# Uninflectable/uninflected verbs

Typological trends and a corpus-based comparison of two Nakh-Dagestanian languages

Katherine Walker & Eva van Lier
University of Amsterdam

AMSTERDAM CENTER FOR LANGUAGE AND COMMUNICATION

NWO

# Overview

- Intro: Uninflectability/uninflectedness *in verbs* as part of Differential Argument Indexing* (DAI)
- Typological survey
- Uninflectable verbs in discourse: Sanzhi Dargwa & Chechen

* Indexing = “agreement”

# Uninflectability/Uninflectedness as DAI

- DAM (Witzlack-Makarevich & Seržant 2018)
  - typological work shows that at least for P DAM is typologically dominant (e.g. Sinnemäki 2014 for flagging, Haig 2018 for indexing)
  - mostly focuses on contextual factors:
    - argument-related factors (inherent, e.g. humanness, person; non-inherent, e.g. definiteness, topicality)
    - TAM-based splits
  - often ignores lexically determined splits (Haig 2018)
  - traditionally mostly about flagging (although more recently see e.g. Just 2022)

# Our focus

- Differential Argument Indexing (DAI)
  - (non-)indexing on the verb of person/number/gender features of S/A/P
- Asymmetrical splits (i.e. zero/uninflected vs. overt indexing)
  - cf. symmetrical splits (e.g. Iemmolo 2010)
- Splits involving lexical verb classes
  - in interaction with other DAI-triggering (contextual) factors

# Typological survey

# Data set

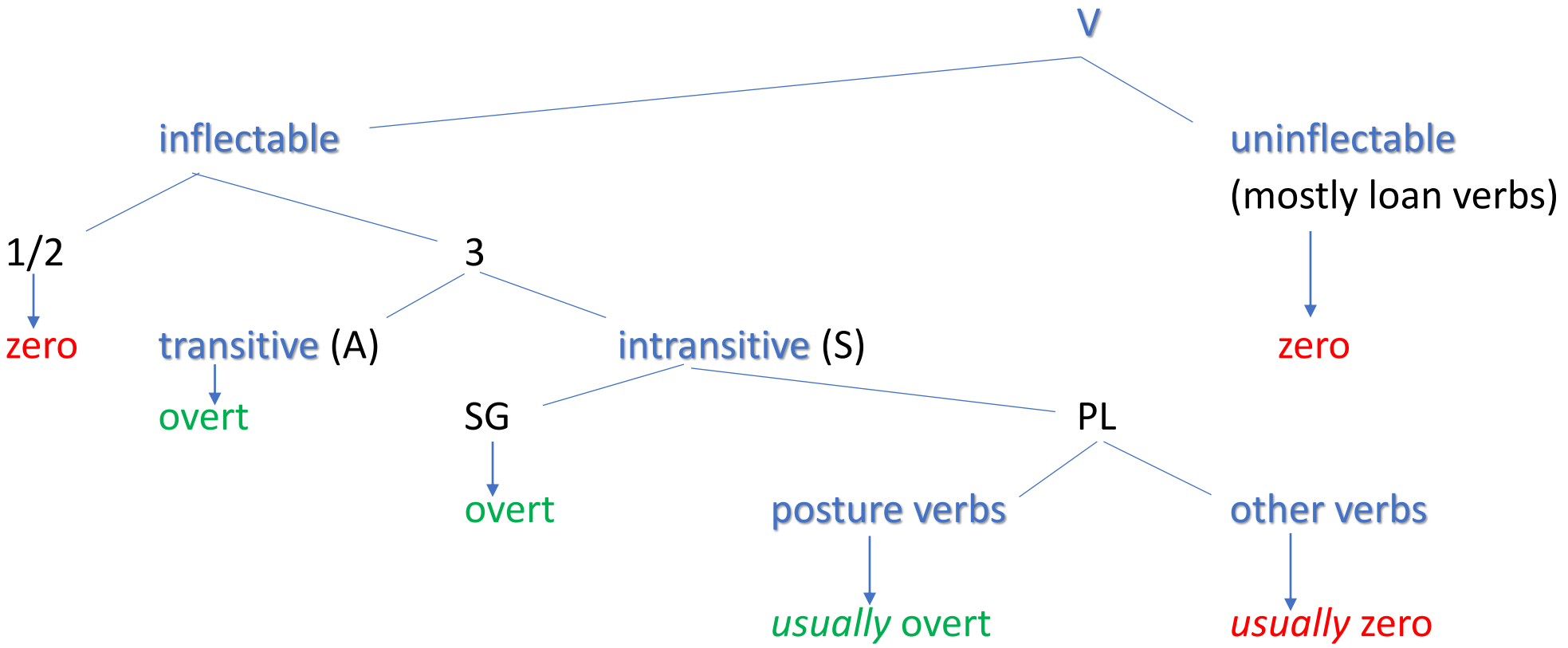
- 90 language sample selected from AutoTyp (https://www.autotyp.uzh.ch/)
- Currently 40 languages covered
- Half of them (20/40) have DAI involving verb class

