

Mechanisms behind standardization of written Estonian 1880-1920

Peeter Tinitis
(Tallinn University/University of Tartu)

18.03.2021



HiSoN 2020/2021 ERLANGEN: INTRA-WRITER VARIATION

Intra-Writer Variation in
Historical Sociolinguistics

HiSoN 2021 Conference, March 17-19
FAU Erlangen (fully online!)

Intra-Writer Variation in Historical Sociolinguistics

Individual variation and language change

- Style, register, audience: a source of change
- Micro ~ macro: a window on larger patterns

In the context of standardization:

- Prescription & attention to language
- Accommodation & audience design
- Register, style, etc



**Intra-Writer Variation in
Historical Sociolinguistics**

HiSoN 2021 Conference, March 17-19
FAU Erlangen (**fully online!**)



A language in transition

Estonian language around 1800:

Spoken in the Baltic German governed area in Russian Empire

Rural, peasant language

Mostly oral and low prestige

Some writing tradition

Mostly non-native clerics

20% natives read, less write

Migration limited to local counties and parishes



Map: Estonia and Livonia in the Russian Empire 3

Estonian language 1850-1920

Freedom to relocate and move

1826, 1849/1856, 1863 tsarist reforms

Growth in literacy, schools

1800: 20% read, less write

1897: 91% read, 77% write

Urbanization

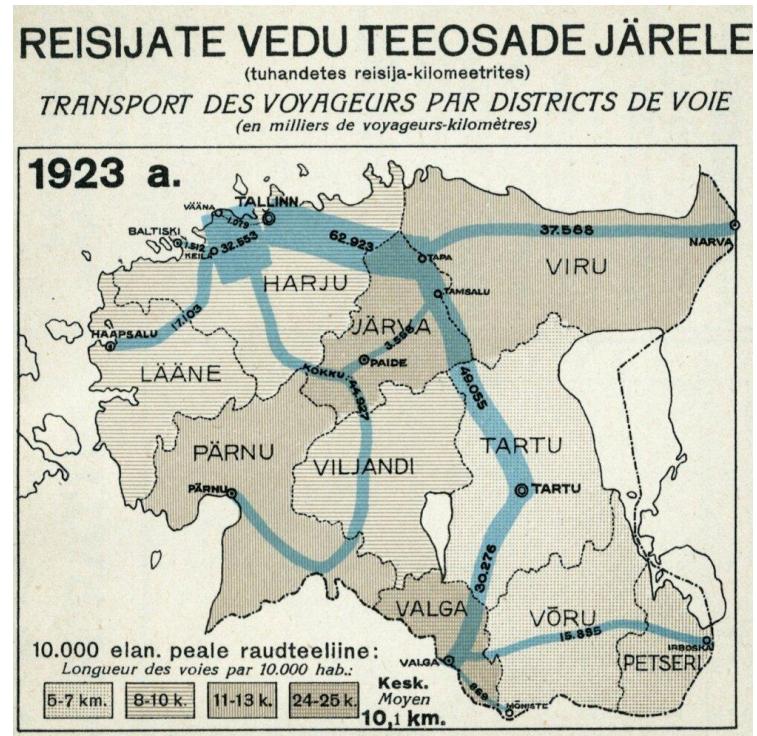
6% 1825 -> 19.2% 1897

Railroads

1870, 1876, 1889, 1896

Newspapers from 1857 onwards

Dictionaries, high literature



Map: Railroad passengers in 1923

Spoken to written language community

1850s

central figures promote „write as you speak“, but unified standard also discussed

1870s

transition from Germanic to Finnic spelling

1872, 1886, 1908-1911

dictionaries & discussions on standard

1920s

Institutionalized language policies in newly independent Estonia

1930s

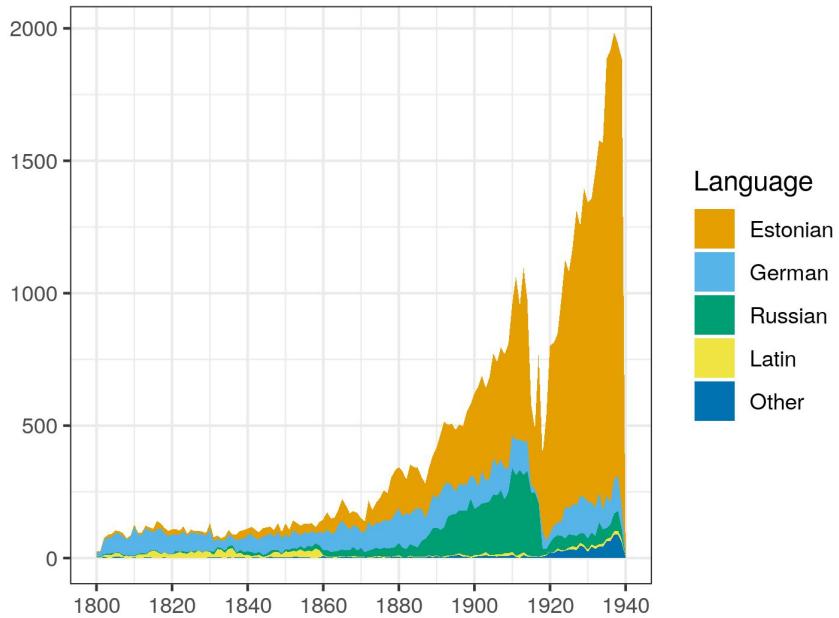
a standard language culture (Milroy 2001)



