

Status update

Determinants of Dialectal
Variation, Syntax subproject

Marco René Spruit

<http://marco.info/pro>

Overview

- Education
 - Attended schools & courses
- Software
 - sync, dimp, finc, basc
- Data, based on nearly definitive SAND1
 - complementisers, subject pronouns, expletives, subject doubling, subject clitisation following yes/no, reflexive and reciprocal pronouns, fronting

Education

- Past months
 - Summer School, ESSLLI, Nancy
 - Minimalist syntax course, ULCL, Leiden
- Next months
 - LOT Winterschool, Groningen
 - Writing English for Publication, KNAW
 - Statistics courses (...)

Main software

- Core program (*sync*)
 - The SYNtactic measurement Console (*sync*) program is the interface between the SAND data and the dialectometric RuG/L04 software
- Automatic data preparation tool (*dimp*)
 - The *Data IMport Program* transforms SAND output files into an appropriate XML file which can be processed by *sync*

Additional software

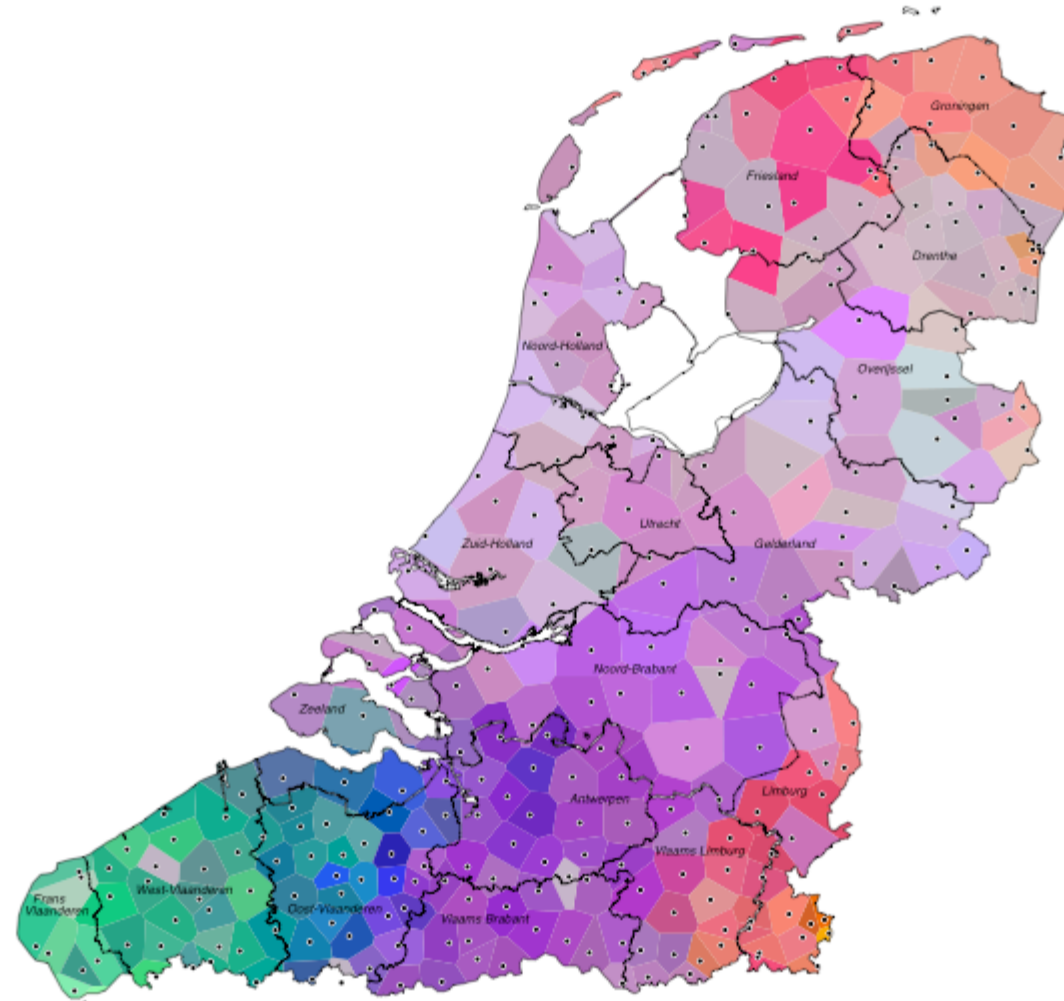
- Automatic map correction tool (*finc*)
 - The *FINe-tune postscript map Console* program rearranges all SAND location labels on the RuG/L04 maps so they don't overlap
- Uniform batch processing tool (*basc*)
 - The *BATch Script Console* program is a cross-platform script processor so only one test script needs to be maintained

Data visualisation

- Preliminary maps
 - Available SAND data not definite
 - Errors in current software

