Evidentiality as an areal feature

an East Caucasian perspective on typology vs. reality

Samira Verhees

LCL / NRU HSE Moscow

Workshop Language Sampling

Institute of Linguistics RAS

St. Petersburg – 20.12.2019

What is evidentiality?

▶ Evidential markers indicate where the speaker's information comes from (personal experience, hearsay, inference)

▶ "highly diffusable" (Aikhenvald 2004: 302)

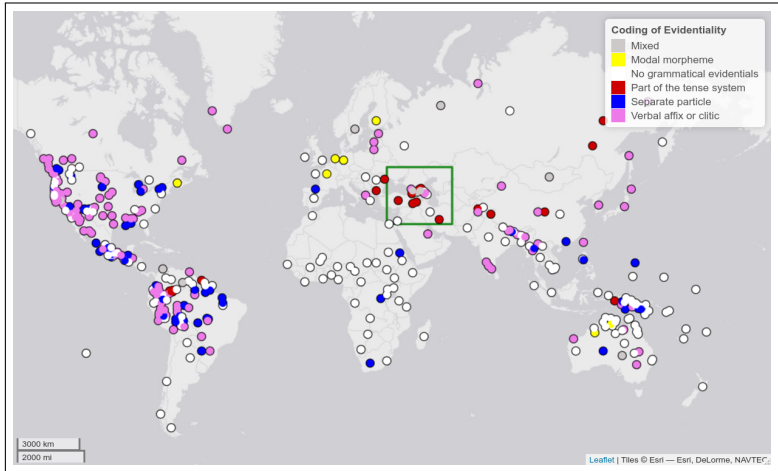▶ Several evidential areas worldwide (Plungian 2010: 19–21)

What is evidentiality?

- Evidential markers indicate where the speaker's information comes from (personal experience, hearsay, inference)
- "highly diffusable" (Aikhenvald 2004: 302)
- Several evidential areas worldwide (Plungian 2010: 19–21)
  - The "Evidential Belt" of Eurasia
  - Perfect - Inferential implicature - Unwitnessed past
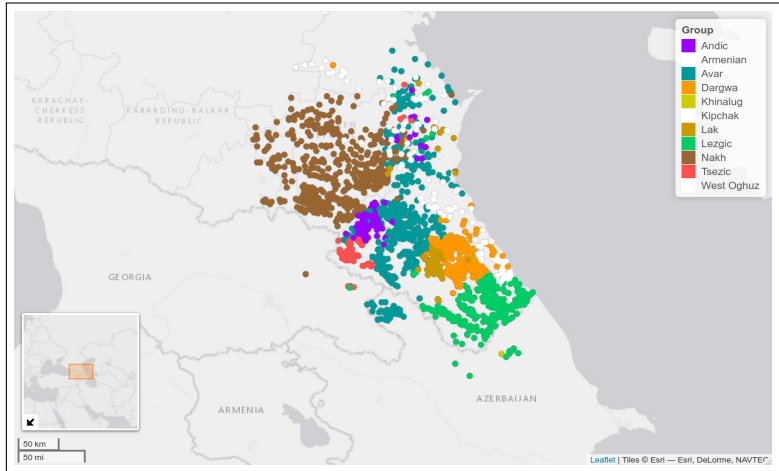 - A rather common and natural development which *can* arise language-internally

What is evidentiality?

- Evidential markers indicate where the speaker's information comes from (personal experience, hearsay, inference)
- "highly diffusable" (Aikhenvald 2004: 302)
- Several evidential areas worldwide (Plungian 2010: 19–21)
  - The "Evidential Belt" of Eurasia
  - Perfect - Inferential implicature - Unwitnessed past
  - A rather common and natural development which *can* arise language-internally
  - ... but is suspiciously frequent in a specific area.

Evidential coding in WALS (de Haan 2013)

Coding of Evidentiality
- Mixed
- Modal morpheme
- No grammatical evidentials
- Part of the tense system
- Separate particle
- Verbal affix or clitic

The eastern Caucasus

Aim of this talk

How much detail does areal typology require
to paint an accurate picture?