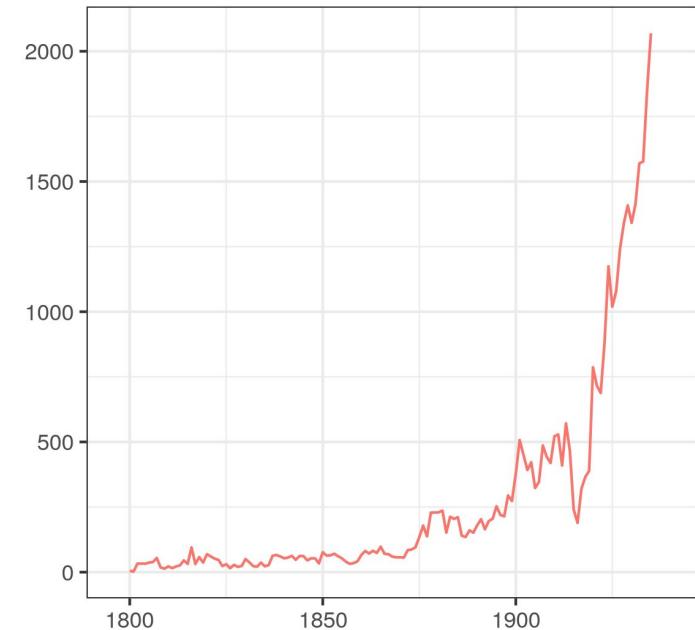
„Battles on field of language“
1914 cartoon commenting on
polemics around language

Estonian print publications

Publications in Estonian National Bibliography



Authors per 1M native speakers



Data from Estonian National Bibliography and Reconstructed Demography

Questions that arise

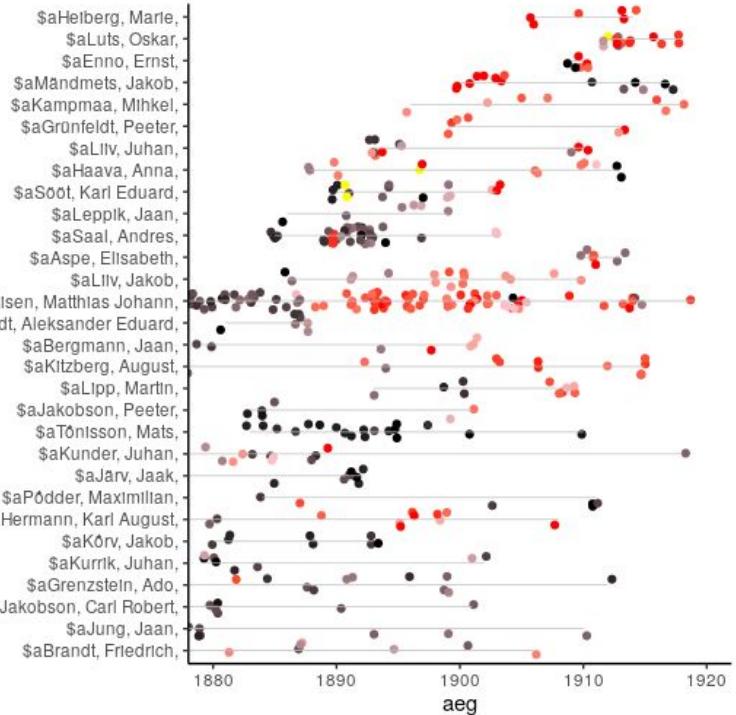
How did the language standardize?

- Standardization vs dialect mixtures? Domination of majority?
- Was the standardization guided by travel, moving homes?
 - Influence of railroads?
 - Influence of freedom of movement?
- Were there literary networks that guided the preferred forms?
- Did the transition follow generational lines?
 - Possible role of new schoolbooks.
- Were the speakers aware of an enregistered “standard”, and competing for this?

