

Reduction in variation in written Estonian in 1880–1920

Peeter Tinitis
University of Tallinn

Variation

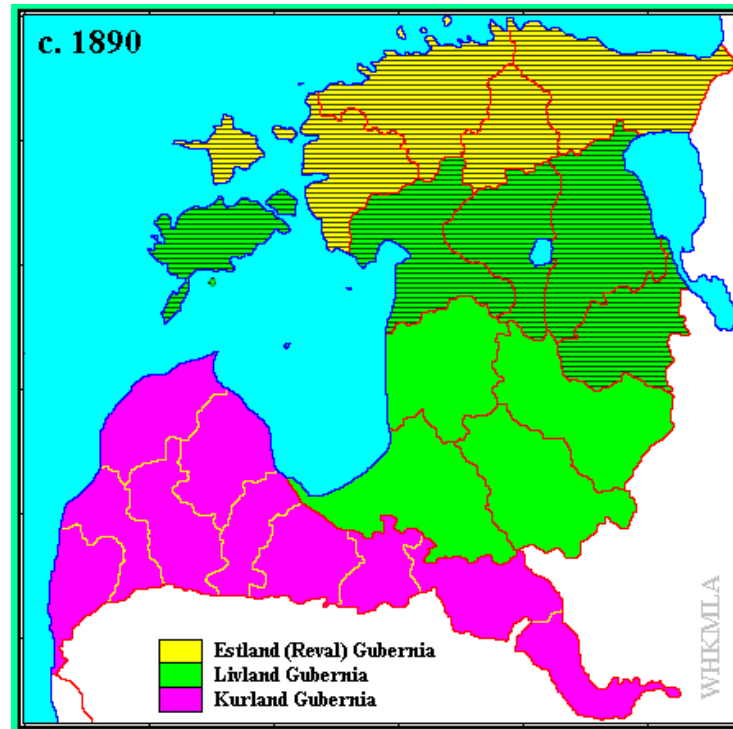
- Sociolinguistic variation that can be analysed as diverse linguistic forms used for same purpose, often enregistered as such.
- Amount of variation as a typological variable
 - Early literary cultures seem to have more variation in written forms (e.g. Rutkowska & Rössler 2012, Raag 1999, Scragg 1974)
 - Later it seems to decrease
- Why?

Language change

- Language standardization with two theoretically differentiable drivers (Haugen 1972; Deumert & Vandenbussche 2003)
 - contact influence, or linguistic accommodation
 - intentional change, or cultural intervention on language
- Standard language culture (Milroy, 2001) – imposition of uniformity on variation with prestige, legitimacy, correctness, formality

The case of Estonian

- Estonian language community 1880-1920:
- Special part of Russian empire run by the Germans
- Purely peasant community until the 1850s (reforms of 1802, 1804, 1809, 1816/1819, 1849/1856 changed this)
- Quickly urbanizing and modernizing
- Spread of romanticism, 1860s – 1885 known as national awakening
- 1885 – 1914 Russian policies of assimilation significantly changed the governates for both Germans and Estonians.



Language situation for Estonian

- German language community, Estonian language community
- Writing traditionally for German clerics and romanticists
- Since 1857 an increase in Estonian to Estonian written communication (journalism and then other fields)
- Emerging Estonian to Estonian communication, urbanization, social mobility increase variation in writing
- Some implications of a 'lack of stigma' on language use initially
- 'Language improvement' discussions emerge in 1870s

Variation

- Described to be rather high by later linguistic assessments (of course their baseline is whatever happened next) (Kask 1984; Laanekask 2004; Raag 2008)
 - Used to be Tallinn and Tartu enregistered written varieties, this changed into one with one central variety sought after
 - Influx of dialect speakers from the country, language contacts in the cities
- Schooling for most basic levels in Estonian (interfered 1885-1914), but teachers are complaining they don't know what to teach (1876, 1886, 1912)
 - Notably quick changes possible: (e.g. transition from 'old' to 'new' writing: 1869 20% of books use 'new system', 1873 50%, 1879, 90%)
 - High literacy (1897, 96% of over 9-year olds can read)

Conscious intervention

- Note, linguist was a very high prestige job in late 19th century
- First emerges for vocabulary extension and graphization reforms
- The 'problem' of variance and parallel forms emerges in the awareness of language enthusiasts by 1900.
- A number of public and publicized discussions on particular linguistic variants (notably 1872, 1886, 1908-1911)
- Language arguments sometimes turned political
- 1920s was heavily prescriptive, 1930s seems to have reached a common standard considered as fixed

Discourse on variation

- Societies and committees formed to improve language
- No legislative power only informal influence
- Lots of arguments with various argumentation
 - comparative-historical (old is good, related is good) & other linguistic theories
 - own dialectal background
 - if it sounds nice

Decrease in variation?

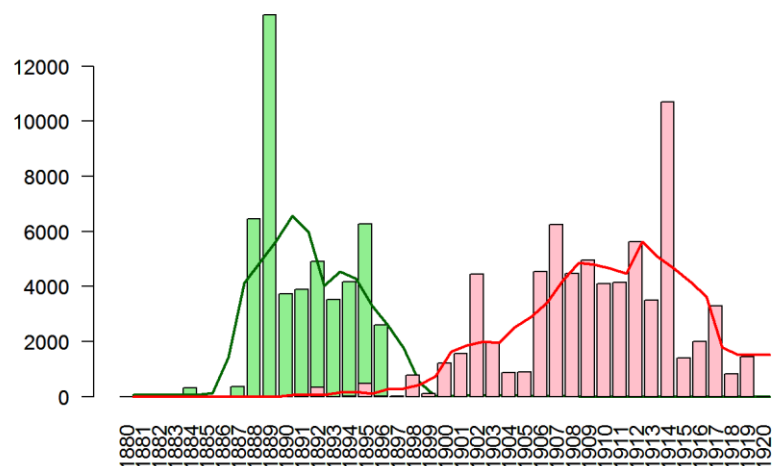
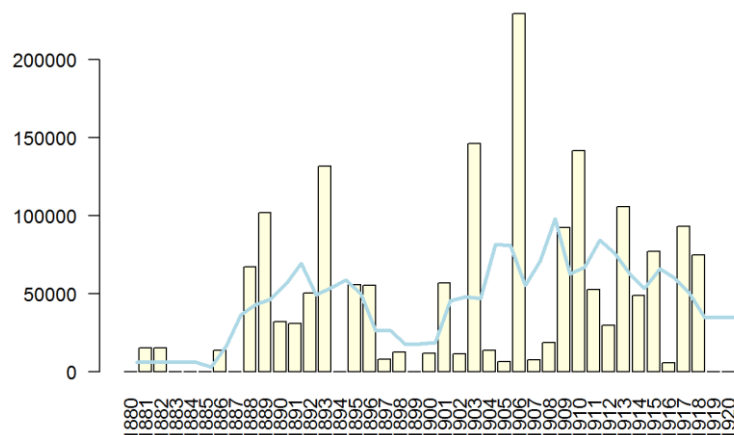
- It seems to be an overly dominant paradigm to assume that there can be only one. Variance was assumed to be bad for the language.
- At the same time, choosing between parallel forms was one part of the story. In language elaboration many new word forms were borrowed, derived and created which increased the variation.
- People were very reflexive on this issue, some wanted more:
- „Language must not only be organized, but improved and made greater. Now is the time for it. Ten-twenty years from now, when the language has become solid and fixed through common use, it will be much more difficult, almost impossible. Strike the iron while it's hot!“ (Aavik 1914a: 2)
- „Some people, especially older ones, complain that too many novelties will bring confusion into the language in some type of an interregnum. [...] We don't see a danger in that, the bigger the confusion the better [...] language will come out cleaner, clearer, prettier and more complete from this purgatory.“ (Aavik 1913m: 99)

Variables

- 4 slightly different ones
 - ‘ea’ and ‘ää’ (*pea-pää* [head], *hea-hää* [good])
 - ‘sti’ and ‘ste’ (*kenasti-kenaste* (nicely), *armsasti-armsaste* (dearly))
 - ‘ai’ and ‘ae’ (*naine-naene* (woman), *laine-laene* (wave))
 - ‘au’ and ‘auu’ (*honour*), ‘õu’ and ‘õuu’ (*nõu* – advice)
- All mentioned in prescriptive works
- Most have a connection to dialectal background in pronunciation, but individual backgrounds are not always clear on their early life communities.

Corpora in use

- Novel corpora in compilation
 - printed publications in digitization (left, yellow)
 - Literary classics ~1879-1920 (~33 authors, ~90 publications, 1,233,170 strings)
 - 2 collections of personal letters (1 – right red, 2 – right green)
 - 1) Urbanizing writers ~1900-1920 (~10 writers, ~243 letters, 71,224 strings)
 - 2) Letters from folklore collectors ~1888-1896 (~269 wr., ~290 let., 51,908 strings)



Language change

- How did particular linguistic variables travel within the community?
- Why did people choose the variants they chose?
- Change in time
- Age-gradient, education, chosen domicile

Calculations

- Basis is a set of synonym pairs (e.g. *naine-naene, naisele-naesele*) where absolute frequencies are counted.
- How much (%) modern standard forms vs non-standard

$$\hat{u} = \frac{\sum_{i=1}^N a_i}{\sum_{i=1}^N (a_i + b_i)}$$

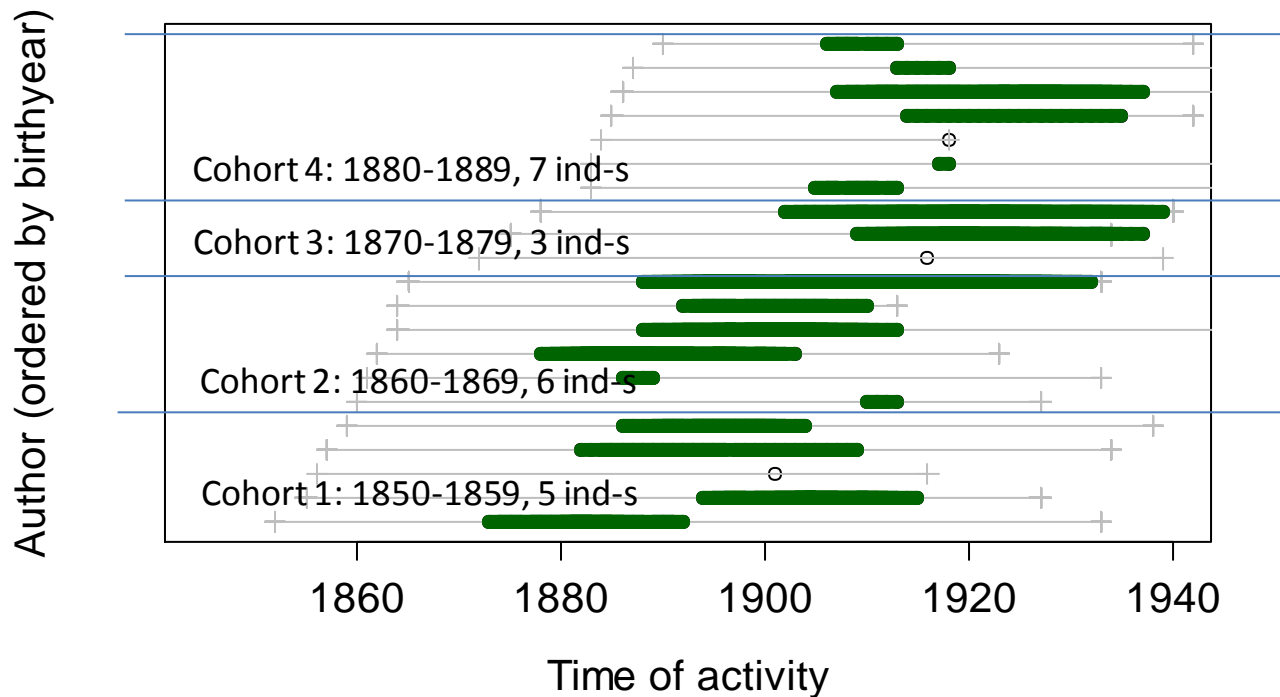
- If many lemmas, take each as independent & equal source of evidence: e.g. (ratio of *naine* + ratio of *laine*)/N

$$\hat{u} = \frac{1}{N} \left(\sum_{i=1}^N \frac{a_i}{a_i + b_i} \right)$$

- For authors in group, the same, each author has a ratio (does not depend on corpus size in its contribution)

