

Microcomparative Morphosyntactic Research Tool

MIMORE

Sjef Barbiers

Meertens Instituut & Utrecht University

The CLARIN MIMORE team

Developers

- Matthijs Brouwer, Meertens Instituut
- Jan Pieter Kunst, Meertens Instituut
- Folkert de Vriend, Meertens Instituut

Researcher

- Sjef Barbiers, Meertens Instituut

Research object and type of data

- Variation in Natural Language
- Systematically collected data on > 600 Dutch dialects spread evenly over the language area, i.e. The Netherlands, Frisia, Dutch speaking part of Belgium, French Flanders, between 1975 and 2005
- Data on variation at the levels of sounds, words, phrases and sentences
- Sound recordings, transcriptions, POS-tagging, English glosses and translations

Research field: Linguistics

- **Phonetics:** sound system, physical properties
- **Phonology:** sound system, cognitive properties
- **Syntax:** combinatory system at the phrasal and sentential level
- **Morphology:** syntax of complex words
- **Theoretical linguistics**
- **Dialectology:** geographic distribution of language properties
- **Historical linguistics:** language change

MIMORE

- www.meertens.knaw.nl/mimore
- **Three databases**
 - (i) **SAND**: (morpho-)syntax
600MB (database) + 50 GB (sound recordings)
 - (ii) **GTR**: phonetics / (morpho-)phonology
1 GB (database) + 17 GB (sound recordings)
 - (iii) **DIDDD**: (morpho-)syntax at the level of noun phrases
120 MB (database) + 7,7 GB (sound recordings)

MIMORE: functions

- Common search engine (+ individual search engines)
- Search with text strings, POS-tag strings, linguistics properties, and combinations thereof
- Two presentation modes for search results:
 - tables
 - geographic maps
- Export as Excel, CSV and PDF
- Add to Virtual collection
 - union, intersection, complement
 - map with 2 or more linguistic variables

MIMORE demo

- **Two phenomena**

- (i) Subject Pronoun Doubling

Ik wil da **ge** **gij** gaat.
I want that you you go.2s

- (ii) Complementizer Agreement

Ik wil da-st to gaa-st
I want that.2s you go.2s

- **Research questions**

- Which dialects have these properties?
- Are there dialects that have both?
- Which of these dialects have pronoun *gij*?

MIMORE demo

(i) Search with tag constructor

Pron 2sg Pron 2sg

Add to virtual collection

(ii) Search with Syntactic phenomena

C 2sing agreement da-st

Add to virtual collection; combine in one map

(iii) Combine Pron 2sg and 'gij' in one search.

Add to virtual collection,

(iv) Create union, intersection and complement of (i) and (iii) and show maps.