Language change and contact in southern Finnic

An areal typological study

MIINA NORVIK, ULDIS BALODIS, HELLE METSLANG,
KARL PAJUSALU, EVA SAAR, PIRE TERAS, JANEK VAAB

AREAL LINGUISTICS AND LANGUAGE CONTACT OF URALIC LANGUAGES

JUNE 12–14, 2025. RIGA

This work was supported by the Estonian Research Council grant (PRG2184) – "From East to West: Typological shift in Estonian and the Southern Finnic languages against the background of Uralic".

Outline

- Finnic among the languages of the Baltic Sea area
- Aims
- Datasets & Method
- Results
  - Phonology
  - Morphosyntax
  - Phonology + Morphosyntax
- Discussion & Conclusions

Finnic among the languages of the Baltic Sea area

- Longterm contacts between Finnic, Baltic, Slavic, and Germanic languages
- Common features have spread inconsistently
- A contact area – not a proper Sprachbund (Koptjevskaja-Tamm & Wälchli 2001)
- Southern Finnic languages have contacts with unrelated (Indo-European) and related languages
- Language history and language contacts, genealogical and areal factors intertwine
- Study of the large Grambank language dataset concludes that genealogical factors have a greater influence on languages than areal factors (Skirgård et al. 2023)

Some common typological features

Finnic languages (Grünthal 2022, Laakso 2022)

- main word stress on the 1st syllable
- complex quantity oppositions
- agglutination
- stem alternations
- many cases
- SVO word order
- possessive constructions not based on *habeo*-verbs

The languages of the **Baltic Sea area** (Koptjevskaja-Tamm & Wälchli 2001):

- main word stress on the 1st syllable
- SVO word order
- special verb forms for reported evidentiality
- possessive constructions not based on *habeo*-verbs
- an interrogative particle in polar questions

Aims

To make use of phonological and morphosyntactic datasets coded for Finnic languages/varieties in order to discuss the following questions:

- What do the analyses of phonological and morphosyntactic features say about the (dis)similarities among Finnic languages/varieties?
- To what extent are the analyses of these two domains similar/different?
- How is continuity and change reflected in the figures?
- How to disentangle the role of areal and genealogical relationship and what are the results of contacts between cognate and non-cognate contact languages?

An additional aim is to discuss the advantages and pitfalls of datasets.

Datasets & Method

UraTyp (see UraTyp)

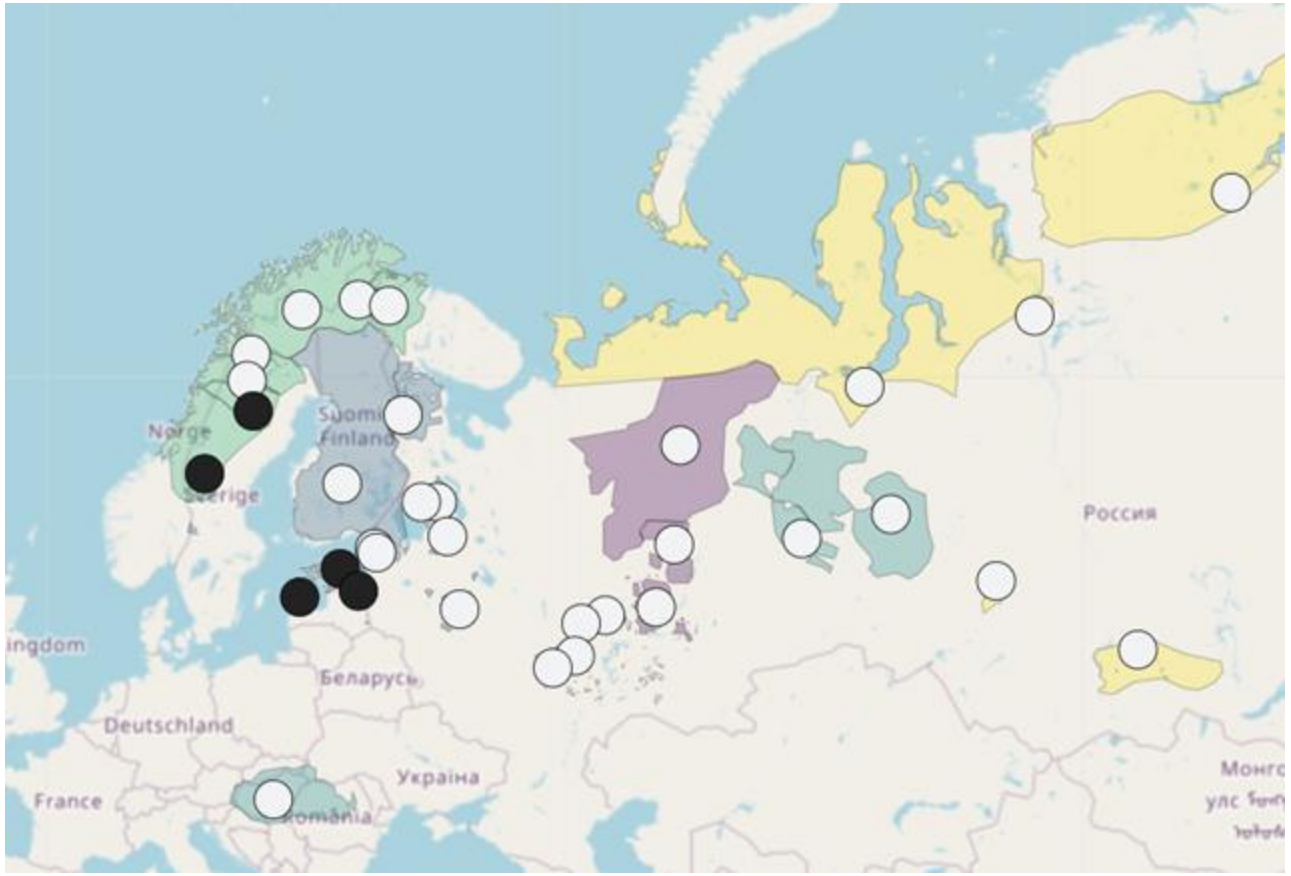
- 360 features, incl. 51 phonological features and 309 morphosyntactic features (195 originate from Grambank, i.e., GB features; the rest are called UT features)
- 35 languages, incl. 11 Finnic languages
- values 1 ‘yes’ vs. 0 ‘no’ (+ a few multi-state features)