Publication corpus

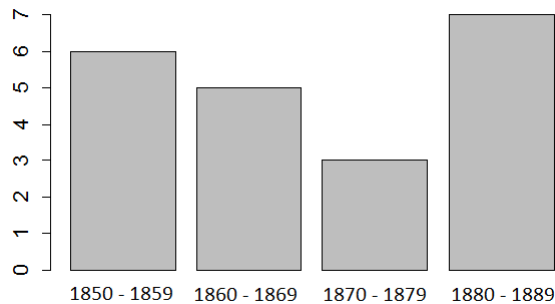
- Cohorts in the corpus



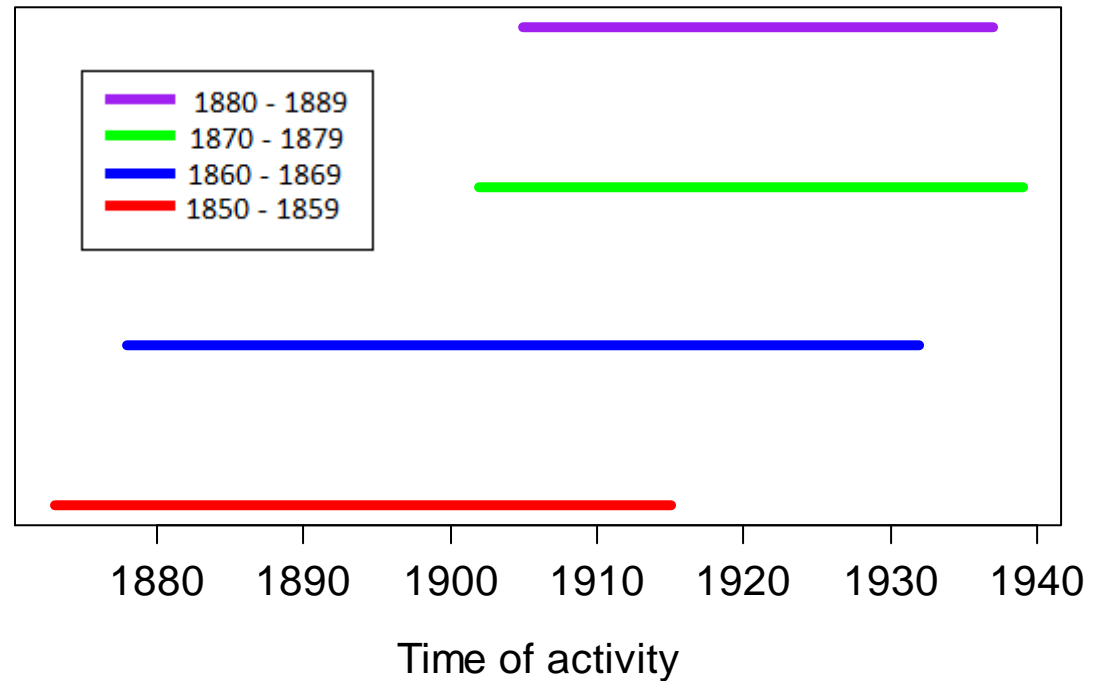
Grouping by birth

- Time of activity by age

Number of authors per cohort



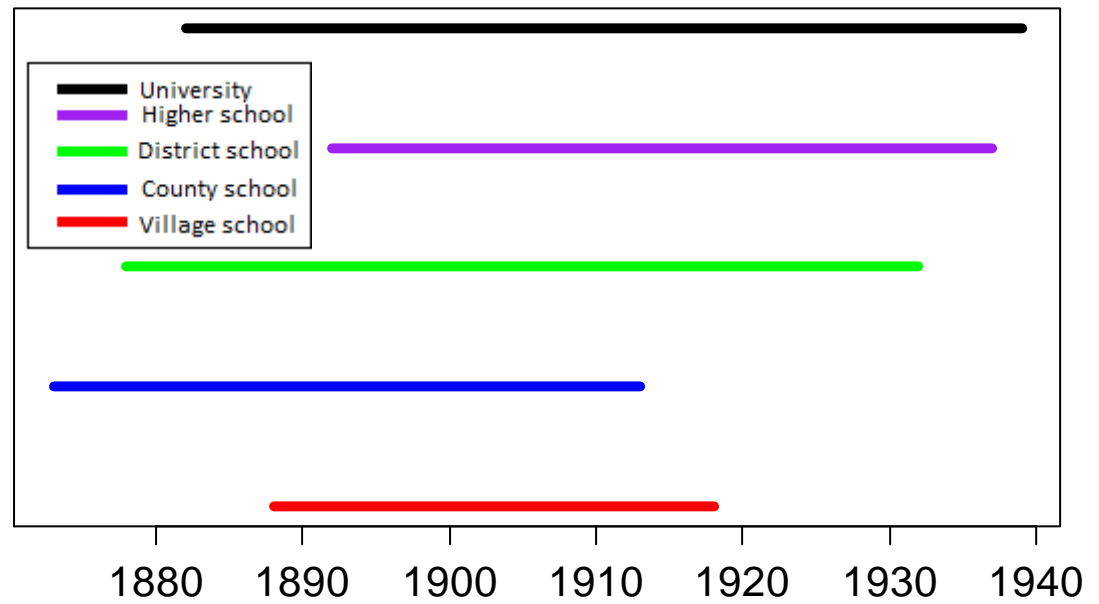
Time of activity per cohort



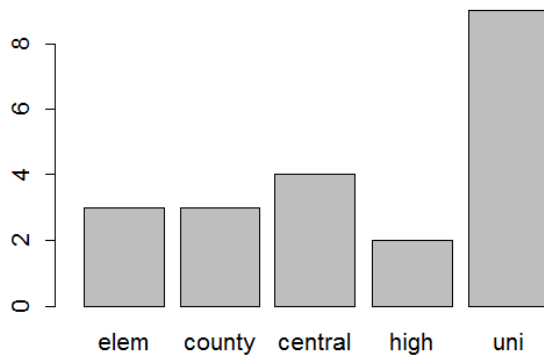
Grouping by education

- Education cohort, time of activity

Time of activity per cohort

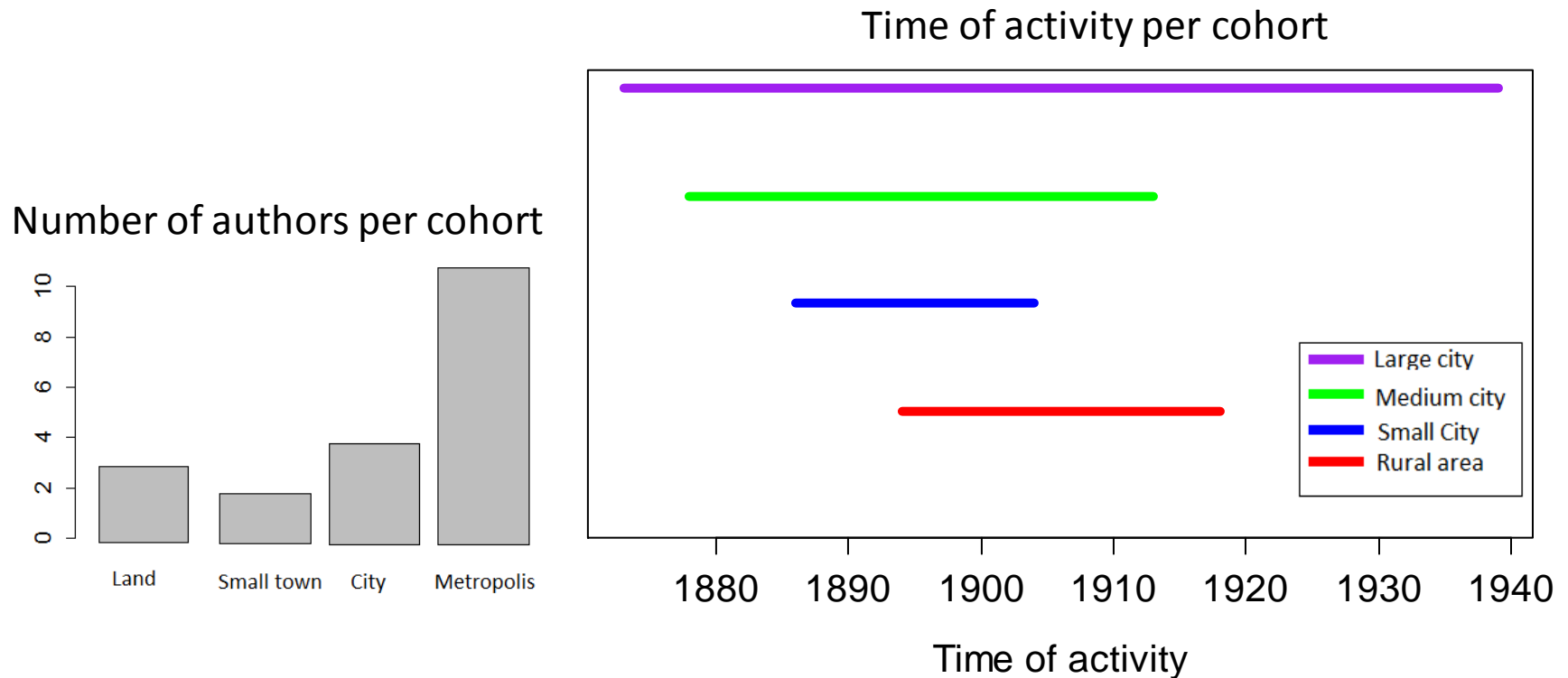


Number of authors per cohort



Grouping by domicile

- Activity in time by domicile



Some issues with the data

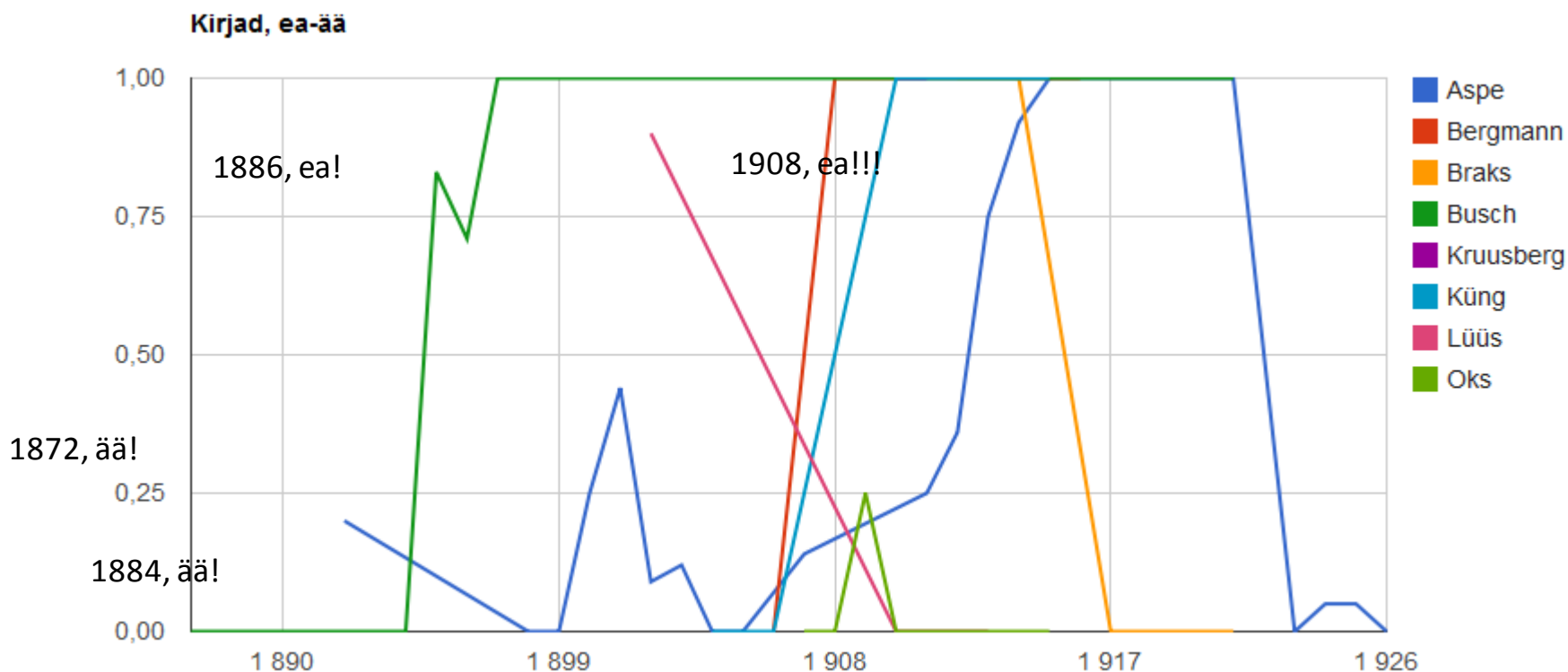
- Sparse and uneven data
- Representativity depends on availability (but what should it be representative of?)
 - Overrepresentation (?) of higher education levels is typical in historical sociolinguistics
- How to calculate across time – to give each piece of evidence a fair value?

‘ea’ – ‘ää’

- Includes lemmas *hea*, *pea*, *seal*
- Quite frequent debates, recommendations in opposite directions; has notable dialectal background (*ea* is in biggest dialect northern central one; *ää* is common elsewhere)