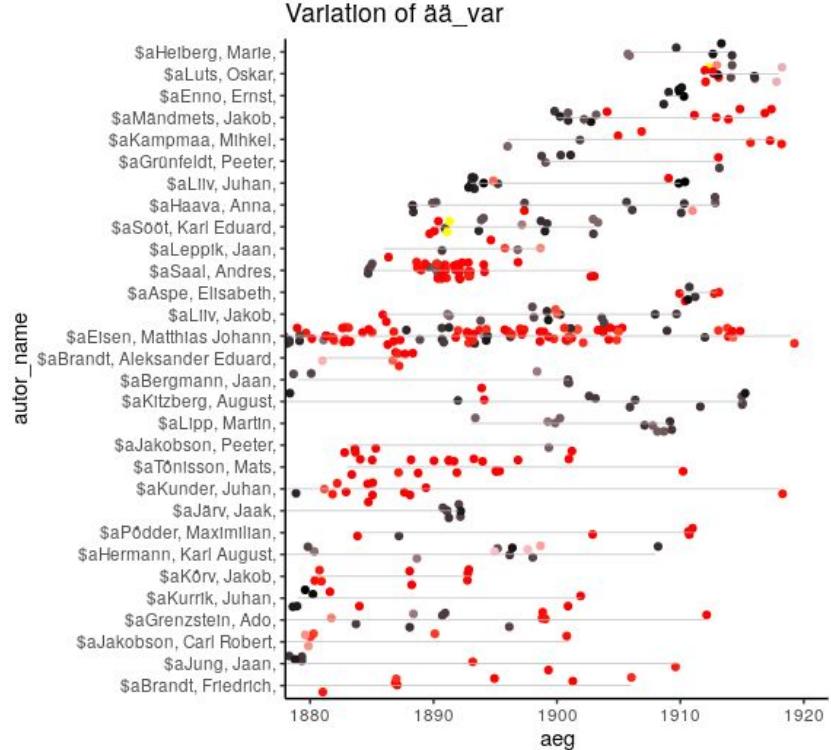
Variation between and within authors

author_name

Variation of ste_var

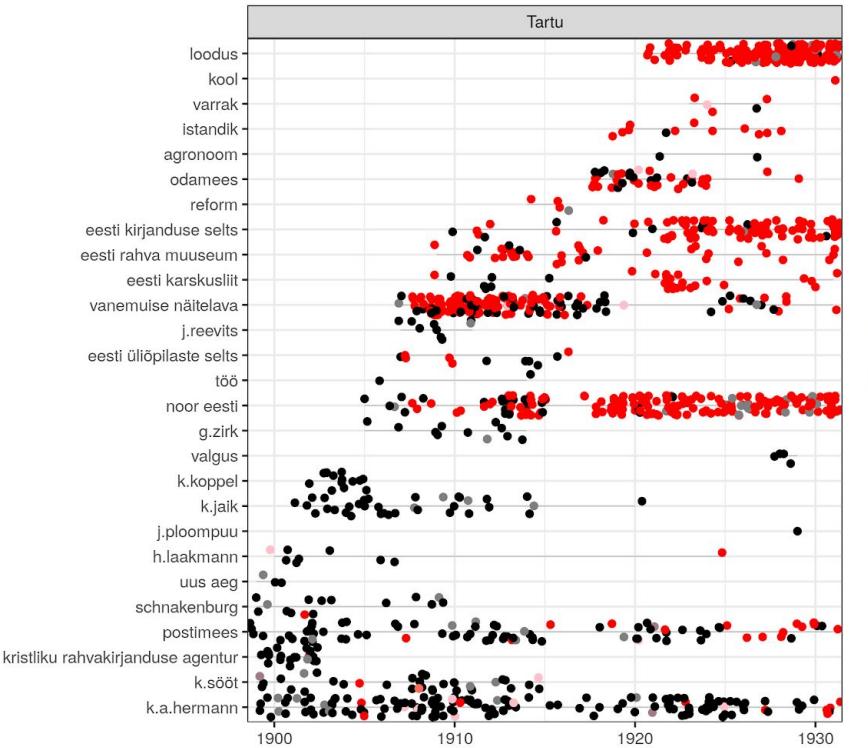
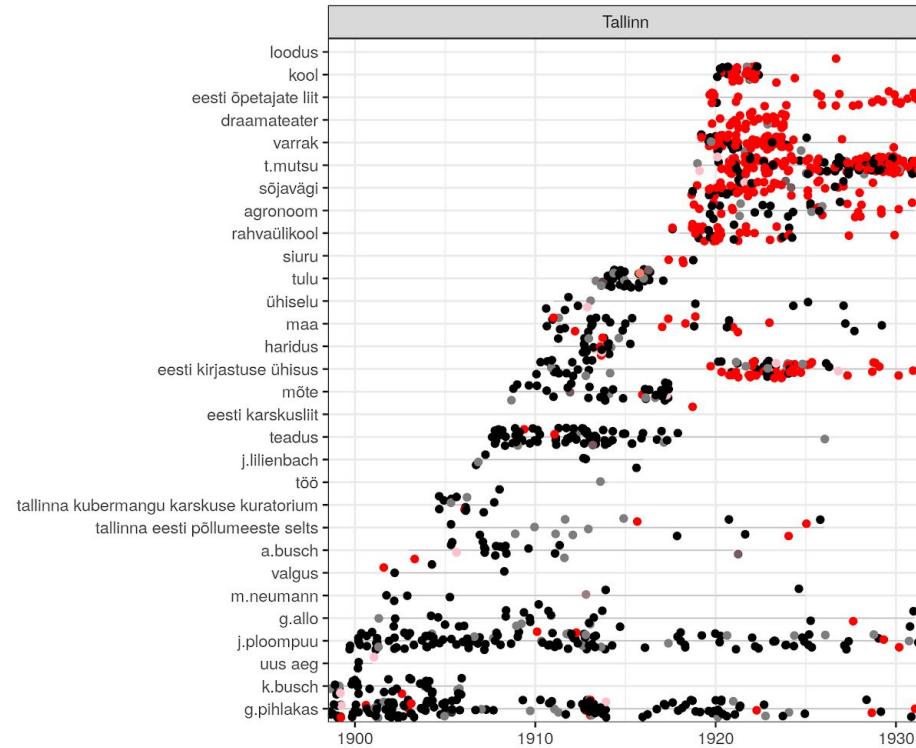


Variation of ää_var



Variation between and within publishers

Variation of W vs V in print titles by publisher



Studying the variation



HiSoN 2020/2021 ERLANGEN: INTRA-WRITER VARIATION

Intra-Writer Variation in
Historical Sociolinguistics

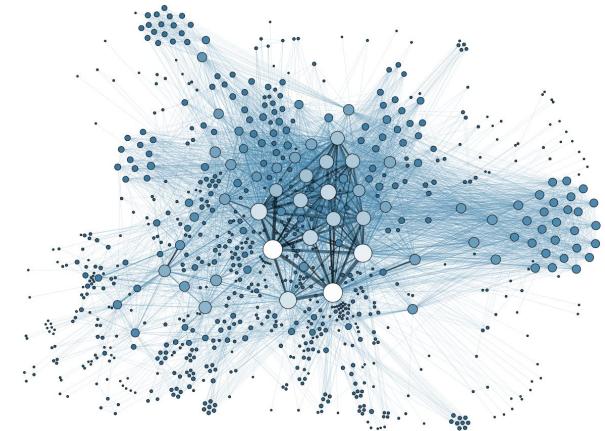
HiSoN 2021 Conference, March 17-19
FAU Erlangen (fully online!)