EstTyp

- A selection of GB and UT features + additional FT features = 58 phonological features and 47 morphosyntactic features
- 38 Finnic languages/varieties (incl. 11 Finnic languages already in UraTyp)
- values 1 ‘yes’, 0 ‘no’, 0.5 ‘limited occurrence’

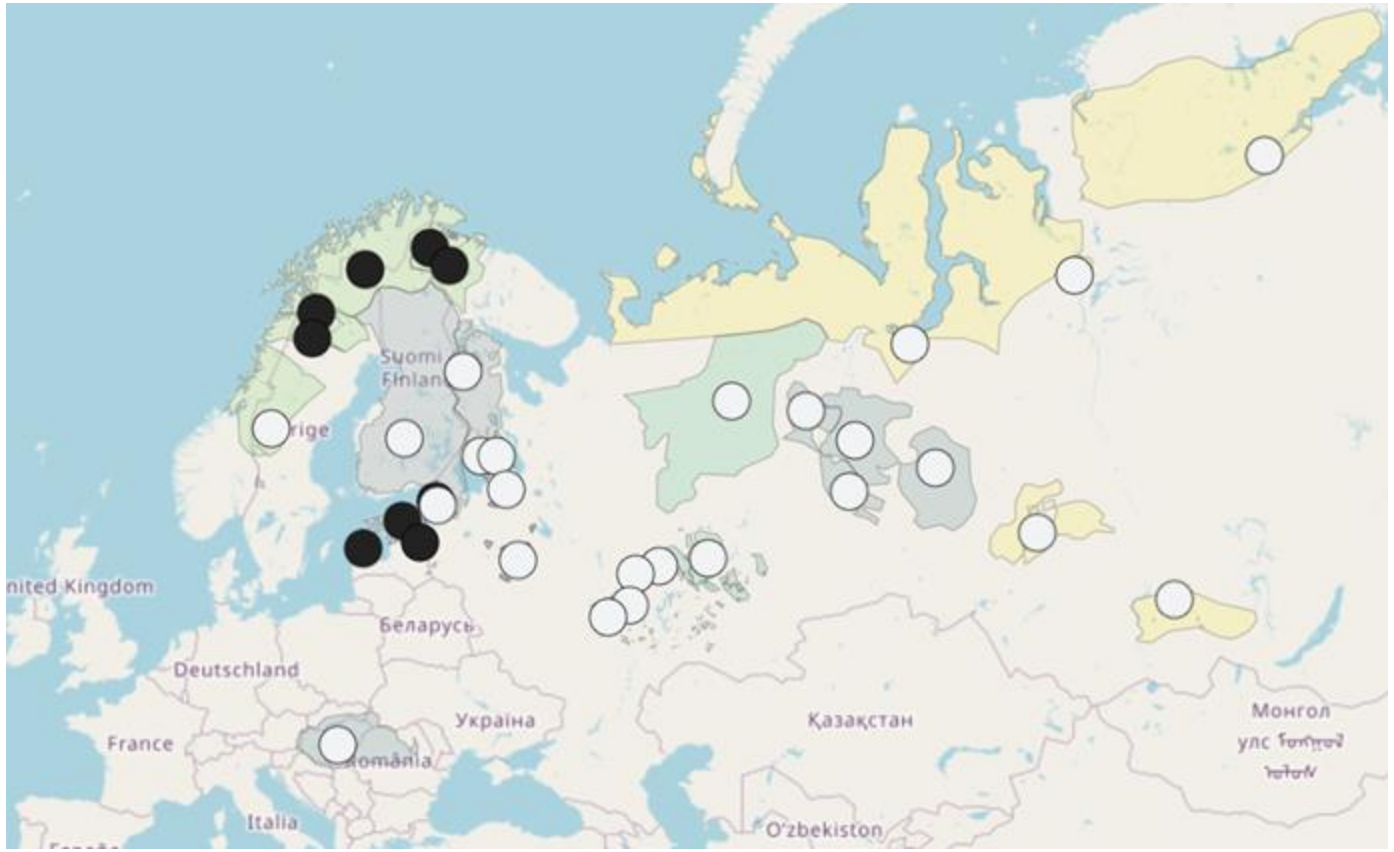
Examples of features

GB262: Is there a clause-initial polar interrogative particle?