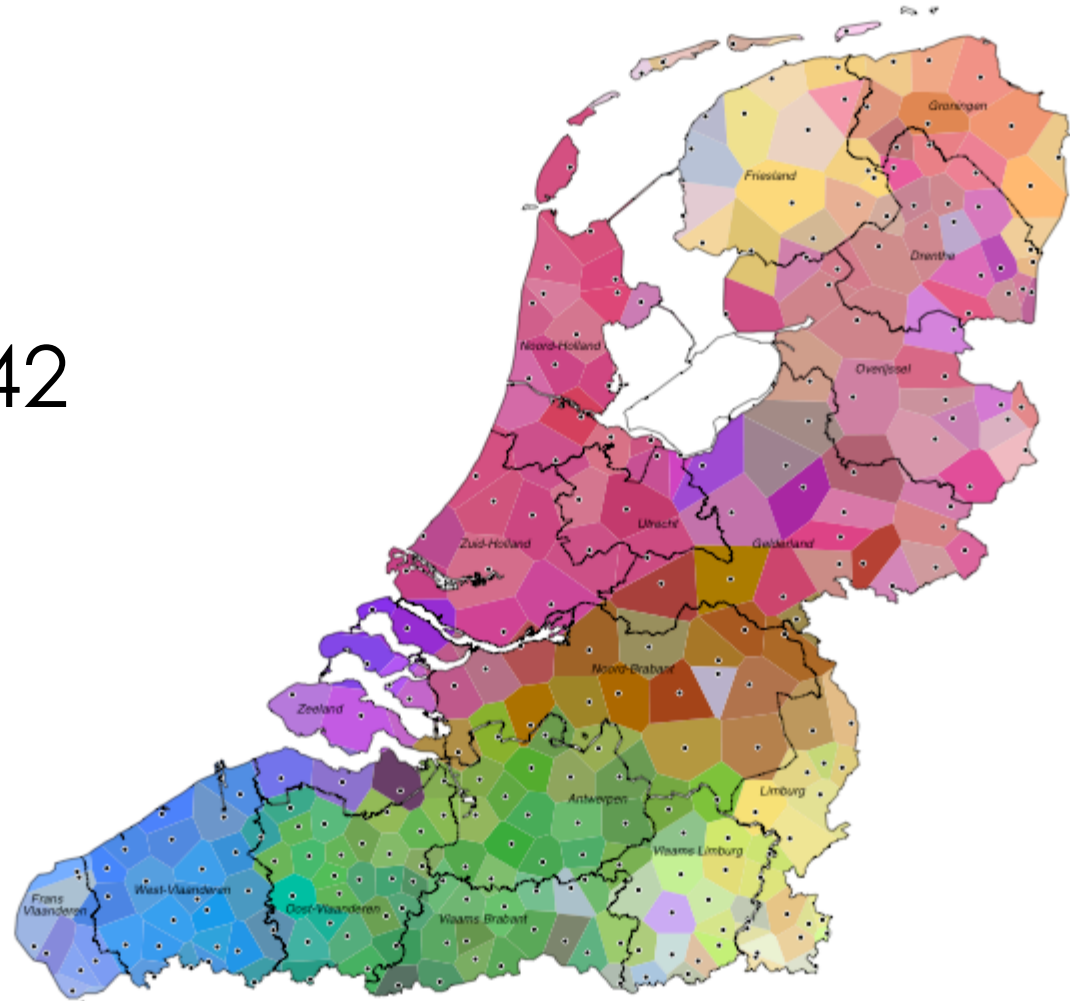
Complementisers

- 117 variants
- 279 locations
- $r = 0.9419835$



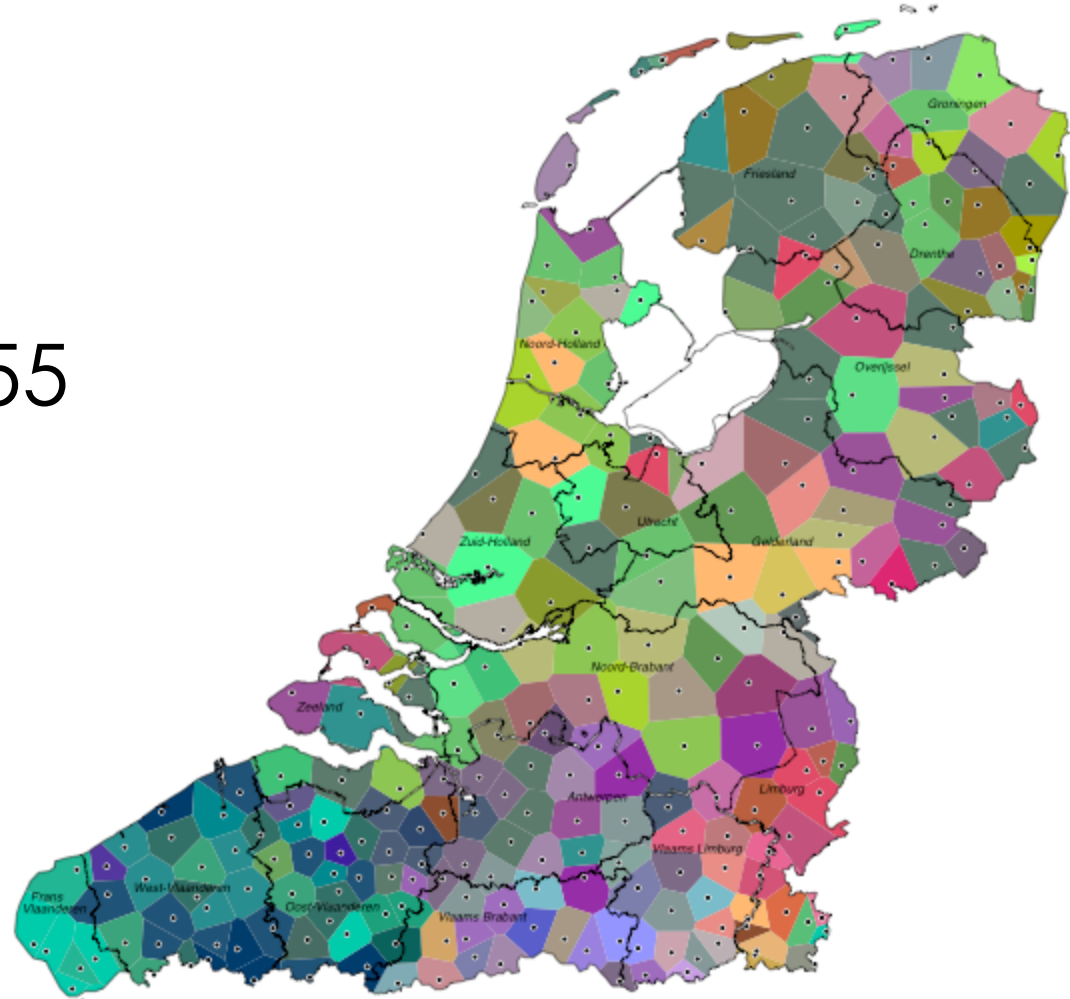
Subject pronouns