Competing hypotheses

Different primary drivers (what drives the variant taken up)

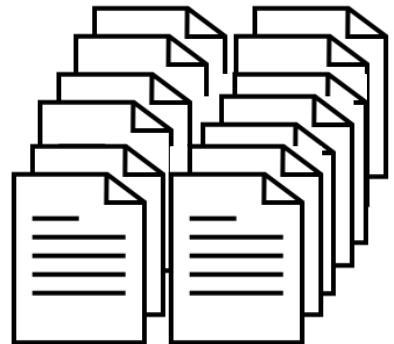
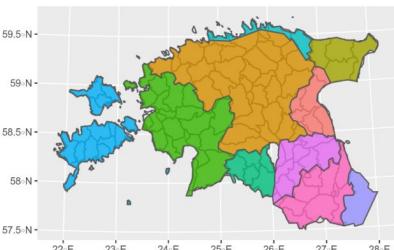
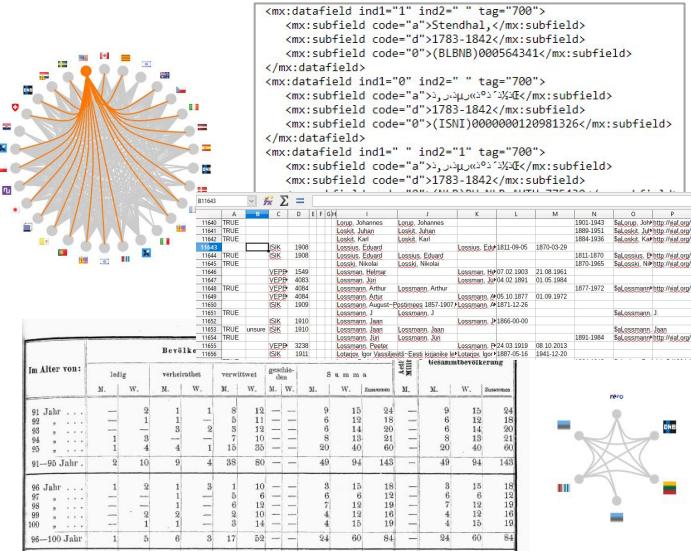
- Prescription
- Simple majority (home dialects)
- Urban dialects
- Educated circles
- New generations (schools, role models)



Data sources

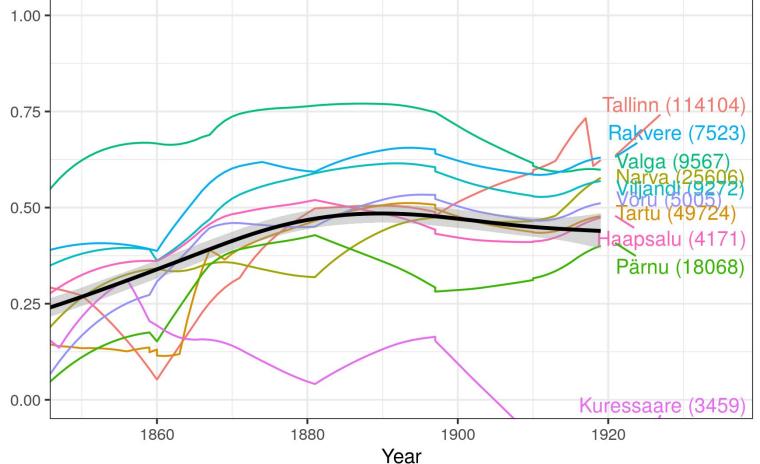
Assembling datasets and corpora

- Digitized publications, linked to authors
- Estonian National Bibliography
- Linked author metadata and biographies
- Aggregated historical census data
- Dialect maps



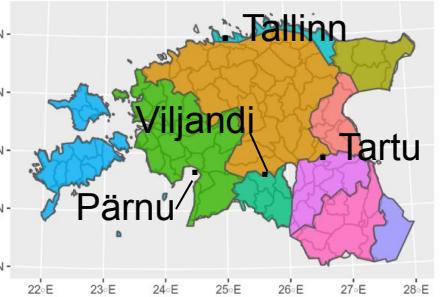
Data 1: Demographics

Proportion of people not born locally in 10 largest cities

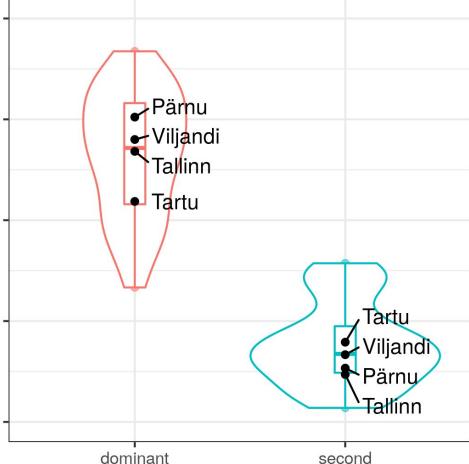
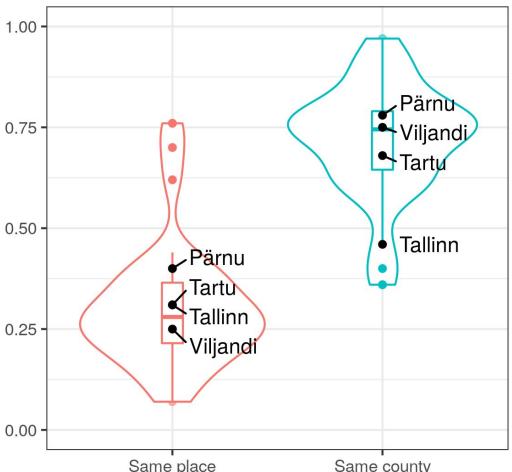


