the Khinalug word.)

	'sun'	'moon'
Ingush	<u>m</u>aalx	<u>b</u>utt
Avar (Chadakolob)	<u>b</u>aq'	<u>m</u>oc'
Lak	<u>b</u>argh	<u>b</u>arz
Dargwa (standard)	<u>b</u>erħi	<u>b</u>ac
Lezgi (Axty)	ragh	<u>w</u>arz
Khinalug	ynq'	<u>v</u>ac'

In the right phonological environment, these gender initials can create clusters:

		Chechen	Ingush	Batsbi	Khinalug	
'vein, tendon'	(B gender)	*pxa*	*pxa*	*pxa*		
'scapula'	(B)	*pxanar*	*pxandar*	*pxaner*		
'wool fleece'	(D)	*txa*	*txa*	*txe*		*kxa*

Other than sets like these (a minority in every language, and always simplex root words), nouns in East Caucasian do not ordinarily begin with a consonant echoing their gender marker.

In a survey of 50 nouns most prone to show this pattern, within branches the individual words are quite consistent as to whether they have it or not, though the words that have it differ from branch to branch; and the pattern is most frequent in the west (Nakh, 79%) and least frequent in the east (Khinalug, 48%). This suggests that Bantu-like gender existed in ancestral East Caucasian, branch protolanguages varied in which words had it, the variation was consistent within branches, and the overall frequency in each branch took shape later and formed a cline running west to east following the modern linguistic geography.

I also surveyed other morphological paradigmatic features that can be expected to be very old: frequency of derived noncausals in causal-noncausal pairs; pronoun root consonantism; bipartite locative and directional preverbs; and others (survey ongoing). All of these have distributions that are useful in reconstructing etyma and subgrouping but appear to tell us nothing about origins and dispersal routes.

In contrast, sets 1-3 above have distributions in space and across upper branches that are revealing of likely early prehistory: a west-east distribution of branches, consistency of patterns within branches, and the Nakh and Avar-Andic-Tsezic languages were the most peripheral and the first to move away. All this points to a very early dispersal, early enough that features expected to be stable have been lost or changed beyond recognition; that was followed by movement to present branch locations. This picture is consistent with Schulze's dispersal plan but not his homeland, more consistent with Nichols's homeland and dispersal, but the archaism of non-peripheral, highland-locked Lak points to a central role of the highlands in the larger picture.

Amirkhanov, Kh. A. 2022. Chronology of cultural deposits of the Chokh multilayered settlement (according to 2022 data). *History, archaeology, and ethnography of the Caucasus* 5:18.3.715–728.

Friedman, Victor A. 2024. Pre-root verb morphology in Lak: Preverbs, preverboids, and pre-root gender carriers. *Languages and cultures of the Caucasus: A Festschrift for Kevin Tuite,* 35–55. Wiesbaden: Reichert Verlag.

Gudava, T. E. 1959. *Sravnitel'nyj analiz glagol'nyx osnov v avarskom i andijskix jazykax*. Maxachkala: IIJaL Dag. Fil. AN SSSR.

Nichols, Johanna. 2003. The Nakh-Daghestanian consonant correspondences. *Current Trends in Caucasian, East European, and Inner Asian Linguistics: Papers in Honor of Howard I. Aronson,* ed. Dee Ann Holisky and Kevin Tuite, 207–251. Philadelphia: John Benjamins.

Nichols, Johanna. 2024. The East Caucasian homeland and dispersal: A preliminary model. *Languages and cultures of the Caucasus: A Festschrift for Kevin Tuite*, 57–71. Wiesbaden: Reichert Verlag.

Schulze, Wolfgang. 2008. Toward a history of Khinalug. *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag. Band II: Chomolangma,* ed. Brigitte Huber, Marianne Volkart, and Paul Widmer, 703–744. Halle (Saale): International Institute for Tibetan and Buddhist Studies GmbH.

Schulze, Wolfgang. 2018. Khinalug in its genetic context: Some methodological considerations. Part 2: Some words about grammar and conclusions. *Iran and the Caucasus* 22.366–376.

LEXICALLY CONDITIONED SPATIAL MARKERS IN (HERITAGE) LAZ

Balkız Öztürk Başaran, Boğaziçi University & Ömer Eren, University of Chicago

Aim: Laz—an endangered South Caucasian language spoken in Türkiye—features an intricate system for encoding spatial relations, employing 27 spatial prefixes (SPs) on the verb. Kutscher (2011) categorizes Laz as a satellite-framed (s-framed) language in Talmy's (2000) typology. This study reexamines the typological classification of Laz, comparing varieties spoken by proficient older generation "baseline" speakers with those of younger heritage speakers. Our analysis of baseline Laz reveals that, unlike typical s-framed languages, spatial prefixes (SPs) are incompatible with manner-of-motion verbs. This incompatibility suggests that SPs do not function as satellites that introduce a path to motion verbs lacking inherent path information. Thus, we argue that Laz does not meet the criteria for an s-framed language. Expanding on Acedo Matellán's (2010) distinction between weak and strong s-framed languages, we propose that verb-framed (v-framed) languages vary along a weak-strong continuum. We classify Laz as a weak v-framed language, where the verb root and path are realized as a single unit via the underassociation of the verb root for the path layer (Ramchand 2008), rather than being integrated as a single verb root as in strong v-framed languages. Turning to Heritage Laz, we observe a typological shift, which involves a reduction in the spatial prefix inventory, accompanied by a substitution with postpositions. We argue that this change stems from the lexically conditioned nature of spatial prefixes as previously reported in the aspectual system in Heritage Laz (Eren To appear). The relevant changes, albeit in two different domains, are in line with the changes in other heritage languages (Polinsky 2018).

Baseline Laz: SPs in Laz, which can be monomorphemic or bimorphemic, are dependent on the verbal root morphologically. They convey information regarding, vertical/horizontal orientation (1), the thither-hither orientation (deixis) (2), and the axial part (Svenonious 2006) expressed by the second morpheme in bimorphemic SPs (3). Person agreement and valency markers occur in between the verbal root and SPs (3). Following the literature (Den Dikken 1995, Asbury *et al* 2007, Gehrke 2008), we also assume that SPs belong to the category P that extends beyond adpositions including affixes and particles. As for the internal structure of these prefixes, we adopt the fine-grained PP structure in Svenonius (2006) (4).

With respect to Talmy (2000)'s typology, Kutscher (2011) argues that Laz should be classified as an s-framed language, as SPs denote Path in the context of motion-verb roots. However, we see a clear split in the motion verbs which can take SPs in Laz. While SPs are compatible with verbs of directed motion (DMV) with nominative subjects (5b-c), they are incompatible with manner of motion verbs (MV) with ergative subjects (6b), (7b). Building on Ramchand (2008), Son and Svenonius (2008) focus on the Directed Manner of Motion Constructions (DMMCs), where MVs yield a directed motion interpretation via a goal denoting PP. Based on the structure in (8), they divide motion verbs into three groups: (i) DMVs (*come*, *go*), which freely form DMMCs cross-linguistically (9a), (ii) MVs denoting directionality (*run*, *walk*) which form DMMCs only in some languages (9b), (iii) MVs without any directionality (*dance*, *limp*) which form DMMCs only in languages like English, but not in others (9c). In Laz, we also assume that the Path head is always present in the lexical specification of DMVs in the first group, but not in that of MVs in the second and third groups.

SPs in Laz are only compatible with DMVs involving a Path layer, which is a goal or source oriented transitional path in the sense of Pantcheva (2011). If the verb lacks such a Path layer, SPs cannot independently introduce Path information. With MVs in the second group it leads to ungrammaticality (6b), (7b), and with the ones in the third group, SPs simply denote Ground information and do not yield DMMCs (10b). This pattern is also observed with transitive predicates which inherently denote an oriented transitional path, such as, *bring, take, send, forward,* as illustrated in (11). Thus, only when Path is introduced by the verbal root in Laz, SPs can be used to specify the content of the Path information licensed by the verb. In this respect, Laz behaves differently from s-framed languages, where spatial P-like elements do not exhibit selectivity with respect to motion verbs mainly in the first group. Thus, we conclude that Laz cannot be an s-framed language.

Acedo Matellán (2010) groups s-framed languages into two: (i) strong s-framed languages where Path is introduced as an independent DP/PP and (ii) weak s-framed languages where Path and the verb form a single word. In Laz, too, Path and the verb forms a single word, however, given the fact that MVs cannot form DMMCs it is not possible to classify it as a weak s-framed language. Furthermore, Laz lacks the type of resultative constructions derived from activity verbs, typically observed in s-framed languages like English (e.g. *John hammered the metal flat*), which are assumed to be correlated with the availability of DMMCs in a given language (Beck and Snyder 2001). We also observe lexicalized SP+verb combinations in Laz, where the spatial meaning of the SP is no longer transparent as in (12). Given these, we argue that v-framed languages should also be split into two as weak vs. strong. Strong v-framed languages are like Spanish, where Path and the verb are realized as a single verb root. In this respect, we propose that Laz is a weak v-framed language where the verb and Path form a single word, but not a single verb root. We argue that the formation of a single

word is achieved via the underassociation of the verbal root in the sense of Ramchand (2008), which allows for the realization of the Path information via the SP.

Heritage Laz: The findings above specifically concern highly proficient speakers from older generations. In Heritage Laz, the grammatical encoding of spatial relations reveals an erosion of bound spatial markers, characterized by a reduced inventory and the omission of these markers in favor of free postpositional constructions (Eren 2023). Heritage Laz shows a phonologically conditioned reduction in spatial case allomorphy, with phonologically heavier (and thus more salient) markers becoming overgeneralized (13b), replacing specific spatial markers found in the baseline variety (13a). Additionally, while bound spatial markers remain obligatory in baseline Laz, where postpositional phrases act as adjuncts to intensify or specify spatial relations (14a), Heritage Laz displays a shift, with prefixes frequently omitted in the vicinity of postpositions (14b). An analysis of these divergences suggests an emerging typological realignment in Laz, as it transitions from a weak v-framed language toward an intermediary typology between verb- and satellite-framed structures. According to the lexical specification and underassociation framework presented here, heritage speakers' decreased sensitivity to the lexical properties of verbal roots provides insight into the gradual erosion of spatial markers. This erosion appears linked to the dependence of these markers on the lexical specification of verbal roots, particularly given the aspectual reanalysis occurring in Heritage Laz, where speakers exhibit reduced sensitivity to the selection of imperfective marker allomorphs conditioned by lexical features in the verbal root (Eren To appear). These two shifts impact bound morphemes that rely on the lexical properties of verbal roots, aligning with cross-linguistic patterns (Montrul 2016, Polinsky 2018).

(1) K'at'u **ey**-u-l-u-n / **gol**-u-l-u-n.
cat.nom up-val-go-TS-impf.3ps/forward-val-go-TS-impf.3ps
'The cat is going up (vertically) / forward (horizontally)'

(2) Ayşe **m(o)**-u-l-u-n / **me**-l-u-n.
Ayşe.nom towards-val-go-TS-impf.3ps/away-appl-go-TS-impf.3ps
'Ayşe is coming/ going.' Literal: 'Ayşe is going towards or away from the speaker.'

(3) Ma otva-şe e-**yo**-v-ul-ur.
I.nom roof-all up-top-1p-val-go-TS-impf.1ps
'I am going up onto the roof.'

(4) [PathP [PlaceP [AxPartP [KP [DP]]]]

(5) a. Ali marketi-şe u-l-u-n.
Ali.nom market-all val-go-TS-impf.3ps
'Ali is going to the market.'

b. Ali marketi-şe am-u-l-u-n.
Ali.nom market-all into-val-go-TS-impf.3ps
'Ali is going into (entering) the market.'

(6) a. Ali-k marketi-şe i-gzal-s.
Ali-erg market-all val-walk-impf-3ps
'Ali is walking to the market.'

b. Ali-k marketi-şe *(am)-i-gzal-s.
Ali-erg market-all into-val-walk-impf-3ps
Intended: 'Ali is walking into the market.'

(7) a. Çxomi-k mosa-şe i-nçir-s.
fish-erg net-all val-swim-imf-3ps
'The fish is swimming into the net.'

b. Çxomi-k mosa-şe *(am)-i-nçir-s.
fish-erg net-all into-val-swim-impf-3ps
Intended: 'The fish is swimming into the net.'

(8) [InitP [ProcP [DirP [PathP [(XP)]]]]]

(9) **a.** 1st group, e.g., *go*, *come*: [Init, Proc, Dir, Path]
b. 2nd group, e.g., *run*, *walk*: [Init, Proc, Dir]
c. 3rd group, e.g., *dance*, *limp*: [Init, Proc]

(10) a. Bere-pe-k i-xoron-an.
child-pl-erg val-dance-impf.3ppl
'The children are dancing.'

b. Bere-pe-k ce-i-xoron-an.
child-pl-erg down-val-dance-impf.3ppl
'The children are dancing in the same spot.'

(11) K'oçi-k puci axiri-şe am-iyon-am-s.
man-erg cow stall-all into-bring-TS-impf.3ps
'The man is bringing the cow into the stall.'

(12) a. gama-dvalu
out-put
'to spread out'

b. e-k'o-zdalu
up-back-pull
'perceive'

c. gama-k'otu
out-fold
'to slap'

(13) a. Mcvabu kavanozi-şa {ama-/gama-}-xt-u.
frog.nom jar-all {into/out of}-go-past.3sg
'The frog went into/out of the jar.'

b. Mcvabu kavanozi-şa menda-xt-u.
frog.nom jar-all away-go-past.3sg
Int: 'The frog went into/out of the jar.'

(14) a. Deluği-şi (doloxe) dolo-tz-e-y.
hole-all inside inside-look-impf-3ps.
'He is looking down inside the hole.'

b. Kavanozi-şi doloxe. coğori-k o-tzk-e-n
jar-gen inside dog-erg val-look-impf-3sg
Int: 'Inside the jar, the dog is looking.'

Pasko Iada

Reported Speech in Hunzib

Problem. The fact that reported speech in Northeast Caucasian languages proposes a challenge to the standard distinction between direct and indirect speech has been addressed in many studies. It is usually viewed from one of the two perspectives. The main focus of the first one is the use of long-distance reflexive pronouns, which refer to the speaker, whose speech is being reported — these reflexives are analyzed as logophors. Logophoric pronouns have been discussed on the material of Akhvakh (Creissels 2007), Archi (Daniel 2015), Avar (Rudnev 2017), and many others. This approach considers direct speech as a starting point, treating the use of a long-distance reflexive as a shift from the original perspective. The second possible line of analysis, on the contrary, suggests that the direction of the shift in the 'mixed' reported speech constructions is the opposite — they fall under the Indexical Shift (IS) phenomenon. IS occurs when indexical expressions in indirect speech refer to reported speech act participants, as they do in direct speech. IS has been observed in a number of Northeast Caucasian languages, including Hinuq (Forker 2013: 662–664), Tabasaran (Bogomolova 2024), Tanti Dargwa (Sumbatova 2019) and Tsez (Polinsky 2015a, 2015b). In our talk, we will present an overview of the 'mixed' reported speech strategies in Hunzib[1] (Tsezic < Northeast Caucasian).

Data. Both phenomena of 'mixed' reported speech, characteristic of Northeast Caucasian languages, are observed in Hunzib. First, indexicals of different types (1st and 2nd person pronouns, temporal deictic expressions) can undergo IS when they occur in the finite clausal complements of verbs denoting speech (e.g. 'say', 'ask'). IS is not obligatory for simple personal pronouns: they can refer to either reported or current speech act participants, (1). However, the pronouns must receive a 'shifted' interpretation if they host the intensifier particle *=ž*, (2). In most cases, it is not clear whether the examples include direct speech or indirect speech with IS: the quotative particle *=ƛe* is used with both citation and unambiguously indirect speech, and the complements of speech verbs are obligatorily finite. However, the application of some syntactic diagnostics can prove that indexicals are indeed able to refer to reported speech act participants when the reported speech clause is unambiguously non-citation. These diagnostics are based on the fact that citation has to be opaque to long-distance syntactic processes (see Shklovsky & Sudo 2014: 384). In Hunzib, NPIs can be licensed in the 'shifted' reported speech clause by negation in the matrix clause, which confirms the presence of IS — compare (3a) and (3b), which is ungrammatical due to the absence of an NPI-licensor. Second, in Hunzib, the simple 3rd person reflexive *žu* can be used in a logophoric context, (4). When both the reported speech speaker and the addressee are mentioned in the reported speech clause, one of the arguments can be expressed by a reflexive, while the other one — by a personal pronoun, (5).

Discussion. In the talk, we will compare the Hunzib facts with those obtained for other Northeast Caucasian languages. Unlike Tsez (Polinsky 2015a: 17), with which Hunzib is closely related, Hunzib allows IS only with a narrow class of speech verbs, but does not allow it in the complements of verbs denoting thought, knowledge and attitude. At the same time, the obligatory IS of pronouns with an intensifier particle constitutes a similarity between Hunzib and Tsez (Polinsky 2015b: 260). The status of the intensifier is unclear. On the one hand, personal pronouns hosting it can be used in the reflexive contexts, both local and distant (that is why they are considered to be reflexive pronouns in Van den Berg 1995). On the other hand, the intensifier can be used to denote focus on the non-pronominal constituents. Moreover, this particle is in complementary distribution with other elements marking focus, such as particle *=bəs* 'even', (6). Nevertheless, we cannot conclude that IS is caused solely by the focus interpretation: first, focus markers other than *=ž* do not impose obligatory IS, and second, 'shifted' pronouns without intensifier are not always prosodically prominent, which would be expected under the focus interpretation. As for logophors, the fact that a 3rd person reflexive pronoun and a personal pronoun referring to the reported speech act participants can be used in the same clause unites Hunzib with Akhvakh (Creissels 2007), Archi (Daniel 2015), and Tanti Dargwa (Sumbatova 2019), but distinguishes Hunzib from Hinuq (Forker 2013: 664). Furthermore, it is problematic to analyze the Hunzib 3rd person reflexive pronoun *žu* as a logophor, as its use is not restricted to reported speech contexts: it can be bound locally and distantly in a non-reported context, or occur without any binder emphatically.

[1] The data were obtained by means of elicitation during the author's fieldwork in 2024 in Stal'skoe (Republic of Dagestan, Russian Federation), and from the corpus in Van den Berg 1995.

Examples

(1) pat'imat-lo nɨsə-r **də** ken b-uw-aa=ƛe
Patimat-ERG say-PRET **I.ERG** food.ABS III-do-GNR=QUOT
'Patimat$_i$ said she$_i$ / I would cook the food.'

(2) pat'imat-lo nɨsə-r **də=ž** ken b-uw-aa=ƛe
Patimat-ERG say-PRET **I.ERG=ITS** food.ABS III-do-GNR=QUOT
'Patimat$_i$ said she$_i$ / *I would cook the food.'

(3) a. maħamad-lo nɨs-iis dii suk'u=loddon aħ-li-ʔ ac'e-r=ƛe
Mahomed-ERG say-PRET.NEG I.DAT who.ABS=NPI garden-OBL-IN.ESS see.I-PRET=QUOT
'Mahomed$_i$ did not say he$_i$ / I saw anyone in the garden.'

b. *dii suk'u=loddon aħ-li-ʔ ac'e-r
I.DAT who.ABS=NPI garden-OBL-IN.ESS see.I-PRET
*'I saw anyone in the garden.'