Three types of data:

▶ Subset of a large-scale typological study (WALS maps on evidentiality)

▶ A more detailed study of the eastern Caucasus (my dissertation research)

▶ A dataset of all East Caucasian villages and the languages spoken there

Introducing the data

The WALS maps

▶ Sampling philosophy: maximizing genealogical and areal diversity (Haspelmath et al. 2005)

▶ Balanced core sample of 100 languages (handpicked, motivating parameters largely unspecified), 100 more "recommended" languages

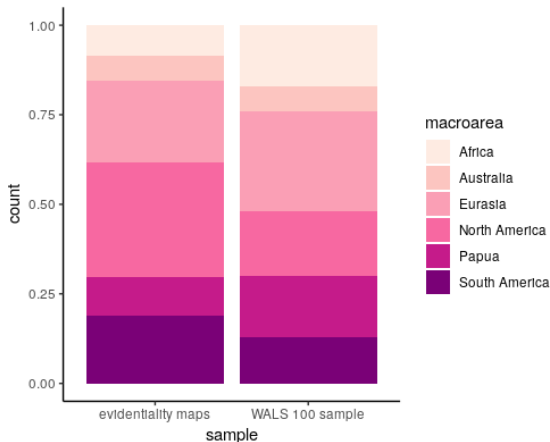▶ Authors of individual chapters could expand on the core sample

Introducing the data

The WALS maps

▶ Maps on evidentiality feature 418 languages, of which 192 core (97 from the core 100 list)

▶ Based on descriptive grammars → bias of available literature

Introducing the data

Figure 1: Proportion of languages by macroarea

Introducing the data

The WALS maps

- ▶ Coverage in the eastern Caucasus:
  10 East Caucasian languages
- ▶ Two maps on evidentiality: "Coding of
  evidentiality" and "Semantic distinctions of
  evidentiality" → Coding

Introducing the data

My data

- ▶ Descriptive grammars, papers on evidentiality, some fieldwork
- ▶ At least one datapoint for each recognized language, data on dialects if available (maximize coverage rather than linguistic diversity)
- ▶ 29 East Caucasian languages, 3 Turkic languages (no Tat, Armenian)
- ▶ Data on additional dialects for specific features
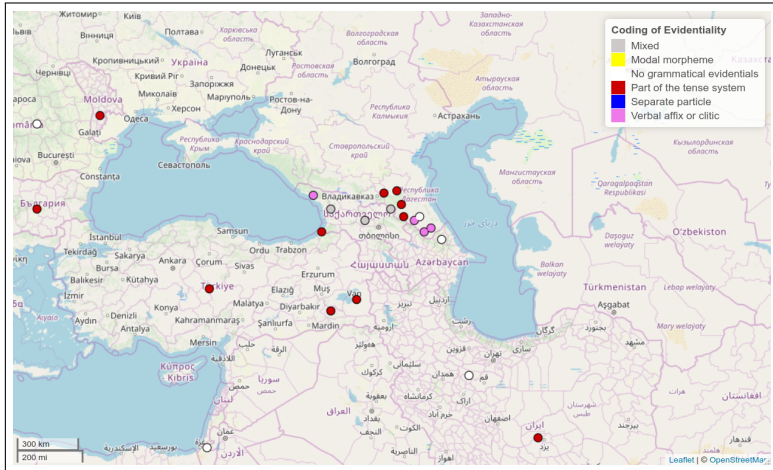
Introducing the data

East Caucasian villages

► Datapoint for each village where an East Caucasian language is spoken, and for villages where a non-East Caucasian language is spoken in the republics of Dagestan, Chechnya and Ingushetia

Introducing the data

East Caucasian villages

► Datapoint for each village where an East Caucasian language is spoken, and for villages where a non-East Caucasian language is spoken in the republics of Dagestan, Chechnya and Ingushetia

