

MIMORE

Demonstratieproject CLARIN-NL

Sjef Barbiers, Norbert Corver, Folkert de Vriend, Jan Pieter
Kunst



MIMORE

- MIMORE = **M**icrocomparative **M**orphosyntax **R**esearch Tool.
- Samenwerking tussen Meertens Instituut (KNAW) en Utrecht institute of Linguistics OTS.
- Doel: de ontwikkeling van onderzoeksgereedschap waarmee taalkundigen (morfo-)syntactische verschijnselen uit de resources MAND, DiDDD en DynaSAND gecombineerd kunnen interpreteren.

Resources in MIMORE

- MAND (Morfologische Atlas van de Nederlandse Dialecten)
 - Morfologische variatie op **woord**niveau.
 - Voorbeeld: *boekje* vs. *boekske*
- DiDDD (Diversity in Dutch DP Design)
 - Syntactische variatie op het niveau van **zelfstandig-naamwoordgroepen**.
 - Voorbeeld: *hem zijn boek* vs. *zijn boek*
- DynaSAND (Dynamische Syntactische Atlas van de Nederlandse Dialecten)
 - Syntactische variatie op **zins**niveau.
 - Voorbeeld: variatie in woordvolgorde op het einde van bijzinnen: *(dat hij) moet kunnen zwemmen* vs. *zwemmen kunnen moet*.

Voorbeeld van onderzoek dat ondersteund wordt met MIMORE (1)

- Er is aanleiding te denken dat de voornaamwoorden van de tweede persoon (*jouw, jullie*) geen onanalyseerbare **woorden** zijn, maar dat het **woordgroepen** zijn die bestaan uit twee of meer woorden of morfemen in een hiërarchische structuur (vergelijkbaar met de structuur van een zelfstandig-naamwoordgroep).
- Zijn daar bewijzen voor? En hoe zit die hiërarchische structuur dan in elkaar?

Voorbeeld van onderzoek dat ondersteund wordt met MIMORE (2)

Met MIMORE wordt het mogelijk om de variatie die voornaamwoorden van de tweede persoon vertonen **in isolatie** (MAND) te vergelijken met de variatie die ze vertonen in verschillende **syntactische contexten**, namelijk:

- Binnen zinnen (DynaSAND)
- Binnen zelfstandig-naamwoordgroepen (DiDDD):

jouw opa

Vorstenbosch: *ouw opa*

Zoutleeuw: *oere bompa*

Axel: *joen opa*

jullie opa

Vorstenbosch: *hullie én opa*

Zoutleeuw: *oerlingze bompa*

Axel: *judder opa*

Welke technologie willen we hiervoor ontwikkelen binnen MIMORE?

- Zoekmachine
- Kaartapplicatie
- Statistiekmodule^{Light}

De technologie zal via een **web interface** aan eindgebruikers beschikbaar gesteld worden. Waar wenselijk zal de technologie ook als **web service** aangeboden worden.

Welke databewerkingen zijn hiervoor nog nodig?

- MAND en DiDDD worden automatisch van **POS-tags** voorzien en vervolgens handmatig gecorrigeerd. De POS-tagging van de DynaSAND is op dit moment al in de afrondende fase. Voor alle drie de resources wordt dezelfde tag set gebruikt (ISOcat)
- We bekijken of het voor de MAND en DiDDD haalbaar is om de opnames met dialectspraak (semi-)automatisch **op te lijnen** met de transcripties. De opnames van de DynaSAND zijn al opgelijnd.
- De resources zullen voorzien worden van **Persistent Identifiers**.
- Zowel de resources als de tools en web services zullen voorzien worden van **CMDI-metadata**

Tijd en mens

- Looptijd project: 1 april tot 1 november 2010
- Betrokkenen:
 - Sjef Barbiers (Projectcoördinator, Meertens Instituut)
 - Norbert Corver (Universiteit Utrecht, UiL-OTS)
 - Jan Pieter Kunst (Meertens Instituut)
 - Matthijs Brouwer (Meertens Instituut)
 - Folkert de Vriend (Meertens Instituut)

Dank