- **1872 – use *ää*!** [public and published speeches (J. Hurt): incl. „use new writing system“, use *ää* instead of *ea* always, *ää* was argued to be more archaic and having wider presence in dialects]
- **1884 – use *ää*!** [K. A. Hermann published an influential grammar]
- **1886 – use *ea*!** [Committee for making an improved new testament; use *hea* and *pea*]
- **1908 – use *ea* in the words in focus, *ää* almost everywhere else!** [One of the main points in first Language Conference 1909: Solution was compromise: use long vowels generally: e.g. *hää*l, *jää*ma, *pää*stma; *hüü*dma, *nüü*d, *tüü*tama; *võõ*ras, *mõõ*k, *rõõ*m, but *ea* in *hea*, *pea*, *teadma*, *seadma*, *seal*, *teal*]
- 1953 *ää* is removed from dictionaries, still in use nowadays as part archaism

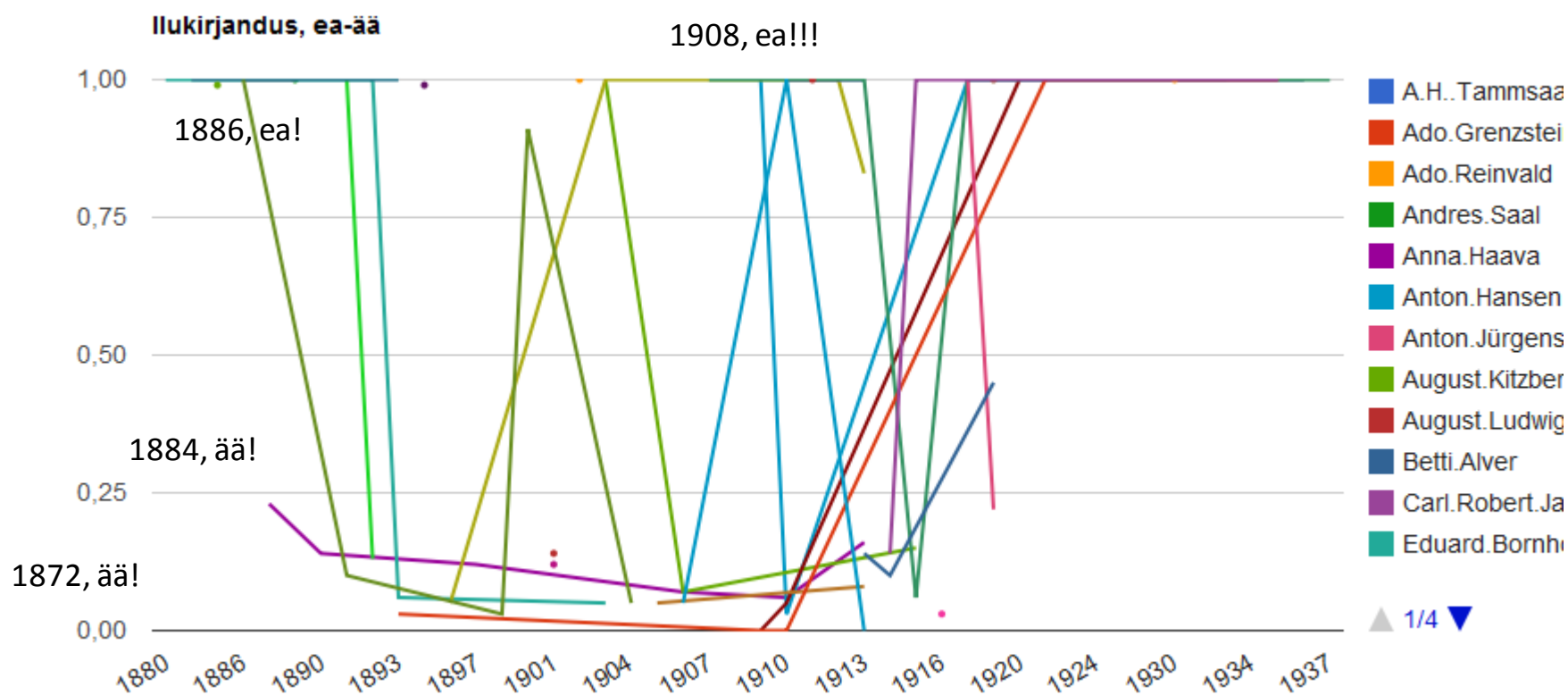
ea-ää, individual variation in letters

- Letter corpus shows some change in time but is fairly consistent



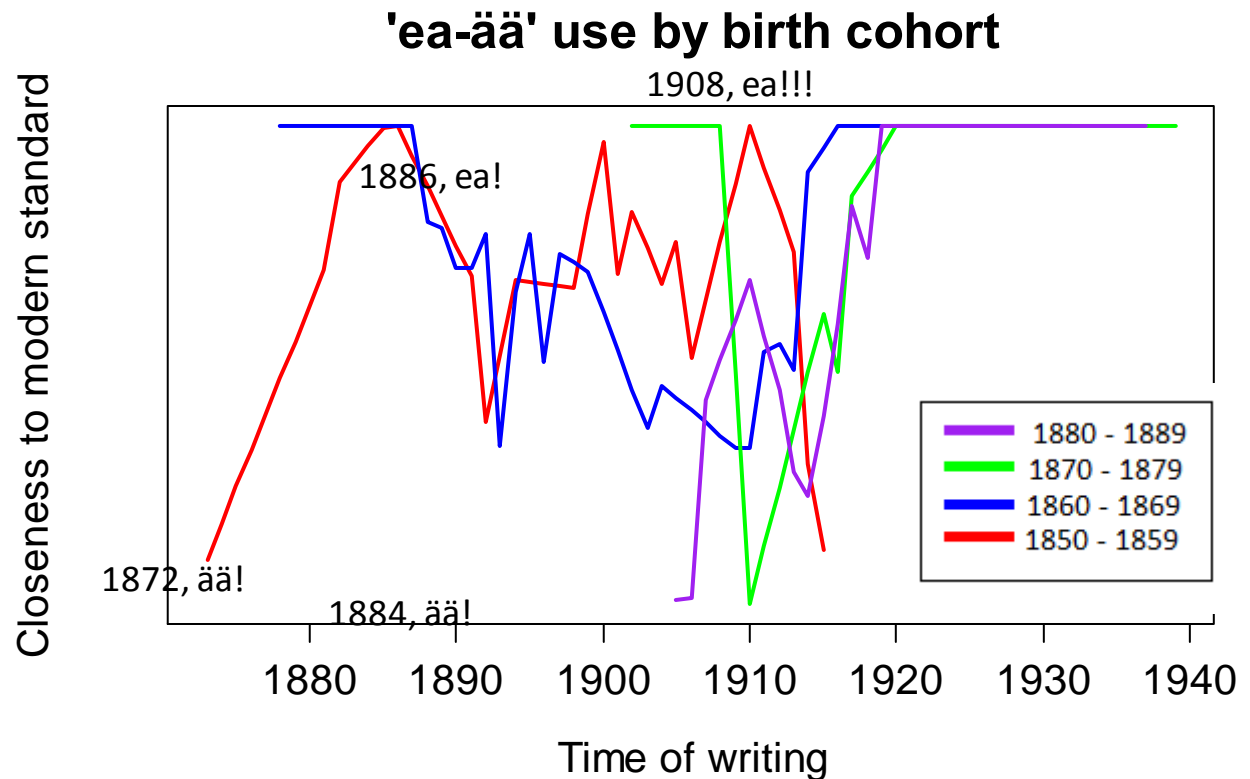
ea-ää, individual variation in pub-s

- Printed publications shows a lot of variation particularly in time, inconsistent at first with a sharp conclusion