- 130 variants
- 272 locations
- $r = 0.91136942$



Expletives

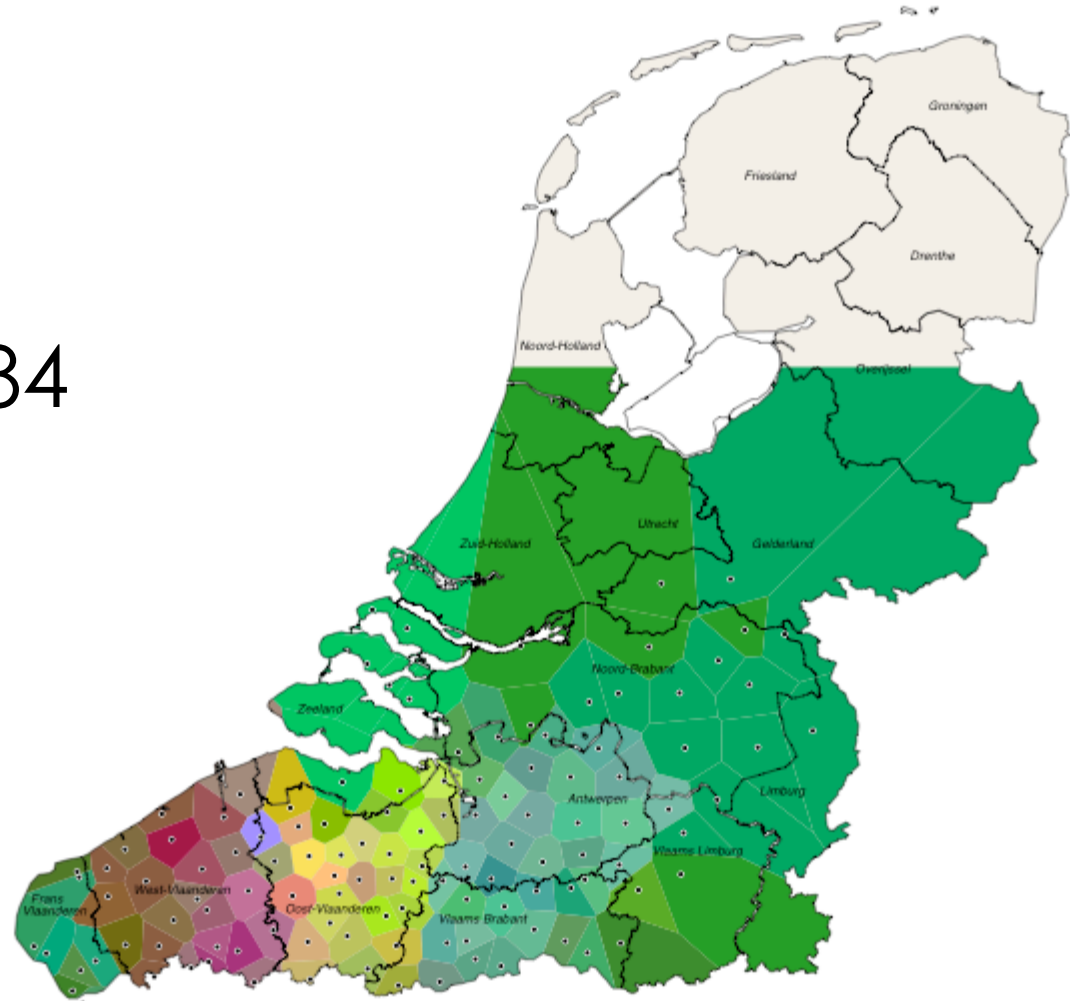
- 12 variants
- 263 locations
- $r = 0.87039055$