(4) sɨd-ƛαr-o-1 nɨsə-n li əgi ƛ'odo č'eq-i-d bαk'-lα-α-s qoqla
one-FOC-OBL-ERG say-CVB be.IV there above bird-OBL-GEN2 nest-OBL-IN-EL egg.III
b-oxče-č **žin-lo**=ƛe
III-take-PRES **REFL-ERG**=QUOT
'One (thief) said that he could take the bird's egg from the nest above <…>' (Van den Berg 1995: 152)

(5) pat'imat-lo rasul-go nɨsə-r **žin-lo** **mə** buƛi ij-ojs
Patimat-ERG Rasul-AD.ESS say-PRET **REFL-ERG** **you.ABS** in.house пускать.I-FUT.NEG
'Patimat$_i$ told Rasul$_j$ that she$_i$ would not let him$_j$ / you in the house.'

(6) *pat'imat-lo nɨsə-r **də=ž** Rasul**=bəs** ok'e-r=ƛe
Patimat-ERG say-PRET **I.ERG=ITS** Rasul.ABS**=even** invite.I-PRET=QUOT
Int. 'Patimat$_i$ said she$_i$ invited even Rasul.'

List of abbreviations

I — I agreement class, III — III agreement class, ABS — absolutive case, AD.ESS — adessive, CVB — converb, DAT — dative, EL — elative, ERG — ergative, FUT — future tense, GNR — general tense, IN.ESS — inessive, ITS — intensifier, NEG — negation, NPI — negative polarity item, OBL — oblique, PRET — preterite, QUOT — quotative.

References

Creissels, D. (2007, December 7–9). *Intensifiers, reflexivity and logophoricity in Axaxdərə Akhvakh* [Conference presentation]. Conference on the Languages of the Caucasus, Leipzig.

Daniel, M. (2015). Logophoric reference in Archi. *Journal of Pragmatics, 88*, 202–219.

Forker, D. (2013). *A Grammar of Hinuq.* Berlin: de Gruyter.

Ganenkov, D. (2016, April 15). *Of what person are long-distance reflexive pronouns?* [Conference presentation]. Workshop on Nakh-Daghestanian languages, Bamberg.

Polinsky, M. (2015). Embedded finite complements, indexical shift and binding in Tsez. *Languages of the Caucasus, 1*(1), 1–37.

Polinsky, M. (2015). *Tsez syntax: A description.* Ms., Harvard University.

Rudnev, P. (2017). Minimal pronouns, logophoricity and long-distance reflexivisation in Avar. *Studia Linguistica, 71*, 1–2, 154–177.

Shklovsky, K., & Sudo, Y. (2014). The syntax of monsters. *Linguistic Inquiry, 45*, 381–402.

Sumbatova, N.R. (2019). Personal deixis in the reported speech: the case of Tanti Dargwa. *Tomsk Journal LING & ANTHRO, 3*(25), 72–84.

Van den Berg, H. (1995). *A grammar of Hunzib.* München: Lincom Europa.

Exploring transitivity prominence variation in Armenian

Carlo Maria Pertica, Università per stranieri di Siena

This paper aims to present the initial results of our ongoing investigation of transitivity prominence in Armenian languages. This contribution will help enhance their typological profile and provide, at the same time, a few considerations on how this feature can develop in time and space.

Transitivity prominence can be defined, after Creissels (2018b, p. 180), as the language-specific "tendency to use basic transitive coding with reference to situations other than prototypical actions involving an agent and a patient". That is a recent and rather unstudied linguistic topic (cf. Bossong, 1998; Say, 2014; Haspelmath, 2015, etc.). Almost nothing is known about how and why this feature varies over time, space, and society.

Given a short theoretical background, we will propose a preliminary methodology for measuring it cross-linguistically and, also by taking a look at some known cases, provide some initial hints for a diachronic study of this feature (cf. Creissels, 2018a; 2018b; Creissels & Mounole, 2017). We will, then, present the results of our investigation of Biblical Classical (CA), Modern Eastern (MEA), and Western Armenian (MWA) degrees of transitivity prominence. Relevant data has been collected through annotated corpora (https://arak29.org/bible/ for CA, http://www.eanc.net for MEA, and http://www.nayiri.com/text-corpus for MWA), online databases (Khurshudian & Daniel, 2013; Kocharov, 2022), physical and digital dictionaries (Ciackciack, 1837; Decours, et al., 2014, etc.), grammars, and by interviewing native speakers. To conclude, present what variations in transitivity prominence can be detected in the Armenian *continuum* and in what measure these relate to the morphosyntactic macro-changes that took place in the meantime.

This study will provide a first picture of how this understudied typological trait can evolve within a group of languages and its relations with other morphosyntactic features.

Works cited

Bossong, G., 1998. Le marquage de l'experient dans les langues d'Europe. In: J. Feuillet, ed. *Actance et valence dans les langues de l'Europe.* Berlin: De Gruyter Mouton, pp. 259-294.

Ciackciack, E., 1837. *Dizionario armeno-italiano.* Venezia: Tipografia mechitarista di San Lazzaro.

Creissels, D., 2018a. The Obligatory Coding Principle in diachronic perspective. In: F. Zúñiga, ed. *Typological hierarchies in synchrony and diachrony.* Amsterdam: John Benjamins, pp. 59-109.

Creissels, D., 2018b. Transitivity prominence in typological perspective: the case of Basque. *Anuario Del Seminario De Filología Vasca "Julio De Urquijo"*, 52, pp. 175-187.

Creissels, D. & Mounole, C., 2017. Non-canonical valency patterns in Basque, variation and evolution. In: L. Hellan, A. Malchukov & M. Cennamo, eds. *Contrastive studies in verbal valency*. Amsterdam-Philadelphia: John Benjamins, pp. 151-174.

Decours, A., Ouzounian, A., Riccioli, T. & Vidal-Gorene, C., 2014. *Project Calfa.fr, dictionary of Classical Armenian.* [http://calfa.fr/].

Haspelmath, M., 2015. Transitivity Prominence. In: A. L. Malchukov & B. Comrie, eds. *Valency classes in the world's languages: A Comparative Handbook. Vol. 1.* Berlin: De Gruyter Mouton, pp. 131-147.

Khurshudian, V. & Daniel, M., 2013. Eastern Armenian (standard Eastern Armenian). In: I. Hartmann, M. Haspelmath & B. Taylor, eds. *Valency Patterns Leipzig.* Max Planck Institute for Evolutionary Anthropology: Leipzig. [https://valpal.info/contributions/east2283].

Kocharov, P. 2022. Classical Armenian. In: S. Luraghi, M. Cennamo & A. Palmero Aprosio, eds. *PaVeDa – Pavia Verbs Database*. [https://paveda.unipv.it/contributions/clas1256#tabout].

Say, S., 2014. Bivalent verb classes in the languages of Europe. A quantitative typological study. *Language Dynamics and Change*, 4, pp. 116-166.

Petrova Varvara

Towards the Acoustic Description of Sibilants in Two Varieties of Digor Ossetic

1. Introduction

Digor is one of the two major dialects of Ossetic (‹ Iranian ‹ Indo-European), alongside Iron. Instrumental research of Digor spoken data is scarce, and the differences exhibited by its subdialects have never been investigated from an acoustic standpoint before. Two main varieties of Digor Ossetic are Digora and Chikola (see [Tsallagova 2015]). These varieties differ in realisation of sibilants before /i/ and /e/. Chikola exhibits a more posterior place of articulation (roughly [ʃ], [ʒ], [ʧ], [ʤ]) as opposed to Digora (roughly [s], [z], [ts], [dz]). Present study is an investigation of the voiceless fricative and affricate before the front vowels in both varieties from the viewpoint of their spectral properties, which, along with the amplitude of vowel and frication, serve as the main cue to the place of articulation of fricatives (see e.g. [Jongman et al. 2000]).

2. Theoretical background

The first four spectral moments of fricative spectra are some of the features most commonly investigated for correlation with the place of articulation of sibilant sounds, see e.g. [Kochetov 2017] on fricatives and [Żygis et al. 2012] on affricates. They include spectral Center of Gravity (CoG; the intensity-weighted main frequency of the spectrum), Standard Deviation (SD; the dispersion of the spectrum), Skewness (the tilt in energy distribution), and Kurtosis (the peakedness of the spectrum).

Additional measurements of the first spectral moment at the onset, midpoint and offset of the noise spectra were introduced in [Kochetov 2017], which provided the acoustic characterization of Russian sibilant fricatives /s/, /sʲ/, /ʂ/ and /ʃʲ/ with a dynamic perspective. All of these fricatives were shown to have a higher CoG at midpoint than at offset or onset.

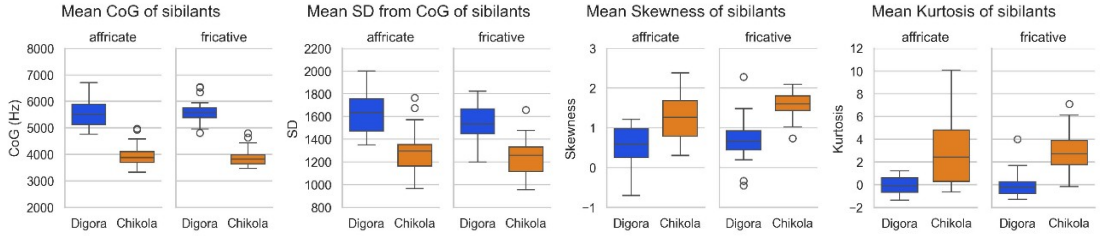
3. Data

Data for the study were collected in Vladikavkaz (Republic of North Ossetia-Alania, Russia) in July 2024. The utterances were obtained from four speakers: two female native speakers of Digora Digor, one male and one female speaker of Chikola Digor.

The questionnaire consisted of the words with /s/ and /ts/ in the word-inital position before front vowels. The consultants were instructed to pronounce each word which they were familiar with in a natural manner and then repeat it twice with a pause after each utterance. They were then asked to pronounce the same word in a carrier phrase “Say [word] three times” (also thrice with pauses between the sentences). The recordings were made in a quiet room, using a Zoom H5 recorder and a WH20XLR microphone.

The segmentation was carried out manually in Praat [Boersma & Weenink 2024]. Mean values of the four spectral moments throughout the frication and CoG at its onset, midpoint and offset were measured for each sound.

4. Results



a. b. c. d.

Figure 1. Boxplots for (a) CoG (Hz), (b) SD (Hz), (c) Skewness and (d) Kurtosis by variety and manner of articulation

As expected, all four spectral moments serve as cues for place of articulation, the correlation with anteriority being positive for CoG and SD from CoG and negative for Skewness and Kurtosis.

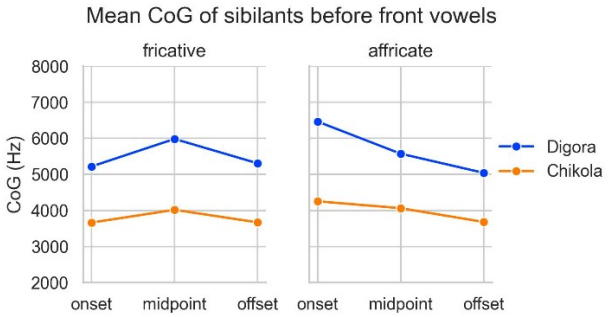


Figure 2. Mean CoG of sibilants before front vowels at onset, midpoint and offset of fricative spectra

From the dynamic viewpoint, fricative sounds of both varieties were found to follow the general pattern demonstrated in [Kochetov 2017]. Both Digora and Chikola affricates, however, contravened it, having the onset point as the highest value of CoG. The tendency for the preceding dental burst to positively affect the initial CoG values of the frication regardless of the presence of phonological boundary (i.e. both in affricates and plosive-fricative sequences) will be discussed in the talk. Since anterior Digora affricates have a significantly steeper slope compared to the posterior Chikola sounds, a new correlate for place of articulation, the difference in (Onset-Midpoint) CoG values, is proposed.

5. References

Boersma, Weenink 2024 — Boersma, Paul & Weenink, David (2024). Praat: doing phonetics by computer. URL: http://www.praat.org

Jongman et al. 2000 — Jongman, Allard & Wang, Yue & Sereno, Joan. (2000). Acoustic and perceptual properties of English fricatives. *Proc. 6th International Conference on Spoken Language Processing (ICSLP 2000)*, Vol. I. P. 536–539.

Kochetov 2017 — Kochetov, Alexei. (2017). Acoustics of Russian voiceless sibilant fricatives. *Journal of the International Phonetic Association*, 47(3). P. 321–348.

Żygis et al. 2012 — Żygis, Marzena & Pape, Daniel & Czaplicki, Bartłomiej. (2012). *Dynamics in sibilant systems: Standard Polish and its dialects*. Jena: Friedrich-Schiller-Universität Jena.

Tsallagova 2015 — Цаллагова И. Н. (2015). О говорах переходного типа (на осетинском языковом материале). *Известия СОИГСИ*, 16(55), с. 77–84.

Semantic transitivity and the Circassian ablaut

Rainer Feer
(SIL Global)

The Northwest Caucasian language West Circassian, together with the closely related East Circassian, makes extensive use of the antipassive to derive monovalent forms from transitive verbs. Quite unusually, this even extends to deriving monovalent forms from bivalent intransitive verbs, called the "indirect antipassive" (Arkadiev and Letuchiy, 2021). Many have pointed out that the antipassive in Circassian is sometimes, or usually, accompanied by a change of vowel (Rogava and Keraševa, 1966; Gishev, 2008). Some call this quite regular vowel change "ablaut" (Kumakhov, 1989; Chirikba, 1996). In this system, transitive verbs end on the high vowel [ə] while the derived monovalent forms end on the low vowel [ɐ], henceforth written *e*. Thus, the antipassive form of the verb in (1a) comes with the *e*-vowel, while the transitive form has /ə/:

(1) (elicited)

a. *č'emə-r ma-šx**e***
cow-ABS DYN-eat
'The cow is eating.'

b. *č'emə-m wəcə-r j-e-šx**ə***
cow-OBL grass-ABS 3SG.A-DYN-eat
'The cow is eating grass.'

However, Lander and Letuchiy (2017) raised cautious objections against explaining the Circassian ablaut purely in terms of antipassive marking. They call into mind the earlier observation about Circassian by Arkadiev and Letuchiy (2008) that the P-argument of a transitive verb is sometimes not deleted but demoted to an indirect object, heeding the same vowel changes. Furthermore, the vowel changes go both ways: we can assume the intransitive verb *ma-ḳʷe* 'goes' to be the base form from which the transitive *je-ḳʷə* 'goes a certain distance' is derived. Still, the vowel distribution is the same. Their conclusion is that the Circassian ablaut displays semantic transitivity in the spirit of Hopper and Thompson (1980), where "*ə*-verbs are more semantically transitive than *e*-verbs" (Lander and Letuchiy, 2017, p. 26).

In this paper, I test the idea that the Circassian ablaut mainly expresses semantic transitivity by using new research data. Analysing well over five hundred unique verb roots I find three observations in its favour:
I) There is a correlation between the three basic construction types (monovalent intransitive, bivalent intransitive, and transitive) and the vowel they occur with: while monovalent verbs generally prefer *-e* and transitive ones prefer *-ə*, bivalent intransitive verbs hold the middle ground by having a less clear preference for *-ə*. This hints at a gradual instead of a categorical phenomenon, and thus at degrees of transitivity.
II) Looking at verbs that occur only in the monovalent construction, i.e. verbs that are clearly underived from a more transitive verb, we see a strong preference for the *e*-vowel in homogenous verbs. These are verbs that "go on in time in a homogeneous way" and where "any part of the process is of the same nature as the whole" (Vendler, 1957, p. 146). Also, there is a slight corelation between the *-ə* final vowel and verbs of change. For example, homogenous verbs like *me-ŝəʁʷe* 'grieves', *me-gʷəse* 'is offended' or *ma-ble* 'shines, burns' usually come with the *e*-vowel while verbs of directed activity like *me-tḳʷə* 'melts' and *me-čə* 'hardens' have a slight preference for the *ə*-vowel. In this case, there is no antipassive derivation involved and still atelic verbs prefer the *e*-vowel. And since telicity is related to higher semantic transitivity, we again conclude that the *e*-vowel correlates with lower semantic transitivity.
III) Perhaps most crucially, there are at least four verbs that distinguish degrees of semantic transitivity just by their final vowel, without any change in their morphosyntactic alignment. Compare the superficial impact on the P-argument in (2) with the *e*-vowel (changed to /*a*/ under stress) on the verb and the total impact on the object in (3) with *-ə*:

(2) ["Adyghe Maq" (adygvoice.ru) 20.04.2010]

ç̣ə-r	*avtomobil'*	*š'ereχ-xe-m*	*a-wəba-ʁ*
ground-ABS	car	tire-PL-OBL	3SG.A-pound-PST

'The car tires pounded the ground.'

(3) [untitled corpus example]

bž'ənəf	*š'əʁʷə-r-jə*	*ə-wəbə-ʁ*
garlic	salt-ABS-ADD	3SG.A-pound-PST

'She pounded the garlic salt.'

The clause in (2) describes the current state of a once beautiful clearing where now cars ruin the plants. The situation entails superficial impact on the object. The verb in (3) can also be translated 'crush' since this is how garlic salt is being produced. It speaks of total impact and accordingly has the more transitive *ə*-vowel. That is, without any morphosyntactic derivation involved and without a change in valency this small group of verbs has the *e*-vowel for the version with lower transitivity and the *ə*-vowel for the higher transitivity version.

Interestingly, this is not the only area in Circassian grammar where sensitivity to semantic transitivity surfaces. Already Yakovlev (1946) showed that verbs that completely impact the object tend to stand in the ergative construction while verbs with partial impact on the object often stand in the bivalent intransitive construction. In other words, verbs with higher semantic transitivity prefer the ergative construction. Therefore, I conclude that the Circassian ablaut is embedded in a broader system that pays attention to degrees of transitivity.

Lastly, the Circassian ablaut and the antipassive may still be related, just the exact nature of that relationship needs to be re-examined. Synchronically speaking, the function of the Circassian ablaut is much broader in that it signals semantic transitivity regardless of derivation being involved or not.

References

Arkadiev, P. and Letuchiy, A. (2008) 'Derivacii antipassivnoj zony v adygejskom jazyke', in Plungian, V. & Tatevosov, S. (eds.) *Issledovanija po otglagol'noj derivacii*. [Antipassive derivation in the Adyghe language]. Moscow: Jazyki slavjanskix kul'tur, pp. 77–102.

----- . (2021) 'Indirect Antipassive in Circassian', *Antipassive: Typology, Diachrony, and Related Constructions, ed. Katarzyna Janic–Alena Witzlack-Makarevich, Amsterdam–Philadelphia: John Benjamins* pp. 483–514.

Arkhangelskiy, T., Bagirokova, I., Lander, Y.A. and Sorokina, A. *West Circassian (Adyghe) Corpus. (Available online at: adyghe.web-corpora.net)*.

Chirikba, V.A. (1996) *Common West Caucasian. The Reconstruction of Its Phonological System and Parts of Its Lexicon and Morphology*. Leiden: Research School CNWS.