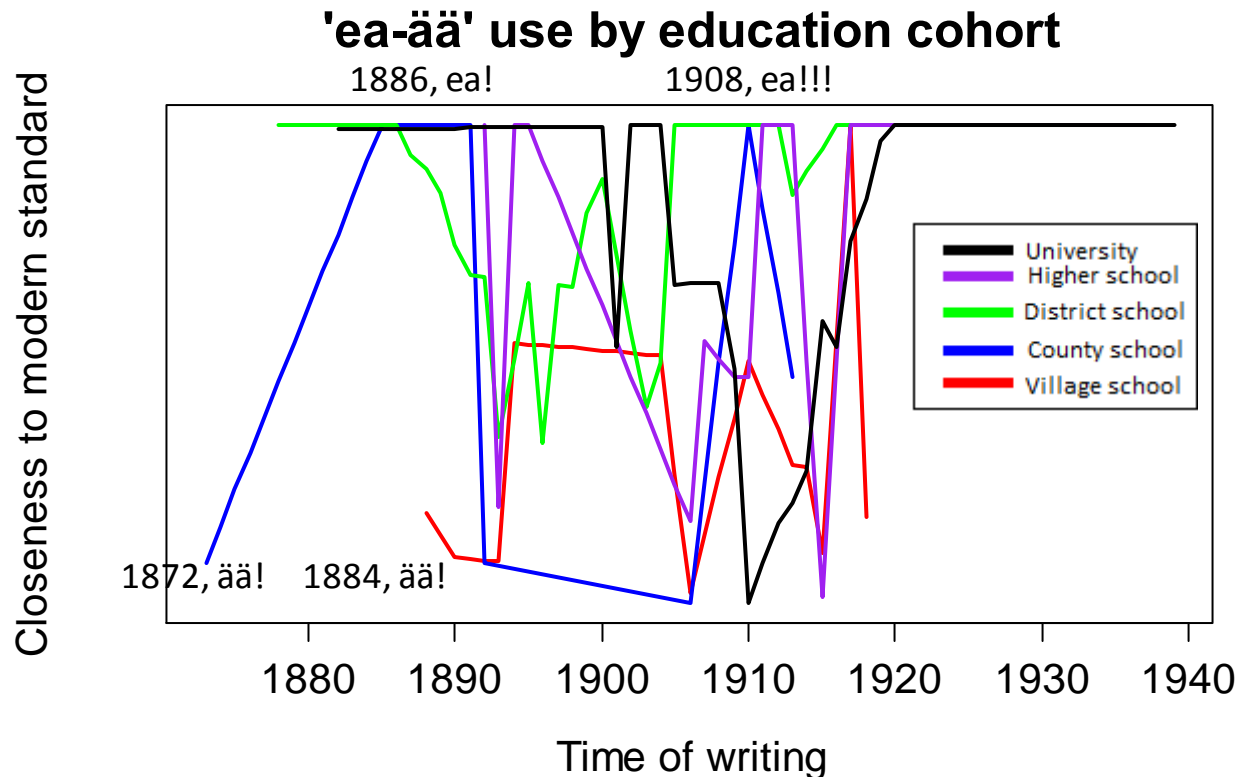
ea-ää, by time of birth

- Birth cohorts show the conclusion, but the differences seem small.



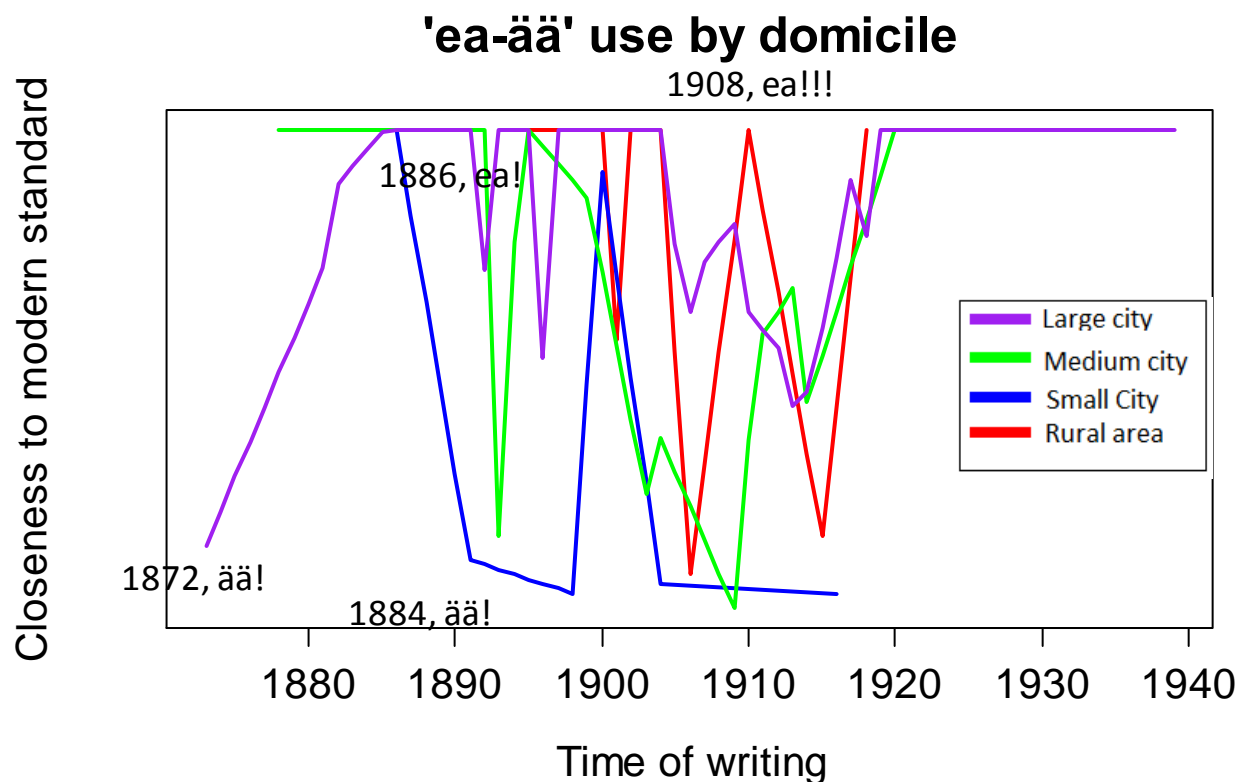
ea-ää, by education

- Village schools slightly lower, university slightly higher, inconsistency may be due to varying authors



ea-ää, by chosen domicile

- Large cities seem to be leading the *ea* variant

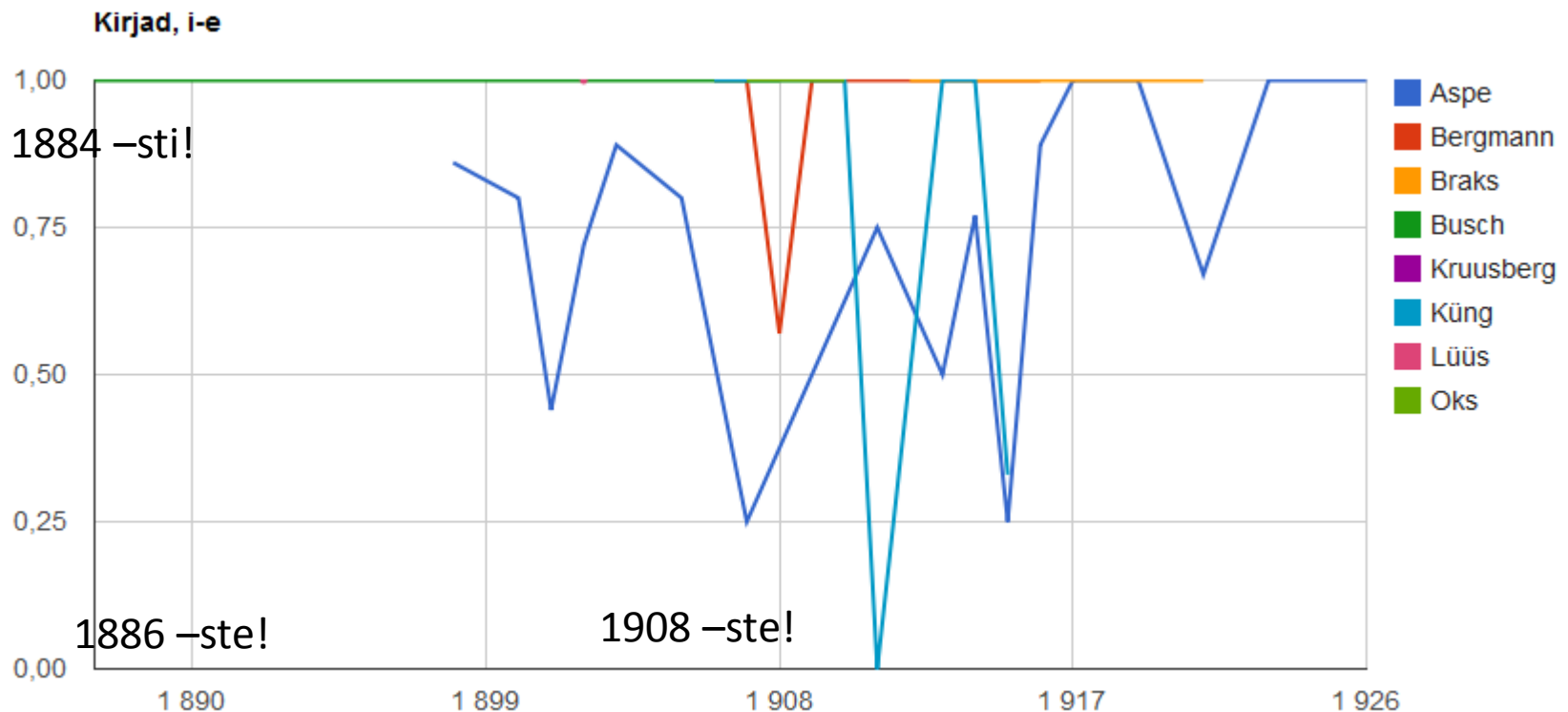


i-e variable

- -sti - -ste variable; 19 lemmas in total; for example täieste/i, õiete/i, vististe/i, lahkeste/i, uueste/i
- Known topic of polemic
- **1884 – Use –*sti*!** [K.A. Hermann recommends in his influential grammar; also Hans Einer recommends it so in 1885]
- **1886 – Use –*ste* in adverbs; use –*sti* in *õieti*, *kõigiti*!** [Committee on bible publication]
- **1908 – Use –*ste* in adverbs!** [First language conference]