► Mined from Wikipedia by George Moroz and students of the School of Linguistics at NRU HSE Moscow (maintained by me and available at: github/sverhees/master_villages)
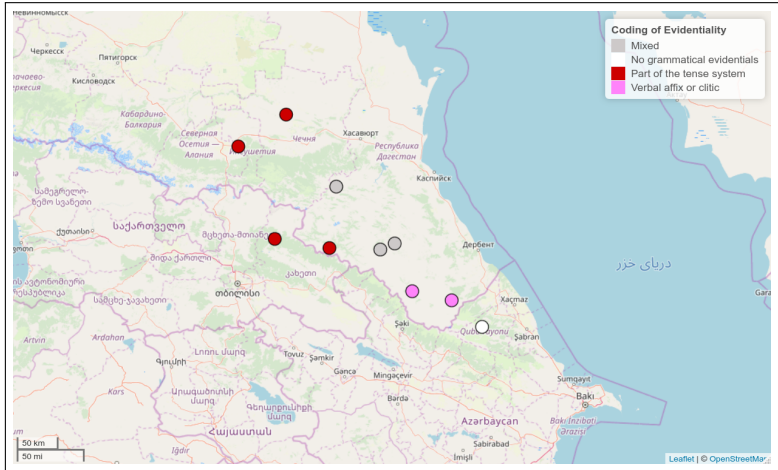
Introducing the data

East Caucasian villages: shortcomings

► Still incomplete

► May still contain errors

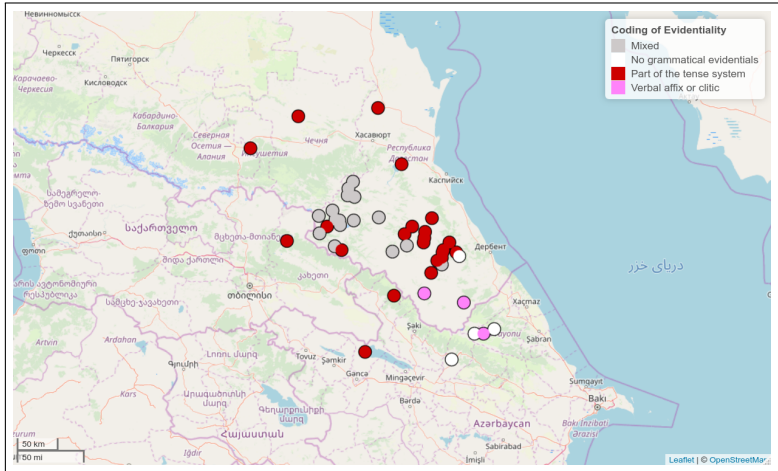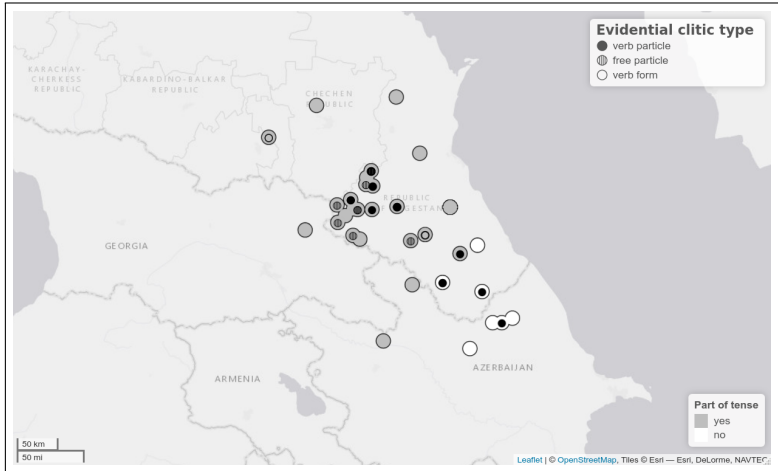► Multi-ethnic villages, towns and cities not included