Gishev, N.T. (2008) *Izbrannyje trudy po jazykoznaniju*. [Selected works in Linguistics]. Maykop: Kačestvo.

Hopper, P.J. and Thompson, S.A. (1980) 'Transitivity in Grammar and Discourse', *Language*, 56(2), pp. 251–299.

Kumakhov, M. (1989) *Sravnitel'no-Istoričeskaja grammatika adygskix (čerkesskix) jazykov*. [Comparative-Historical Grammar of Adyghe (Circassian) languages]. Moscow: Nauka.

Lander, Y.A. and Letuchiy, A. (2017) 'Decreasing Valency-Changing Operations in a Valency-Increasing Language?', in Alvarez, A. & Navarro, Í. (eds.) *Valency changes. Theoretical and typological perspectives*. Amsterdam: John Benjamins Publishing Company, pp. 285–304.

Rogava, G.V. and Keraševa, Z.I. (1966) *Grammatika adygejskogo jazyka*. [Grammar of the Adyghe language]. Krasnodar: Krasnodarskoe knižnoe izdatel′stvo.

Vendler, Z. (1957) 'Verbs and Times', *The philosophical review*, 66(2), pp. 143–160.

Yakovlev, N.F. (1946) 'Drevnye jazykovye svjazi Evropy, Azii i Ameriki. [Ancient Linguistic Ties between Europe, Asia, and America]', *Iz. Akad. nauk SSSR, Otdel Lit. i jaz.*, V(2), pp. 141–148.

Demonstratives, Referentializers, Tense formants – A comparison between Khinalug and Lezgic

Monika Rind-Pawlowski, Goethe University Frankfurt

Khinalug and several languages of the Lezgic branch have some striking paralells in their morphology of demonstratives, the use of demonstratives as referentializers, and the use of referentializers as tense formants. The productive Khinalug demonstratives are *dV* (PROX) and *hV* (DIST). Besides, there are relics of an older system with *e* (PROX) and *o* (DIST). The bare productive proximal demonstrative *dä* is used as an adjectivizer, and *dä* with traces of an ancient class marking system I *-w-* (hence /ä/ > /u/), IV *-y-* (hence /d/ > /ʒ/, /ä/ > /i/) as a nominalizer.

The petrified system *e* (PROX) vs. *o* (DIST) reminds of Schulze's reconstruction for Proto-Udi **-i* 'region of the speaker' (modern Udi /i/ > /e/) vs. **-a* 'beyond' (Schulze 2008: 252ff.) as well as Lezgian *i* 'this' and *a* 'that' (Haspelmath 1993: 190). In Caucasian Albanian, there is only one demonstrative, which makes no distinction between proximal or distal. Here, the attributive forms are Masc. *o*, Fem. *a*, Neutr. *e*, nominalized as Masc. o(o), Fem. *aġ*, Neutr. *e(e)*. Just like the Khinalug proximal demonstrative *dV*, the nominalized pronouns are used as referentializers in CA (Gippert & Schulze 2023: 186-187). Demonstratives consisting of only one distinct vowel are not attested anywhere else in the ND languages.

The Khinalug distal demonstrative *hV* is certainly a cognate of the Lezgic emphatic particle *ha*. Notably, Khinalug uses both bare *hä* as emphatic particle and *hä* + traces of an ancient class marking system: I *-w-* (hence /ä/ > /u/) as a demonstrative (for *ha* in Lezgic see Alekseev 1985: 75 for PL, Schulze 39 for Tsakhur, Alekseev 1994: 225 for Rutul, Babaliyeva 2013: 73 for Tabasaran, Haspelmath 1993: 190 for Lezgian, Authier 2009: 62 for Alik, Hummatov & Rind-Pawlowski 2020: 66 for Kryz, Authier unpublished: 13 for Budugh). *hV* is not used as a referentializer in any of the languages. Most probably, *hV* is inherited from Proto-ND into modern Khinalug, independently from Lezgic. *hV* as a demonstrative, usually with vowel variation to distinguish distances to the speaker (or the listener), is attested particularly in the Avaro-Andic branch, e.g. for Avar (Forker 2020: 251), Godoberi (Kibrik 1996: 42), Akhvakh (Creissels (forthcoming): 14), Andi (Sulejmanov 1957: 324ff.), and Karata (Magomedbekova : 104:ff), and demonstratives that have integrated the emphatic particle occur in some Tsesic languages, e.g. in Hinuq (Forker 134 ff), Bezhta (Comrie et al. 301 ff), and Khwarshi (Khalilova 143 ff.), and *ha-* as a prefix of demonstratives also occurs in Dargwa varieties, e.g. Sanzhi (Forker 2020: 90 ff.), Icari (Sumbatova & Mutalov 2003: 37 ff.), and Tanti (Sumbatova & Lander 2014: 76 ff).

The Khinalug proximal demonstrative/referentializer *dV* is certainly related to what Schulze reconstructs as Proto-South-East-Caucasian **dV-* (distal, horizontal) (next to **mV-* (proximal), **ɢu-* / **ʁu-* (distal, higher) and **gu* (distal, lower), cf. Schulze 1997:38). Notably, the phonetic process triggered by the class marker *-y-* in Khinalug (i.e. dä > *ǯi*) is identical to the process observed at the referentializer *-d(V)* in the South Samur languages, i.e. in languages that nowadays do not have any demonstrative **dV*.

Also in the Kryz varieties, the suffixed referentializers have -d in CL1-3, and -ǯ in CL4. In the Kryz variant of Kryz village, these serve as nominalizers at adjectives, while attributive adjectives end in an attributive marker *-ä* (*-äʕä* after vowel). E.g. "wet": Predicative *č'äb,* Attributive: *č'äb-ä,* Nominalized (absolutive): CL1-3: *č'äb-ä-d*, CL4: *č'äb-ä-ǯ.* The same pattern is also attested for any other nominalized word class, including independent (i.e. nominalized) demonstratives, e.g. Attributive: *lu* (PROX), *lä* (DIST), Independent (nominalized): CL1-3: *lu-d* (PROX), *lä-d* (DIST), CL4: *lu-ǯ* (PROX), *lä-ǯ* (DIST).

In the Ismayilli Haput variant of Kryz, both attributive and nominalized (absolutive) adjectives (including participles) carry the markers *-d* for CL1-3 and *-ǯ* for CL4 reference. The mere nominalizing function of these elements is only maintained in demonstratives, which distinguish attributive *lu* (PROX) and *lä* (DIST) from independent (nominalized) CL1-3 *lu-d*, *lä-d* CL4 *lu-ǯ*, *lä-ǯ* like in Kryz proper.

For attributive and nominalized adjectives in Alik, which mostly follow the Kryz proper pattern, see Authier (2009: 67 ff.) In Budugh, adjectives are attributed in their bare stem form. The suffixes CL1-3 *-d* vs. CL4 *-ǯ* function as nominalizers of adjectives and also occur at independent demonstratives (CL1-3: *ud* (PROX), *ad* (DIST), CLIV *uǯ* (PROX), *aǯ* (DIST) (Authier: unpublished: 12).

The Khinalug Perfect and the Aorist of all South Samur languages obviously go back to participles nominalized according to this pattern, and they all show the *-d* vs. *-ǯ* variation. E.g. the Khinalug Perfect of "die, kill", formed from the perfective participle: CL1: *k'i-du*, CL2: *zi-k'i-dä*, CL3: *bi-k'i-dä*, CLIV *k'i-ži* (where /ʒ/ is abraded into /ž/). Likewise the Aorist e.g. of "roll" in Kryz proper: CL1: *qızır-d*, CL2+CL3: *qı<b>zır-d-u* (where *-u* < CL2/3 *-v*), CL4: *qızır-ǯ*, and the Aorist e.g. of 'bring' in Ismayilli-Haput: CL1: *ʕäʁä-d*, CL2: *ʕäʁo-d-u* (where *-u* < CL2/3 *-v* and *-o-* due to regressive assimilation), CLIV: *ʕäʁä-ǯ* (see also Authier 2009: 141 for Alik, Authier unpublished: 26 for Budug.)

The Khinalug Future is expressed by the nominalized imperfective participle, e.g. the Future of "die, kill": CL1: *k'li-du*, CL2: *k'li-dä*, CL3: *k'li-dä*, CLIV *k'li-ži* (/ʒ/ > /ž/). Ismayilli Haput has a dialectal distinction

between a future tense based on /d/ (Mollaisaqli) and /ʒ/ (Hajihatamli). E.g. the future of "go out, leave" in the Mollaisaqli variant: CL1,4: *qärč'är-d-e*, CL2,3: *qärč'är-d-o*, vs. in the Hajihatamli variant: CL1,4: *qärč'är-ʒ-e*, CL2,3: *qärč'är-ʒ-o.* Also here, similar to the Khinalug future tense, we may assume an underlying nominalization pattern. The former class specific distinction between CL1-3 /d/ and CL4 /ʒ/ must have been neutralized by the petrification of either the one or the other form in each dialect.

Regarding the Lezgian language, which has abolished its noun class system, we cannot expect to find class specific variation. Still, we find forms with /d/ that remind of Khinalug, Kryz, and Haput patterns. As for the standard language, the invariable future/habitual suffix *-da* corresponds to the Mollaisaqli suffix *-dä*. The suffix used for nominalization in the singular is *-di* (*-da* in the oblique cases) (Haspelmath 1993: 110). Notably, a suffix *-da* can also be used at adjectives to give them predicative function – simliar to the proximal demonstratives in Khinalug at predicative adjectives (and participles). Notably, the Yargun dialect of Lezgi distinguishes a categorical future in *-da* and an uncertain future in *-d(i)* (Babaliyeva 2023), which reminds of Schulze's theory on *-i* 'region of the speaker' vs. *-a* ('beyond') in Proto-Udi, i.e. the distinct vowel that combined with **m-* (prox), **k-* (med) and **t'-* (dist) in Proto-Udi (Schulze 2008: 252).

Hypothesis: The petrified demonstratives *o* (DIST) and *e* (PROX) in Khinalug as well as the productive demonstrative *hV* (DIST and EMPH. PTCL) are inherited and go back to Proto-ND. The productive demonstrative *dV* (PROX, -ADJ, -NMLZ) is borrowed from Lezgic. Particularly the forms of PROX and -NMLZ have been borrowed from, or developed under the influence of, South Samur. Hence, at some point, South Samur (and also Lezgian) must have had a demonstrative **dV*, which is now extinct.

Literature

Alekseev, M.E. (1994): *Rutul*. In: The indigenous languages of the Caucasus. Vol. 4, North East Caucasian languages, part 2, presenting the three Nakh languages and six minor Lezgian languages. New York, Caravan Books.

Alekseev, M.E. (1985): Voprosy sravnitel'no-istoricheskoj grammatiki lezginskix jazykov, Morfologija, Sintaksis. Moscow, "Nakuka".Authier (unpublished): Budugh. Sketch version available at: https://www.academia.edu/66065917/Budugh_Mai_2021_vuMS_SENT_TM_YK_correctedGA_Pour_Academia

Authier, Gilles (2009): Grammaire Kryz (Langue Caucasique d'Azerbaïjan, Dialekt d'Alik). Lieuven-Paris: Peeters.

Babaliyeva, Ayten (2023): Formes imperfectives dans les dialectes lezgi. Paper presented at: Imperfective Modalities in the Caucasus and Beyond; Projet ANR IMMOCAL –11-12 septembre 2023, Paris.

Babaliyeva, Ayten (2013): Études sur la morphosyntaxe du tabasaran littéraire. Doctoral Thesis.

Creissels, Denis (forthcoming): A sketch of Northern Akhvakh. In Yuri Koryakov and Yury Lander and Timur Maisak (eds.), The Caucasian Languages: An International Handbook, 1-36. Berlin: DeGruyter Mouton.

Forker, Diana (2020a): Avar. In: Maria Polinsky (ed.): The Oxford Handbook of Languages of the Caucausus. Oxford University Press, pp. 243-280.

Forker, Diana (2020b): A Grammar of Sanzhi Dargwa. Berlin: Language Science Press.

Gippert, Jost & Schulze, Wolfgang (2023): The language of the Caucasian Albanians. In: Jost Gippert and Jasmine Dum-Tragut: Caucasian Albania. De Gruyter Mouton, Berlin – New York; pp. 167-229.

Haspelmath; Martin (1993): A Grammar of Lezgian. Mouton de Gruyter, Berlin – New York.

Hummatov, Israfil & Rind-Pawlowski, Monika (2020). Qrızca-azərbaycanca-ingiliscə lüğət (Kryz-Azerbaijani-English Dictionary). Baku: AFPoligrAF.

Kibrik, Aleksandr E. & Tatevosov, Sergej G. & Eulenberg, Alexander (1996): Godoberi. München/Newcastle: Lincom Europe.

Magomedbekova, Z.M. (1971): Karatinskij Jazyk. Tbilisi, Mecniereba.

Schulze, Wolfgang (2008): Deictic Strategies in Udi. In: M.E. Alekseev, T. A. Majsak, D.S. Ganenko, Ju. A. Lander (red.koll.). Udinskij sbornik: Grammatika, leksika, istorija jazyka. Moskva: Academia, 241-310.

Schulze, Wolfgang (1997): Tsakhur. München/Newcastle: Lincom Europa.

Sulejmanov Ja.G. (1957): Grammatičeskij očerk andijskogo jazyka. Doctoral Thesis. Moscow: Institut Jazykoznanija Akademii Nauk Sojuza CCR.

Sumbatova, Nina R. & Mutalov, Rasul (2003): A Grammar of Icari Dargwa. Languages of the World/Materials 92. LINCOM.

Sumbatova, Nina R. & Lander, Juri A. (2014): Darginskij govor selenija Tanty: Grammatičeskij očerk. Voprosy sintaksisa. Moscow: JaSK.

Kabardian adaptation of loanwords: the case of Russian /k/

Chaeli Rule
The Ohio State University

Following the Russian conquest of Circassia in the 19th c., Russian bilingualism among speakers of Kabardian (East Circassian) living in Kabardino-Balkaria has increased rapidly. As a result, the Kabardian language has seen a massive influx of Russian loanwords, and these loanwords have undergone adaptation, or nativization, upon entering the lexicon. The present study aims to identify the strategies employed by Kabardian in adapting Russian /k/, which is not native to the language, and to determine the factors that condition these adaptations. This investigation is supported by an original database consisting of loanwords borrowed directly from Russian into Kabardian, sourcing from Sukhunov & Sukhunova's (1998) *Kabardino-Circassian-Russian-English-Turkish Picture* Dictionary, Kardanova's (1957) *Kabardinsko Russkij Slovar'*, and Lopatinskij's (1890) *Russko-Kabardinskij Slovar'*. In the database, Russian /k/ yields six segments in Kabardian: (1) voiceless velar stop /k/ (imported); (2) labialized voiceless velar stop /kʷ/; (3) labialized voiceless uvular affricate /q͡χʷ/; (4) voiceless palato-alveolar affricate /č/; (5) ejective voiceless palato-alveolar affricate /č'/; and (6) voiceless uvular stop /q/. A large portion of the loans were entirely unadapted, i.e., they were imported directly from Russian into Kabardian with the segment /k/ as in their Russian form. The seemingly peculiar adaptations of /k/ as /kʷ/ and /q͡χʷ/ can be understood as misperceptions of Russian segment sequences that are illicit in Kabardian. The variation in single segment adaptations of /k/ as /č/, /č'/, and /q/, however, is not readily explained by the phonology. Discussions around loanword adaptation typically concern the nature of the process itself and argue for an understanding of the phenomenon as phonetic or phonological (or as a combination of both), but it is clear based on the lack of pure linguistic evidence that this kind of approach cannot account for the variation in Kabardian. The lack of convergence on a single adaptation strategy for /k/ can be explained by the low level of community bilingualism at the time of borrowing and insufficient time for these adaptations to regularize before speakers began wholesale importation of Russian words. The Kabardian adaptation of Russian /k/ provides evidence that loanword adaptation is best understood when both linguistic and extralinguistic factors are taken into consideration.

Alina Russkikh, HSE University

The additive particle =*lo* in the Upper Andi variety of Zilo: on the combination of the additive particle with other markers

Additive focus particles, such as English *also* or German *auch*, are widely discussed in typological literature (König 1991; Krifka 1998; Forker 2016). One of the significant questions related to describing the functions of additives is whether an additive particle is used in a specific function independently or together with other obligatory markers. König (2017) discusses that, unlike some other functions (e.g., simple additivity), for the formation of concessive clauses the additive particle is always used with an additional marker: "Additive focus markers often show up as components of concessive markers (Engl. *even though, even if, even so*), but concessivity is never expressed by an additive marker alone, and requires the compositional effect of a scalar marker and a conditional adverbial" (König 2017: 40). Therefore, it is of particular interest to describe the functions of the additive particle with special attention to contexts in which it is used together with other markers and those in which it is used on its own. This talk will provide such a description of the additive particle *=lo* in the Zilo variety of the Andi language (< Nakh-Dagestanian).

The additive particle *=lo* in Zilo Andi is a multifunctional clitic used both in typologically frequent contexts for additive particles and in more peripheral ones. Table 1 lists the functions in which the additive *=lo* is used and specifies additional markers when it is combined with them.

Table 1. Functions of =lo

Function / combination with other markers	**Marker**	**Mandatory / optional**
Additivity	*=lo*	mandatory
Scalar additivity		
Coordination		
Converbal clauses		
Collective numerals with an exhausting meaning	*-gu=lo*	mandatory
Universal quantifiers		optional
Comitative	*-loj*	is included
Complex numerals		
Concessive clauses	*=bolo*	is included
Indefinite pronouns	*=lo / =gulo, =bolo*	optional / mandatory, is included

As seen in Table 1, the additive particle is used independently in less than half of its functions. In three functions, the additive particle is used in combination with the emphatic particle *=gu*. Notably, the markers *-loj* and *-bolo* likely originate diachronically from a combination of *=lo* with the uncertain marker *-j* and the conditional marker *-bor* (Kaye et al., forthcoming), respectively.

Thus, all the functions of *=lo* in Zilo Andi can be divided into three groups based on the combination of the additive with other markers:

I. The additive clitic is used independently and does not require additional components, see (1).

(1) **Additivity**

aħmadi	cʼːudur=lo
Ahmad	intelligent=ADD

‘Ahmad is also intelligent’.

II. The additive clitic is used together with other markers that should be analyzed as separate markers. In example (2), the additive particle is attached to a cardinal numeral form constructed using *-gu* [Kaye et al., forthcoming] and cannot be directly attached to the numeral stem.

(2) **Collective numerals with an exhausting meaning**

čʼe-gu=lo	/	* čʼe=lo	močʼi
two-EMPH=ADD		two=ADD	child

‘Both children’.

III. The additive clitic is used together with other markers. However, a distinct marker has been formed as a result of fusion. For instance, the marker *-bolo* contains *=lo* diachronically, but on a synchronic level, there is no separate marker *-bo*.

(3) **Concessive Clauses**

cʼːa	retɬi-bolo	sore-ri	w-oʔinn-ija
rain	go-CONC	walk-INF	M-PL.go-FUT

“Even if it rains, we will go for a walk.”

The talk will discuss in detail these highlighted cases of the combination of the additive particle with other markers and their semantic contribution. In addition, the proposed language-specific analysis is expected to contribute to the methodology for describing additives from a typological perspective.