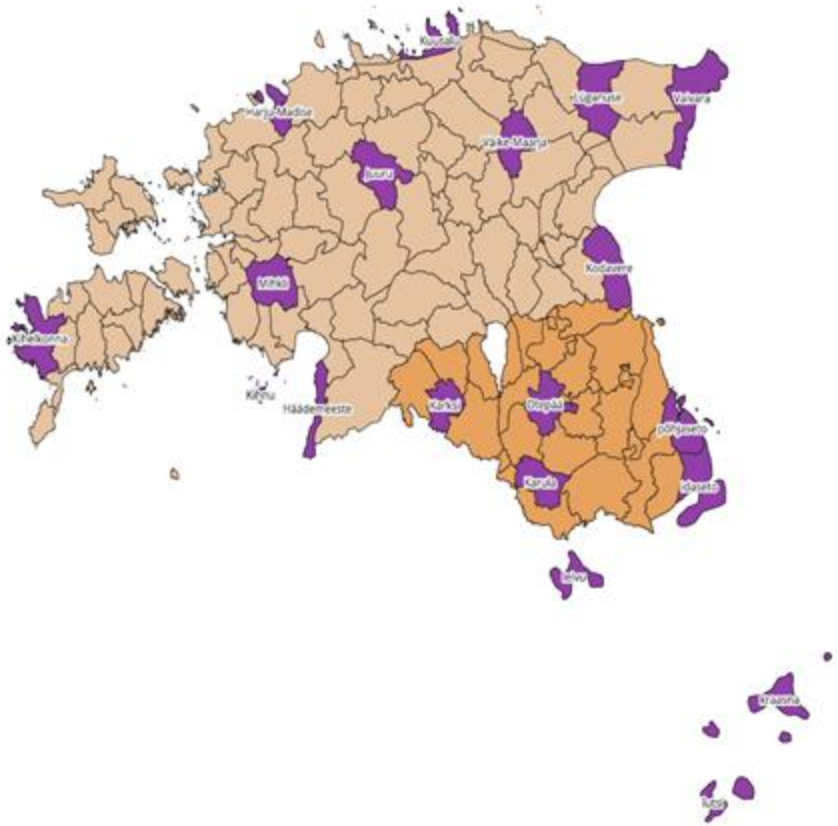
	Value
○	no
●	yes

UT120: Are there more than two phonological quantities?



	Value
○	no
●	yes

Languages/varieties in EstTyp



= 38 languages/ varieties

EstTyp datasets

Dataset I. Phonology – a selection of UT features + additional features

Final dataset: 51 UT features and 7 additional features

41 features show variation

38 Finnic languages/varieties

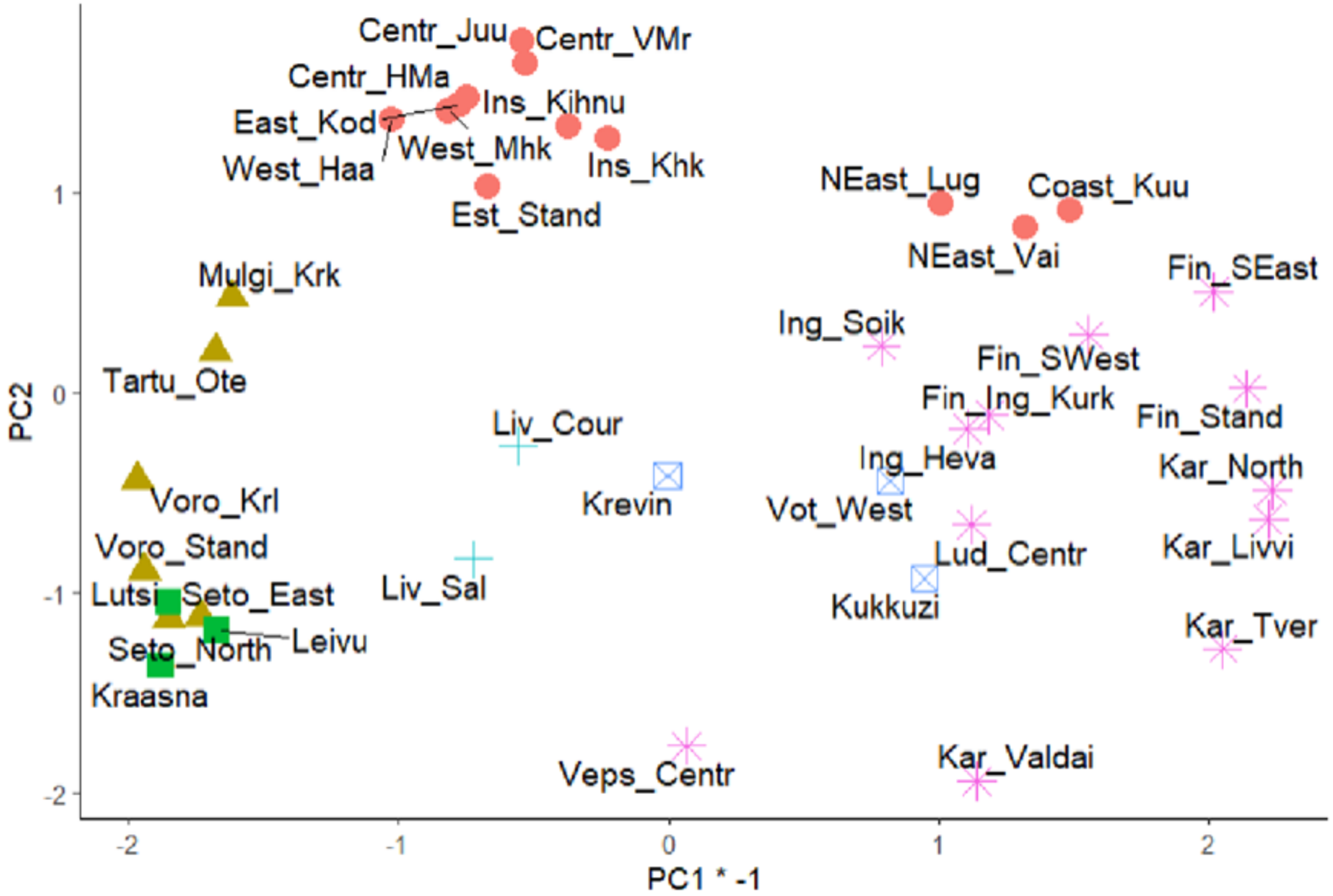
Dataset II. Morphosyntax – a selection of GB and UT features + additional features

Final dataset: 20 GB features, 25 UT features, and 2 additional features

all 47 features show variation

35 Finnic languages/varieties (Krevin, Kukkuzi, Kurkola excluded – not enough data)

Results: Phonology



Finnic phonology:

41 features, 38 varieties;

PC1 – 32%, PC2 – 18%, PC3 – 11%

Group

- I Nth Estonian varieties
- II Sth Estonian varieties
- III SthEst language islands
- IV Livonian
- V Votic
- VI Nth Finnic

Comments on the groupings

- Northeastern Estonian group together with the Northern Finnic varieties (PC1) but they also have several similarities to North Estonian dialects (PC2)
- Western Votic and Kukkusi varieties are close to the Northern Finnic varieties, however, Krevin is more similar to southern Finnic
- Krevin is somewhat closer to Livonian than to Votic ← hundreds of years of separation from the other Votic varieties; contact with Latvian.
- Mulgi and Tartu dialects differ from the rest of the South Estonian dialects and South Estonian language islands.
- Veps differs from the rest of the Northern Finnic varieties.