Cf. Kerswill 2018:
50% immigrants ---> new dialect formation

| TARTU ja VALGA m.kd. | | | | | | | | | | | | | | | | | | Tab. 8 (järg) — | | | | | | | | | | | | | | | | | | | | | | | |
|---------------------------------|-----|---------------|----|-----------------|-----|------------|----|------|----|------|----|-------|----|-------|----|-------|----|--------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|---|---|---|---|---|---|---|
| Sama maakond -- Le mõne erondis | | | | | | | | | | | | | | | | | | Teised maakonnad — | | | | | | | | | | | | | | | | | | | | | | | |
| Sama adm. lükus | | Linn ehk alev | | Alevik ehk vald | | Adm. üksus | | Viro | | Järv | | Harju | | Lääne | | Saare | | | | | | | | | | | | | | | | | | | | | | | | | |
| M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | M. | N. | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | | | | | | | | | | | | | | | | | | | | | | | | |
| VALLAD—Com. rurales: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 53. Roela | 605 | 658 | 9 | 11 | 241 | 254 | — | — | 8 | 15 | 4 | 2 | 3 | 4 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |
| 54. Rõuge | 239 | 241 | 95 | 116 | 380 | 441 | 3 | 1 | 15 | 11 | 1 | 3 | 10 | 8 | 2 | 3 | 1 | 3 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | | |
| 55. Rõngu | 490 | 548 | 21 | 22 | 303 | 372 | — | — | 3 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |
| 56. Saare | 534 | 564 | 3 | 9 | 223 | 305 | — | — | 4 | 7 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 57. Saare | 380 | 386 | 55 | 69 | 318 | 408 | — | — | 4 | 6 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 58. Soorin | 435 | 442 | 53 | 51 | 262 | 327 | — | — | 5 | 10 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |
| 59. Tõstamaa | 222 | 189 | 44 | 53 | 365 | 458 | — | — | 10 | 12 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |
| 60. Türi | 293 | 289 | 19 | 26 | 190 | 265 | — | — | 9 | 12 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |
| 61. Udens | 422 | 419 | 19 | 14 | 154 | 197 | — | — | 9 | 12 | 2 | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |
| 62. Üllik | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| 63. Vainastvere | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |



Origin of inhabitants (at 1922)

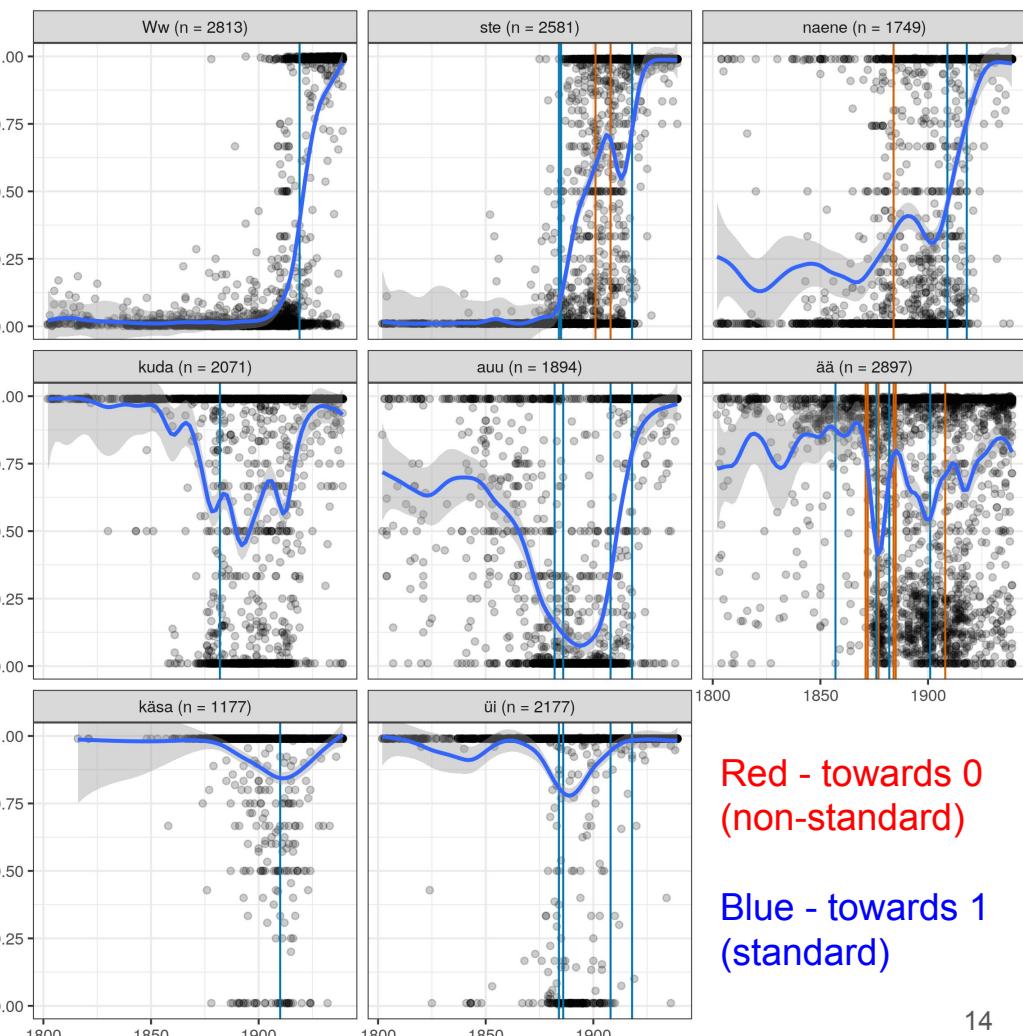


Data 2: Prescription

8 orthographic variables

- some dialect-related

- Prescriptive advice (dates & directions)
- 0-1 proportion of later standard variants
- Change points in population norms