List of Abbreviations

ADD — additive particle, COM — comitative, CONC — concessive marker, EMPH — emphatic particle, FUT — future tense, INF — infinitive, M — masculine gender, PL — plural.

References

Forker 2016 — D. Forker. Toward a typology for additive markers. Lingua. 2016. Vol. 180. P. 69–100.

Kaye et. al., forthcoming — S. Kaye, G. Moroz, N. Rochant, S. Verhees, A. Zakirova. Andi (Zilo dialect) // Y. Lander, T. Maisak, Yu. Koryakov (eds.). The Caucasian Languages. An International Handbook. De Gruyter Mouton, Berlin/New York, forthcoming.

König 1991 — E. König. The Meaning of Focus Particles: A Comparative Perspective. London: Routledge, 1991.

König 2017 — E. König. Syntax and semantics of additive focus markers from a cross-linguistic perspective // A.-M. De Cesare, C. Andorno (eds.). Focus on additivity: Adverbial modifiers in Romance,

Germanic and Slavic languages [Pragmatics & Beyond New Series 278]. Amsterdam & Philadelphia: John Benjamins Publications, 2017. P. 23–44.

Krifka 1998 — M. Krifka. Additive particles under stress // Semantics and Linguistic Theory. 1998. Vol. 8. P. 111–128.

Word Order in a free word order language: A corpus study of Ossetic

Emine Şahingöz, Goethe University Frankfurt

Ossetic is commonly characterized as having a relatively free word order, predominantly left-branching, with both postpositions and a limited number of prepositions supplementing the case forms. While the unmarked order of constituents is often described as SOV, previous studies suggest that all possible orderings of S, O, and V can be grammatical (cf. Erschler 2012: 676; Lowe & Belyaev 2015: 231). This suggests that Ossetic exhibits discourse-configurational characteristics, with constraints on word order choice primarily influenced by information structure.

This variability can be seen in three similar sentences, each with a distinct constituent order: (1) subject-verb-object, (2) object-subject-verb, and (3) object-verb-subject. Despite the different orderings, all three sentences are considered grammatical, highlighting the absence of a fixed word order in Ossetic:

(1) Iron Ossetic

Respublikæ-jy	*prezident*	*ba-fys-ta*	*ukaz*	*radon*	*æfsad-mæ*	*ærsid-ty*	*tyxxæj.*
republic-GEN	president	PV-write-3SG.PST	decree	regular	army-ALL	call-GEN.PL	for

"The president of the republic signed a decree for the call to the regular time at the army." (ONC)

(2) Iron Ossetic

Æmbælon	*ukaz*	*Xussar*	*Iryston-y*	*prezident*	*ba-fys-ta*	*majræmbon-y.*
respectable	decree	south	Ossetia-GEN	president	PV-write-3SG.PST	Friday-INESS

"The president of South Ossetia signed a respectable decree on Friday." (ONC)

(3) Iron Ossetic

Æmbælon	*ukaz*	*ærtyccædž-y*	*ba-fys-ta*	*prezident*	Tybyl-ty	*Leonid.*
respectable	decree	thirty-GEN	PV-write-3SG.PST	president	PN-GEN.PL	PN

'The president Tybylty Leonid signed a respectable decree on the 30th.' (ONC)

However, while constituents may not adhere to a fixed order, certain tendencies emerge in specific contexts. For instance, clitic pronouns in tmetic constructions tend to follow a consistent order, with the indirect object appearing before the direct object, as can be seen in the following examples:

(4) Digor Ossetic

je	*dær*	*ra-somi*	*kod-ta*	*čidær*	*kor-aj*
3SG	also	PV-oath	do.PST-3SG.PST	whatever	ask-3SG.SUBJ

ra-din-æj-æt-dzænæn	*zæğ-gæ*
PV-CLIT.2SG.DAT-CLIT.3SG.GEN-give-1SG.FUT	say-CV

"He also swore and said: 'whatever you ask for, I will give it to you.'" (Narty)

(5) Digor Ossetic

ærba-jimæ-'j-xas-toncæ	*fæstæ-mæ*	*'sk'el-i*	*[...]*
PV-CLIT.3SG.ALL-CLIT.3SG.GEN-take.PST-3SG.PST	back-ALL	heel	
ni-jin-æj-zild-toncæ	*je*	*'sk'el-i*	*[...]*
PV-CLIT.3SG.DAT-CLIT.3SG.GEN-turn.PST-3SG.PST	3sg.GEN	heel	

"They took it to him (and) turned his heel back to him […]." (Narty)

Further analysis of direct object marking reveals additional tendencies, particularly concerning pronominal elements. In a corpus-based study analyzing over 4000 examples, direct objects were categorized based on animacy and markedness. Results indicate a tendency for pronominal direct objects to appear before the verb, regardless of markedness or animacy. This observation aligns with the consistent pronominal word order observed in tmetic constructions.

In conclusion, while Ossetic exhibits a generally free word order, tendencies emerge in specific linguistic contexts, particularly concerning pronominal elements. This study contributes to our understanding of word order variability in Ossetic and highlights the importance of considering both structural and functional factors in linguistic analysis. In this talk, I will present instances where tendencies for a fixed word order can be seen in Ossetic language that is described to be a language with a free word order.

References

Erschler, David. 2012. From preverbal focus to preverbal "left periphery": The Ossetic clause architecture in areal and diachronic perspective. *Lingua* (122(6)). 673–699.

Lowe, John J. & Oleg Belyaev. 2015. Clitic Positioning in Ossetic. In Miriam Butt & Tracy H. King (eds.), *Proceedings of LFG15* (CSLI Publications), 228–249. Stanford.

Narty = Xamicaeva, A. A. & Bjazyrov, A. X. 1990. *Narty*. *Osetinskij geroičeskij épos 1*. Moskva: Glavnaja Redakcija Vostočnoj literatury.

Vydrin, Arseniy P. & Belyaev, Oleg et al. 2011. *Ossetic National Corpus* (ONC).

Available online at http://corpus.ossetic-studies.org/search/index.php?interface_language=en, accessed 11.05.2024.

Locative Forms in Kaytag Dargwa: Overview and Grammaticalization

In Dargwa languages nominal paradigm consists of so-called grammatical cases and locative forms. While grammatical cases have a single marker, Tanti and Kaytag locatives include three ordered inflectional slots: localizaion (LOC), orientation (OR) and direction (DIR).

In most Dargwa languages locative forms can be used in non-spatial contexts (see Sumbatova, Lander 2014; Chechuro 2016). Some Tanti and Chirag locative forms according to (Ganenkov, Lander 2011) lost their spatial meaning, however, continue to be used in non-spatial contexts, and thus this leads to the emergence of new non-locative cases. For instance, in Tanti DIR can be expressed with LAT and ELAT orientations, in combination with former it is optional, with latter – obligatory, except for three grammaticalized forms with SUPER, ANTE and INTER localizations. Those three unmarked for direction ELAT forms only have non-spatial usage, for example INTERELAT can be used with verb 'to take from X' (1). Ganenkov and Lander (2011) suggest that these forms are new non-locative cases.

(1) *rasul-li* *durħaˁ-li-cːe-r* *žuž* *s-asː-ib*
PN-ERG guy-OBL-INTER-EL book HITHER-take.PF-PRET
'Rasul took the book from the guy' (Sumbatova, Lander 2014: 74)

In this talk, we argue that in Kaytag (data collected during fieldwork in the village of Javgat in June 2024) there is evidence that similar process is taking place, however, it is accompanied by other significant changes in the nominal paradigm: reduction in the number of LOC category values, increase in the number of allomorphs of SUPER and IN localizations and loss of ESS marker.

In Kaytag, according to (Temirbulatova 2004), there are four DIR values: *-k'in* 'up', *-χin* 'down', *-tin* 'away from the speaker', *-žin* 'towards the speaker'. However, our data suggests that *-žin* is not used in Javgat (2).

(2) *atːa* *dubur-cːi-r-(*žin)* *h-ag-ur*
father.ABS mountain-IN-ELAT-DIR UP-become.PF-AOR
'Father went up the mountain'

DIR in Kaytag can only be expressed with ELAT orientation (3) and is not obligatory (4).

(3) *atːa* *dubur-cːi-r-k'in* *h-ag-ur*
father.ABS mountain-IN-ELAT-DIR UP-become.PF-AOR
'Father went up the mountain'

(4) *du* *čːemi-j-tːar* *r-eg-ur-da*
1SG.ABS bridge-SUPER-ELAT F-go.PF-AOR-1
'I got off the bridge'

However, DIR suffix *-χin* 'down' is required when expressing temporal value 'from some time' (5):

(5) *ijun-cːi-r-χin* *nejgi* *qil* *her* *d-irh-u-d*
june-IN-ELAT-DIR new house.IN live PL-stay.IPF-FUT-1
'We will be living in a new house since June'

Another interesting tendency regarding DIR category is that it is frequently used in contexts with PROLAT orientation meaning, cf. (6)[1].

(6) *erk'-la* *xar-ži-r-k'in* *waˁ* *it'in* *burχːa* *čibagar-el*
river-GEN bottom-SUPER-ELAT-DIR go.IMP red roof.ABS see-TEMP
'Walk along the river to the house with the red roof!'

[1] Most of marked for direction instances in our data can be interpreted as having a PROLAT meaning. However, it should be noted, that for Dargwa and other Caucasian languages native speakers the difference between ELAT and PROLAT is not that distinct. At the same time, they are more sensitive to level changes, which might be the cause for such tendency.

So, the data suggests that these forms are in the process of a semantic shift since they are less actively used in their original deictic and gravitational function and required in non-spatial context.

As mentioned, this shift in Kaytag is accompanied by other changes in locative forms paradigm. Main related changes occur in the LOC category. According to (Ganenkov, Lander 2011) and (Chechuro 2016) INTER and SUPER are the most used in non-spatial contexts across all Dargwa languages. It seems that it is also true for Kaytag, and the high frequence of usage might be the cause for the increase in quantity of SUPER and IN localizations markers.

In the village of Javgat SUPER localization can be expressed with four markers: *-ž(i), -j(a), -čːi* and *-e*. Former two are highly productive, latter two are used with a limited set of nouns. Suffixes *-ž(i)* and *-j(a)* are used actively in their spatial meaning (7-8) and they also function as DAT markers see (9).

(7) *hel* *dubur-ži* *h-ag-ur*
DEM.ADR.CNTR mountain-SUPER(LAT) UP-become.PF-AOR
'He went up the mountain'

(8) *du* *čːemi-ja* *k-Ø-icː-ul* *da*
1SG.ABS bridge-SUPER(ESS) DOWN-M-stop.IPF-CNV COP:PRS.1
'I am standing on the bridge'

(9) *b-ekː-a* *ila* *χʷaˁ-l-ži* *dig*
N-give.IPF-IMP 2SG.GEN dog-OBL-DAT meat.ABS
'Give your dog meat!'

According to our data, *-čːi* (10) is idiosyncratically used with a few nouns: *mikːeri* 'stair', *daˁʁi* 'fence', *maˁla* 'window', *daˁʔ* 'face' and *-e* (11) with *bek'* 'head', *bazar* 'bazaar' and *erk'* 'river'.

(10) *rirsi-la* *daʔ-a-čːi* *pinc'a* *biʔira*
girl-GEN face-OBL-SUPER(ESS) dirt.ABS BE:PRS.3
'The girl has dirt on her face'

(11) *maˁč* *dila* *bek'-e* *ka-b-ič-iw*
ball.ABS 1SG.GEN head-SUPER(LAT) DOWN-N-fall.PF-AOR
'The ball fell on my head'

In most Dargwa languages *-cːi* is used to mark INTER, however, in some dialects, including Kaytag, it has widened its usage, also marking IN. Thus, IN has two markers *-cːi* and *-(a)n*, see (12).

(12) *ʔaˁšak-an* *χink'-i* *čajnik-{cːi/an}* *šin*
pot-IN(ESS) hinkal.ABS-PL kettle-IN(ESS) water.ABS
'There is hinkal in the pot, water in the kettle'

In the talk, morphology and both spatial and non-spatial usage of Kaytag Dargwa locatives will be discussed more thoroughly. Grammaticalization prosses of Kaytag marked for DIR forms will be described and compared with Ganenkov and Lander's (2011) results for Tanti and Chirag.

References

● Chechuro I. Yu., Comparative Study of Locative Systems of Northern Dargwa Dialects. 2016 ● Ganenkov D. S., Lander Yu. A. Locative Forms as a Source of Non-locative Cases: Dargwa Data // Acta Linguistica Metropolitana. Proceedings of the Institute of Linguistic Research. 2011. No.3. ● Sumbatova N. R., Lander Yu. A. Dargwa Dialect of the Village of Tanti: An Essay on Grammar, Questions of Syntax: Monography. M.: Languages of Slavic culture, 2014 ● Sumbatova N.R., Mutalov R.O. A Grammar of Icari Dargwa. [Languages of the World Series 92]. München Newcastle: LINCOM Europa, 2003. ● Temirbulatova S.M.: Khaydak Dialect of the Dargwa language, Monography. Makhachkala, 2004

Satina Maria, Rasiukevich Anna, Nerush Varbara

The Ossetic Attenuative Marker *-gomaw*: beyond Scalarity

Introduction. The present study is concerned with the description and analysis of a number of remarkable features in the distribution and functioning of the attenuative marker *-gomaw* in Ossetic (Eastern Iranian < Iranian < Indo-European), a language of the Northern Caucasus. Based on the descriptions presented in Ossetic grammars (Abaev 1950, p.118; Bagaev 1965, p.383, ; Axvlediani 1963, p.140), one is led to believe that this suffix acts as a scalar modifier, its main function commonly formulated as "denoting the reduced degree of a quality". However, our field data suggest that the functions of *-gomaw* go beyond just attenuation in its canonical understanding (see e.g. Grandi & Körtvélyessy 2015), but are rather concerned with encoding the speaker's evaluation of a certain quality, therefore the semantics of the discussed marker are highly subjective in nature. The data also reveal certain facts on the distribution of *-gomaw* (discussed in more detail below) that were previously overlooked in grammars and that make a significant contribution to our understanding of the marker's semantics and functions.

Data and methods. The data for this study were obtained by elicitation, as well as from the Ossetic National Corpus (ONC). The elicited data were collected from nine speakers of the Iron and Digor dialects during two field trips to Vladikavkaz (July 2024, October 2024) and one speaker of the Kudar dialect in Moscow (October 2024). We express our gratitude to Eleonora Izmailova (a speaker of the Iron dialect) for the assistance with glossing of the examples.

Findings and analysis. The main phenomena under discussion in the present study are the following:
1. As already stated above, our data show that the semantics of *-gomaw* in modern Ossetic do not exactly match the descriptions presented in the existing literature. We believe that *-gomaw* has undergone the development from serving as an attenuative scalar modifier in the classical sense, as documented in the grammars, to becoming a subjective evaluation marker the semantics of which can be described as 'according to the speaker's perception, not exactly matching the description that is encoded in the base word', see (1). In other words, synchronically, *-gomaw* seems not to modify the degree of a certain quality *per se*, but to denote the speaker's skepsis regarding the accuracy of the description of this quality. Thus emerges the grammaticality of (2) and (3), where *-gomaw* occurs on adjectives modified by *ӡgӕr* 'overly' and *tӕng* 'very' that both encode the amplification of a quality; we assume that if *-gomaw* was to serve as a canonical scalar modifier, a contradiction would arise and these examples would be considered ungrammatical. We must also note that whenever a speaker aims to express the reduced degree of a certain quality without giving it his/her subjective evaluation, the strategy of encoding it by analytical rather than synthetic means is preferred; compare (4) and (5), where the word *gӕc:ӕl* 'little' is used to convey the idea that Amina is only a little taller than Alan.

2. It is stated in (Axvlediani 1963, p.140) that *-gomaw* occurs solely in the adjectival/adverbial[1] domain; however, we found that it can also occur on participles, both active (6) and passive (7). Cross-linguistically, attenuative markers commonly occur on numerals, normally yielding the meaning 'roughly being in the quantity of X'; however, our data also shows that *-gomaw* cannot occur on numerals, which appears to correlate with the idea that the marker denotes the speaker's stance regarding the description of a quality: it seems to be counterintuitive to express skepsis towards an accurate numerical quantity.
The use of *-gomaw* in the verbal and nominal domains has been shown to be impossible.

3. The discussed marker can modify adjectives with the meaning 'possessing an object or a number of objects' when the aforementioned objects are uncountable or non-discrete (*qʷӕn-ӝӕn-gomaw* [hair-PROPR-ATT] 'kind of hairy', *ӡnc'ӕlt:ӡg-gomaw* [wrinkle.PTCP-ATT] 'kind of wrinkly'); whereas the use of *-gomaw* with adjectives denoting the possession of a particular, fixed number of objects is impossible (**sӕšt-ӝӕn-gomaw* [tooth-PROPR-ATT], intended meaning: 'kind of toothy'; **qʷӕšӕn-gomaw* [big.eared-ATT], intended meaning: 'kind of eared/having big ears'; **dӕndag-ӝӕn-gomaw* [eye-PROPR-ATT], intended meaning: 'kind of big-eyed')[2]. This fact seems to correlate with the ungrammaticality of the use of *-gomaw* on numerals described above.

[1] There is no clear morphological distinction between adjectives and adverbs in Ossetic.
[2] Russian stimuli: *глазастый, ушастый, зубастый* (respectively).

4. According to (Axvlediani 1963, p.140), the discussed marker can be used in a word form in combination with the derivational marker *-bən* that conveys the meaning 'the weaker quality' (Vydrin 2019). The recently collected field data contradicts the point made in (Vydrin 2019) that native speakers consider the use of both suffixes in a single word form ungrammatical; all of our consultants found such words as *bur-bən-gomaw* 'kind of yellowish' and *wɜž-bən-gomaw* 'kind of a bit lazy' grammatical. It is crucial to note that the reverse linear order of the markers is impossible: **bur-gomaw-bən, *wɜž-gomaw-bən*. The fact that *-bən* seems to always linearly precede *-gomaw* presents another argument for the claim that *-gomaw* is not really a derivational attenuative marker in its canonical understanding, as, according to (Haspelmath 2024), derivational markers tend to linearly gravitate towards the root. However, we acknowledge that this fixed order of affixes in this case can be due to some factors (for example, prosodic) that we might yet not be aware of.

Examples

(1) *wəj* *bənton* *žɜrond-**gomaw*** *nɜ* *wəd-i* *kʷərəχon* *qʷədə-t-ən*
that.DEM quite old-**ATT**[3] NEG be-PST.3SG wise thought-PL-DAT
'That woman is **kind of** not old enough to have wise thoughts'; the speakers consistently noted that they interpret *žɜrond-**gomaw*** not as 'younger than the prototypical old person', but as 'not really old in terms of age, but kind of resembling an old person'.