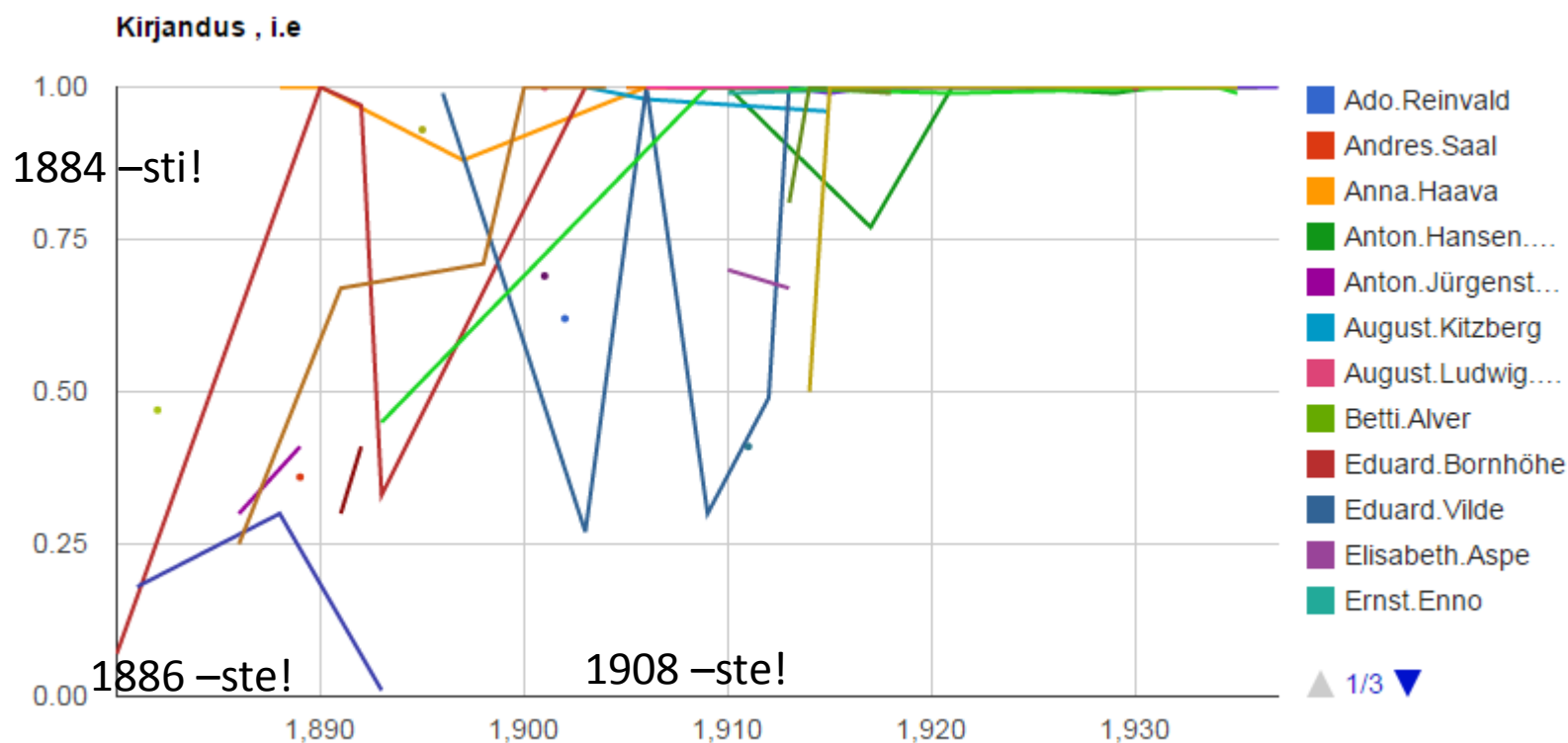
i-e, individual variation in letters

- Authors seem fairly consistent, *-sti* seems standar to begin with



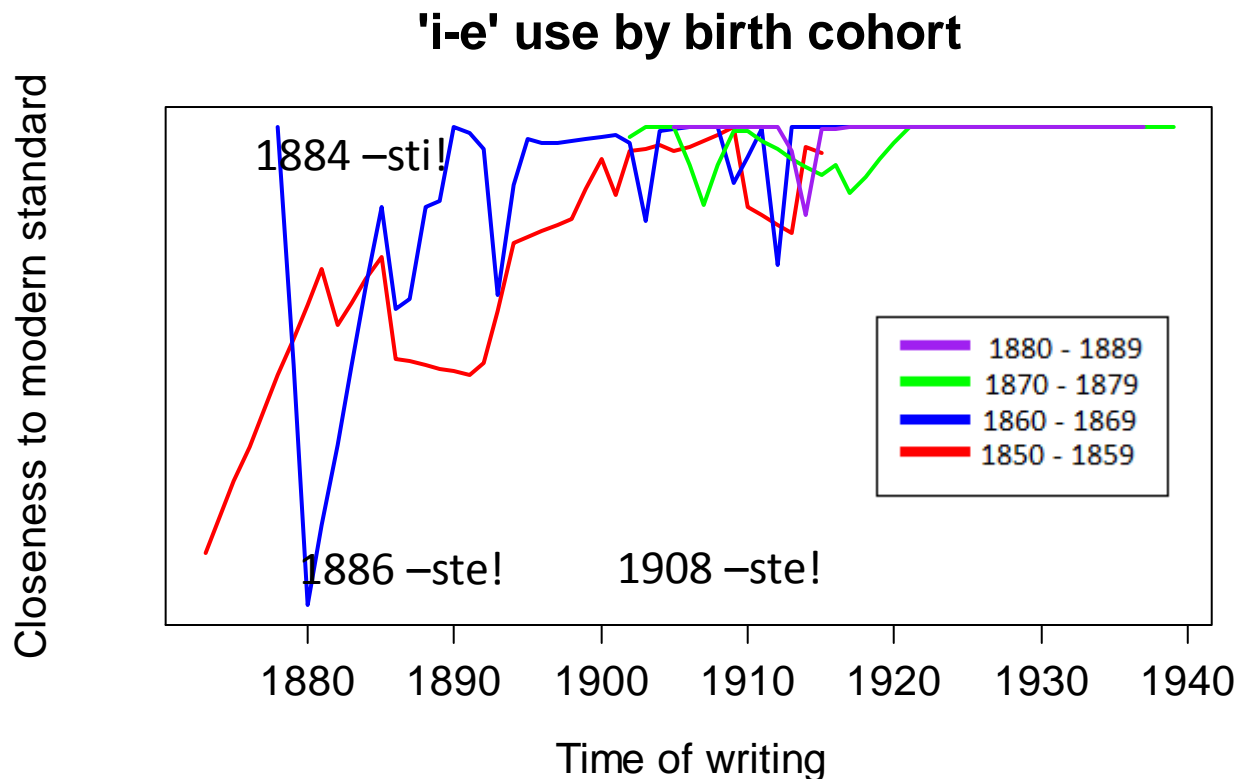
i-e, individual variation in pub-s

- Some variation and inconsistency, quick resolution



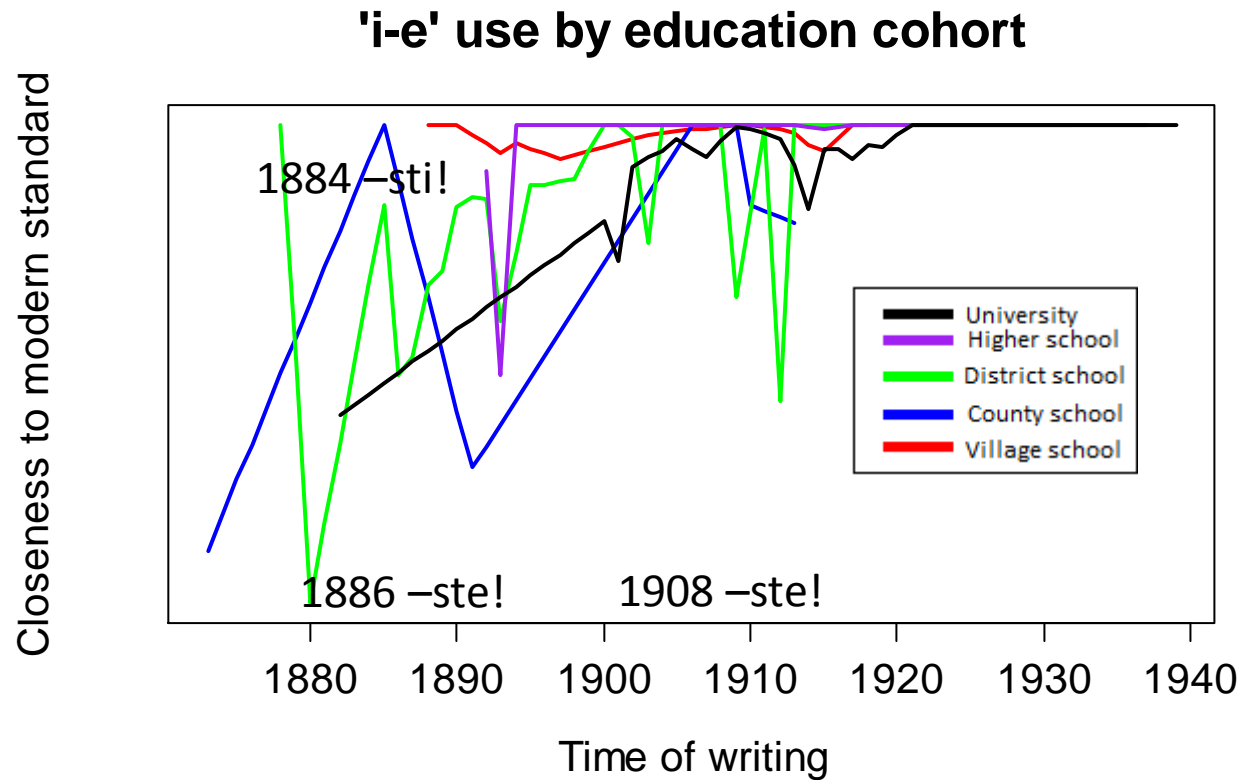
i-e, by time of birth

- Seems a rather smooth transition for all