Data 3: Author biographies

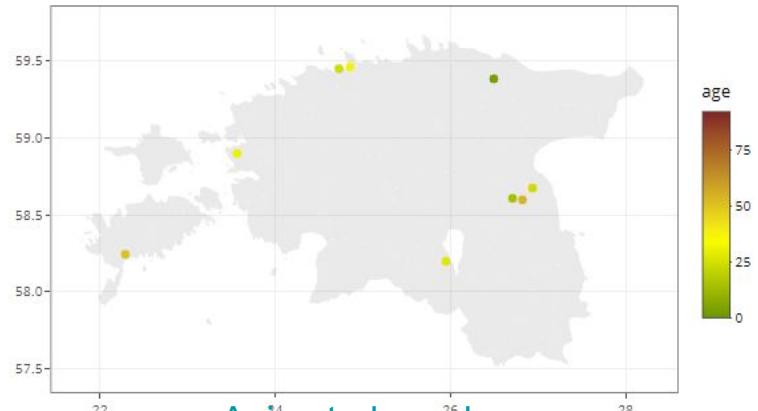
Combining databases

- Publication authors
- Biographical databases
- Geolocations



The screenshot displays several interconnected components:

- Wikidata**: A knowledge base showing the item Q1124581, which links to Julius Aamisepp.
- Deutsche Nationalbibliothek**: A library catalog entry for Julius Aamisepp.
- Wikimedia Commons**: A collection of images and files, including a portrait of Julius Aamisepp.
- Eesti biograafiline andmebaas IISIK**: An Estonian biographical database entry for Julius Aamisepp, listing his birth and death dates (1883-1960), residence (Tartu), and other details.
- Julius Aamisepp (Q1124581)**: A detailed Wikidata page for Julius Aamisepp, showing his life events, education (University of Tartu), and work at the University of Tartu.
- Animated graph**: A network visualization showing connections between various entities, including Julius Aamisepp and other people.



Animated graph

Author & language used in each publication

Make into very static predictors:

Variant used (0,1) ~ predictor1 + predictor2 + predictor3

Publication information:

- Tradition
(during last 5 years)
- Big city, small town, rural
(log population size)

Author information:

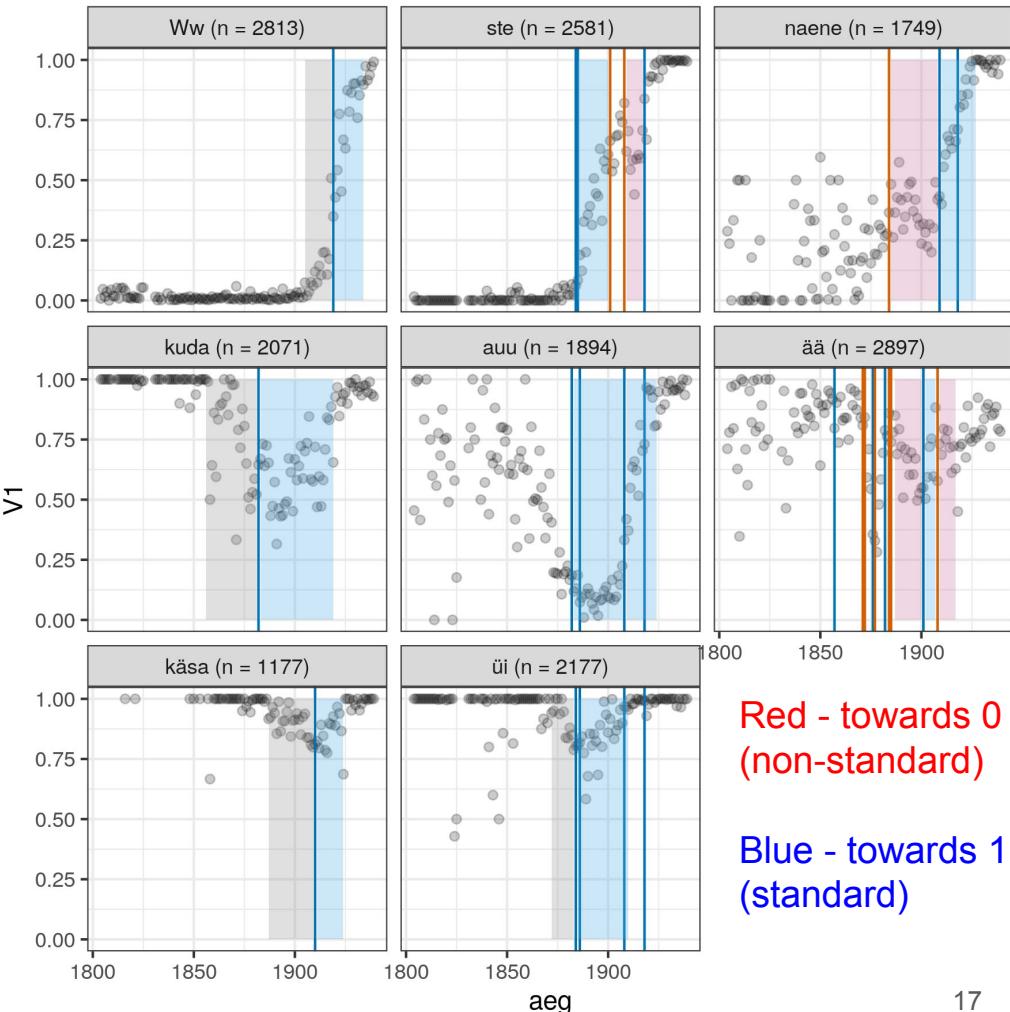
- Age (birthyear)
- Dialect (birth geolocation)
- Accommodation (later home geolocation)
- Years of education