(2) *aj* *ɜgɜr* *gəc:əl-**gomaw*** *χəžən* *u*
this.DEM overly little-**ATT** bag be.PRS.3SG
'This bag is kind of too little'

(3) *nɜ* *raž-mɜ* *šəlgojmag* *təng* *darʁ-**gomaw*** *k'aba-jə*
1PL.POSS front-ALL woman very long-**ATT** dress-IN
'A woman in a dress that is kind of too long walked in front of us'(ONC)

(4) *amina* *alan-ɜj* *bɜržond-**gomaw**-dɜr* *u*
Amina Alan tall-**ATT**-COMPAR be.PRS.3SG
'Amina is kind of taller than Alan'

(5) *amina* *alan-ɜj* *gəc:əl* *bɜržond-dɜr* *u*
Amina Alan-ABL little tall-COMPAR be.PRS.3SG
'Amina is a little bit taller that Alan'

(6) *tad-**gomaw*** *šaldʒg* *χɜr-ən*
melt.PTCP-**ATT** ice.cream eat-PRS.1SG
'I eat ice-cream that is kind of melted'

(7) *p'ol-əl* *χɜrd-**gomaw*** *fɜtkʷə* *lɜw:-ə*
floor-SUPER eat.PTCP-**ATT** apple lay-PRS.3SG
'An apple that is kind of bitten at lies on the floor'

References. 1. Ossetic National Corpus (ONC). http://corpus.ossetic-studies.org **2. Abaev, Vassiliy I.** 1950. Grammaticheskij ocherk osetinskogo jazyka [Grammar sketch of Ossetic]. Ordzhonikidze: Severo-Osetinskoe knizhnoe izdatelstvo. **3. Axvlediani, Georgiy S.** (ed.) 1963. Grammatika osetinskogo jazyka [Ossetic grammar]. Vol. 1: Fonetika i morfologija [Phonetics and morphology]. Ordzhonikidze: NII Severo-Osetiskoj ASSR. **4. Bagaev, Nikolay K.** 1965. Sovremennyj osetinskij jazyk [Modern Ossetic]. Part 1: Fonetika i morfologija [Phonetics and morphology]. Ordzhonikidze: Severo-Osetinskoe knizhnoe izdatelstvo. **5. Grandi, Nicola & Körtvélyessy, Lavinia** (eds.) Edinburgh Handbook of Evaluative Morphology. Edinburgh: Edinburgh University Press. **6. Haspelmath, M.** Inflection and derivation as traditional comparative concepts. // Linguistics, №1, 2024. p. 43-77. **7. Vydrin, Arseniy P.** Ossetic color terms system // Raffaelli, Ida, Daniela Katunar and Barbara Kerovec (eds.). Lexicalization patterns in color naming. A cross-linguistic perspective [SFSL 78]. John Benjamins Publishing Company. 2019. P. 401-426.

[3] We gloss ***-gomaw*** as an attenuative marker, following the existing convention, although we believe (and we make it our point in this abstract) that it acts as more of a metalinguistic evaluation marker.

Grammatical gender assignment of loanwords in East Caucasian languages: Towards a typological overview

Jesse Wichers Schreur[1], Kate Bellamy[1] & George Moroz[2]
[1]Leiden University, [2]HSE University

Introduction: Multilingualism is a long-standing feature of life in the Caucasus, often between structurally complex languages that may or may not be genealogically related. One characteristic of most East Caucasian is the presence of grammatical gender (GG). In this talk we will bring together these two threads, to investigate how loanwords from donor languages with and without GG are integrated into languages with systems of varying levels of complexity.

Any language possessing a GG system must assign a gender class to a noun, borrowed or otherwise, in order for it to be appropriately integrated into its grammar, through patterns of *agreement* (Corbett, 1991). This assignment can proceed on the basis of semantic features of the referent (notably humanness or animacy), or formal features, namely, phonological and/or morphological properties of the noun itself. Data from some studies of Spanish-English speakers in parts of the US with more balanced bilingualism and dense code-switching patterns indicate that a default gender may be employed for all inserted nouns. In other situations, the gender of the translation equivalent in the recipient language determines the GG assigned (see Bellamy & Parafita Couto, 2022 for an overview). When both languages have GG, the gender from the donor language may potentially be applied to the noun in the recipient language, a process known as gender copy (e.g. Stolz, 2009).

Previous work: A number of studies have reported gender assignment strategies for loanwords in several East Caucasian languages. It is clear that in all these languages, borrowed nouns with human referents are assigned masculine or feminine, while borrowed non-human nouns show a high degree of variation. Beginning with the Nakh branch, Wichers Schreur (2021: 30) demonstrates that borrowed nouns in Tsova-Tush (five genders) "follow the same collection of semantic and phonological rules when it comes to assigning their gender. Competition between these types of rules results in a picture that is less than transparent." In contrast, "in Chechen (six genders), 90% of all nouns borrowed from Russian that have a non-human referent are assigned J gender" (Bellamy & Wichers Schreur, 2022: 276).

Turning to Daghestanian languages, Hinuq non-human borrowed nouns are distributed across genders III to V, sometimes on the basis of semantic analogy, sometimes through phonetic similarity, but for most borrowed nouns, there are no obvious explanations for the assignment observed (Forker, 2016: 95). Similarly, patterns are hard to identify for non-human loanwords from Russian into Archi (Lezgic; four genders) (Levkovych, 2024: 254). Similarly, in Bezhta (Tsezic; four genders) "[t]he rules (if there are any) for the distribution of Avar and Russian loans in GGs III and IV in Bezhta are as in the case of Archi rather opaque and are governed both by semantic and formal aspects" (Levkovych, 2024: 255). Andi speakers also demonstrate more variation in noun class assignment with loanwords than with native terms (Moroz & Verhees, 2019). It should be clear, therefore, that all three assignment strategies (semantic, formal, default gender) are observed in loanwords in East Caucasian and more data, both linguistic and sociolinguistic, is needed to try and better understand the opaque patterns observed.

Methodology: In order to investigate how East Caucasian languages assign GG to loanwords from various donor languages, we have identified a sample of recipient and donor languages, where the former must have a GG system with one or more non-human gender values (since languages with only three genders have transparent semantic assignment), and the latter may or may not possess a GG system, as follows:

Recipient language (affiliation)	No. of GG values		Donor language (affiliation)	No. of GG values
Tsova-Tush (Nakh)	5		Russian (Indo-European)	3
Chechen (Nakh)	6		Arabic (Semitic)	2
Ingush (Nakh)	6		Avar (Andic)	3
Chamalal (Andic)	5		Persian (Indo-European)	0
Andi (Andic)	5-6		Azerbaijani (Turkic)	0
Bezhta (Tsezic)	6		Kumyk (Turkic)	0
Hunzib (Tsezic)	5		Turkish (Turkic)	0
Hinuq (Tsezic)	5		Georgian (Kartvelian)	0
Tsez (Tsezic)	4			
Khwarshi (Tsezic)	5			

Our starting point for creating as exhaustive as possible a list of loanwords in these ten recipient languages is the DAG < APT database (Balahanov et al., 2024), which contains Arabic, Persian and Turkic loanwords into Daghestanian languages, and Khalilov (2004). We are currently adding the gender value to all of these loanwords, using existing dictionaries and grammatical descriptions, which also enable us to expand the lists where necessary. In so doing, we will be able to identify gender assignment patterns, per donor language, semantic field, and/or sociolinguistic situation.

Implications: This study will offer a finer-grained analysis of how loanwords are treated in East Caucasian languages with one or more non-human gender values, including the treatment of nouns from languages with no GG (cf. Levkovych, 2024). These findings thus have implications for our understanding of languages in contact, especially gender systems in contact, as well as the importance of the sociolinguistic situation, type of bilingualism or degree of code-switching on contact outcomes. Moreover, it helps us to reflect further on the notion of gender copy, and whether we can really say that the GG systems of the donor and the recipient language can ever be straightforwardly equated, or congruent.

References

Balahanov, T., T. Dedov & S. Verhees (2024). *DAG < APT, An online database of borrowed vocabulary in the languages of Daghestan, v. 1.0.0*. Moscow: Linguistic Convergence Laboratory, NRU HSE. https://lingconlab.github.io/DAG-APT/.

Bellamy, K. & M. C. Parafita Couto. (2022). Gender assignment in mixed noun phrases: State of the art. In: D. Ayoun (Ed.), *The Acquisition of Gender: Crosslinguistic perspectives*, Amsterdam/Philadelphia: John Benjamins, pp. 14-48.

Bellamy, K. & J. Wichers Schreur. (2022). When phonology and semantics collide: Gender assignment in mixed Tsova-Tush–Georgian nominal constructions. *International Journal of Bilingualism, 26*(3): 257-285.

Forker, D. (2016). Gender in Hinuq and other Nakh-Daghestanian languages. *International Journal of Language and Culture,* 3(1):90-114. https://doi.org.10.1075/ijolc.3.1.05for.

Khalilov, M. Sh. (2004) *Gruzinsko-dagestanskie jazykovye kontakty [Georgian-Daghestanian language contact].* Moscow: Nauka

Moroz, G., & Verhees, S. (2019). Variability in Noun Classes Assignment in Zilo Andi: Experimental Data. *Iran and the Caucasus*, *23*(3), 268-282. https://doi.org/10.1163/1573384X-20190306

Levkovych, N. (2024). Gender assignment in language contact. *STUF*, 77(2): 235-260.

Stolz, C. (2009). A different kind of gender problem: Maltese loan-word gender from a typological perspective. In: B. Comrie, R. Fabri, E. Hume, M. Mifsud, T. Stolz & M. Vanhove (eds.), *Introducing Maltese linguistics* (pp. 321-353). Amsterdam/Philadelphia: John Benjamins.

Wichers Schreur, J. (2021). Nominal borrowings in Tsova-Tush (Nakh–Daghestanian, Georgia) and their gender assignment. In L. Grenoble & D. Forker (Eds.), *Language contact in the territory of the former Soviet Union* (pp. 15–33). Amsterdam: John Benjamins.

Relative chronology of sound changes in Avar-Andi-Tsezic: short affricates and word-final nasals

Peter Schrijver
Universiteit Utrecht

In Indo-European historical linguistics, it is usual to determine a relative chronology of prehistoric sound changes on the basis of two considerations:

- If the output of sound change A forms the input of sound change B, assume that A chonologically preceded B
- If related languages underwent the same sound change under the same conditions, assume that the sound change occurred in the common ancestor of those languages, unless there is evidence to think otherwise.

At present, the relatively rudimentary state of knowledge of the prehistory of the East Caucasian languages rarely allows application of the former consideration, but there are ample instances to which the latter consideration could be applied. The presentation discusses two (complexes of) sound changes in Avar-Andi-Tsezic (AATs) that show that the latter consideration would lead to incorrect conclusions concerning the relative chronology.

The first sound change concerns the development of the Proto-AATs short voiceless affricates **q* and **ƛ*. In all Andic languages, with one exception, **q* and **ƛ* develop into the fricatives **χ* and **ɬ*, respectively. The exception is Northern Axwax and the Southern Axwax dialect of Ratlub, where the Proto-AATs affrcates are preserved. Had we not known about these Axwax dialects, the development into fricatives would have been attributed to the Proto-AATs stage.

In Tsezic, developments were rather complex: in all Tsezic languages the affricates became fricatives, but with many differences of detail:

PAATs **q*	Tsez	Hinuq	Xwarshi	Inxoqwar	Bezhta	Hunzib	correspondence agrees with other PAATs phonemes?
/#_	*χ-*	*χ-*	*h-*	*h-*	*χ-*	*χ-*	unique
(/#_**r*C	*ʁ-*	*ʁ-*	*ʁ-*	*ʁ-*	*h-*	*h-*	~ **χ*)
/V_	*-χ-?*	*-h-*	*-h-*	*-h-*	*-ʁ-*	*-h-*	unique
/C_	*-ħ-*	*-χ-*	*-χ-*	*-χ-*	*-χ-*	*-χ-*	~ **χ:, *q:*

PAATs **ƛ*	Tsez	Hinuq	Xwarshi	Inxoqwar	Bezhta	Hunzib	correspondence agrees with other PAATs phonemes?
/#_	*ɬ-*	*ɬ-*	*ɬ-*	*h-*	*ɬ-*	*ɬ-*	unique
/V_	*-ɬ-*	*-ɬ-*	*-ɬ-*	*-l-*	*-ɬ-*	*-l-*	unique
(/_ᴺ				*-n-*			)
/C_	*-ɬ-*	*-ɬ-*	*-ɬ-*	*-ɬ-*	*-ɬ-*	*-ɬ-*	~ **ɬ:, *ƛ:*

An analysis of the details reveals that at the Proto-Tsezic stage the affricates were still intact and that the development into fricatives occurred independently in all Tsezic languages (except perhaps Tsez-Hinuq, which agree with one another). This result is contrasted with the analyses by Gudava 1979: 120-121 and by Nikolayev-Starostin 1994: 58, 112.

The second sound change involves the retraction of a word-final nasal into the first syllable, where it appears as nasalization, which in turn is capable of affecting word-initial class indicators (**r- > *n-*; **b- > *m-*). This development is widespread in Andic languages according to an erratic geographic pattern (e.g. North Axwax is affected but South Axwax is not; incidental instances are

found in Lower Andi but not in Upper Andi; e.g. Gudava 1964: 74; Nikolayev-Starostin 1994: 42). The change also occurs in all Tsezic languages, which it affects in identical fashion (although analogical leveling has ousted the class indicators **n-* and **m-* from some languages). Hence one might think that the development belonged to Proto-Tsezic, if it were not for general doubts about the applicability of the second consideration in East Caucasian.

On the basis of a detailed description and analysis of the two sound changes under review, hypotheses will be formulated about their chronology, and suggestions will be made about the possible mechanisms that may explain why an identical sound change affected languages long after they had split from one another.

References:

Gudava, T.E. 1964: *Konsonantizm andijskix jazykov* (Tbilisi: Izdatel'stvo Akademii Nauk Gruzinskoj SSR)
Gudava, T.E. 1979: *Istoriko-sravnitel'nij analiz konsonantizma didojskix jazykov* (Tbilisi: Mecniereba)
Nikolayev, S., S. Starostin 1994: *A North Caucasian Etymological Dictionary* (Moscow: Asterisk; reprint Ann Arbor: Caravan Books 2007)

Shushurin Philip

Suppletive adjectives in Nakh languages: patterns of allomorphy and implications for theories of suppletion

Introduction In this paper, I want to present novel data from Ingush, which provide a remarkable pattern of number-conditioned allomorphy in adjectives. I also present data from Melkhi Chechen, for which no elaborate grammatical description has existed until now. I provide a novel generalization for suppletive adjectives and discuss several implications that this data present for our understanding of adjectival suppletion.

Core data **Ingush** The data presented here was collected with three speakers in Moscow, Russia and in Ingushetia. In Ingush, number (SG vs. PL) in adjectives is encoded either (a) on the class prefix (4a), or (b) in a suppletive way. One (perhaps the sole) example of full number-conditioned suppletion in Ingush adjectives is the pair *zwamiga~kegii* 'small'. The pair is usually taken to represent a strict number opposition: *zwamiga SG~kegii PL*. (Nichols 2011, Norris 2022). However, speakers report that in most plural cases, both forms can be used with no tangible difference in meaning. In contrast, in strict singular contexts only the singular form can be used.

(1) a. zwamiga/kegii cisk-až[1] «small cats»
b. zwamiga/*kegii cisk «small cat»

Furthermore, the distribution of the two suppletive forms is different in idiomatic expressions, most prominently with the noun *sag/nax* 'person/people'. The adjective *zwamiga~kegii* can form idiomatic relation with this noun yielding the meaning 'young person'/'young people'. However, in plural contexts, only the *kegii* form can be used (3). The pattern seems to be robust as all the consulted speakers report that the string *zwamiga nax* can only mean 'small people' and cannot mean 'young people'. In the singular, only the *zwamiga* form can be used for both idiomatic and non-idiomatic readings.

(3) a. kegii nax «young people», «small people»
b. zwamiga nax *«young people», «small people»
c. zwamiga/*kegii sag «small person», «young person»,

Next, a similar pattern is observed with the adjective CL-*oakkxa*/CL-*oakkxii* 'big'. Similarly to the adjectjve *zwamiga~kegii* 'small', the second form CL-*oakkxii* is reserved for strictly plural contexts, while the form CL-*oakkxa* can be used both in plural and singular contexts. Furthermore, similarly to *zwamiga~kegii* 'small', the CL-*oakkxii* form should be used in plural contexts on idiomatic readings, while both forms can be used in plural contexts on non-idiomatic readings. The class prefix tracks the number of the noun regardless of the idiomatic/non-idiomatic distinction (The plural prefix for 'bull' is *d-*, the plural prefix for *sag/nax* is *v-*).

(4) a. d-oakkxii ust-až «big bulls»
b. d-oakkxa ust-až «big bulls»

(5) a. v-oakkxii nax «old people», «big people»
b. v-oakkxa nax *«old people», big people»
c. v-oakkxa/*v-oakkxii sag «old person», big person»

While the pattern of distribution of *small* vs. *big* is remarkably similar, two key differences between the two adjectives can be highlighted. First, the latter adjective has a slot for a class/number prefix, unlike *zwamiga~kegii*. Next, while the paradigm of *small* is fully suppletive, the two forms of *big* share the same stem (*-oakkx(a)-*). The question can be raised of whether the segment *-ii* in CL-*oakkx**ii*** is a plural suffix (or part of the stem). While I remain agnostic on this question, I want to propose, based on a similar distribution

[1] *-až* is a plural morpheme is Ingush and *-aš* is a plural suffix in Melkhi Chechen. All other glosses are self-explanatory and are omitted for space reasons.

of SG vs. PL forms in idiomatic and non-idiomatic expressions in the two adjectives that in both cases we are dealing with the same phenomenon of **suppletion**. To that effect, the language seems to make no distinction between **full suppletion** (as in *zwamiga~kegii* 'small') and **partial suppletion** (as in CL-*oakkxa*/CL-*oakkxii* 'big').

Melkhi Chechen In this section I provide data from Melkhi Chechen – a underdescribed and highly endangered Nakh dialect spoken in several villages in Western Chechnya and North-Eastern Ingushetia. The data in this section was obtained during a fieldwork trip to Arshty, Ingushetia (August 2024). The dialect has been found to contain a pattern of allomorphy which mimics in very close detail the one found in Ingush. More specifically, the following generalizations hold for Melkhi Chechen.

(6) The adjective *žwamig* has a plural form *kegii* which can be used interchangeably with *žwamig* in plural non-idiomatic contexts

a. žwamig/kegii cisk-aš «small cats»
b. žwamig/*kegii cisk-aš «small cat»

(7) In idiomatic contexts, only the *kegii* form can be used with the plural

a. žwamig nax *«young people», «small people»,
b. kegii nax «young people», «small people»

(8) The adjective CL-*oakx* has a plural form CL-*oakxii* which can be used interchangeably with -*oakx* in plural non-idiomatic contexts. The plural prefix must occur in all plural contexts.

a. d-oakx/*d-oakxii cisk «big cat»
b. d-oakx cisk-aš «big cats»
c. d-oakxii cisk-aš «big cats»

(9) In plural idiomatic contexts, only -*oakxii* form can be used

a. v-oakx sag – «old person»
b. b-oakxii nax – «old people», big people»
c. b-oakx nax – *«old people», big people»

<u>Generalization (for both Ingush and Melkhi Chechen)</u> For suppletive adjectives, specialized plural form is obligatory in idiomatic plural contexts and optional in non-idiomatic plural contexts.