MRS 10/12/2004

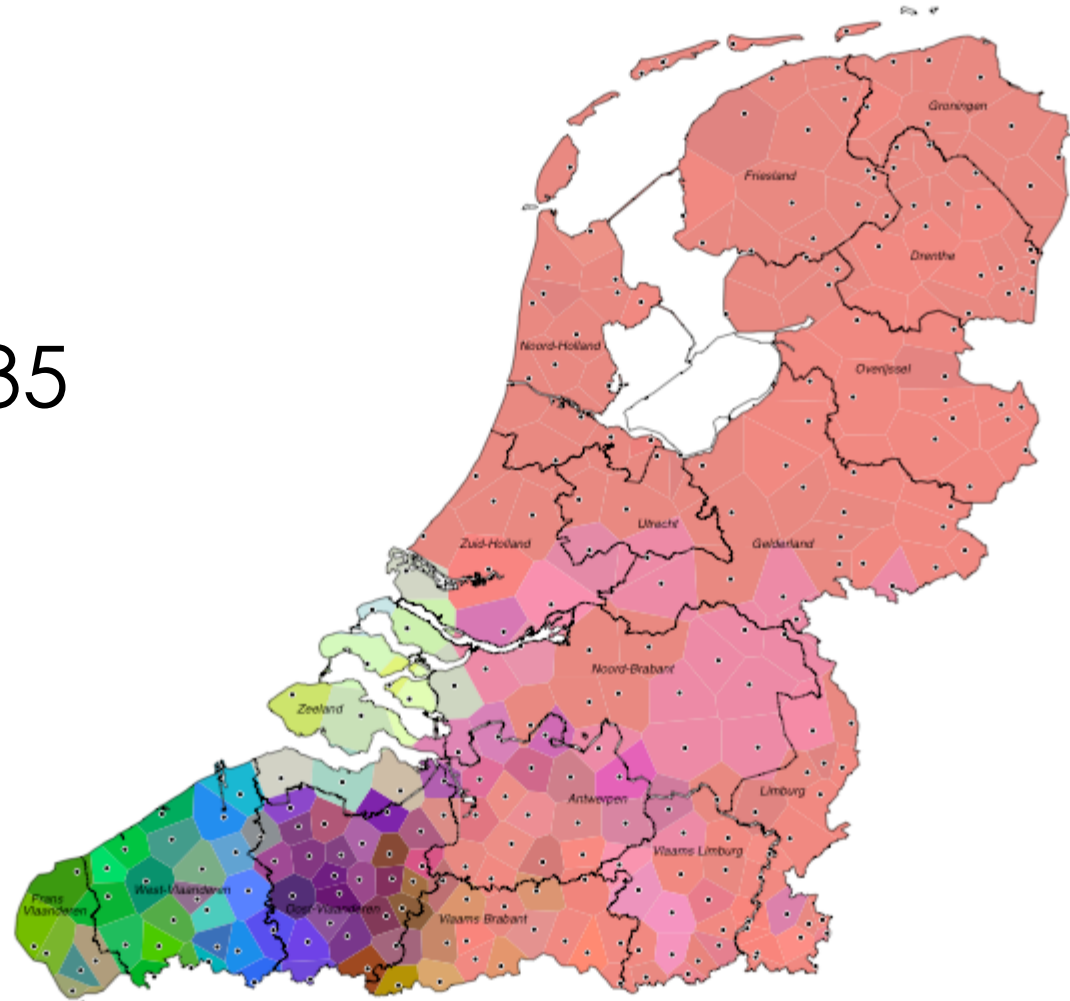
Subject doubling

- 46 variants
- 121 locations
- $r = 0.96260884$



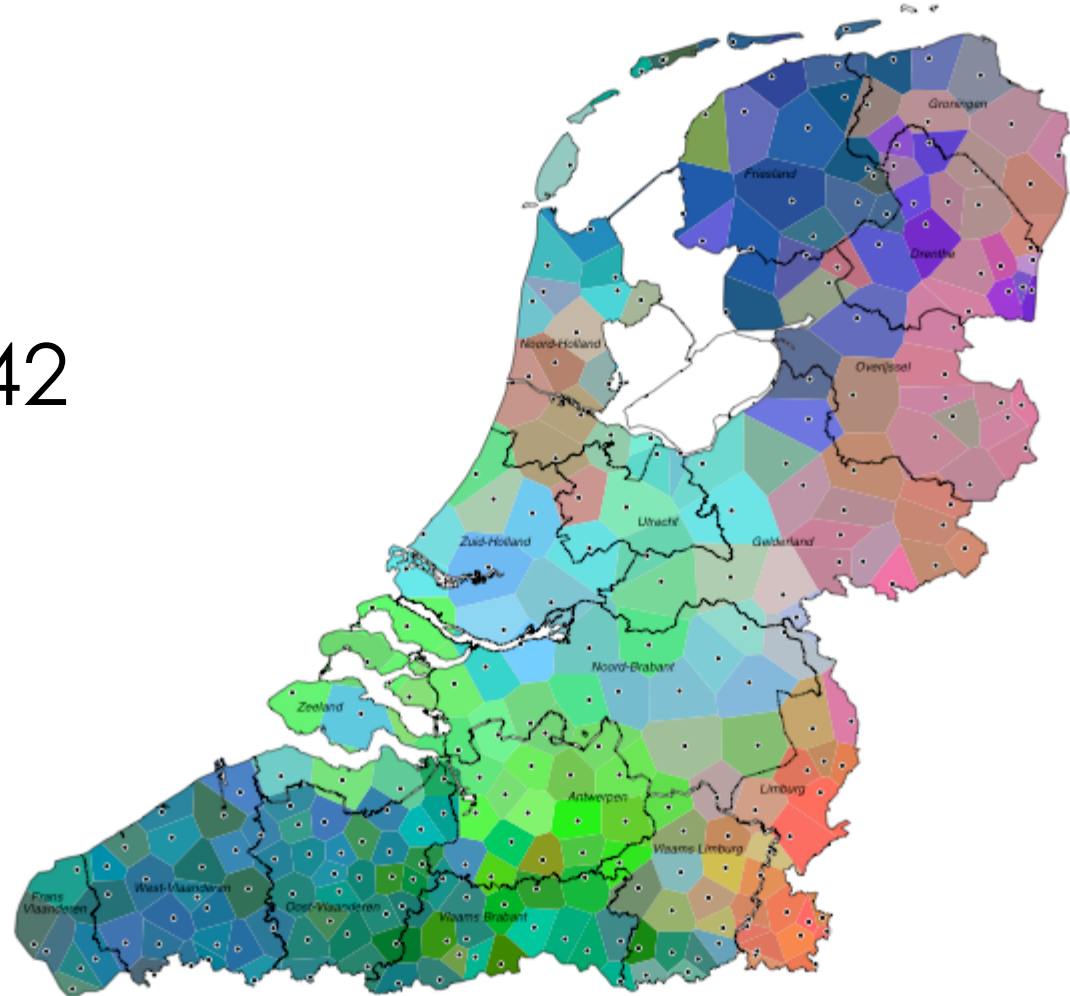
subject clitisation « yes/no

- 46 variants
- 266 locations
- $r = 0.96801635$