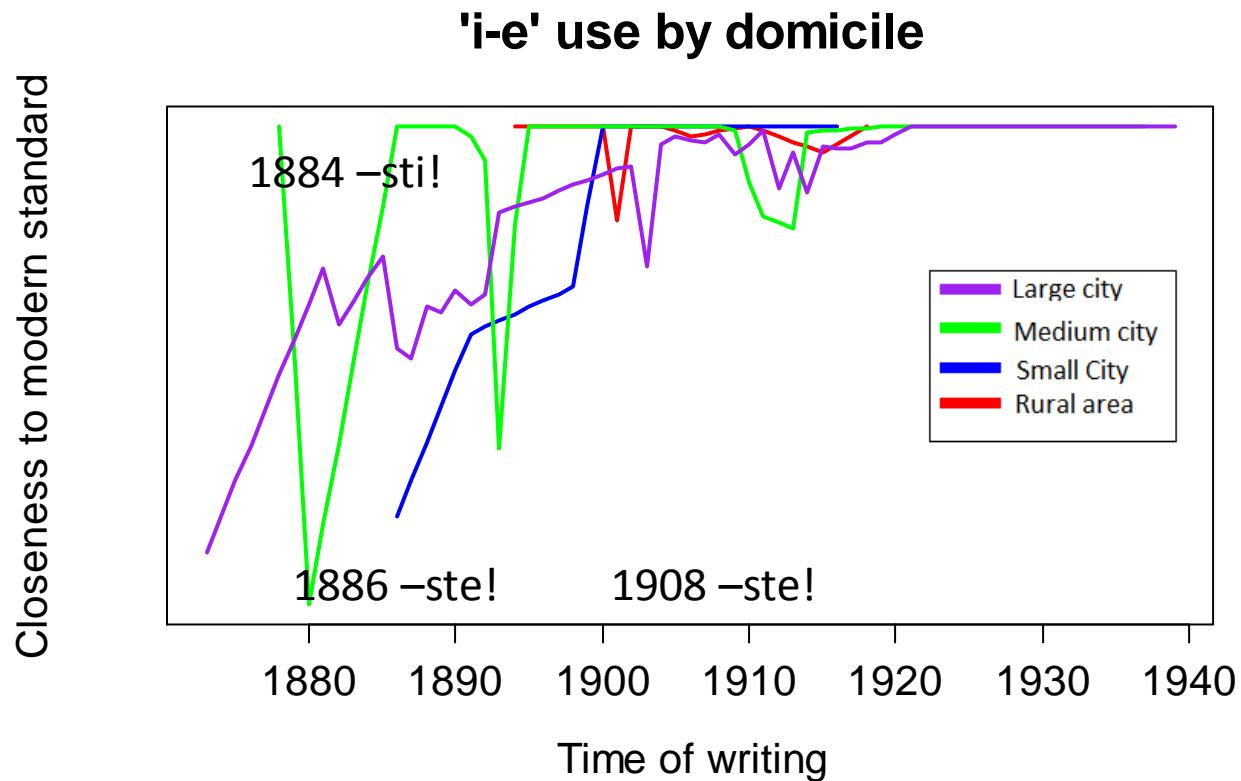
i-e, by education

- Seems smooth transition



i-e, by chosen domicile

- Seems a smooth transition

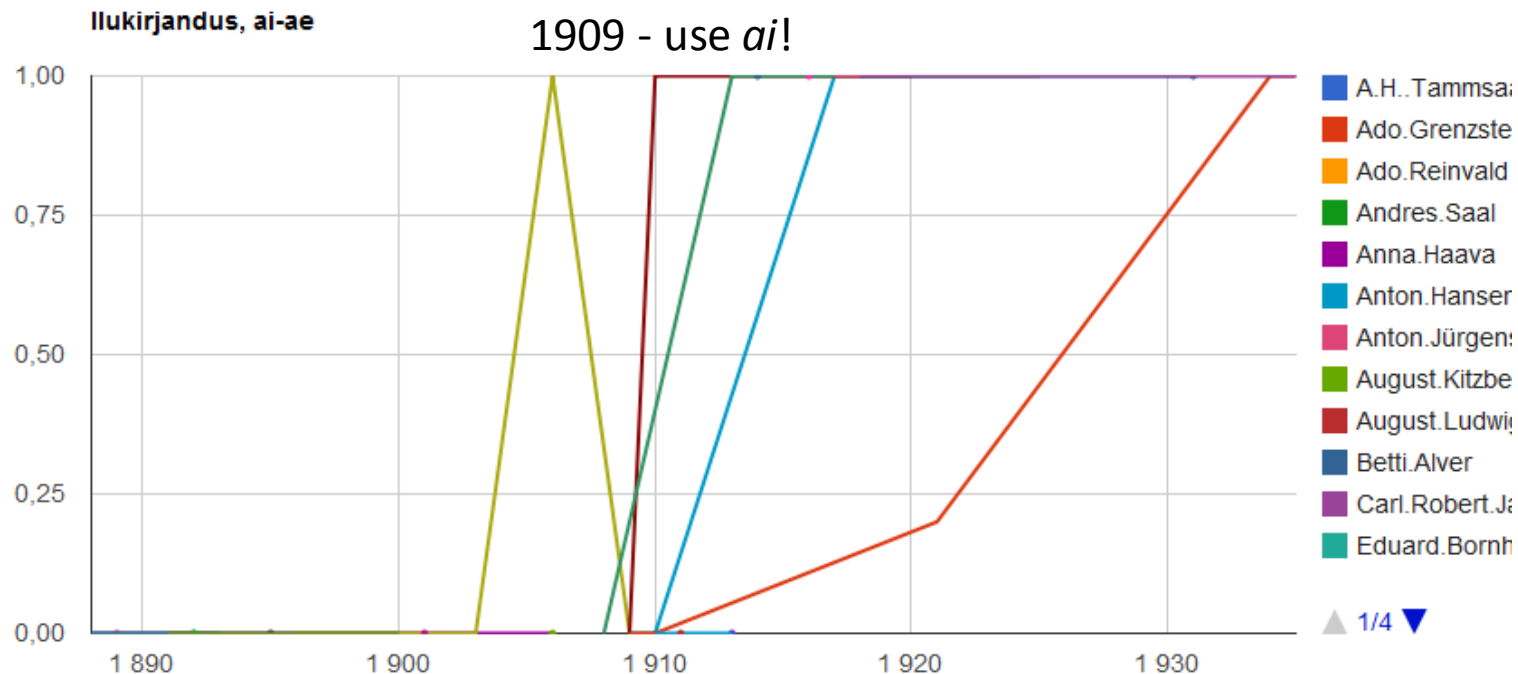


ai-ae variable

- lemmad: naine-naene, laine-laene
- Has some dialectal background, but was not really in focus
- **1909 – Use *naine* & *laine*** [Second language conference 1909]