Leaders and followers

Fit a logistic regression for each phase

- Prescribed direction
- Metadata
- Observed direction

Logistic regression effects for trends in who leads, who follows



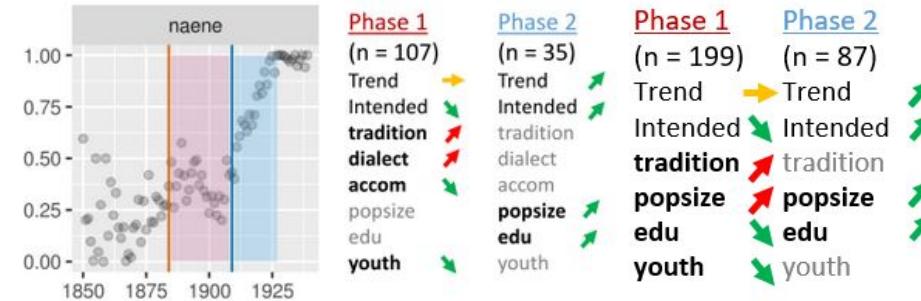
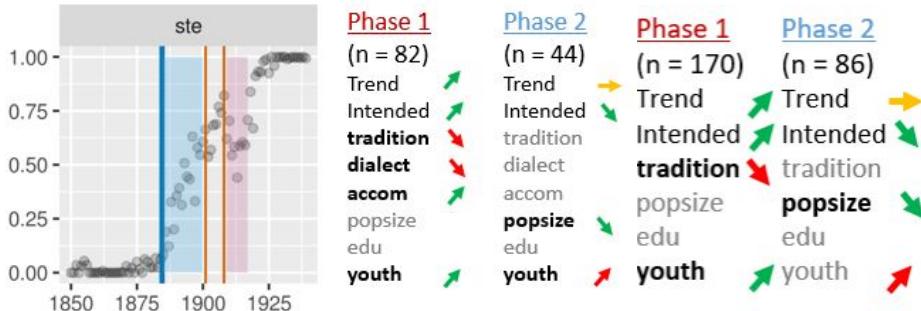
Assembling models

Model results (sig predictors in bold):

Model1: authors with birth years

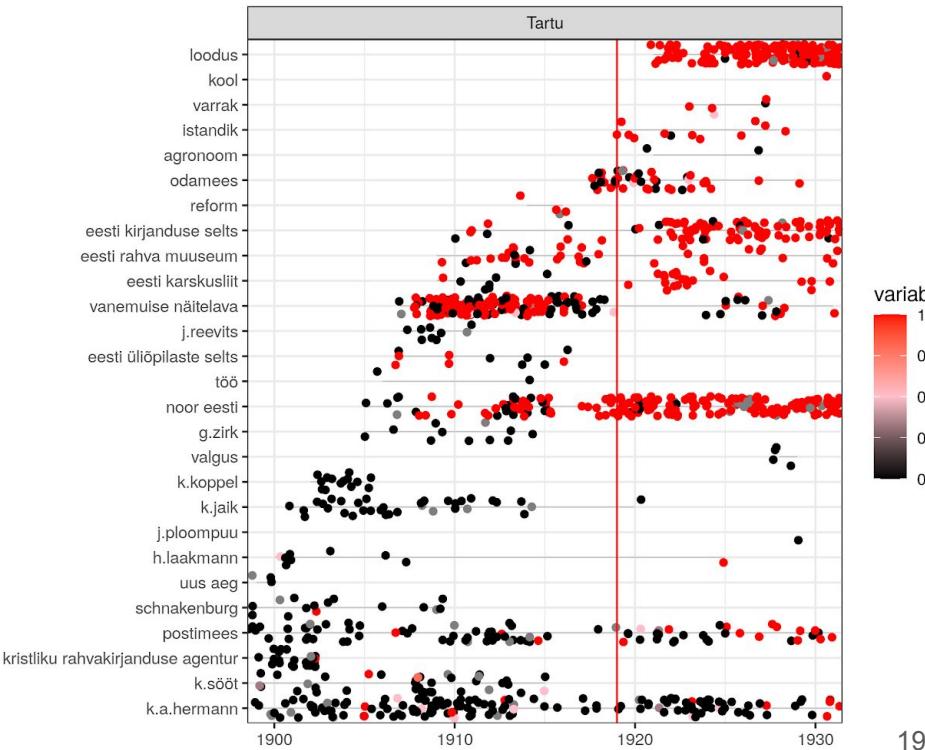
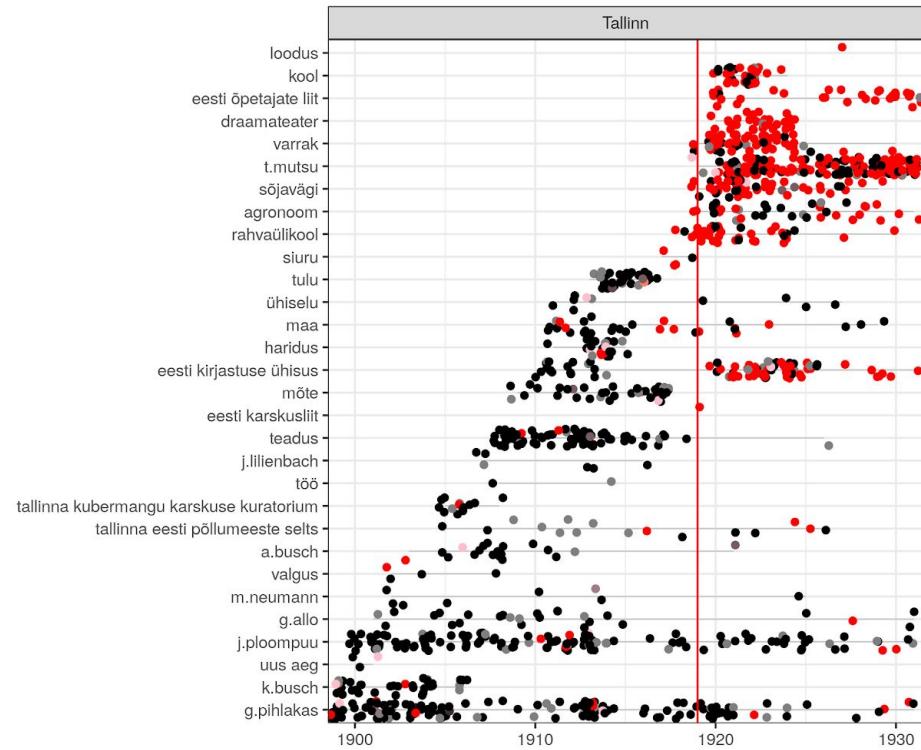
Model2: authors with geotags

- Effects vary
- A trend noticed: For successful prescription, either cities or youth had to be significant predictors.