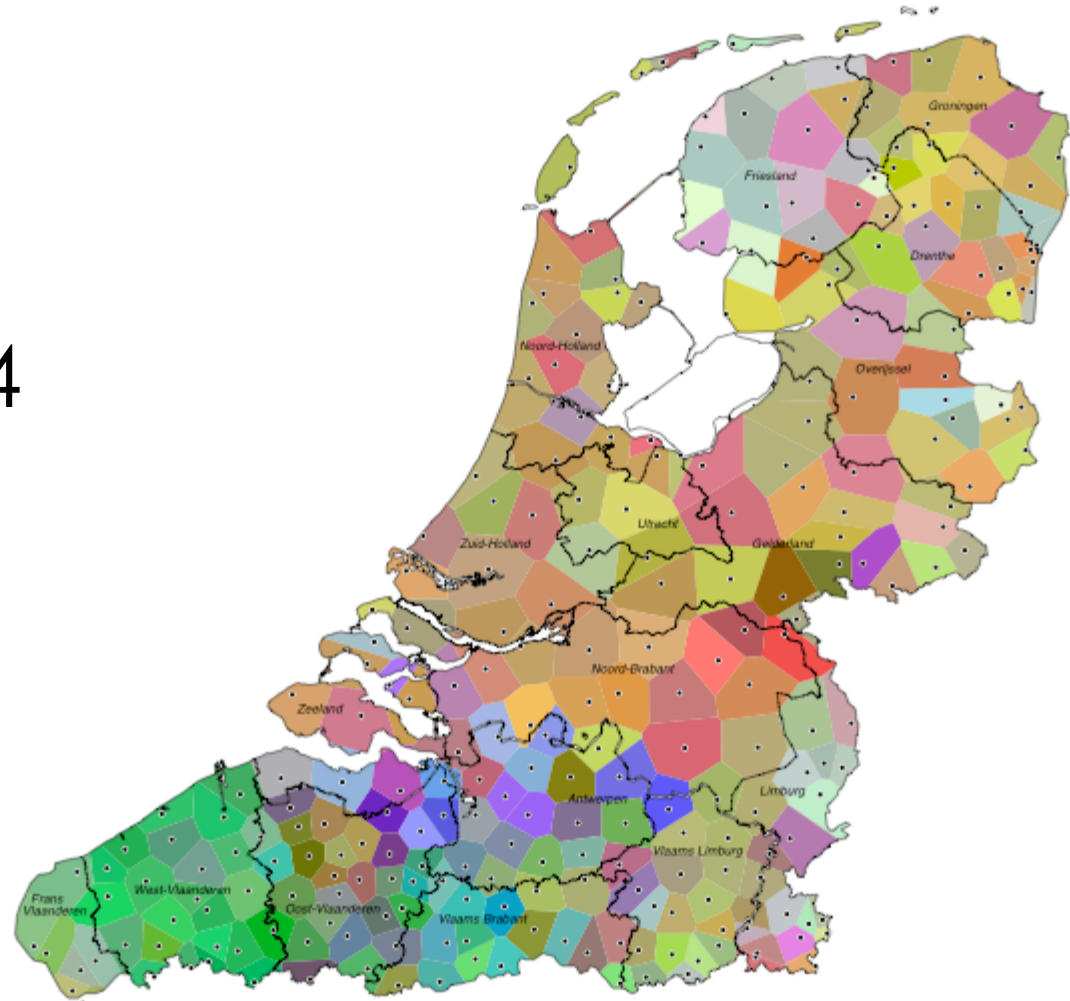
reflexive & reciprocal pronouns

- 130 variants
- 272 locations
- $r = 0.91136942$



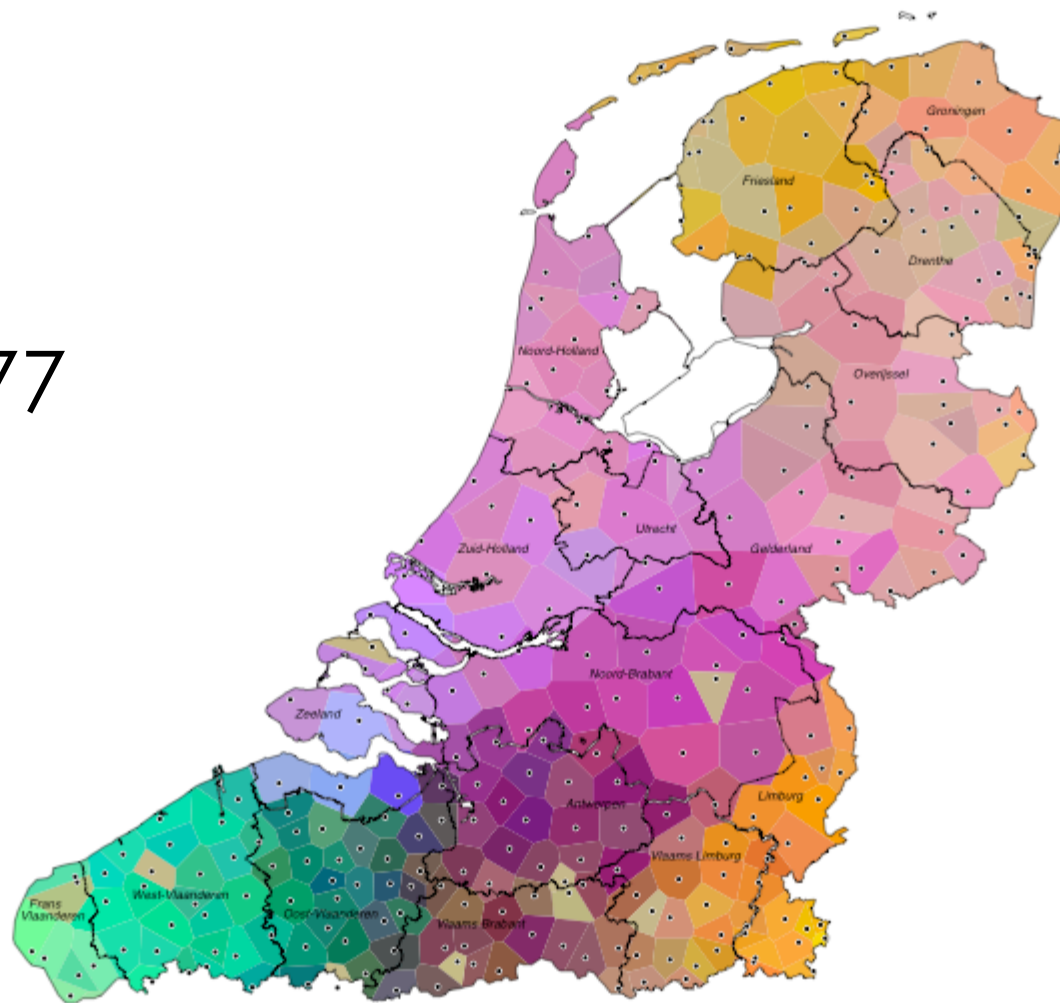
Fronting

- 66 variants
- 273 locations
- $r = 0.8146074$

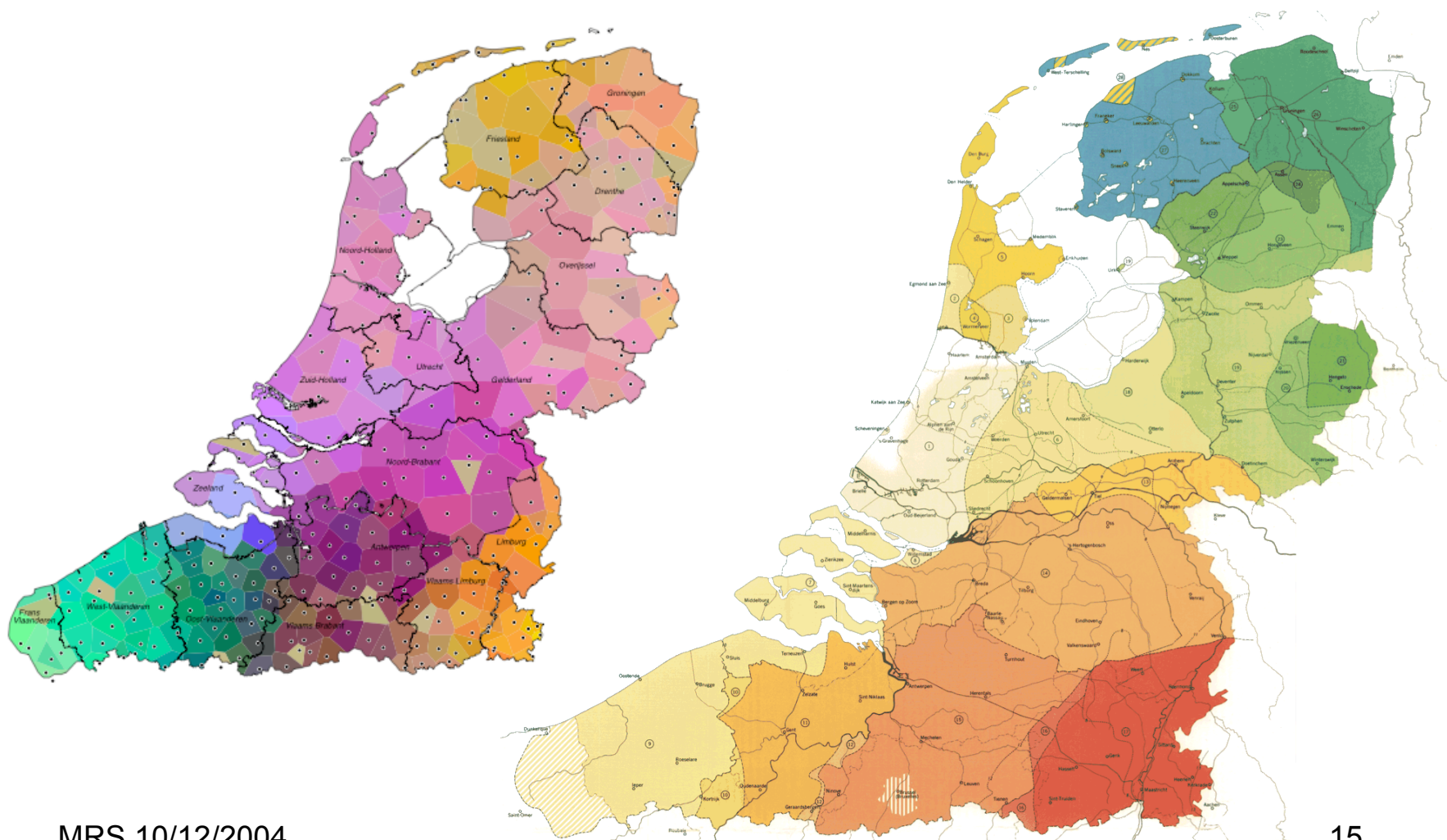


All data