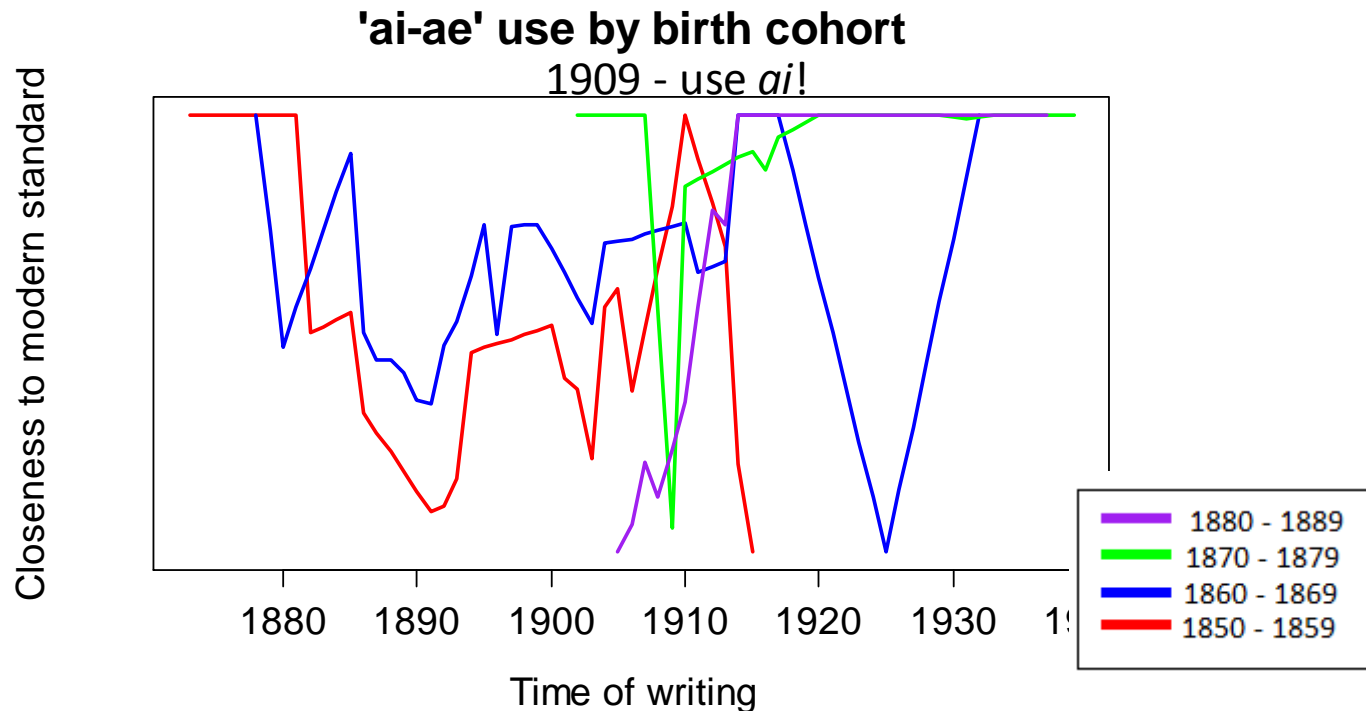
ai-ae, individual variation in pub-s

- Almost everyone is consistent throughout
- change happens once



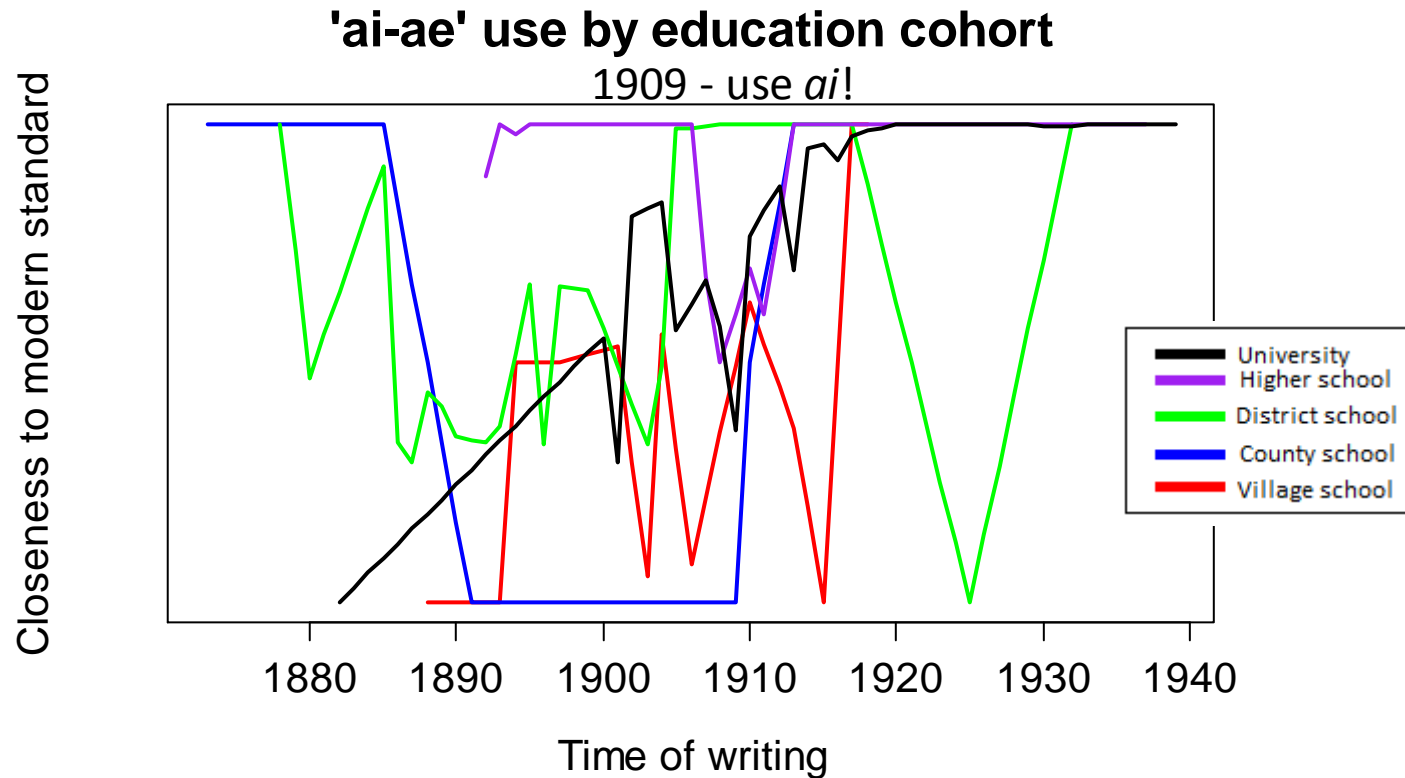
ai-ae, by time of birth

- Much interpersonal variation 1880-1914
- Crucial period seems 1910-1914



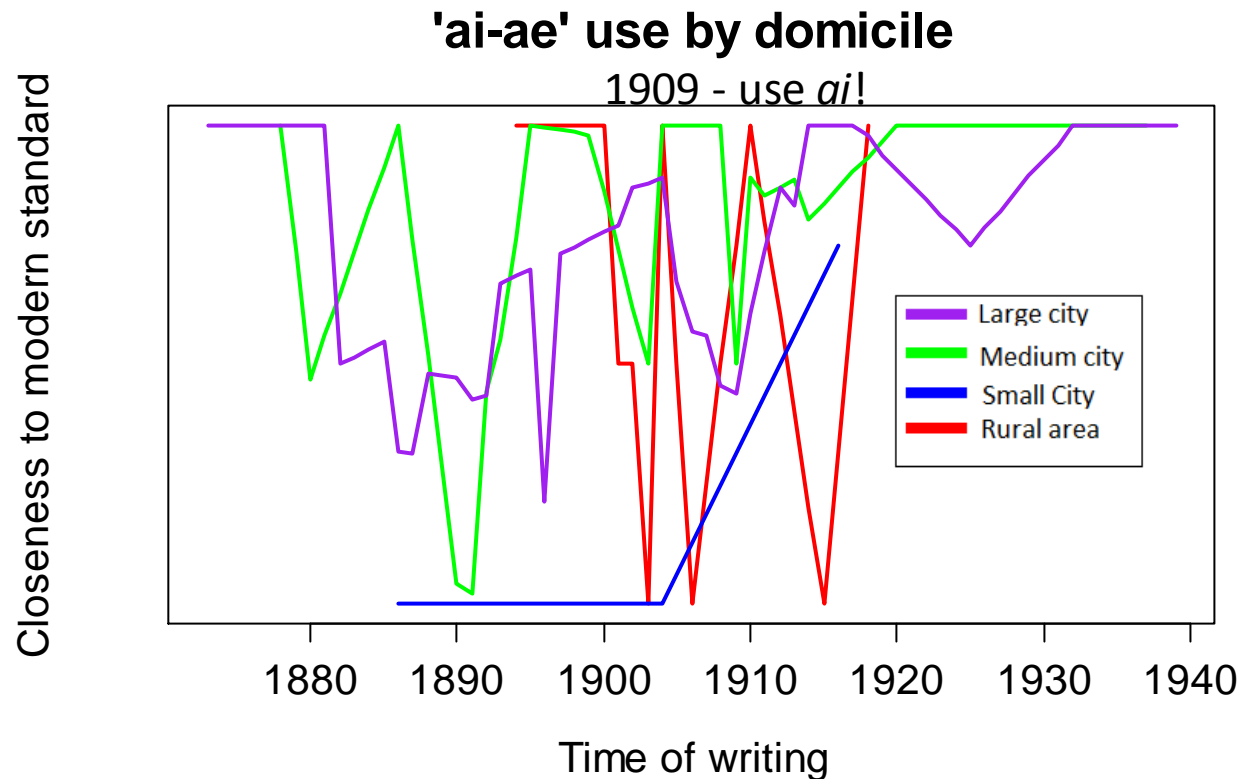
ai-ae, by education

- Some lag visible in lower education levels



ai-ae, by chosen domicile

- Rural areas seem somewhat slower in use, inconsistency may be due to varying writers.

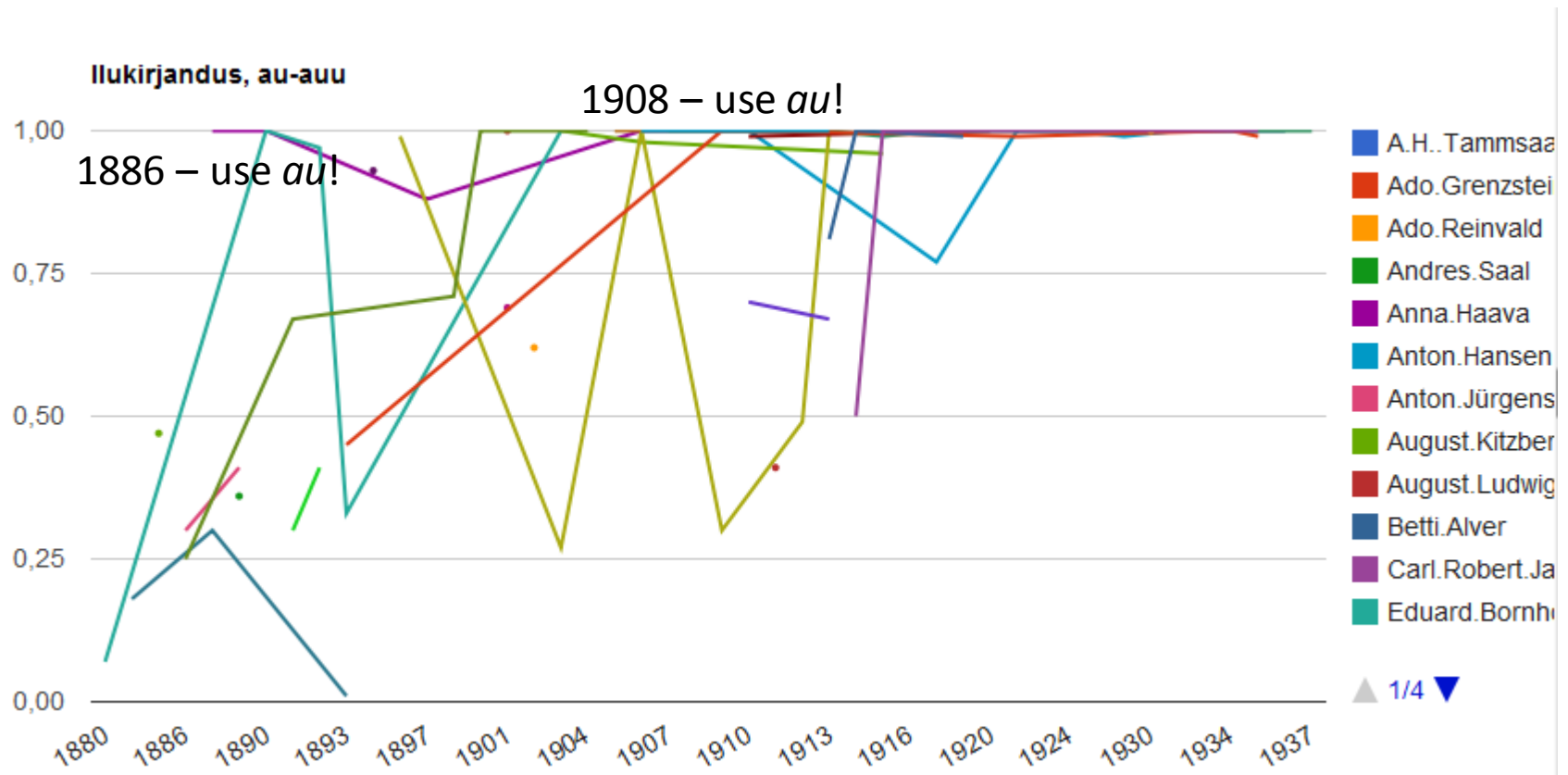


au-auu variable

- Here included as a single lemma: *auu* was a mostly archaic pronunciation *auvu* stored in the 'old written system of language; *auu* was standard during the 19th century
- Under discussion in some events of prescription, 'new writing system'
- **1886 – Use *au*, *nõu*** in nominative; *auu*, *nõu* in other cases [New bible publication committee]
- **1908 – use *au*, *nõu*** in all case forms [First language conference]