# Dimensions of variation

- Interaction between verb class and other factors in deciding between indexing/zero
  - Examples
- Types of splits within languages (cf. Spencer 2020):
  - (Always) inflecting verbs vs. uninflected verbs (conditioned)
  - Uninflectable verbs (never) vs. uninflected verbs (conditioned)
  - (Always) inflecting verbs vs. uninflectable verbs (never)
- Default vs. minority patterns
  - Grammar-based typology vs. corpus-based case studies

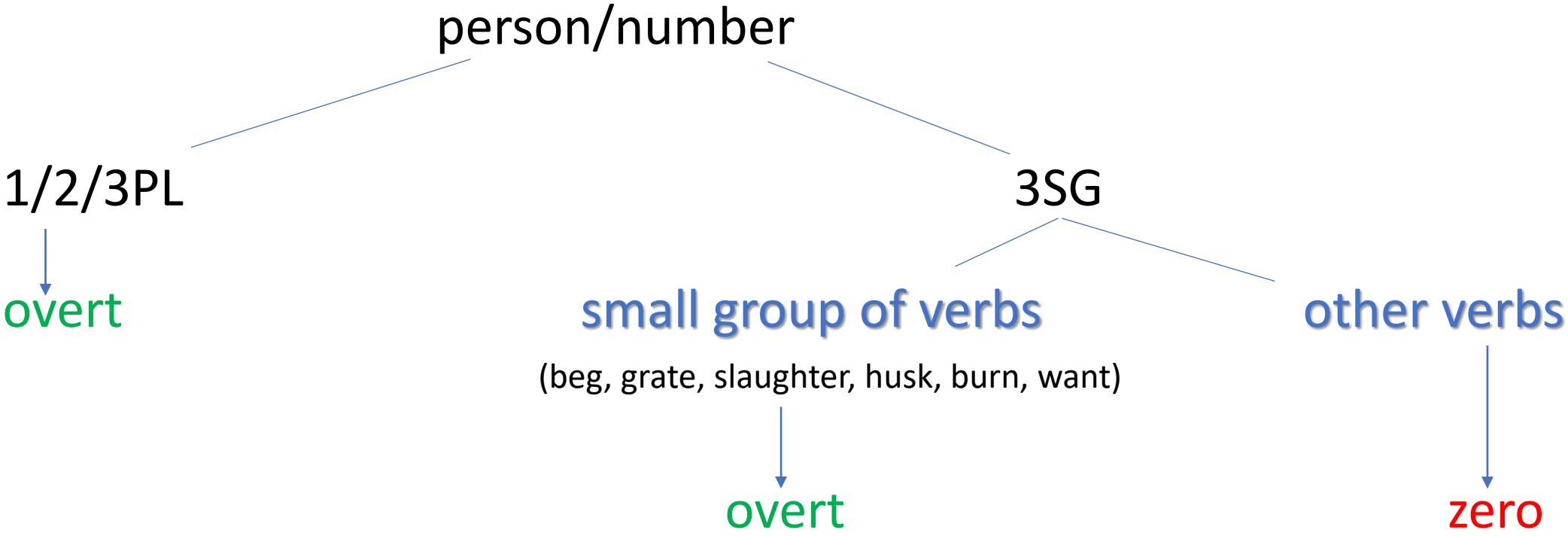
# Verb classes in DAI systems (i)

- Ese Ejja: S/A person indexing (Paco-Tacanan, Bolivia; Vuillermet 2012)



# Verb classes in DAI systems (ii)

- South Efate: P person/number indexing (Oceanic, Vanuatu; Thieberger 2006)



# Types of verb class splits within languages

| Example languages | ALWAYS | NEVER (uninflectable) | CONDITIONED (uninflected) |
|---|---|---|---|
| e.g. South Efate (P) | √ | | √ |
| e.g. Hua (P) | | √ | √ |
| Goemai (S/A/P), Tariana (S) | √ | √ | |