- 489 variants
- 283 locations
- $r = 0.95519777$

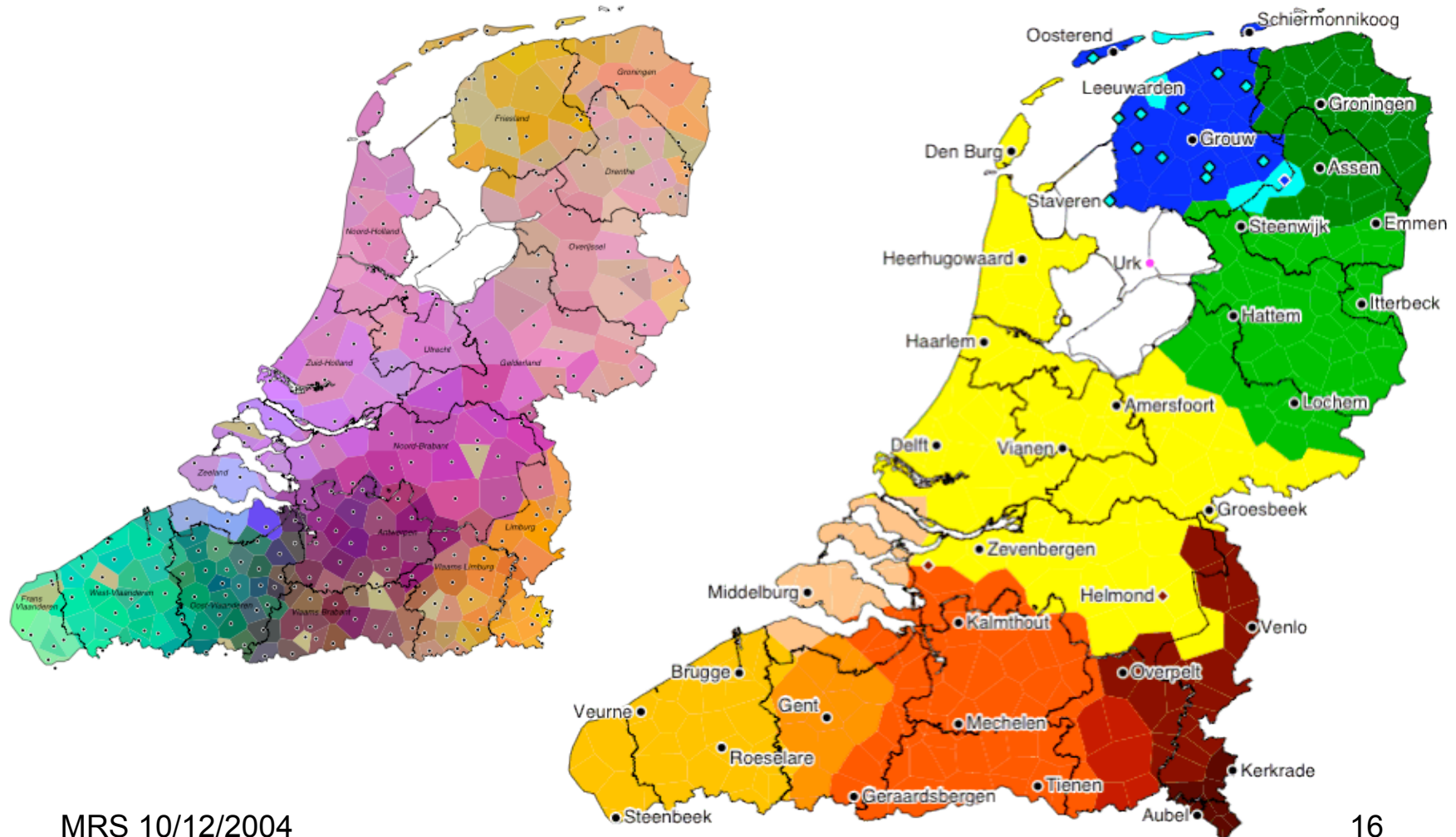


Jo Daan dialect map



MRS 10/12/2004

Heeringa dialect map



Presentations

- January 29, 2005: TIN-dag, Utrecht
- June 25, 2005: With John, at iCLaVE3 workshop *Quantitative Analysis of Language Variation in Time and Space*, Amsterdam
- August 1-5, 2005: Methods 2005 workshop *Progress in Dialectometry: Toward Explanation*, Moncton, Canada

Attention Areas

- More advanced measurements
 - Incorporation of variant frequency (GIW)
 - Accounting for variant probability (alternative variants for a feature)
 - Distance matrix generation for qualitative cluster maps
- Papers
 - Technical report, target audience: Dialectometrists
 - Overview, target audience: Linguists
- Appropriate statistics courses