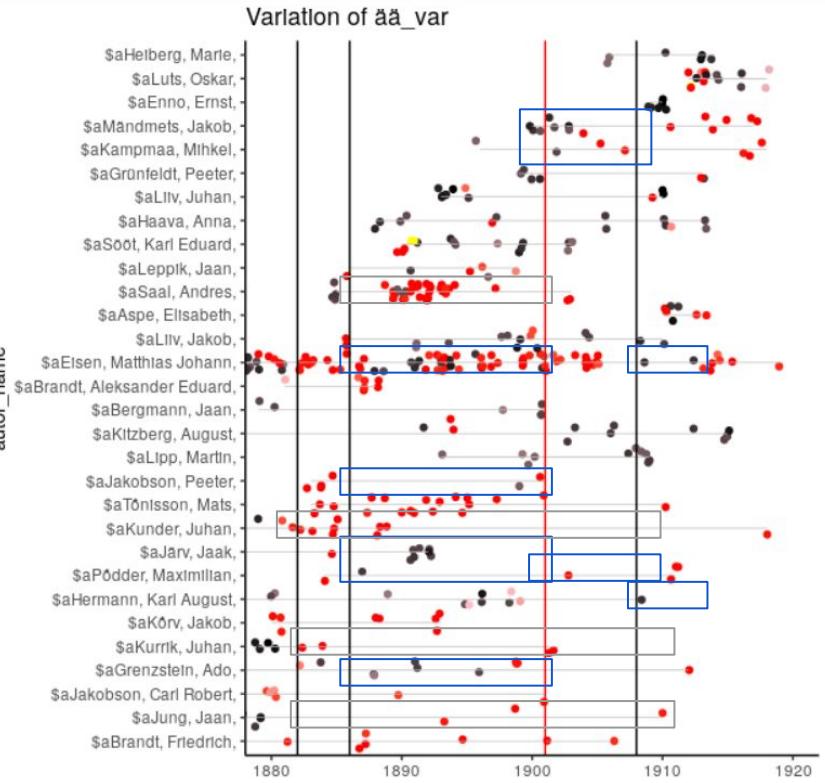
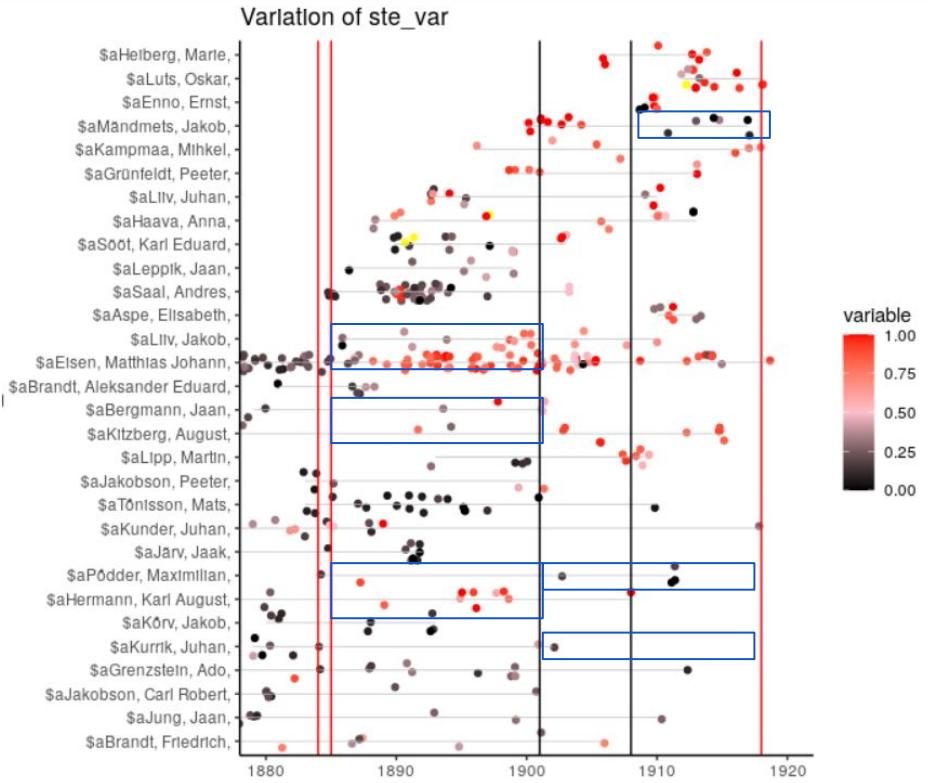


Intra-individual variation again

Variation of W vs V in print titles by publisher



Intra-individual variation again



Summing up

Society undergoing standardization can be an interesting case

- Migration, ideology, schooling, network effects
- From language to society

Estonian case:

- Urbanization remained mostly local (non-intense dialect contacts)
- At least some prescription efforts quickly carried to the community
- Individual variation in response: but cities and youth led successful changes

Combining databases allows for more detailed historical study

- Especially for investigating individual variation
- More study needed (of course)