6-7 most significant phonological features

DIM1:

FT171 – Is there syncope?

UT145 – Are there word-final clusters of three or more consonants?

UT120 – Are there more than two phonological quantities?

FT174 – Is there quantitative alternation between long and overlong vowels?

FT170 – Is there apocope?

UT119 – Is there a contrast in the length of geminate consonants?

DIM2:

UT141 – Are there word-initial consonant clusters?

UT155 – Are there voiced plosives (e.g., /b g d/)?

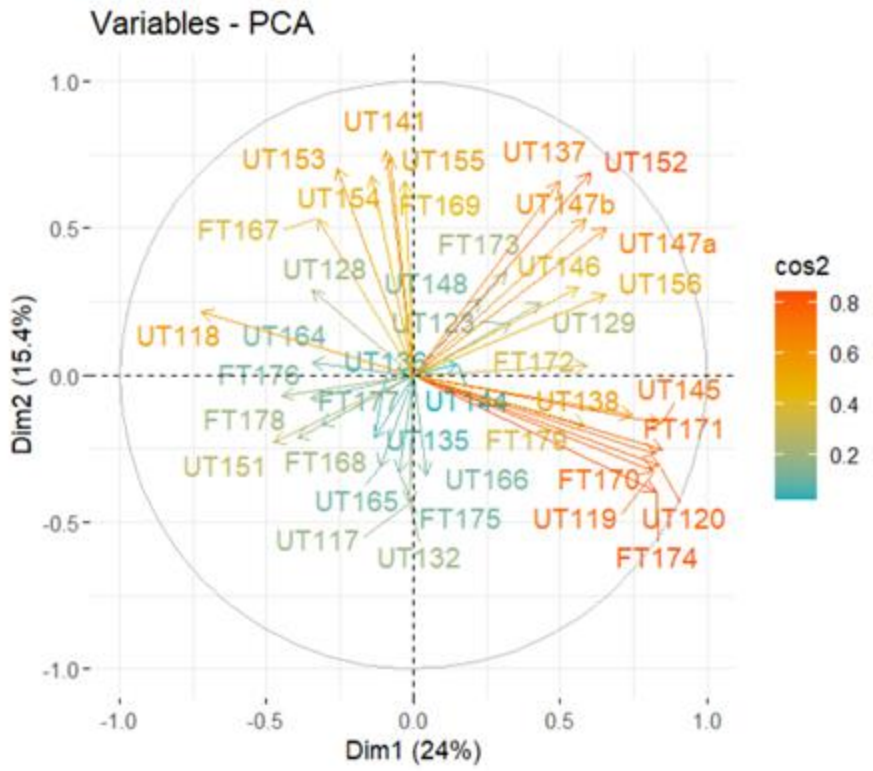
UT153 – Is there a non-alveolar affricate?

UT152 – Is there an alveolar affricate?

UT154 – Is there a phonological contrast between an unvoiced alveolar sibilant /s/ and an unvoiced postalveolar sibilant /ʃ/?

FT169 – Is there word-final laryngeal fricative ([h])?

UT137 – Is there a high back unrounded vowel /ɯ ɨ/?



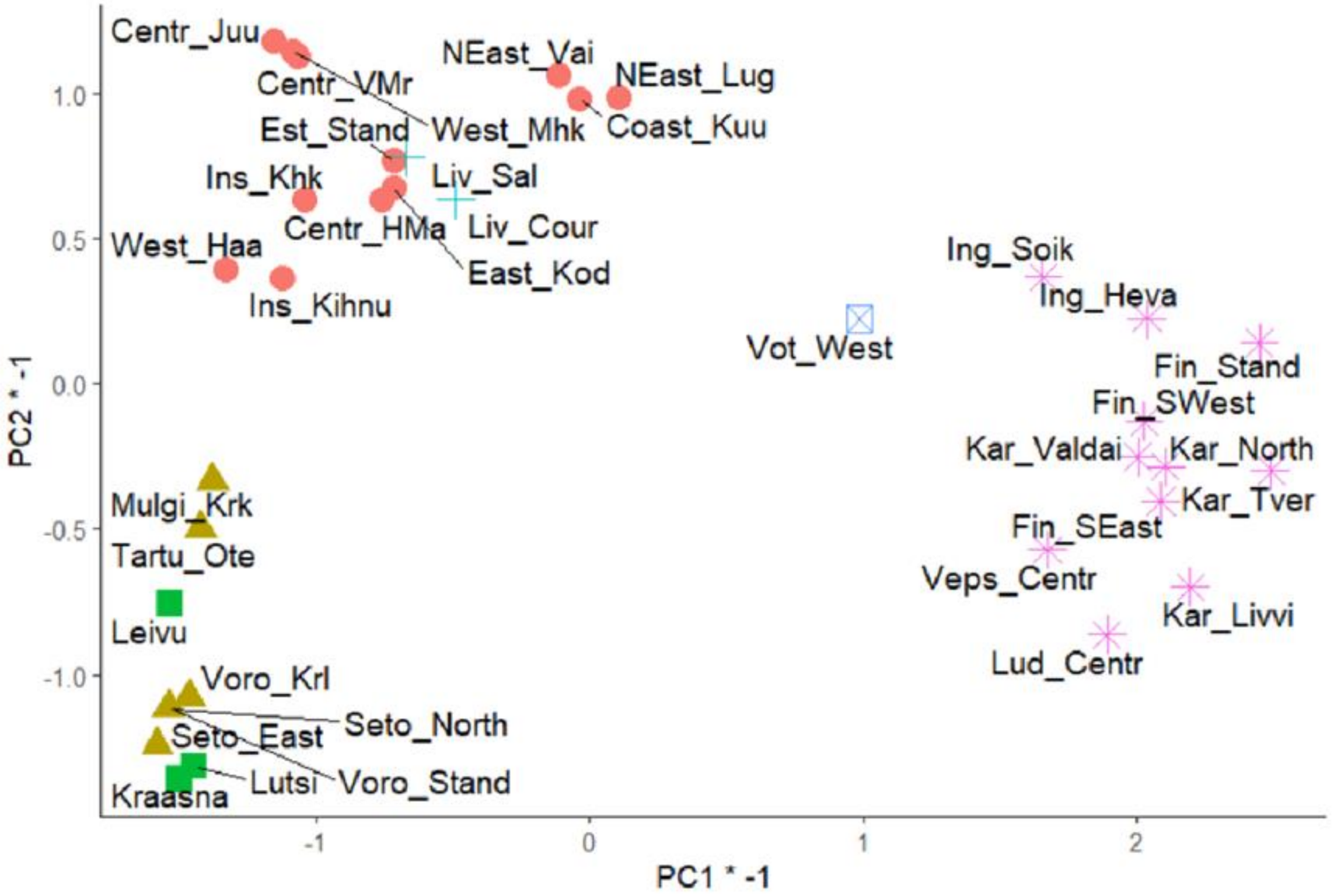
Zooming in on Veps (Central)