# Types of verb class splits within languages

| Example languages | ALWAYS | NEVER (uninflectable) | CONDITIONED (uninflected) |
|---|---|---|---|
| e.g. South Efate (P) | √ | | √ |
| e.g. Hua (P) | | √ | √ |
| Goemai (S/A/P), Tariana (S) | √ | √ | |

- Hua P (p/n): minority ('put', 'eat', 'bash') never vs. default (other verbs) conditioned by humanness/person (lexical) **reverse so far unattested**
- Goemai S/A/P (n): minority (10%) always vs. default (90%) never (etymological)
- Tariana S (p/n/g): minority (small closed class of stative/physical state verbs) never vs. default (other verbs) always (semantic)

# Always-never splits based on verb class

- Seem to be relatively rare, compared to cases where there is a verb class with DAI conditioned by other factors
- Apart from Goemai and Tariana few candidates
- Apparently, the uninflectable class can be the majority (Goemai), or the minority (Tariana)
- Role of discourse frequency in maintaining minority pattern (cf. Fedden 2019)?

# Uninflectable verbs in discourse

Comparing Sanzhi Dargwa and Chechen (Nakh-Daghestanian)

# The role of frequency

- ‘Frequency of usage is often implicated in the stability of irregularities in language’ (Fedden 2019)
- → Irregular patterns should be overrepresented in discourse (low type frequency = high token frequency).

# Why Nakh-Daghestanian?

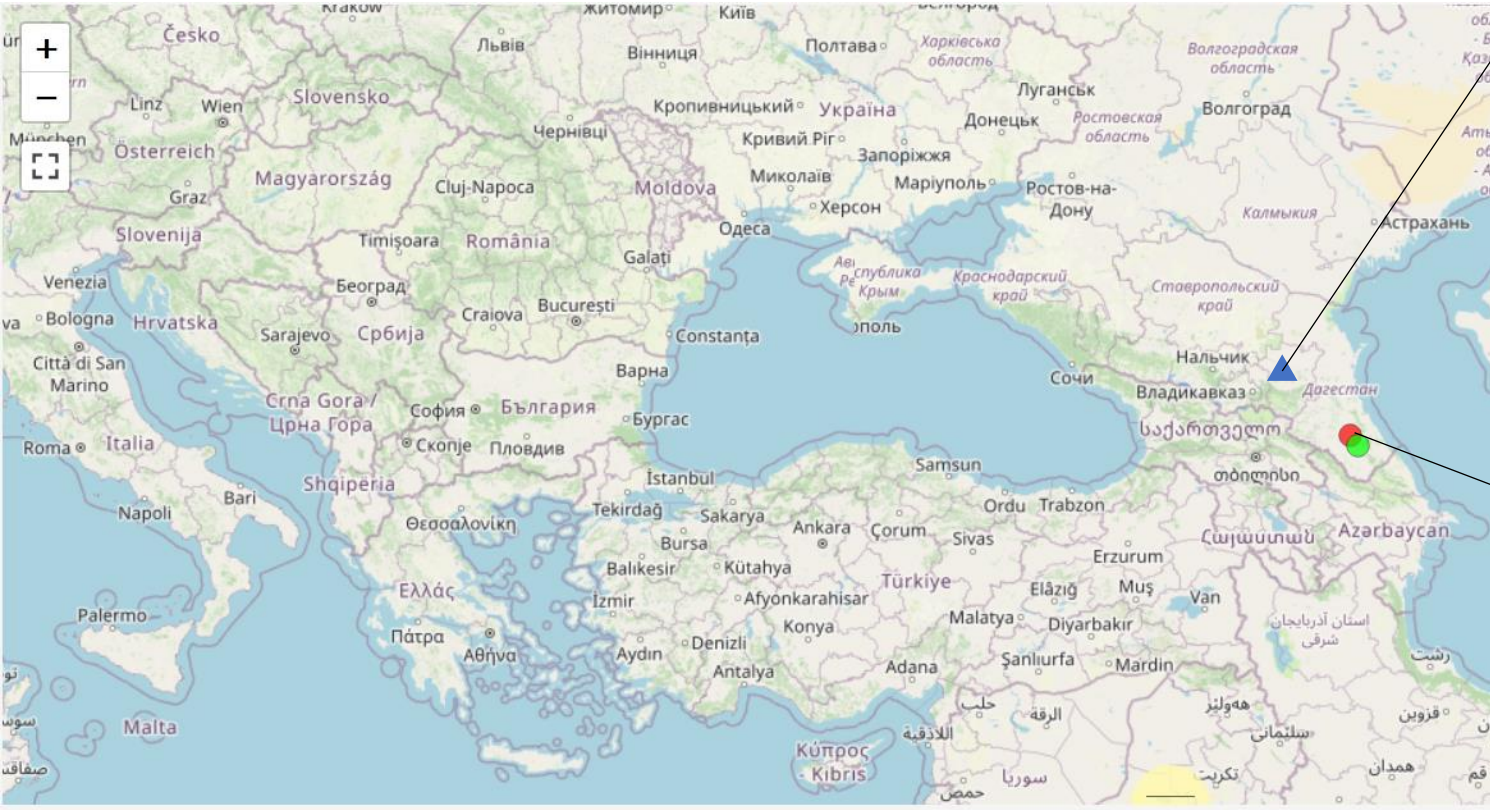
‘there is no Nakh-Daghestanian language with gender agreement in which all verbal roots have agreement exponents’ (Forker 2018: 867)