Results

Results

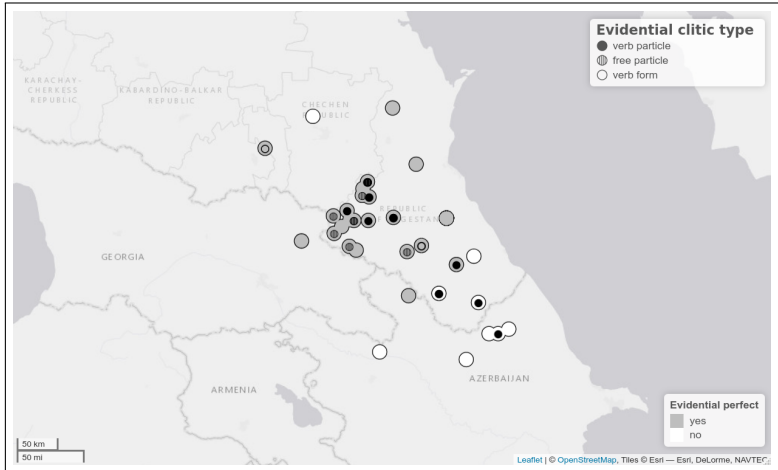
With some corrections, see issue.
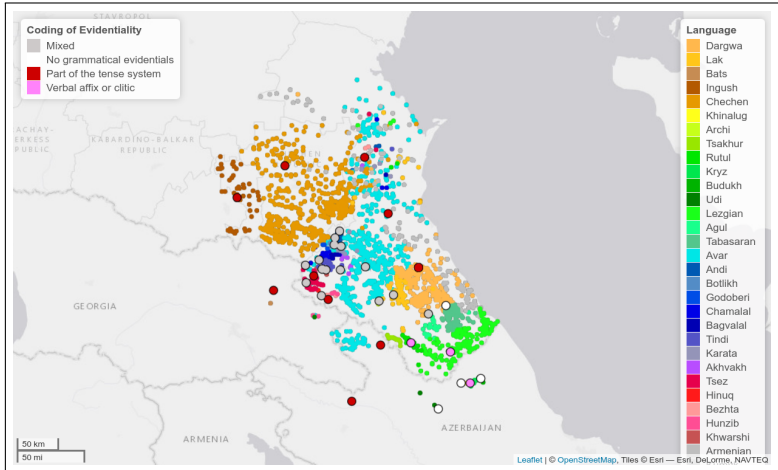
Results

Results

Results

Intermediate summary

► Both samples show a similar picture for the eastern Caucasus: evidentiality as part of the tense system is very widespread, but notably absent in the southeast

► The pattern seems to comprise both East Caucasian and Turkic languages, but further research is necessary

► The larger sample also reveals a mixed transitional zone that is still waiting for an explanation

Typology vs. reality

Discussion

- George Moroz argues that sampling is useless for linguistics, due to various biases (including descriptive bias)
- It is better to take all the available data for a specific family
- My data show that typological samples do not necessarily perform worse in revealing areal patterns

Discussion

But...

- ▶ That is all they can do

Discussion

But...

- ▶ That is all they can do
- ▶ Specialist knowledge is necessary to:
    - ▶ Corroborate their findings
    - ▶ Refine their patterns
    - ▶ Ultimately *explain* where the observed patterns come from

Discussion

So...

- ▶ Typological samples are fine for exploratory analysis
- ▶ Specialists should aim to falsify such studies and elaborate their observations

Discussion

So...

- ▶ Typological samples are fine for exploratory analysis
- ▶ Specialists should aim to falsify such studies and elaborate their observations
- ▶ ... Rather than sulk about the inaccuracies of typological studies in private, and aiming to carry out a "better" study independently
- ▶ Also, they should perspectivize their own findings by contrasting data used with data lacking

References I

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.

de Haan, Ferdinand. 2013. Coding of evidentiality. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/78.

References II

Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie. 2005. Introduction. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 1–9. Oxford: Oxford University Press.

Plungian, Vladimir A. 2010. Types of verbal evidentiality marking: an overview. In Gabriele Diewald & Elena Smirnova (eds.), *Linguistic realization of evidentiality in european languages*, 15–58. Berlin/New York: De Gruyter Mouton.