	Veps	Stand_Est
DIM1:		
FT171 – Is there syncope?	1 ‘yes’	1 ‘yes’
UT145 – Are there word-final clusters of three or more consonants?	1 ‘yes’	1 ‘yes’
UT120 – Are there more than two phonological quantities?	0 ‘no’	1 ‘yes’
FT174 – Is there quantitative alternation between long and overlong vowels?	0 ‘no’	1 ‘yes’
FT170 – Is there apocope?	1 ‘yes’	1 ‘yes’
UT119 – Is there a contrast in the length of geminate consonants?	0 ‘no’	1 ‘yes’
DIM2:		
UT141 – Are there word-initial consonant clusters?	1 ‘yes’	1 ‘yes’
UT155 – Are there voiced plosives (e.g., /b g d/)?	1 ‘yes’	0 ‘no’
UT153 – Is there a non-alveolar affricate?	1 ‘yes’	0 ‘no’
UT152 – Is there an alveolar affricate?	1 ‘yes’	0 ‘no’
UT154 – Is there a phonological contrast between an unvoiced alveolar sibilant /s/ and an unvoiced postalveolar sibilant /ʃ/?	1 ‘yes’	0.5 ‘limited’
FT169 – Is there word-final laryngeal fricative ([h])?	1 ‘yes’	0.5 ‘limited’
UT137 – Is there a high back unrounded vowel /ɯ ɨ/?	0.5 ‘limited’	0 ‘no’

Interpreting the results of the phonological analysis

- Based on the differences in the phonological system, the Finnic language area is first divided into southern and northern groups (DIM1), then into western and eastern groups (DIM2).
- The most important phonological features that distinguish northern and southern Finnic varieties are related to word prosody (most features for PC1);
  northern Finnic languages have simpler syllable structure (cf. UT145),
  southern Finnic languages have more complex syllable structure (FT171, FT174) and more than two quantity degrees (UT120)
- The features that distinguish western and eastern Finnic languages are mostly segmental (see features for DIM2)
- Innovative phonological features are consistently represented in southwestern Finnic varieties, e.g., the loss of *h*, the presence of mid vowels, schwa, etc.
- Livonian is a part of the southern Finnic dialect continuum, on the other hand, it has certain archaic features due to its position on the periphery of the language area, such as *d* remaining in the weak grade (e.g., *kädūd* ‘hands’).

Results: Morphosyntax



Finnic morphosyntax:
47 features, 35 varieties;
PC1 – 44%, PC2 – 12%,
PC3 – 10 %

Group

- I Nth Estonian varieties
- II Sth Estonian varieties
- III SthEst language islands
- IV Livonian
- V Votic
- VI Nth Finnic

Comments on the groupings

- The studied languages/varieties fall into three fairly distinct groups.
- In terms of morphosyntax, South and North Estonian are more similar to each other than they are in terms of phonology.
- The Northeastern Estonian dialects (Vaivara, Lüganuse, Kuusalu), as well as Livonian, are similar to the North Estonian dialects and Standard Estonian.
- Compared to phonology, the northern group of Finnic is more cohesive.
- Western Votic falls between the northern Finnic languages and North Estonian.

The 10 most significant features of DIM-1

Typical for Northern Finnic:

GB264 – neither a clause-initial nor clause-final polar interrogative particle (*kO*-type)

GB432 – suffix on the possessor (genitive ending)

UT038 – impersonal forms in the personal paradigm

GB298 – inflectional negative marker

GB433 – suffix on the possessed noun (possessive suffix)

UT078 – additional cases with pronouns (accusative)