# Verb classes & frequency

Previous work on Nakh-Daghestanian: (% of inflectable/inflected verbs):

| **Language** | **Types** | | **Tokens** | |
|---|---|---|---|---|
| Lak | Some are uninflectable | | <90%<br>(500 clauses) | *Forker 2018* |
| Tsez | 27%<br>(dictionary) | 60%<br>(discourse) | 84%<br>(3,000 lines, CDS) | *Gagliardi 2012;*<br>*Gagliardi & Lidz 2014* |
| Hinuq | 33% | | 65%<br>(1729 clauses) | *Forker 2013; 2018* |
| Avar | ‘most vowel-initial verbs’ | | 63%<br>(845 clauses) | *Forker 2018; 2021* |
| Chechen | 30% | | 50% | *Komen et al. 2021* |

# Chechen & Sanzhi Dargwa




**Chechen**:

- Speakers: 1.28 million (Dobrushina et al. 2021)
- Official language of the Chechen Republic – printed literature, newspaper

**Sanzhi Dargwa:**

- Approx. 250 speakers
- Critically endangered
- Official language of Republic of Dagestan, but not an official written language
- Description: Forker (2020)

# Gender indexing: Chechen

(1) k’ant-as quor b-u’u ← V initial: agreement
boy(V)-ERG pear(B).ABS B-eat.PRS
‘The boy eats the pear.’ (Molochieva et al. 2022)

(2) k’ant-ana quor go ← C initial: no agreement
boy(V)-DAT pear(B).ABS see.PRS
‘The boy sees the pear.’ (Molochieva et al. 2022)

# Gender indexing: Chechen

(1) k’ant-as quor b-u’u ← V initial: agreement
boy(V)-ERG pear(B).ABS B-eat.PRS
‘The boy eats the pear.’ (Molochieva et al. 2022)

(2) k’ant-ana quor go ← C initial: no agreement
boy(V)-DAT pear(B).ABS see.PRS
‘The boy sees the pear.’ (Molochieva et al. 2022)

(3) hwaahwa’a, t'argh olxu-sh j-olu jaxk ← V initial: no agreement
no wool comb-CVBsim J-be.PTCP comb(J)
‘no, the comb that cards wool’ [witch_383]

# Gender indexing: Chechen

(1) k’ant-as quor b-u’u ← V initial: agreement
boy(V)-ERG pear(B).ABS B-eat.PRS
‘The boy eats the pear.’ (Molochieva et al. 2022)

(2) k’ant-ana quor go ← C initial: no agreement
boy(V)-DAT pear(B).ABS see.PRS
‘The boy sees the pear.’ (Molochieva et al. 2022)

(3) hwaahwa’a, t'argh olxu-sh j-olu jaxk ← V initial: no agreement
no wool comb-CVBsim J-be.PTCP comb(J)
‘no, the comb that cards wool’ [witch_383]

(4) i shie j-aqq-ana j-olu saara’a ← (V-initial, REL clause + AUX)
dem 3sg.refl J-take-CVBant J-be.PTCP stick(J)
‘The sticks he had taken for himself.’ [bear_018]

# Gender indexing: Sanzhi Dargwa

As for Chechen:

- Gender prefix indexes ABS argument
- Appears on most vowel-initial verbs

# Gender indexing: Sanzhi Dargwa

**PLUS:**

- Optional gender marker following **NEG prefix**:

(5) a-arq’-ib=da=jal

NEG-DO.PFV-PRET=1=INDQ (Forker 2020: 528)

(6) a-b-arq’-ib=da,

NEG-N-DO.PFV-PRET=1 (Forker 2020: 306)

# Gender indexing: Sanzhi Dargwa

## PLUS:

- Optional gender marker following **NEG prefix**:

(5) a-arq'-ib=da=jal

NEG-DO.PFV-PRET=1=INDQ (Forker 2020: 528)

(6) a-b-arq'-ib=da,

NEG-N-DO.PFV-PRET=1 (Forker 2020: 306)

- Gender **suffixes/infixes** on a small number of C-initial verbs:

(7) ca-r ca-b ca<b>i …

COP-F COP-HPL COP<HPL>

# Gender indexing: Sanzhi Dargwa

## PLUS:

- Optional gender marker following **NEG prefix**:

(5) a-arq'-ib=da=jal
NEG-DO.PFV-PRET=1=INDQ (Forker 2020: 528)

(6) a-b-arq'-ib=da,
NEG-N-DO.PFV-PRET=1 (Forker 2020: 306)

- Gender **suffixes/infixes** on a small number of C-initial verbs:

(7) ca-r ca-b ca<b>i …
COP-F COP-HPL COP<HPL>

- **Masculine: w-/ø-**

Masc. is obligatorily ø- prior to /u/

(8) *uk:-unnea=da* (masc.)
*r-uk:-unne=da* (fem.)
GM-eat.IPFV-ICVB=1
'I will eat'.

Masc. is optionally ø- prior to /i/

(9) *(w-)ik'-ul* (masc.)
*r-ik'-ul* (fem.)
GM-say.IPFV-ICVB
'saying'. (Forker 2020: 214)

**…and more**

# Data

Multi-CAST

*Multilingual Corpus of Annotated Spoken Texts*

Open-access collection of spoken monologic data (approx. 1,000 clauses; 17+ diverse languages. Haig & Schnell 2022)

**Chechen** (Molochieva & Walker in progress)

- 4 narrative monologues by 2 speakers (female, 70+, different dialects)

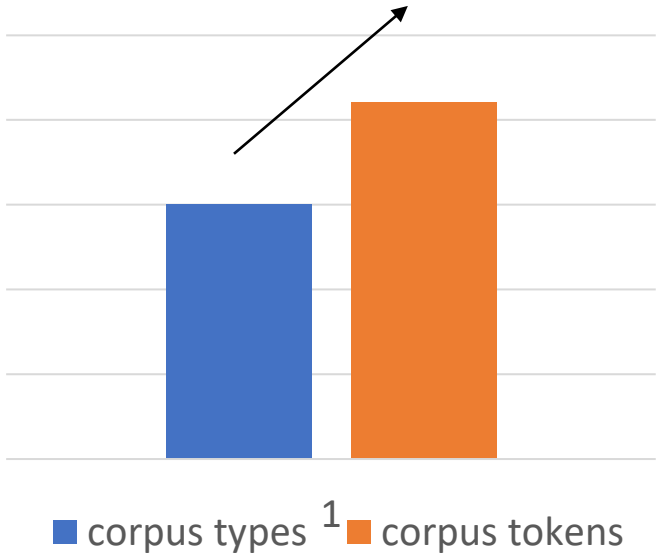
**Sanzhi Dargwa** (Forker & Schiborr 2019)