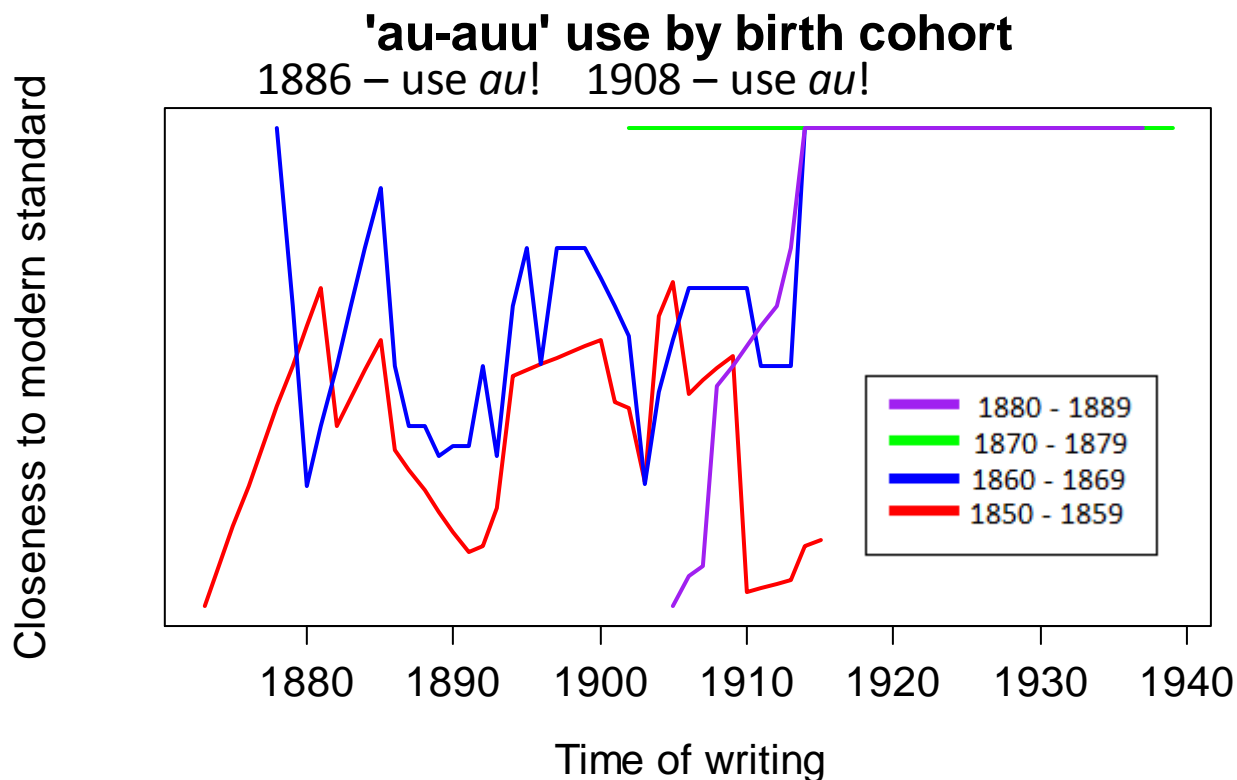
au-auu, individual variation in pub-s

- Some temporal inconsistency over individuals, not a lot, the variable is much rarer though.



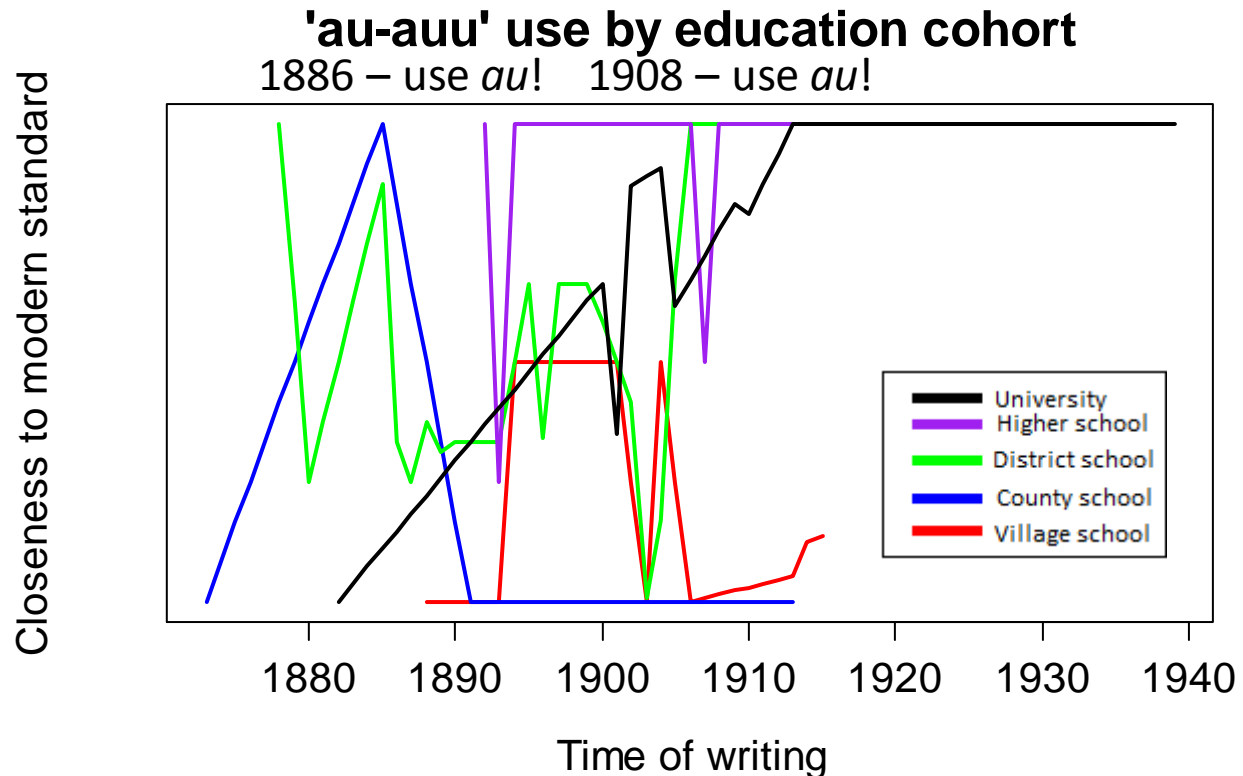
au-auu, by time of birth

- Quick change around 1914, new generations did lead the change slightly?



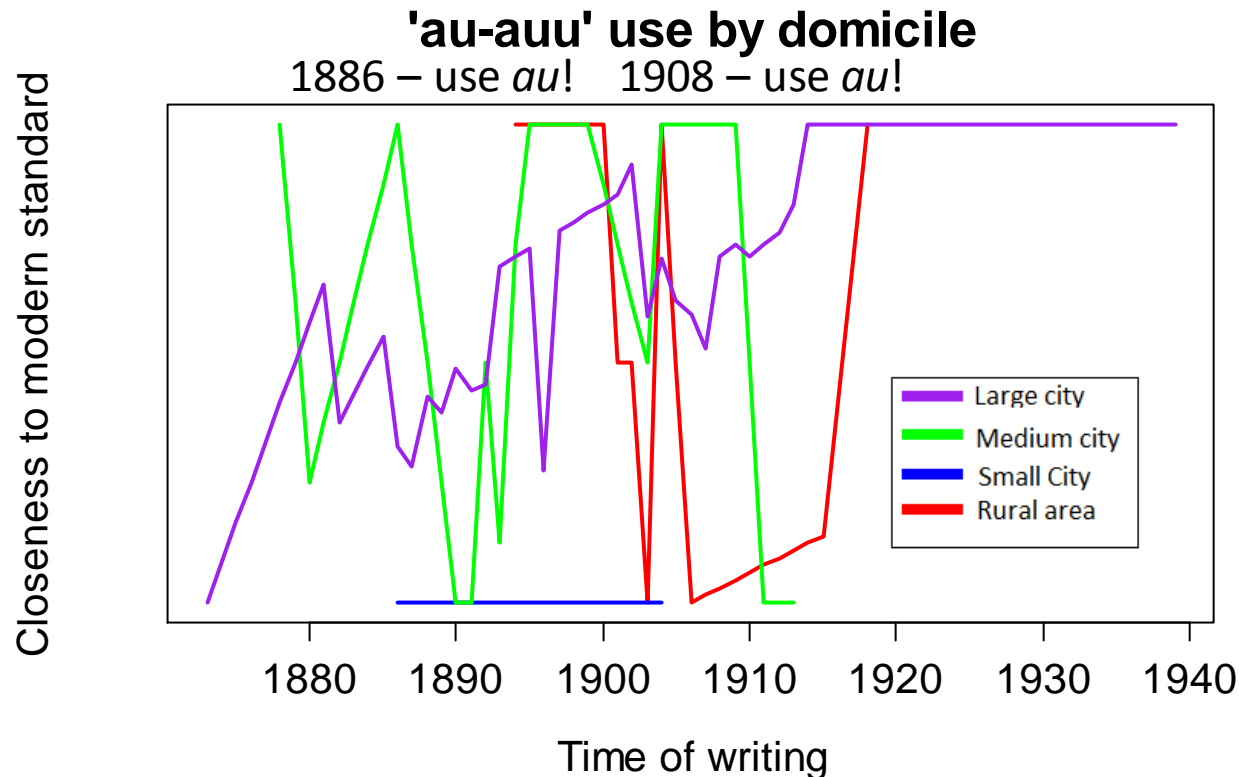
au-auu, by education

- University/higher education is the quickest, village and county is quite slow at this.



au-auu, by chosen domicile

- Here the domicile seems to noticeably impact the average use of the new variant, but it may just be the corpus.



Estonian context in 1880-1920

- Standard language culture in transformation
 - emerges in time
 - different in some details
- The given variables show some trends in the pathways of variables through larger cities and higher education, but not very clear

Future issues

- Standardization in transitionary contexts
- Measuring general variation
- General methodology with history

References

- Aavik, J. 1913m. Aforismid Eesti keele üle. Postimees 02.05.1913, 99
- Aavik, J. 1914a. Eesti rahvusliku suurteose keel : Lisad : Rahvalaulu salmimõõt; Rahvalaulu keel; Proov parandatud "Kalevipoja" teksti. Tartu: Reform
- Deumert, A., & Vandenbussche, W. (2003). Research directions in the study of language standardization. *Germanic standardizations: Past to present*, eds. A. Deumert and V. Vandenbussche, 455-470.
- Haugen, E. (1972). *The Ecology of Language*. Essays by Einar Haugen, selected and introduced by A. S. Dil. Standford: Stanford University Press.
- Kask, Arnold (1984). Eesti murded ja kirjakeel. (Eesti NSV Teaduste Akadeemia Emakeele Seltsi toimetised, 16.) Tallinn: Valgus
- Laanekask, H. (2004). Eesti kirjakeele kujunemine ja kujundamine 16.-19. sajandil. (Dissertationes Philologiae Estonicae Universitatis Tartuensis 14.). Tartu: Tartu Ülikool.
- Milroy, J. (2001). Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, 5(4), 530–555
- Raag, R. (1999). *Fran allmogemal till nationalsprak. Sprakvard och spralkpolitik i Estland fran 1857 till 1999*. (Acta Universitatis Upsalaensis. Studia Multiethnica Upsaliensia 12.) Uppsala: Uppsala universitet.
- Raag, Raimo. (2008). Talurahva keelest riigikeeleks. Tartu.
- Rutkowska, H., & Rössler, P. (2012). Orthographic Variables. *The Handbook of Historical Sociolinguistics*, 211-236.
- Scragg, D.G. (1974). *A History of English Spelling*. (Mont Fellick Series, 3). Manchester: Manchester University Press.

- Thank you!
- Also thanks to the digitization efforts in Estonian museums and archives (particularly Estonian Literary Museum).
- And thanks to European Social Foundation, DoRa programme and Archimedes organization for supporting this presentation.