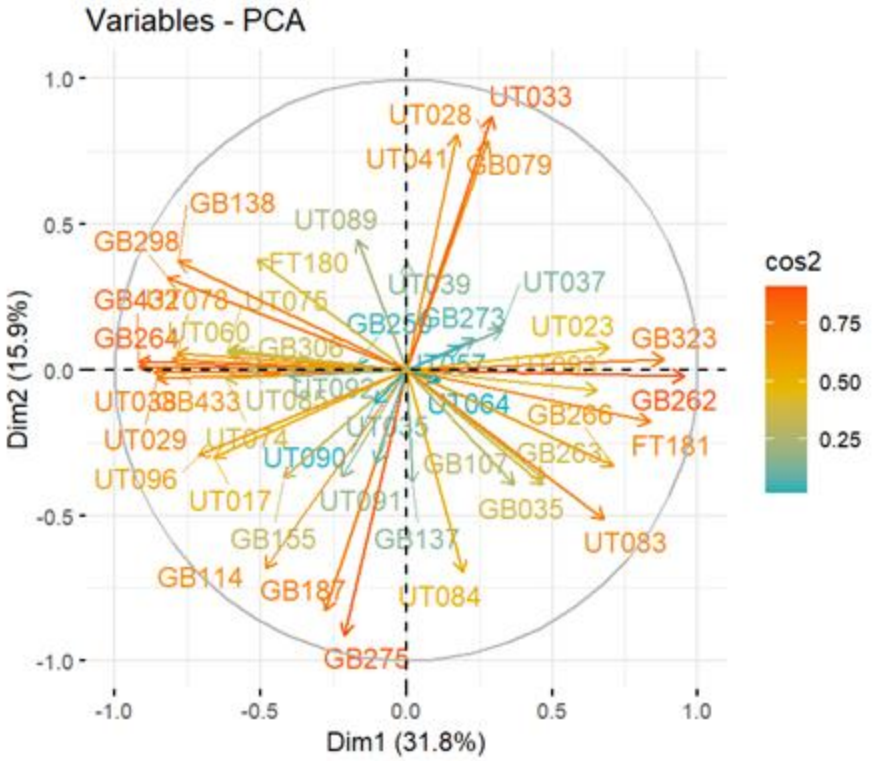
UT029 – potential mood

Typical for Southern Finnic:

GB262 – clause-initial polar interrogative particle

GB323 – grammatical marking for reported evidentiality

FT181 – neutralization of personal endings



Zooming in on Votic (Western)

Typical for Northern Finnic:

	Votic
GB264 – *kO*-type interrogative particle	1 ‘yes’
GB432 – suffix on the possessor	0 ‘no’
UT038 – impersonal forms in the personal paradigm	1 ‘yes’
GB298 – inflectional negative marker	1 ‘yes’
GB433 – suffix on the possessed noun	0 ‘no’
UT078 – additional cases with pronouns (accusative)	0.5 ‘limited’
UT029 – potential mood	0 ‘no’

- *Cf. Standard Estonian gets 0 for all the features*

GB264:

*lugõt**ko** jõka päiv kazettia* (J, VKS)

‘Do you read newspapers every day?’

Typical for Southern Finnic:

GB262 – clause-initial polar interrogative particle	0.5 ‘limited’
GB323 – grammatical marking for reported evidentiality	0 ‘no’
FT181 – neutralization of personal endings	0 ‘no’

- *Cf. Standard Estonian gets 1 for all the features*

GB262:

***vai** tänävä on saunapäivä?* (Li, VKS)

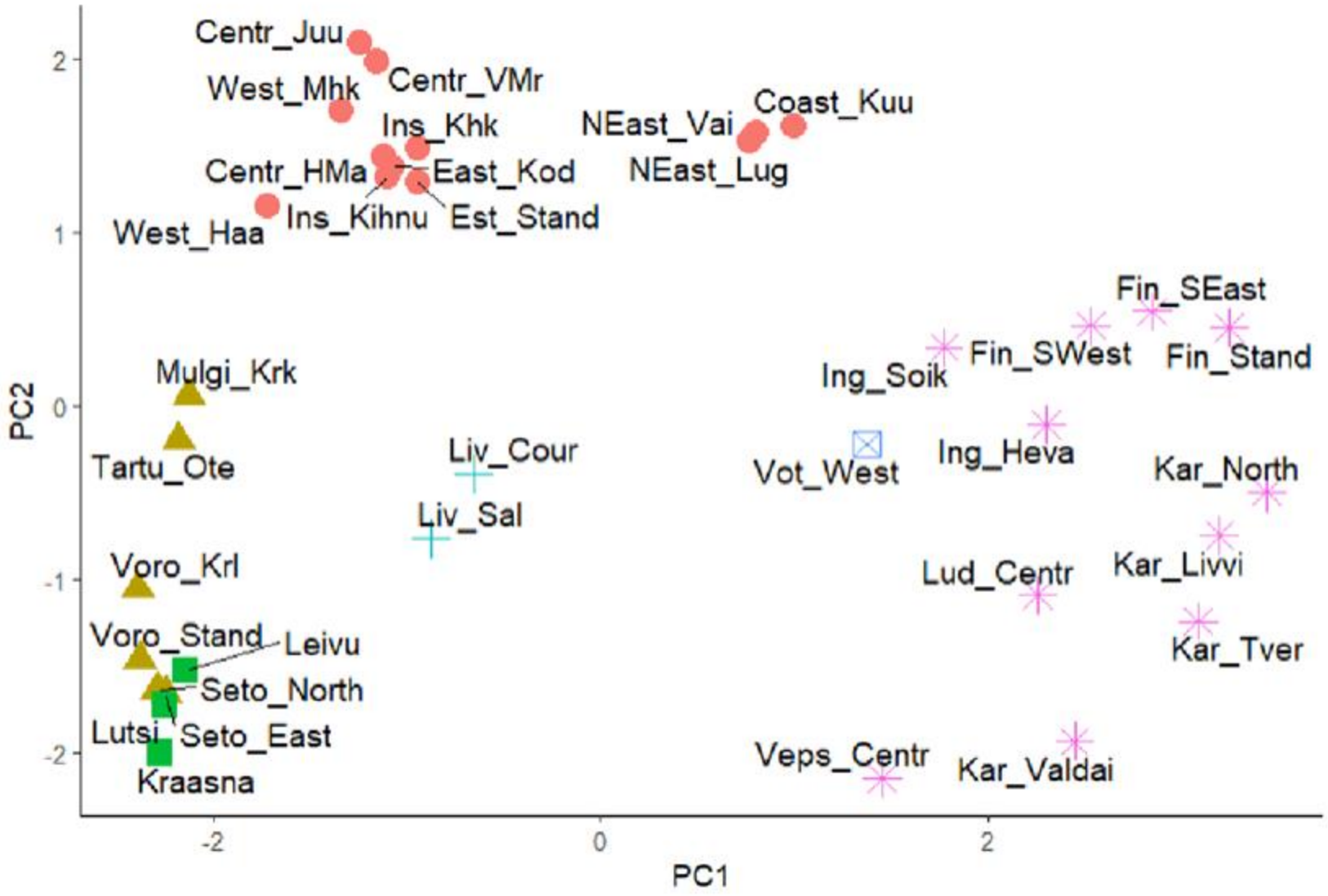
'Is it sauna day today?'

dark red = deviation from the group

Interpreting the results of the morphosyntactic analysis

- Features that are correlated with **PC1** show a distinction between **northern and southern Finnic** languages; however, with Votic quite in the middle and northeastern Estonian varieties closest to Votic and northern Finnic.
- The **northern** Finnic languages have preserved several **Proto-Finnic features**, e.g., a sentence-internal question particle, accusative case on pronouns, genitive *-n*, possessive suffixes, which have disappeared from southern Finnic.
- The **southern** Finnic languages share several morphosyntactic **innovations**, e.g., formation of polar **questions** by particles. Both Courland and Salaca Livonian are typical innovative southern Finnic languages.