Discussion The Nakh data can provide novel evidence for the typological and theoretical studies of number allomorphy in adjectives. First, it confirms the generalization that number-conditioned suppletion in adjectives is most commonly found in size adjectives (Vafaeian 2010).

Second, the Nakh data shows that number-conditioned suppletion in adjectives can be sensitive to idiomatic/non-idiomatic distinction. I suggest that this difference can be analyzed in terms of the relative position of the adjective in the nominal structure. More specifically, I suggest that allomorphy in Ingush and Melkhi Chechen can be sensitive to syntactic locality. Assuming that idiomatic adjectives are lower in structure than non-idiomatic adjectives, I suggest that suppletion triggered by plurality in Nakh is obligatory in lower parts of the nominal structure and optional for higher parts of the nominal structure.

Nichols, J. (2011). *Ingush grammar* (Vol. 143). Univ of California Press.
Norris, M. (2022). Nominal inflection in Distributed Morphology. *Ms., lingbuzz/006516*.
Vafaeian, G. (2010). Breaking paradigms: a typological study of nominal and adjectival suppletion.

Aspectual Reanalysis of Copular Forms in Tat

Murad Suleymanov
(EPHE-PSL / ILARA, Inalco, PROCLAC)

Although Tat, like many Iranian languages, is a language with a two-stem verb system, the variation is not aspectual, as it is in most neighbouring East Caucasian languages. In the present tense, however, Tat makes a clear distinction between the stative identification copula, the dynamic identification copula and the existential copula. The stative copula is expressed by a set of enclitics. The dynamic copula is expressed by a verb of the stem *bir-* 'be, become'. The existential is expressed by the conjugated copular form *häst*. In the negative, there is convergence between the stative and existential copulae.

Abşeron Tat (Balaxanı, Suraxanı)

(1) a. *doxtur=ü.*
doctor=COP:3
'(One) is a doctor.'

b. *doxtur nist(=ü).*
doctor NEG.EXIST(=COP:3)
'(One) is not a doctor.'

(2) a. *doxtur bir-än.*
doctor be$_2$-PRS:3
'(One) becomes a doctor.'

b. *doxtur nä-bir-än.*
doctor NEG-be$_2$-PRS:3
'(One) does not become a doctor.'

(3) a. *doxtur häst(=ü).*
doctor EXIST(=COP:3)
'There is a doctor.'

b. *doxtur nist(=ü).*
doctor NEG.EXIST(=COP:3)
'There is no doctor.'

With the exception of 'be', all the aforementioned forms are defective. It is 'be' that takes over the missing parts of all copular paradigms. Most Tat varieties thus show complete convergence of the copulae in the past (Suleymanov 2020: 146).

Abşeron Tat (Suraxanı)

(4) a. *doxtur bü.*
doctor be$_2$:PST:3
'(One) was a doctor.'
'(One) became a doctor.'
'There was a doctor.'

b. *doxtur nä-bü.*
doctor NEG-be$_2$:PST:3
'(One) was not a doctor.'
'(One) did not become a doctor.'
'There was no doctor.'

So far, the only Tat variety known to have developed a distinction between the static and dynamic uses of the copula is Judaeo-Tat (Authier 2012: 137), which has the verb stem *bistor-* 'become', presumably derived out of the periphrastic construction *birɛ* 'be.PTCP' + **istor-* 'stand' (Authier, p.c.). The distinction is likewise characteristic of the past tense, where both 'to be' and 'to become' possess full conjugations.

Judaeo-Tat (literary, Authier 2012: 117, 97, adapted)

(5) *mɛ=š bɛbɛ=šmu=rɛ xuno ʕošir-ɛ odomi bir-üm.*
I=OBL father=POSS:2PL=OBL as rich-ATTR person be$_2$:PST-1
'I was also a rich man, like your father.'

(6) *omor-ɛ-omor-ɛ mɛ e hɛči=rɛvoz bistor-um selkor...*
come-PTCP-come-PTCP I LOC so=INSTR become$_2$:PST-1 village_reporter
'And this is how, little by little, I became a village reporter.'

Fieldwork carried out recently on the variety spoken in Balaxanı (Abşeron Tat) has uncovered a so-far unattested distinction between the static (including identification and existence) and dynamic uses of the copula, which makes it the only known Muslim variety to do so.

Abşeron Tat (Balaxanı)

(7) *ävväl* *četin* *birü,* *hözüm* *hasand=ü.*
first difficult COP:STAT:3 now easy=COP:3
'It was difficult before (but) now it's easy.'

(8) *ənǰä* *azärbayǰanli-yä* *birü.*
there Azerbaijani-PL COP:STAT:3
'There were some Azerbaijanis there.'

(9) *či* *bü?*
what COP:DYN:3
'What happened?' (lit. 'What became?')

In the first (and second) person, the distinction manifests itself on a prosodic level:

Abşeron Tat (Balaxanı)

(10) *ä* *pensiya* *nådürmar-ä* *doxtur* *biˈrüm.*
from retirement NEG.exit$_2$-PTCP doctor COP:STAT:1
'Before I retired, I was a doctor.'

(11) *häftadihäf-imǰi* *sal* *diplom=mün=ä* *våstär-üm,*
seventy+seven-ORD year diploma=POSS:1.BS=OBL get$_2$:PST-1
doxtur *ˈbir-üm.*
doctor COP:DYN-1
'In 1977, I received my diploma and became a doctor.'

Balaxanı Abşeron Tat has developed a system similar to Judaeo-Tat in its function but with notable differences in form.

The vowel-final third-person past form *birü* '(s/he) was' looks structurally very similar to that of verbs with °*ir*-final stems, e.g. *käši* '(s/he) pulled' (cf. stem *käšir-* 'pull').[1] Furthermore, the unusual word-final stress in the first-person form *biˈrüm* '(I) was' (inherited past tense forms, such as *ˈbir-üm* in (11), are never stressed on the agreement suffix) may be indicative of sound fusion. Such realisations are evidence to believe that *biˈrüm* and *birü* go back to forms such as **birirüm* and **biri* whose hypothetical stem can be reconstructed to **birir-* 'be', a static counterpart of *bir-* with a full past tense conjugation.

To cite some parallels, Tat varieties are known to add an °*ir* extension to verbs with monosyllabic *r*-final stems, either coexisting with the old form, e.g. Ərüsküş–Dağ Quşçu Tat *bir-üm* / *birir-üm* '(I) was, (I) became', *dir-üm* / *dirir-üm* '(I) saw', or replacing it completely, e.g. Şirvan Tat *čirir-um* '(I) shaved' (cf. more archaic Qonaqkənd Tat form *čir-üm*).

Balaxanı Abşeron Tat interprets the inherited past paradigm of *bir-* as a dynamic copula and invents a new one for the seemingly more widely encountered stative semantics. This is a striking difference from Judaeo-Tat, which does the opposite. The stative vs. dynamic copula distinction may potentially have extended to non-past TAM categories but no such forms have been found so far. If they indeed do not exist, this may point to a paradigmatic parallel with Azeri, a common contact language for Tat, where copular forms show stative vs. dynamic distinction only in the present and past tenses.

ATTR = attributive, BS = bound stem, COP = copula, DYN = dynamic, EXIST = existential, INSTR = instrumental, NEG = negative, OBL = oblique, POSS = possessive, PRS = present, PST = preterite, PTCP = participle, STAT = stative

Bibliography

Authier, Gilles (2012). *Grammaire juhuri, ou judéo-tat, langue iranienne des Juifs du Caucase de l'est*. Beiträge zur Iranistik 36 / Bibliothèque iranienne 76. Wiesbaden: Reichert.

[1] Final *ü* in *birü* may be due either to assimilation into the initial labial (the form is often realised as *bürü*) or to analogy with the final vowel of the more archaic *bü*.

Suleymanov, Murad (2020). *A grammar of Şirvan Tat*. Beiträge zur Iranistik 46. Wiesbaden: Reichert.

Elena Sokur, independent researcher
Nina Sumbatova, Institute of Linguistics, RAS

Lexical database of the Dargwa languages

The Dargwa language group (Nakh-Dagestanian family) consists of about 15 separate languages and several dozens of dialects. The situation with their documentation is very different. Speaking particularly on the documentation of the lexicon, we can state that in this respect, the Dargwa lects fall into several groups: (a) for Standard Dargwa and Kubachi, we have big, recently published dictionaries (Abdullaev 2017; Jusupov 2017; Magomedov, Saidov-Akkutta 2017); (b) for several languages, lexical information has been collected and published online (Sanzhi: https://dictionaria.clld.org/contributions/sanzhi#tabout; Shiri: https://www.webonary.org/shiri/overview/; Mehweb: https://lingconlab.ru/MehwebDict/) or as part of grammatical descriptions and text collections (Aqusha: van den Berg 2001, Itsari: Sumbatova, Mutalov 2003, Tanti: Sumbatova, Lander 2014); some languages like Kunki, Mekegi or Kajtag have lexicons collected and published by non-professional authors (Alzhanbekov, Rabadanov 2021; Gasanova 2011; https://www.mekegi.com/dictionary/mekegi/); some more lexicons have been collected but remain unpublished (Tsugni, Khuduts, Muira, Kadar, etc.). Many lects are presented in the recent dialectological dictionary (Temirbulatova 2022); this book contains many words in more than 40 lects, but there is no information on the words except their initial form. Finally, there is a long list of lects whose lexicon we only know by 100-word lists or do not know at all.

At the same time, many of the Dargwa lects are highly endangered so that their documentation is urgent; for many dialects, any data is precious – even the data that is unpublished, not full, or not thoroughly checked. The lexical database that we present in this talk aims at collecting and systematizing as much information on the lexicon of the Dargwa lects as possible and, further on, on comparative studies of phonology, derivational morphology and lexicon across the Dargwa group. The first version of the database was released at the end of 2023; the second version is being prepared; it is going to be published by the end of 2024.

BASIC INFORMATION AND DATA. The database contains lexical units of different languages and dialects of the Dargwa group. It is a principally open resource that is meant to be supplemented and corrected in course of its usage. We are also ready to implement non-full data if certain types of information are absent.

The database contains different types of data, both the data taken from dictionaries and other published works and unpublished field data (unpublished field data are always used with the permission of their owners and with appropriate references). The source of the data is explicitly provided for each entry.

By now, the database contains the lexicons of the following lects: Aqusha, Itsari, Tanti, Muira, Kadar, Mehweb. The size of these lexicons lies between 1200 and 2500 entries. The lexicons of Kunki, Urakhi and Kubachi are being prepared.

ENTRY STRUCTURE. The entry structure is principally the same for different lects, but differs for nouns, verbs and other parts of speech. The words in Dargwa are presented in two different graphics: (1) the Cyrillic script traditionally used for Standard Dargwa, (2) phonological transcriptions. In the Cyrillic orthography, we use several additional rules to represent the phonemes that are absent in Standard Dargwa, but present in other Dargwa lects (geminated consonants, pharyngealized vowels, etc.).

The lemmas are the absolutive singular forms of the nouns and other declinable words; two aspectual forms of the infinitive of the verbs. The entries contain diagnostic forms: for the nouns, these are the absolutive plural and the ergative or dative singular; for the verbs, this is the aorist. If a lexicon contains more forms (e.g. genitive, locative, present converb, imperative, prohibitive), these forms can also be included in the entry.

The grammatical part of the entry may also contain information on the syntactic class of the verb (transitive/intransitive/affective), the nominal gender, and paradigmatic irregularities.

The definition is given in Russian and English. Unfortunately, for most entries, because of the lack of data, the description of the lexical semantics is not detailed.

Finally, the entry may contain a record of the pronunciation and, in some cases, also a photo illustrating the corresponding object.

SEARCH OPTIONS. The database provides several search options. The most usual variant is looking for an entry by the lemma. In this case, you can use either its transcription or orthographic form in Dargwa. Another common variant is looking for a word by its translation into Russian or English. We can choose one or several lect(s) where the lemma is looked for.

There are also additional search options:

- search for synonyms – the database looks for all synonyms of a given entry
- search for cognates – for a previously chosen entry, the database yields cognate words in other lects
- search for words with certain morphemes – the system finds words with a certain root or derivational affix.

The database contains interior information on the morphological structure of words and on cognate roots, which is used to realize these search options. Of course, this information is not full; in many cases, it will require corrections and additions. The search for synonyms is based on translations.

REALIZATION. The website of the project is developed using Python and the Django framework. Data is stored in a database MySQL. For convenient data moderation, the Django administrative panel has been modified, which allows us to correct specific information for individual words, and the function of uploading and downloading Excel files, where the lexicons are originally stored, has been implemented. You can also completely update the data by deleting one or more lects. The database data is stored as relational tables. Information that is repeated for a group of words (for example, lect, part of speech, morphology, source) is stored in separate tables, which speeds up searching in the database. The website contains a feedback form that can be used to contact the project administrators.

Several automated processes are implemented in the backend. When loading a dictionary into the database, the code automatically creates links between the perfective and imperfective forms of one verb. Also, the Dargwa verb often contains a gender marker. In order to be able to search by any gender form of the verb, we have created a special program that, when processing the dictionary, searches for a gender marker in the given verb lexeme and creates a special field in the database, where the verb forms with all possible gender markers, written in the Cyrillic script and phonological transcription, are saved (however, this method can yield non-existent verbforms and requires further testing).

References

Alzhanbekov, M. G., Rabadanov, K.O. 2021. *Russko-darginskij slovar'. Kunkinskij govor.* Makhachkala.

Abdullaev, Z. G. 2017. *Darginsko-russkij slovar': 12 000 slov, terminov i vyrazhenij*. Makhachkala: In-t JaLI DNC RAN.

Gasanova, Uzlipat. 2011. *Khajadaksko-russkij slovar'.* Makhachkala.

Jusupov, Kh. A. 2017. *Dargan-urus slovar'. Darginsko-russkij slovar': okolo 40 000 slov I frazeologicheskikh vyrazhenij*. Makhachkala: In-t JaLI DNC RAN.

Magomedov, A. Dzh., N. I. Saidov-Akkutta. 2017. *Kubachinsko-russkij slovar'.* Moscow: Nauka.

Sumbatova, N. R., Yu. A. Lander. 2014. *Darginskij govor seleija Tanty. Grammaticheskij ocherk. Voprosy sintaksisa*. Moscow: Jazyki russkoj kul'tury.

Sumbatova, Nina & Rasul Mutalov. 2003. *A Grammar of Icari Dargwa*. München, LINCOM.

Temirbulatova, S. M. 2022. *Dialektologicheskij slovar' darginskogo jazyka.* Makhachkala: Alef.

van den Berg, Helma. 2001. *Dargi folktales. Oral stories from the Caucasus.* With an introduction to Dargi grammar. Leiden: Universiteit Leiden, 2001.

Nina Sumbatova, Institute of Linguistics, RAS
Naida Vagizieva, Institute of Language, Literature and Art, Daghestan Federal Research Centre, RAS

Plural formation in Dargwa: a cross-linguistic study

In the languages of the Dargwa group (Nakh-Dagestanian), the plural forms of nouns are marked by suffixes; the plural formation can be accompanied by vowel omission and/or alternation. The rules of plural formation are relatively complex in several aspects.

First, the sets of plural markers in the Dargwa lects contain 10–25 suffixes. For example, in Muira Dargwa, the suffixes *-ti, -ni, -i, -ri, -bi, -mi, -uni, -urbi, -umi, -ani* are relatively frequent; besides, there are several words with the markers *-up꞉i, -urmi, -lumi, -p꞉i, -ubi, -rbi, -kuri, -ruk꞉i, -rti, -urti* and *-li*. As seen from this list, there are one-syllable (i.e. primary) suffixes (*-ti, -ni, -i, -ri, -bi, -mi*) and two-syllable suffixes, most of which originate from a combination of an unknown element with a primary suffix (*-u-ni, -ur-bi, -u-mi, -a-ni, etc.*).

Second, in most nouns ending in a vowel, this vowel is being omitted in the plural forms, as in *t'um**a*** 'owl' – PL *t'um-ri*, *k꞉urt꞉**i*** 'dress' – PL *k꞉urt꞉-urmi*, *jab**u*** 'horse' – PL *jab-ni*[1]. In some nouns (in most cases, these are derived words containing three or more syllables), we observe the syncope of the stem vowel followed by a sonorant (cf. *bart꞉ih**a**n* 'rope' – PL *bart꞉ihn-i*) or even of two vowels (*k'ak'**a**ri* 'mite' – PL k'ak'r-umi). Generally, vowel omission cannot be predicted, cf. the nouns where the final vowel is retained in the plural: *muq'ula* 'hoarfrost' – PL *muq'ula-bi*, *nazmu* 'poem' – PL *nazmu-rti*, *č꞉ibʔaˁ* 'chicken' – PL *č꞉ibʔaˁ-ni*.

Third, in many nouns, plural formation is accompanied by vowel alternation. The root vowel is usually replaced by a narrow vowel – *u* or, less frequently, *i*. The alternations are not predictable from the form or meaning of the noun, cf. the nouns with alternation in (1) vs. the nouns without alternation in (2):

(1) *daʔ* 'wind' – PL *duʔ-ri*, *barħi* 'day' – PL *burħ-ni*, *sinka* 'bear' – PL *sunk-bi*

(2) *ʁamʁ* 'plum' – PL *ʁamʁ-i*, *wawa* 'flower' – PL *waw-ni*, *t'irxa* 'stick' – PL *t'irx-ni*

Suffixation, vowel omission and alternation are ultimately lexical: neither of these processes seem to consistently follow certain regular rules. Of course, some tendencies in the distribution of the suffixes and the presence/absence of vowel omission and alternation can be noticed. For some languages, these tendencies have been found and described; cf., for example, [Abdullaev 1954: 94–102] for Standard Dargwa, [Forker 2020: 44–50] for Sanzhi, [Chechuro 2019: 41–55] for Mehweb, [Temirbulatova 2004: 72–86] for Kajtag, etc.

In this paper, we present a cross-linguistic study of plural formation in Dargwa. First, we are trying to check whether the tendencies in the distribution of suffixes are common or at least similar for different lects of the Dargwa group. Second, we check if the distribution of the suffixes correlates with vowel omission and alternation. The final goal is the reconstruction of the plural formation in the Dargwa nouns.

The first step of our research is a pilot study of six lects of Dargwa, i.e. Kadar, Muira, Aqusha and Mehweb (North Dargwa), Tanti and Itsari (south Dargwa). For each lect, we systematized full data on plural formation based on a lexicon of more than 500 nouns (except for Aqusha, where the list is smaller). The lexical data are taken from [Vagizieva Ms.] (Kadar), [Sumbatova et al. Ms.] (Muira), [Mutalov 2020] and [Sumbatova, Mutalov 2003] (Itsari), [Sumbatova, Lander 2014] (Tanti), [van den Berg 2001] (Aqusha), [Musaev, Morozova, Daniel 2020] (Mehweb).

[1] All examples in the text are from Muira Dargwa (as spoken in the village of Kalkni, field data).

The lists of plural markers in these three lects are different, but the primary markers are obviously cognate. Table 1 shows the most common primary markers (each column represents cognate suffixes; the number in the brackets is the frequency rank of the marker in our sample).