- mixture of 8 autobiographical and traditional narratives

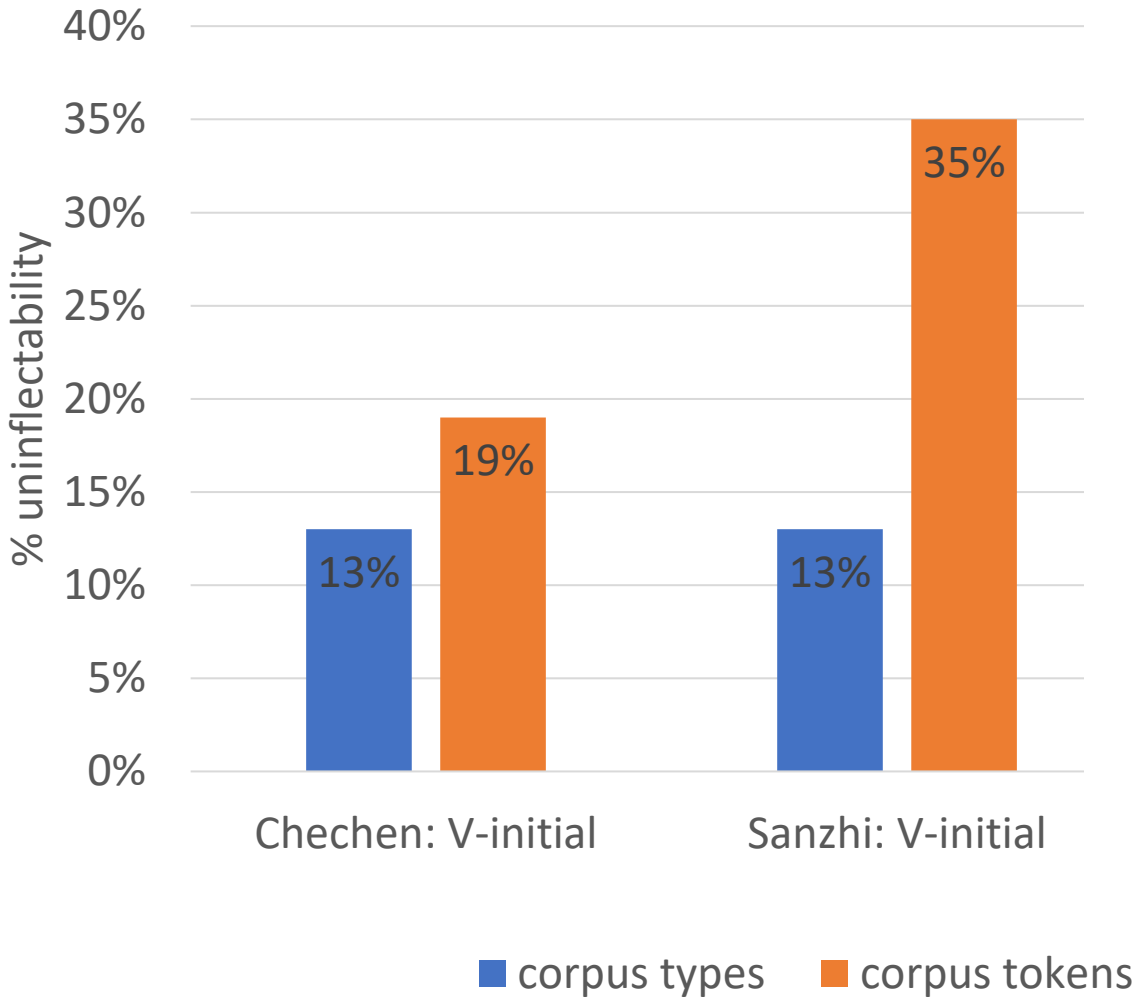
https://multicast.aspra.uni-bamberg.de/

# Expected results

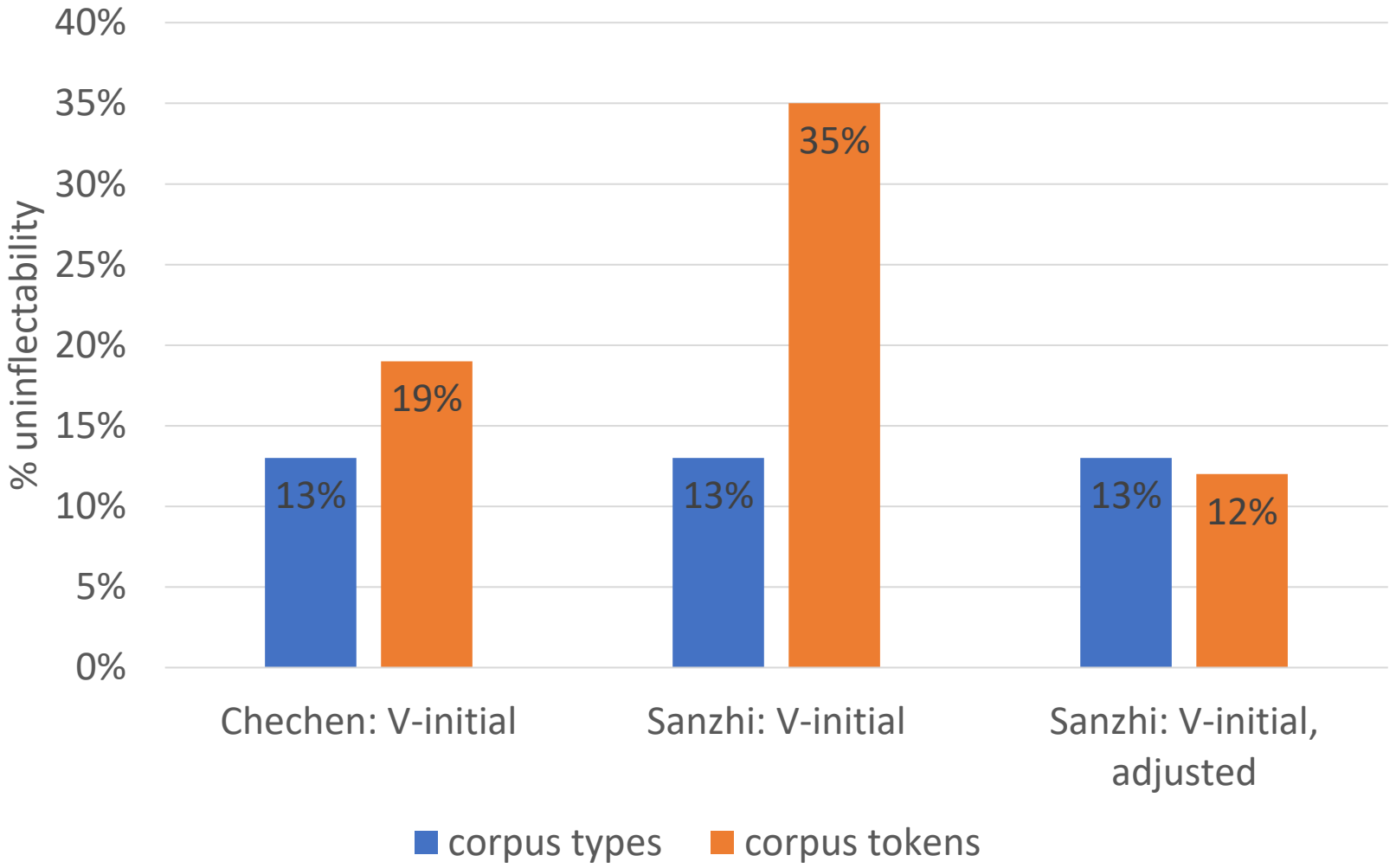
- Irregular pattern: Low type frequency, high token frequency
- For vowel-initial verbs, uninflectability is the minority pattern



# Results: vowel-initial verbs



# Results: vowel-initial verbs



# Possible explanations

Sanzhi Dargwa: Gender indexing as aspect marker?

| IPFV | PFV | preterite | translation |
|---|---|---|---|
| *iC* vs. *b-iC* | | | |
| *it-* | *b-it-* | *-ib* | 'beat up' |
| *irš:-* | *b-irš:-* | *-ib* | 'mow' |
| *ik:-* | *b-ik:$^{w}$-* | *-ub* | 'burn' |

(Forker 2020: 207)

# Possible explanations

Sanzhi Dargwa: Gender indexing as aspect marker?

| IPFV | PFV | preterite | translation |
|---|---|---|---|
| *iC* vs. *b-iC* | | | |
| *it-* | *b-it-* | *-ib* | 'beat up' |
| *irš:-* | *b-irš:-* | *-ib* | 'mow' |
| *ik:-* | *b-ik:$^{w}$-* | *-ub* | 'burn' |

(Forker 2020: 207)

→ Possible explanation for 1/16 verbs

# Possible explanations 2

- ‘while consonant-initial verbs [in Tsez] never agree, there are a few vowel initial verbs – Polinsky and Comrie (1999: 111) list ten – for which **one assumes the presence of an underlying laryngeal** which blocks agreement prefixes, just like any other consonant (Maria Polinsky, personal communication).’ (Fedden 2019)

# Possible explanations 2

- ‘while consonant-initial verbs [in Tsez] never agree, there are a few vowel initial verbs – Polinsky and Comrie (1999: 111) list ten – for which one assumes the presence of an underlying laryngeal which blocks agreement prefixes, just like any other consonant (Maria Polinsky, personal communication).’ (Fedden)

- Possible cognates? (very speculative!)
  - **Chechen**: 1/9
  - **Sanzhi Dargwa**: 1/16

# Case Study Summary

- Does discourse frequency aid in maintaining irregular uninflectability?
- Method
  - Chechen & Sanzhi Dargwa: a minority of vowel-initial verb types are uninflectable
  - gather type and token frequencies to establish whether the minority pattern is more frequent in spoken discourse
- Results
  - Chechen: Uninflectable verbs a little more frequent in discourse
  - Sanzhi Dargwa: Uninflectable verbs slightly less frequent in discourse
- What else is going on?
  - Aspectual distinctions (maybe partially for Sanzhi Dargwa)
  - Etymology (maybe…)

# References


Dobrushina, N., M. Daniel & Y. Koryakov. 2021. Languages and Sociolinguistics of the Caucausus. In M. Polinsky (ed), *The Oxford Handbook of Languages of the Caucasus*. OUP.