Results: Phonology + Morphosyntax



Finnic phonology + morphosyntax:

88 features, 35 varieties

PC1 – 37%, PC2 – 16%, PC3 – 10 %

Group

- I — Nth Estonian varieties
- II — Sth Estonian varieties
- III — SthEst language islands
- IV — Livonian
- V — Votic
- VI — Nth Finnic

Comments on the groupings & Comparison of the separate analyses of phonology and morphosyntax

- Livonian emerges as a separate grouping.

  Cf. the morphosyntactic analysis where Livonian falls in the same group with North Estonian varieties, while phonological features take them further apart.

- Northeastern Estonian varieties and Votic fall on the border between what are traditionally called the southern and northern Finnic languages.

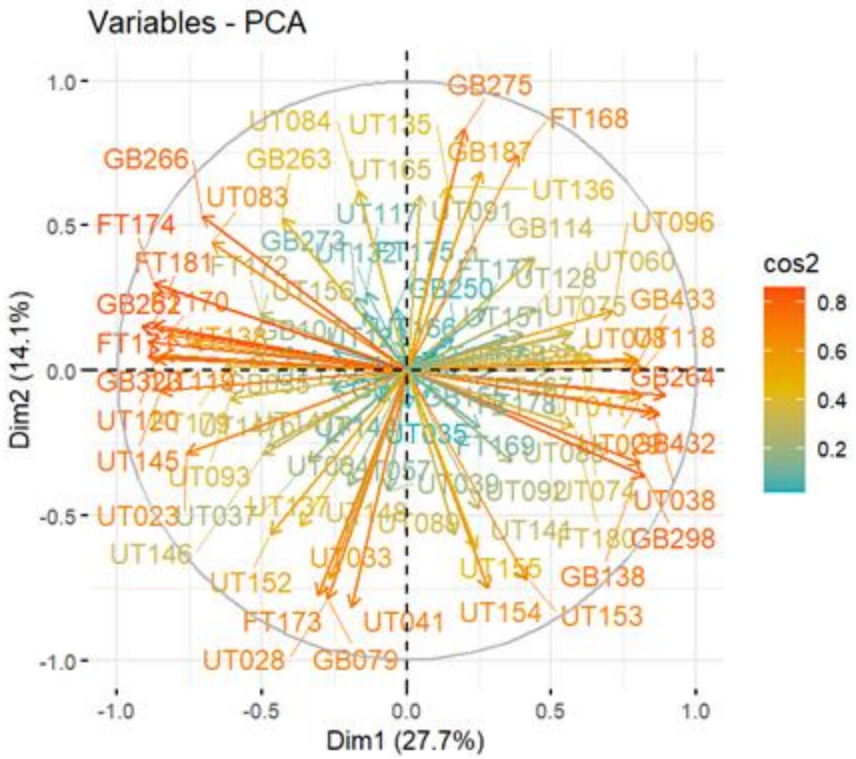
  Cf. the morphosyntactic analysis where Northeastern Estonian dialects are closer to North Estonian varieties, while phonological features place them in the northern Finnic group.

  → Phonological and morphosyntactic features are responsible for somewhat different groupings

The most significant features of DIM1 and DIM2

DIM1
- GB262 – clause-initial question particle
- GB323 – grammatical marking of indirect evidence
- GB264 – *ko*-type question marker
- FT171 – syncope
- FT181 – neutralization of personal endings
- GB432 – suffix on the possessor
- FT174 – quantitative alternation between long and overlong vowels
- UT120 – more than two phonological quantities
- UT145 – word-final clusters of three or more consonants
- UT038 – impersonal instead of a finite form