Table 1. Primary plural markers

Kadar	*-t* (1)	*-me* (2)	*-ne* (3)	*-be* (6)	*-re* (7)	*-e* (9)
Muira	*-ti* (3)	*-mi* (6)	*-ni* (1)	*-bi* (4)	*-ri* (5)	*-i* (2)
Itsari	*-ti* (3)	*-mi* (2)	*-ni* (4), *-in* (7)	*-bi* (6)	*-ri* (5)	*-i* (1)
Tanti	*-te* (5)	*-me* (4)	*-ne* (1)	*-be* (3)	*-re* (6)	*-e* (2)
Aqusha	*-ti* (1)	*-mi* (7)	*-ni* (4)	*-bi* (3)	*-ri* (6)	*-i* (9)
Mehweb	*-t* (2)	*-me (5)*	*-ne* (1)	*-be* (6)	*-re* (9)	*-e* (7)

Our data showed that some of these suffixes obviously tend to be used in certain types of nouns and that these tendencies are the same for cognate markers across our sample, as shown in Table 2 (the numbers in the six last columns show what is the proportion of nouns of the specified type among all nouns with a certain suffix).

Table 2. Plural markers: the tendencies in distribution

Suffix		Kadar	Muira	Aqusha	Mehweb	Itsari	Tanti
-t, -ti, -te	nouns ending in a sonorant	76%	84%	84%	70%	94%	82%
-ni, -ne, -in	nouns ending in a vowel (vowel omitted in the plural)	88%	93%	100%	93%	97%	92%
-e, -i	nouns ending in a(ny) consonant	48%	85%	86%	53%	74%	80%

The suffixes *-mi* (*-me*), *-bi* (*-be*), *-ri* (*-re*), which are not shown in Table 2, show similar distribution: all of them (with different frequencies) are used either in two-syllable nouns ending in a vowel or in one-syllable consonant-final nouns.

On the other hand, some types of nominal stems prefer a certain plural formation type. The biggest group of nouns are two-syllable words ending in a vowel, which is omitted in the plural. In all lects, these nouns show a very clear tendency to attach the suffix *-ni* (*-ne*, *-in*). Two-syllables words ending in a sonorant attach the marker *-ti* (*-t*) in all lects. The nouns ending in a vowel, which is retained in the plural, are also consistent: if they have one syllable, they end in *-mi* (*-me*); two-syllable nouns end in *-bi* (*-be*). The nouns ending in an obstruent are less consistent: they attach different suffixes, mainly *-ri* (*-re*), *-bi* (*-be*), *-i* (*-e*); in Kadar, Aqusha and Mehweb also *-ani*/*-ane*.

The obvious tendencies in the distribution of the suffixes make us suggest that Proto-Dargwa had a single plural morpheme with several variants distributed complementarily. Their choice was conditioned by the final phoneme(s) of the nominal stem. In this talk, we are planning to present broader data and a more detailed historical explanation of the plural formation across Dargwa.

References

Abdullaev, S. 1954. *Grammatika darginskogo jazyka (fonetika I morfologija)*. Makhachkala: Dagestanskij filial AN SSSR.

van den Berg, Helma. 2001. *Dargi folktales. Oral stories from the Caucasus.* With an introduction to Dargi grammar. Leiden: Universiteit Leiden, 2001.

Forker, Diana. 2020. *A grammar of Sanzhi Dargwa* (Languages of the Caucasus 2). Berlin: Language Science Press.

Chechuro, Ilya. 2019. Nominal morphology of Mehweb. In Michael Daniel, Nina Dobrushina & Dmitry Ganenkov (eds.), *The Mehweb language: Essays on phonology, morphology and syntax*, 39–72. Berlin: Language Science Press.

Temirbulatova, S. M. 2004. *Khajdakskij dialekt darginskogo jazyka*. Makhachkala: DNTs RAN.
Mutalov, Rasul. 2020. *Itsari Jena wordlist*. Manuscript.
Musaev A., V. Morozova, & M. Daniel. 2020. *Mehweb Jena wordlist*. Moscow: Linguistic Convergence Laboratory, HSE University. https://linghub.ru/mehwebdict/
Sumbatova, N. R., Yu. A. Lander. 2014. *Darginskij govor seleija Tanty. Grammaticheskij ocherk. Voprosy sintaksisa*. Moscow: Jazyki russkoj kul'tury.
Sumbatova, Nina & Rasul Mutalov. 2003. *A Grammar of Icari Dargwa*. München, LINCOM.
Vagizieva, Naida. *Kadar: a wordlist*. Manuscript.
Sumbatova, Nina, Ivan Kalyakin & Diana Pchelintseva. *Muira (as spoken in Kalkni): a wordlist*. Manuscript.

Tsikh Darya

Postpredicative Constituents in Circassian Languages: Corpus Study

This paper investigates the arrangement of constituents in postpredicative positions within Circassian languages. The language family is characterized by sequential polysynthetism, ergative case marking and argument indexing, and left-branching word order [Аркадьев 2009]. While the basic order is SOV, pragmatic functions can influence constituent positioning.

The spoken Kuban corpus of the Kabardian language, the written corpus of the West-Circassian children's fairy tales and songs, and the fiction sub-corpus of the West-Circassian Corpus [Архангельский et al. 2018-2023] were used for the analysis. The material for the study was 72 clauses with postpredicate positions. From these examples were excluded clauses where the constituents in focus caused the appearance of postpredicate position.

Focus at the end of the sentence

Clause-final focus can occur, though it is exceedingly rare. In my research, I encountered only two such instances, which are presented below.

The first example is from a song, so the postpredicate position can be the consequence of a rhyme or rhythm. The second example is quite unusual for West-Circassian languages beyond the postpredicate position.

Reparative

The postpredicate position of some constituents can be justified by reparative, correcting problematic parts of discourse or restating what was previously said [Schegloff, Jefferson, Sacks, 1977: 363]. In our corpus, such constituents are more often grammatically labelled as parts of a clause, although there are a few examples in which the speaker repeats the form of the word again without any modification.

(1) sorek.pjatəj-godə-m ġe-ḳʷe-ž’-rjə jəṭane zə-pŝeŝe-çəḳ ġe-χʷ-a š’əpχʷ

45-year-OBL DIR-go-RE-ADD then one-girl-smaller DIR-born-PST sister

‘In 1945 he came back and then a girl was born, a sister.’

Antitopics

Antitopics challenge the traditional topic/focus dichotomy by introducing a new dimension of listener familiarity with information [Lambrecht 1981; Chafe 1982]. Their primary function is to redirect attention to previously mentioned information crucial for understanding the preceding clause.

Studies by Gordon [2008] and Andersen [2017] on Omaha-Ponca (Siouan family) and Berta (Nilo-Saharan) respectively, highlight key properties of antitopics: information givenness and intraclausality. In West-Circassian languages, the latter criterion is particularly well-defined—the antitopic is unambiguously included within the clause, often marked on the predicate and appearing in the appropriate case form, as exemplified in (2).

(2) s-j-ade fiskul'ture envepe ja-r-jə-ʁe-h-u školə-m

1SG.PR-POSS-father PE NVP 3PL.IO-DAT-3SG.ERG-CAUS-carry-ADV school- OBL

‘My father used to teach PE and NVP lessons at school’

Another important paper for my research is [Dahlstrom 2020]. Dahlstrom examines Meskwaki, a North American polysynthetic language typologically similar to Circassian. She employs the term "episodic structure," originally introduced by Chafe, to describe a construction where the same information both begins and ends a passage of speech. Dahlstrom argues that such a structure would feature a topic at the beginning

and an antitopic at the end. I believe similar constructions exist in Circassian languages as well, as shown in example (3).

(3) qʷaž'e-r q̇ə-ŝə-ṭəs-a-m-dje hable-r, hable-habl-u-re,

aul-ABS DIR-TEMP-seat-PST-OBL–y district-ABS district–district-ADV-CNV

ze-tje-xʷ-a-we habl-jə-txʷ ṭəs-a q̇ʷaž'e-r

REC.IO-LOC-fall-RES-ADV district-LNK-five seat-PST aul-ABS

'When the aul settled down, it divided itself into neighbourhoods, into five districts.'

Other tendencies

Additionally, out of the 72 sentences in the sample, 9 contained personal pronouns in the postpredicate position, while 16 featured indicative pronouns. Notably frequent was a pronoun series with the root -a-, which, according to [Kumakhov 1964], "indicates a visible or nearby object, or an invisible but known, definite object."

Conclusion

There are two major types of postpredicate constituents in Circassian languages: reparative and antitopic. The key distinction lies in the information they present. Antitopics reintroduce previously mentioned information that might have slipped from the listener's mind. Reparatives, on the other hand, either repeat information from earlier in the clause or introduce new (non-focal) information. Both serve to clarify or enhance the preceding discourse.

References

1. Аркадьев П. М. и др. Введение. Основные сведения об адыгейском языке //Аспекты полисинтетизма: очерки по грамматике адыгейского языка. – 2009. – С. 17-120.
2. Кумахов М. А. Морфология адыгских языков. Синхронно-диахронная характеристика. I. Введение. Структура слова. Словообразование частей речи. – 1964.
3. Рогава Г. В., Керашева З. И. Грамматика адыгейского языка. Краснодар; Майкоп: Краснодарское книжное изд-во, 1966.
4. Чейф У. Данное, контрастивность, определенность, подлежащее, топики и точка зрения //Новое в зарубежной лингвистике. – 1982. – Т. 11.
5. Andersen T. Clause-internal antitopics in Berta, a Nilo-Saharan verb-second language //Studies in Language. – 2017. – Т. 41. – №. 1. – С. 99-141.
6. Couper-Kuhlen E., Ono T. 'Incrementing'in conversation. A comparison of practices in English, German and Japanese //Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA). – 2007. – Т. 17. – №. 4. – С. 513-552.
7. Dahlstrom, A. (to appear) Antitopics and episode structure in Meskwaki. To appear in a special issue of Text and Talk in memory of Wallace Chafe, edited by Jan-Ola Östman and Michael Bamberg.
8. Gordon B. J. Antitopic and afterthought //Kansas Working Papers in Linguistics. – 2008. – Т. 30. – С. 74-87.
9. Kumakhov M., Vamling K. Circassian clause structure //Caucasus studies. – 2009. – Т. 1. – С. 1-219.
10. Lambrecht K. Topic, antitopic and verb agreement in non-standard French //Topic, Antitopic and Verb Agreement in Non-Standard French. – 1981. – С. 1-121.
11. Matasović R. A short grammar of East Circassian (Kabardian) //Manuscript. Zagreb. – 2010.

Differential object marking in modern Iron Ossetic: the role of a proto-role

Olga V. Tuzhik, Natalia V. Serdobolskaya
Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia
vikolya2002@gmail.com

Introduction. Our study is devoted to differential object marking (DOM) in Iron Ossetic. In this language, direct objects can either take a marker of the genitive case or occur in the nominative, see Abaev (1959: 153). As shown in the literature, the main factors underlying the choice of a DO marker are animacy and referential properties of DOs, e.g. see Thordarson (2009: 135). Serdobolskaya & Tuzhik (in press) show that, with human animates, genitive marking is obligatory for the referential types presupposing existence of the referent in hearer's mind (from the point of view of a speaker), either concrete (definite) or abstract (e.g. attributive). For other referential types, the nominative is possible on a par with the genitive. By contrast, marking of non-human animate DOs is described in terms of a scale of relative frequency of genitive and nominative (ibid.). Abaev (1959: 153) also states that the nominative is used with non-animate DOs. On the basis of a corpus analysis, Şahingöz (2022) shows that even specific non-animates tend to be non-marked in the modern language. However, there are some contexts which do not fit in this scheme. For instance, in *Mɜj razmɜ sə* ***qug*** *balxɜdta, wəj wɜj kɜnə* "He is selling a **cow** that he bought just a month ago", the direct object "cow" is in the nominative despite the fact that this cow is definite. The question arises to what extent the observed pattern is a lexical property of specific verbs or whether it is a property of DO.

Method. The data has been collected during fieldwork sessions in the Republic of Ossetia-Alania in 2011–2012 and in 2023–2024 by the method of elicitation from 16 speakers (all women, aged 17 to 53, $Mean_{age}$ = 30.68, *SD* = 12.32). An example from the questionnaire is shown below, see (1)-(4).

(1) *jɜ*	*bon*	*u*	*ɜpːɜt-ə*	*sərd-dɜr*	***fəš-ə***	*/*	****fəš***	*ba-jjaf-ən.*
his	force	be.PRS.3SG	all-GEN	fast-CMPR	sheep-GEN		sheep	PV-catch.up-INF

'(I commit my little son to shepherd our sheep, because he can run fast), he can catch up the fastest **sheep**.'

(2) *mɜn*	*qɜu-ə*	*dɜu-ɜj*	*ɜpːɜt-ə*	*nard*	***fəš***	***/ *fəš-ə***	*ba-lxɜn-ən.*
I.GEN	need-PRS.3SG	you-ABL	all-GEN	fat	sheep	sheep-GEN	PV-buy-INF

'(You have a lot of good sheep, but) I want to buy the fattest **sheep** from you.'

Examples (1) and (2) illustrate the difference in agentivity: the same lexeme *fəš* 'sheep' with the same referential properties (namely, definite non-specific) is marked differently. In (1), the DO *fəš* 'sheep' moves in order to escape from its owner, while the situation of selling a sheep in (2) does not presuppose any volitional involvement or movement of the object selled. The factor of agentivity is not new for the discussion of DOM. For example, Primus (2011: 75) illustrates the following hypothesis on the basis of various languages (Spanish, Malayalam etc.): "a differential object marker is licensed by an object whose intrinsic meaning properties qualifies it as a proto-agent in the situation denoted by the predicate".

We propose analysis of agentivity as a proto-role in terms of Dowty (1991: 572). He differentiates between two semantic roles, a proto-agent and a proto-patient. A proto-agent is characterized as a participant that satisfies any of the following: it can move; it exhibits volitional involvement in the event or state; it is sentient. By contrast, a proto-patient can undergo a change of state; it can be stationary relative to movement of another participant; it does not exist independently of the event, or not at all. We claim that, in Ossetic, DOs can be marked differently depending on the proto-role, i.e. on degree of agentivity. Hence, in the case where agentivity of a DO is implied, it is marked with a genitive as in (1). In (2), the object does not have any proto-agent properties, and it occurs in the nominative. The correlation between proto-agent properties and the genitive marking is a tendency rather than a strict rule, see Table 1 for the calculations based on the set of elicited examples compiled during our fieldwork sessions:

	genitive	**nominative**	**total**
proto-agent	74	30	104

proto-patient	42	69	111
total	116	99	215

Table 1. The number of DOs in nominative and genitive depending on DO's proto-role

This correlation is significant, which is confirmed by the Chi-Square Test (χ^2 = 23.989, p = <.001). To understand the exact impact of proto-agentivity on the choice of the DO marking we tested the following hypotheses.

Hypothesis 1. Proto-agentivity is determined by the lexical properties of the DO, i.e. by whether the DO is likely to be interpreted as a perceiving, volitional and moving entity.

Hypothesis 2. Proto-agentivity is determined by the lexical properties of the verb.

The hypothesis 1 is ruled out based on minimal pairs such as (1) and (2), where the same lexeme shows different marking. Moreover, inanimates can be occasionally marked with a genitive, despite the fact that, as mentioned above, they usually occur in nominative:

(3)	*wəj*	*ba-jjaf-ta*	*ɜp:ɜt-ə*	*sərd-dɜr*	***mašinɜ-jə /***	***mašinɜ***
	he	PV-catch.up-PST.3SG	all-GEN	fast-CMPR	car-GEN	car

'He caught up the fastest **car**.'

The genitive with inanimates may be chosen in the case where the DO demonstrates proto-agent properties such as movement in (3). Examples where the DO is a patientive human animate are also of interest. It turned out that even patientive people are marked with a genitive, which means that the proto-role factor is rather irrelevant for human animates:

(4)	*asə*	*šak'adax-ə*	*saržə-tɜ*	*xɜr-ənѕ*	***adɜm-ə /***	****adɜm***
	this-DET	island-GEN	inhabitant-PL	eat-PRS.3PL	people-GEN	people

'The inhabitants of this island eat **people**'.

Hypothesis 2. This hypothesis assumes that objects with certain verbs would always have only one case regardless of other factors. To test this assumption we analyzed about 63 transitive verbs with DOs of different animacy classes. For all of them, both cases are possible, see (1) vs. (3) and (2) vs. (5).

Conclusions. We have studied the influence of proto-role on the DOM in Ossetic and we suggest the following rule: if the agentivity of the DO is implied, then it is marked with the genitive. By contrast, if the DO does not have agentive properties, then it occurs in the nominative. In our talk we would like to show that this factor is not lexically driven by the semantics of the DO or by the semantics of the verb. Therefore, we hypothesize that it is a context-driven factor.

References

1. Abaev V.I. Grammaticheskii ocherk osetinskogo yazyka // Ordzhonikidze: Severo-Osetinskoe knizhnoe izdatel'stvo, 1959. 168 p.
2. Dowty D. Thematic proto-roles and argument selection // Language. 1991. V. 67. № 3. P. 547-619.
3. Primus B. Animacy, generalized semantic roles, and differential object marking // Case, word order and prominence: Interacting cues in language production and comprehension. Dordrecht: Springer Netherlands, 2011. P. 65-90.
4. Serdobolskaya N.V., Tuzhik O.V. Differential object marking in Modern Ossetic: referential properties and animacy. Accepted in Tomsk Journal of Linguistics and Anthropology. 2024.
5. Şahingöz E. Differential Object Marking in Ossetic: A corpus-based analysis // Conference presentation. 2022.
6. Thordarson F. 2009. Ossetic grammatical studies. Veröffentlichungen zur Iranistik: Vol. 48. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.

Anaphora resolution in Russian-Ossetic bilingual adults

Olga V. Tuzhik
Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia, Center for Language and Brain, HSE University, Moscow, Russia, vikolya2002@gmail.com

Introduction. Our study is devoted to anaphora resolution in Russian-Ossetic bilingual adults, which, to the best of our knowledge, has not been previously considered in the literature. According to Kambolov (2007: 38), "the monolingual type, Ossetic or Russian, is not identified among Ossetians", which means that almost all Ossetians in North Ossetia are Russian-Ossetic bilinguals. Previous works examining cross-linguistic influences on anaphora resolution in bilinguals contain a number of insights. For instance, van Dijk et al. (2022) showed that the more Turkish-dominant children were, the more they inhibited Turkish interpretation when hearing Dutch pronouns in ambiguous sentences. Another study by Keating et al. (2014) showed that both monolingual speakers of Mexican Spanish and bilingual Spanish heritage speakers exhibited a bias towards using a null pronoun for the subject, but the bias towards using an overt pronoun for the object was observed only among the monolinguals.