Fedden, S. 2019. To Agree or Not to Agree? A Typology of Sporadic Agreement. In M. Baerman, O. Bond and A. Hippisley (eds.), *Morphological Perspectives: Papers in Honour of Greville G. Corbett*, 303-326. Edinburgh: Edinburgh University Press.

Forker, D. 2013. *A grammar of Hinuq.* (Mouton Grammar Library, 63.) Berlin: De Gruyter Mouton.

Forker, D. 2018. Gender agreement is different. *Linguistics*, 56, pp. 865-894.

Forker, D. 2020. *A grammar of Sanzhi*. Berlin: Language Science Press.

Forker, D. & N.N. Schiborr. 2019. Sanzhi Dargwa. In G. Haig & S. Schnell (eds) *Multi-CAST: Multilingual corpus of annotated spoken texts.*

Gagliardi, A. C. 2012. Input and Intake in Language Acquisition, doctoral dissertation, University of Maryland, <http://drum.lib.umd.edu/bitstream/1903/13173/1/Gagliardi_umd_0117E_13440.pdf>

Gagliardi, A. C. & J. Lidz. 2014. 'Statistical insensitivity in the acquisition of Tsez noun classes', Language, 90 (1), pp. 58–89.

Haig, G. 2018. The grammaticalization of object pronouns: Why differential object indexing is an attractor state. *Linguistics* 56(4). 781 – 818.

Haig, G. & S. Schnell (eds.). 2022. *Multi-CAST: Multilingual corpus of annotated spoken texts*. Version 2211. Bamberg: University of Bamberg. (multicast.aspra.uni-bamberg.de/)

Iemmolo, G. 2010. Symmetric and asymmetric alternations in direct object encoding. *STUF* 66(4). 378 – 403.

Just, E. 2022. *A functional approach to differential indexing.* Utrecht: LOT Publications.

Komen, Erwin R., Zarina Molochieva, and Johanna Nichols. 2021. Chechen and Ingush. In Maria Polinsky (ed.) *The Oxford Handbook of Languages of the Caucasus*. Oxford: OUP.

Molochieva, Z., P. Faghiri & E. van Lier. 2022. Biabsolutive constructions in Chechen. *Folia Linguistica*, 56(2).

Sinnemäki. K. 2014. A typological perspective on differential object marking. *Linguistics* 52(2). 281 – 313.

Spencer, A. 2020. Uninflectedness: Uninflecting, uninflectable and uninflected words, or the complexity of the simplex. In L. Körtvélyessy & P. Štekauer (eds.), *Complex words: Advances in morphology*, 142–158. Cambridge: CUP.

Thieberger, Nick. 2006. *A grammar of South Efate. An Oceanic language of Vanuatu.* University of Hawaii Press.

Vuillermet, M. 2012. *A grammar of Ese Ejja, a Takanan language of the Bolivian Amazon.* PhD dissertation, Université Lumière Lyon 2.

Witzlack-Makarevich, A. & I. Seržant. 2018. Differential argument marking: Patterns of variation. In I. Seržant & A. Witzlack-Makarevich (eds.), *Diachrony of Differential Argument Marking*, 1–40. Language Science Press.

# Results: Vowel-initial verbs

| | Inflectable | Uninflectable | **Total TYPES** | Inflected | Uninflected | **Total TOKENS** |
|---|---|---|---|---|---|---|
| **Chechen** | 86.8% | 13.2% | *100%* | 81.2% | 18.8% | *100%* |
| | 59 | 9 | *68* | 688 | 159 | *847* |
| **Sanzhi Dargwa** | 86.6% | 13.4% | *100%* | 64.8% | 35.2% | *100%* |
| | 103 | 16 | *119* | 653 | 354 | *1007* |

| Inflectable | Uninflectable | **Total TOKENS** |
|---|---|---|
| 88.2% | 11.8% | *100%* |
| 888 | 119 | *1007* |