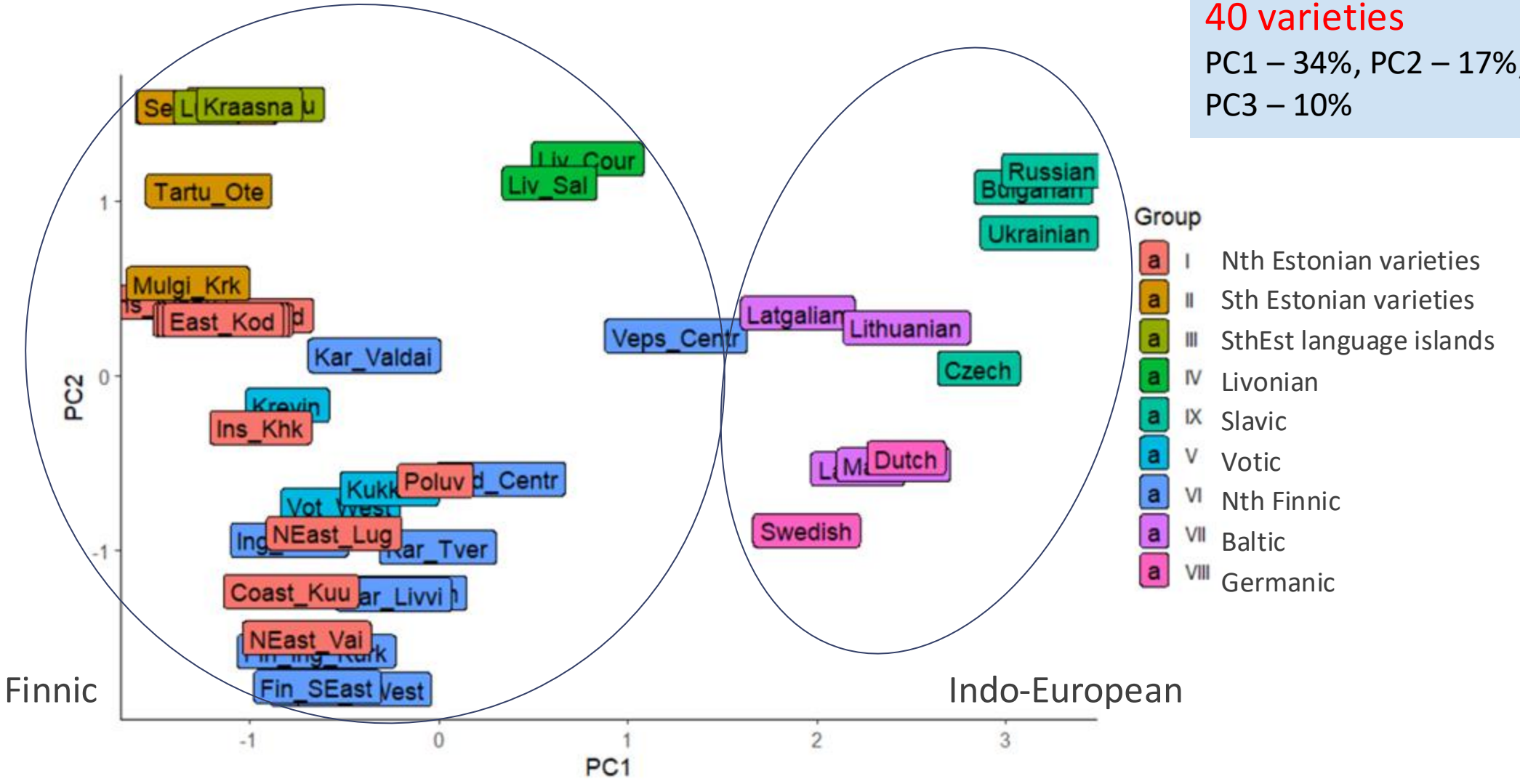
DIM2
- GB275 – bound comparative degree marker
- UT041 – 3Sg for used for impersonal reference
- GB079 – verbal prefixes
- UT028 – grammatical marking of deontic necessity



dark red = morphosyntactic features

A pilot study comparing phonological features of Finnic and Indo-European

49 phonological features,
40 varieties
PC1 – 34%, PC2 – 17%,
PC3 – 10%



Discussion & Conclusions (1)

- Morphosyntactic features group Finnic languages/varieties more strongly, phonological features more diffusely, but on both levels as a whole there is a fairly clear division into **northern and southern Finnic**, followed by North and South Estonian varieties and Livonian.
- With respect to phonology, the most important differentiators are **innovative** features, such as the ternary quantity alternation, syncope, and consonant clusters, rather than conservative features, such as vowel harmony.
- **Livonian** also has some phonological similarities with the eastern Finnic languages – Ingrian, Votic, Veps. These may result from a certain conservatism of peripheral areas or from a partial simplification of the pronunciation system under the influence of contact languages.

Discussion & Conclusions (2)

- The innovative features of southern Finnic morphosyntax are also largely related to **innovations in the Central Baltic** language area, such as grammatical marking of the evidentiality, the marking of polar questions with interrogative particles, etc.
- In the southern Finnic dialect continuum, **Livonian** is morphosyntactically close to the North Estonian varieties and phonologically somewhat close to the to western dialects of South Estonian; however, overall it differs significantly from other Finnic varieties.
- The genealogical origins of a language seem to be a stronger determinant of the language system than areality as seen in the comparison between Finnic and neighbouring Indo-European languages.

References

Grünthal, Riho 2022. The Finnic languages. - Abondolo, Daniel, Riitta-Liisa Välijärvi (toim.), The Uralic languages. Routledge, 433-480.

Laakso, Johanna 2022. Finnic: General introduction. – Bakró-Nagy, Marianne, Johanna Laakso, Elena Skribnik (toim.), The Oxford Guide to the Uralic languages. Oxford University Press, 240-253.

Norvik, Miina; Balodis, Uldis; Ernštreits, Valts; Kļava, Gunta; Metslang, Helle; Pajusalu, Karl; Saar, Eva 2021. The South Estonian language islands in the context of the Central Baltic area. Eesti ja soome-ugri keeleteaduse ajakiri = Journal of Estonian and Finno-Ugric Linguistics, 12 (2), 33–72.

Norvik, Miina; Jing, Yingqi; Dunn, Michael; Forkel, Robert; Honkola, Terhi; Klumpp, Gerson; Kowalik, Richard; Metslang, Helle; Pajusalu, Karl; Piha, Minerva; Saar, Eva, Saarinen, Sirkka, Vesakoski, Outi 2022. Uralic typology in the light of a new comprehensive dataset. Journal of Uralic Linguistics, 1 (1), 4–42. DOI: 10.1075/jul.00002.nor.

Pajusalu, Karl; Kristel Uiboaed; Péter Pomozi, Endre Németh, Tibor Fehér 2018. Towards a phonological typology of Uralic languages. – Eesti ja soome-ugri keeleteaduse ajakiri = Journal of Estonian and Finno-Ugric Linguistics, 9 (1), 187–207.

Skirgård, Hedvig; Haynie, Hannah J.; Blasi, Damián E.; Hammarström, Harald; Collins, Jeremy; Latarche, Jay J.; Lesage, Jakob; Weber, Tobias; Witzlack-Makarevich, Alena; Passmore, Sam; Chira, Angela; Maurits, Luke; Dinnage, Russell; Dunn, Michael; Reesink, Ger; Singer, Ruth; Bowern, Claire; Epps, Patience; Hill, Jane; Vesakoski, Outi ... Gray, Russell D. (2023). Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss. Science Advances, 9 (16), 1–15. DOI: 10.1126/sciadv.adg6175.