In this paper I investigate contexts with referential ambiguity, where available grammatical and semantic information does not allow one to conclude which of the several antecedents the anaphor refers to. In Russian the interpretation of such contexts depends on several different factors. These include the antecedent's grammatical function: it was found that both children and adults are more likely to designate the subject as the antecedent (Gagarina 2007; Delikishkina, Fedorova 2012). Another factor is animacy: an anaphor refers to an animate noun more often than to an inanimate one (Gagarina 2007). If there is cross-linguistic influence from Russian, I assume that the same factors might be relevant for anaphora resolution in Ossetic. Besides, anaphora resolution in Ossetic may be influenced by factors specific to this language. In particular, in Ossetic subject pronouns are usually omitted and overt pronouns are used for additional semantic emphasis, see Akhvlediani (1969: 112). For instance, it is perfectly grammatical to say (1) without a pronoun in the second sentence in Ossetic, but it is impossible to say (2) in Russian. Note that in other contexts Russian allows structures without a subject, see Letuchiy (2022):

(1)	*Swanon*	*ba-jjaf-ta*	*wə-sə*	*birəɣ-ə.*	*Žnon*	*a-mard-ta*	*šag.*
	hunter	PV-catch.up-PST.3SG	that-DET	wolf-GEN	yesterday	PV-kill-PST.3SG	deer

'The hunter caught up with that wolf. Yesterday (he/it) killed a deer.'

(2)	*Oxotnik*	*do-gna-l*	*to-go*	*volk-a.*	*Včera*	**(on)*	*u-bi-l*	*olenj-a.*
	hunter	PV-catch.up-PST.3SG	that-ACC	wolf-ACC	yesterday	he	PV-kill-PST.3SG	deer-ACC

'The hunter caught up with that wolf. Yesterday he killed a deer.'

However, it is still not known whether there is any additional semantics behind this factor, as, for example, in Turkish, which is a null subject language where overt pronouns usually signal a shift in topic or emphasize their antecedent, see van Dijk et al. (2022).

Another possible factor is differential object marking. The main factors underlying the choice of direct object marker are animacy and referential properties of direct objects, see Serdobolskaya & Tuzhik (2024). As for the Russian DOM, it depends on animacy only. I suppose that factors affecting Ossetic DOM might influence anaphora resolution in that a marked object would be more salient, which in turn would lead to linking an anaphor with a direct object, if such a principle is true for Ossetic.

Method. In the summer of 2024, I conducted a field study in Vladikavkaz, eliciting potentially ambiguous examples from eight speakers (all women, aged 17 to 53, $Mean_{age}$ = 34.8, *SD* = 14.2). An example from the questionnaire is given below:

(1)	*alan*	*a-šərd-ta*	*a-sə*	*q'ɜbəla-jə.*	
	A.	PV-chase-PST.3SG	this-DET	puppy-GEN	
	žnon	*fɜ-tɜrš-ən*	*kod-ta*	*gɜdə-jə*	*lɜp:ən-ə.*
	yesterday	PV-fear-INF	do-PST.3SG	cat-GEN	cub-GEN

'Alan chased this puppy away. Yesterday (he) scared a kitten.'

Fisher's exact test showed that most of the examples have been regarded as ambiguous by native speakers, regardless of the type of pronoun:

	unambiguous	**ambiguous**	**total**
overt pronoun	11	24	35
null pronoun	1	10	11
total	12	34	46
***p* =**	*0.118*		

Table 1. The number of unambiguous and ambiguous contexts, depending on the subject

The data also showed that the presence of a modifier (a demonstrative pronoun) before DO might influence anaphora resolution: if there was a demonstrative pronoun before DO, it was chosen as the antecedent of an anaphor more often (18 examples) than in absence of a modifier (1 example). I assume that Ossetic pronouns might refer to a more salient antecedent as in Russian rather than to a less expected one as in Turkish. In the presence of a modifier DO is more salient but still not salient enough to minimize ambiguity (16 ambiguous examples):

	unambiguous	**ambiguous**	**total**
direct object with modifier	18	16	34
direct object without modifier	1	14	15
total	19	30	49
***p* =**	*0.002*		

Table 1. The number of unambiguous and ambiguous contexts depending on the subject pronoun

Discussion. I have been unable to find any influence of factors that are specific to Ossetic (namely presence of subject pronoun and DOM) on anaphora resolution. Nevertheless, it was discovered that direct objects with modifiers (demonstrative pronouns) were chosen as antecedents of anaphors more often than those without modifiers. In my talk I would also like to discuss other factors, in particular, those that influence anaphora resolution in Russian. I will also discuss the results of an additional questionnaire conducted in Russian among native speakers of Ossetic and Russian (as a control group).

References

1. Akhvlediani G.S. Ossetic grammar [Grammatika osetinskogo yazyka]. V. II. Syntax. Ordzhonikidze: Research Institute under the Council of Ministers of the North Ossetian ASSR. 1969.
2. Delikishkina E.A., Fedorova O.V. Influence of syntactic role of an antecedent on the referential ambiguity resolution in Russian [Vliyaniye faktora sintaksicheskoy roli antetsedenta na razresheniye referentsial'noy neodnoznachnosti v russkom yazyke] // Computational linguistics and smart technologies: proceedings of the conference "Dialogue". Moscow: RSUH. 2012. V. 11. PP. 129-137.
3. Gagarina N.V. "The hare hugs the rabbit. He is white... Who is white?": Pronominal anaphora in Russian // ZAS Papers in Linguistics. 2007. V. 48. PP. 139-170.
4. Kambolov T.T. Language situation and language policy in North Ossetia: History, modernity, prospects [Yazykovaya situatsiya i yazykovaya politika v Severnoy Osetii: Istoriya, sovremennost', perspektivy] // Vladikavkaz: NOSU. 2007. 290 p.
5. Keating G.D., Jegerski J., Vanpatten B. Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. Bilingualism: Language and Cognition. 2014. PP. 1–14. doi:10.1017/S1366728914000418
6. Letuchiy A.B. Subject. Materials for corpus description of Russian grammar [Materialy dlya proyekta korpusnogo opisaniya russkoy grammatiki] (http://rusgram.ru). Moscow. 2022.
7. Serdobolskaya N.V., Tuzhik O.V. Differential object marking in Modern Ossetic: referential properties and animacy. Accepted in Tomsk Journal of Linguistics and Anthropology. 2024.
8. van Dijk, C., Aumeistere, A., Brouwer, S., Dijkstra, T., & Unsworth, S. Cross-linguistic influence online: An eye-tracking study on pronoun resolution in simultaneous bilingual Turkish-Dutch children // Proceedings of the 46th annual Boston University Conference on Language Development. Cascadilla, 2022. PP. 153-166.

THE SUFFIXES OF NOMINAL PLURAL IN THE KADAR DIALECT OF DARGWA

Naida A. Vagizieva,
Candidate of Sciences (Philology), research associate,
Institute of Language, Literature and Art
of the Daghestan Federal Research Centre of RAS
Makhachkala, Russia

The study is a description of the suffixes used to form nominal plural in the Kadar dialect of Dargwa. I analyzed more than 700 Kadar nouns taken from the lexicographic database of Dargwa (http://lingconlab.ru/dargwa_dict/) :

-me: for the plural forms of one syllable nouns ending in a vowel: *c'a* 'fire' – PL *c'a-me*, ši 'village' – PL *ši-me*, *u* 'name' – PL *u-me.*

-e: for the plural forms of one syllable nouns ending in a consonant: *bec'* 'wolf" – PL *buc'-e*, *unc 'ox' – PL unc-e*

-be: for the plural forms of two syllable nouns, ending in the vowel: *uzi* 'brother' – PL *uz-be*, *urk'i* heart' – PL *urk'-be.* The root vowel is replaced by a vowel – *u.* The marker *-be* can also be a part of a two syllable suffix: *anda* 'forehead' – PL *and-urbe*, *dugi* 'night' – PL *dug-urbe*, *c'ank'a* "field" – *c'ank'-urbe*.

-t: forms the nominal plural of nouns, ending in sonorants *l, m, n,r, j* or the labial-obstruent *b*: daˤħaˤm 'game' – PL daˤħaˤm-t, ʔaˤmal 'business, behavior, character' – PL 'ʔaˤmal-t'.

-te: forms the nominal plural from participles with the notional meaning; *bak'ib-te* 'visitors (literally. those,who came)', *kabiib-te* 'those people, who sit', *buku-te* 'those, who eat', etc.

-re: is for one or two syllable nouns, ending in a vowel or a consonant: *ʔark'* 'river'-PL ʔurk'-re, *qat* 'palm'- PL *qut-re*, *k'ap'i* – PL *k'ap'-re*. In nouns, ending in a consonant plural formation is accompanied by vowel alternation. The root vowels are alternating with the formation of plural forms. The root vowel is replaced by a narrow vowel – *u.*

The formation of nominal plural in the Kadar dialect is accompanied by some phonetic processes. Suffix *-ne* is attached to two syllable nouns, ending in the vowel. The final vowel is omitted in the plural forms; *durʔaˤ* 'child' – PL *durʔ-ne, čʼik'a* 'chick, chicken' - PL *čʼik'-ne*, *iza* 'disease' – PL iz-ne. The omission of vowel happens when the following suffixes are added to the nouns, ending in a vowel: *-ule*: *diʔebʁa* 'prayer' – PL *diʔebʁ-ule, ispica* 'spoke' – PL *ispic-ule*; -urbe: *gwanza* 'soil' – PL *gwanz-urbe*, *dugi* 'night' – PL *dug-urbe*; *-ube*: *qarʔa* 'bean' – PL *qarʔ-ube, t'erxa* 'stick' – PL *t'erx-ube*; *-ume*: *valri* 'camel' – valr-ume. The suffixes *-re, -me, -ane, -be* form the of plural nouns from monosyllabic words. In the forms of nominal plural with the suffixes *-re, -be* the root vowel is replaced by a vowel – *u:*

-re: *baz* ‘month’ – PL buz-re; bec’ ‘wolf’ – PL buc’-e; -ane: *žuz* ‘magazine’- PL *žuz-ane; k’aw* ‘lake’ – PL *k’aw-ane, gas* ‘natural gas’ – PL *gaz-ane*: *-me*: *u* ‘name’- PL *u-me*; , *ši* ‘village’ – PL *ši-me*: *-be*: deχ ‘cargo’- PL duχ-be, *beʔ* ‘peak’ – PL *buʔ-re*.

There is a number of irregular nouns where the last two phonemes of the singular stem (vowel + sonorant) disappears in the plural: *murʔul* ‘man’ – PL ‘*muʔ-le”, aɬaˁl* ‘guest’ – PL ‘*aɬ-le’* etc.

-lart: the Kadar suffix *-lart*, historically consists of two markers: the Turkic - *lar* and Kadar *-t (*Vagizieva 2021). These markers mainly derives plural forms from the Russian borrowings: *očki* ‘sunglasses’ - PL *očki-lart,* maškara *‘moskit’* - PL *maška- lart, stanciˁa* ‘station’ – PL *stanciˁa -lart, škola* ‘school’ – PL *’škola-lart,* etc. Stress is shifted to the suffix *-lart.*

Thus, the suffixes of nominal plural in the Kadar dialect are as follows; *-e, -re, -be, -me, -ne, -umbe, -une, -t, -ane, -le, -ure, -urt, -ube, -ule, -lart*, etc (Mutalov, Vagizieva 2018, 2019). All these markers (except *-t* and *-lart*) contain the final vowel *-e* . The phonetic processes are identified like that: the alternation of the root vowels with vowel -u-, as well as the simultaneous loss of the final vowels and the alternation of the root vowels, omission of sounds and etc.

References

1. Vagizieva N.A. Leksiko-frazeologicheskiye osobennosti kadarskogo dialekta darginskogo yazyka. - Makhachkala: Institute of Yali DFIC RAS, Alef, 2021. - 200 p.
2. Vagizieva N.A. O sposobakh vyrazheniya kategorii chisla v Kadarskikh imenakh sushchestvitel'nykh // Mir nauki, kul'tury, obrazovaniya. 2023. № 5 (102). Pp. 254-256
3. Mutalov R.O., Vagiziyeva N.A. Zvukovyye sootvetstviya mezhdu darginskim literaturnym yazykom i kadarskim idiomom // Izvestiya Dagestanskogo gosudarstvennogo pedagogicheskogo universiteta. Obshchestvennyye i gumanitarnyye nauki. 2018. T. 12. № 3. S. 60-66.
4. Mutalov R.O., Vagiziyeva N.A. Razryady chislitel'nykh v Kadarskom yazyke // Filologiya: nauchnyye issledovaniya. 2019. № 6. S. 44-49.

Names of maize in the eastern Caucasus

Samira Verhees (Independent researcher)

Maize (*zea mays*) is a staple of Caucasian cuisine and one of the most common crops grown in the area. It is an exotic plant originating from Mesoamerica and was first introduced into the Caucasus in the 17th century (according to Bonavia (2013:255), citing Haudricourt and Hédin (1987: 113)). It seemed well-established by the mid-18th century, when Johann Güldenstädt traveled through the Caucasus and recorded words for maize that are still used today in languages like Georgian, Ingush, Avar, Ossetic and Kabardian (Güldenstädt 1791: 496-552).

What is striking about the terms used in the eastern Caucasus in particular (including languages of the East Caucasian family and local Turkic languages), is that we find a lot of similar terms and strategies, yet the distribution of shared features does not quite follow the usual areal patterns. Borrowed lexemes in Dagestan are generally shared among language communities that historically used the same regional lingua franca (Daniel et al. 2021): Azerbaijani in the south and Avar in the northwest, with a transitional zone in between. Exceptions are general borrowings like 'hour' from Arabic or 'telephone' from Russian, which are more broadly distributed, and incidental local borrowings between neighbors.

In the case of maize, we find oddly similar terms in a variety of East Caucasian lects that are distant both in terms of genealogy and language contact situations, cf. *ħadži-buʁda* (Agul), *ħež-q'ir* (Tukita), *ħažl-ač'e* (Kubachi), *ħæž-k'a* (Chechen),[1] yet they lack the uniformity of general borrowings. All of these words are compounds meaning 'hajj-wheat'. The Agul word is clearly copied from a dialect of Azerbaijani spoken in or near southern Dagestan, while the other languages realize the same meaning with a native word for 'wheat'. It is not uncommon for maize to be referred to as a marked type of known grain like 'wheat' or 'millet', but the name 'hajj-wheat' seems to be specific to the eastern Caucasus. It connects the introduction of maize in the area to the Islamic religion and the Arabic peninsula, and could suggest a different route of introduction than the western and central parts of the Caucasus. Another curious aspect of the term 'hajj-wheat', is that a number of languages appear to have copied the Chechen variant. This includes languages like Lezgian, which do not neighbor Chechen-speaking communities.

My paper will provide an overview of words for maize in 119 distinct lects spoken in the eastern Caucasus, and show the distribution of types on a map. I will also discuss the implications of this distribution.

References

Bonavia, Duccio. 2013. *Maize. Origin, domestication, and its role in the development of culture*. Cambridge: Cambridge University Press.

[1] The Agul and Tukita translations come from the Daghestanian Loans Database (Chechuro et al. 2019), Kubachi is from (Magomedov & Saidov-Akkutta 2017), and Chechen comes from the Chechen Wikipedia page for *zea mays* (https://ce.wikipedia.org/wiki/%D0%A5%D1%8C%D0%B0%D1%8C%D0%B6%D0%BA%D3%80%D0%B0, accessed 21 October 2024).

Chechuro, Ilia, Michael Daniel and Samira Verhees. 2019. *Daghestanian loans.* Moscow: HSE Linguistic Convergence Laboratory. https://lingconlab.ru/dagloans/, accessed 22 October 2024.

Daniel, Michael, Ilia Chechuro, Samira Verhees, and Nina Dobrushina. 2021. 'Lingua francas as lexical donors: evidence from Daghestan'. *Language* 97.3, 520–560.

Güldenstädt, Johann Anton. 1791. *Reisen durch Rußland und im Caucasischen Gebürge 2.* St. Petersburg: Kayserlichen Akademie der Wissenschaften. https://www.digitale-sammlungen.de/de/view/bsb11450090, accessed 22 October 2024.

Haudricourt, André G. and Louis Hédin. 1987. *L'Homme et les plantes cultivées.* Paris: Éditions A.M. Métailié.

Magomedov, Amirbek D. and Nabigulla I. Saidov-Akkutta. 2017. *Kubačinsko-russkij slovar'* [Kubachi-Russian dictionary]. St. Petersburg: Nauka.

Arabic Loanwords in Dagestan Through the Lens of Geography and Linguistic Ecology

Ellie Wren-Hardin
The Ohio State University

Over forty languages are spoken within Dagestan, most of which are members of the Northeast Caucasian language family, resulting in intense language contact over hundreds to thousands of years as well as many lexical borrowings. While there are family-internal borrowings, there has also been contact with many non-Northeast Caucasian languages, such as Arabic. Most contact from Arabic is the result of the conversion of the majority of the population of the area to Islam from the 7th to 15th centuries and its usage for religious purposes (Dobrushina & Kultepina 2021: 341); but it was likely never widely spoken in the region (Daniel et al 2021: 522). As a result, it can be difficult to determine whether Arabic lexical borrowings were borrowed directly from Arabic or were mediated by another language. Studies of loanwords across Dagestan frequently avoid Arabic loanwords due to the difficulty in determining the borrowing path of these items. Consequently, this study attempts to investigate the diffusion of Arabic loanwords by approaching them from a qualitative and geographic perspective.

Utilizing the Daghestanian Loans database (Chechuro et al., 2019), based on a 160-concept lexical list intended to contain items especially susceptible to borrowing, I identified forty-seven concepts with at least one Arabic loan in a Northeast Caucasian language in Dagestan. I then divided the loanword sets according to their geographic placement in Dagestan and examined the loanwords within each concept set individually for evidence of convergence and diffusion. The separation between northwest and south is based on the differing historic linguistic ecologies present in those regions; Avar has served as a lingua franca in the northwest and Azerbaijani in the south. The results of this categorization can be seen in Tables 1 and 2.

Overall, the spread and existence of Arabic loanwords in Dagestan is the result of geographic and linguistic factors: namely, how close the region was to the areas of contact with Arabic and Azerbaijani (which has many Arabic loanwords as well), and what the lingua franca of the region was. The results demonstrate that there are more Arabic loanwords present in the south of Dagestan (closer to Azerbaijani and the historic contact with Arabic) than in the north, as would be expected. The results also demonstrate how the two regional lingua francas served to disseminate and regularize the forms of the loanwords, with many loanwords appearing largely or exclusively in one lingua franca region, and when they are present in both the forms are more similar to one another within a lingua franca region than across it. Division of these items in geographic space as well as across ecologies suggests it is easier for the loanwords to circulate within a lingua franca region that across it. This methodology demonstrates how information can be deduced regarding the diffusion of loanwords based on geography and known linguistic ecologies when the borrowing path of the loanword into a specific language may be otherwise opaque.

Table 1: Geographic Placement of Loans (Strict)

Largely Northwest	**Largely South**	**Both**	**Scattered**	**Few Realizations**
4	23	9	2	10

Table 2: Loanwords in the Northwest and South

Northwest	**South**
24	44

References

Chechuro I., Daniel M., and Verhees S. 2019. Daghestanian loans. Linguistic Convergence Laboratory, HSE.

Daniel, M., Chechuro, I., Verhees, S., & Dobrushina, N. (2021). Lingua francas as lexical donors: Evidence from Daghestan. *Language*, *97*(3), 520–560. https://doi.org/10.1353/lan.2021.0046

Dobrushina, N., & Kultepina, O. (2021). The rise of a lingua franca: The case of Russian in Dagestan. *International Journal of Bilingualism*, *25*(1), 338–358. https://doi.org/10.1177/1367006920959717