



Syntactic Atlas of the Dutch Dialects

Sjef Barbiers,

Meertens Institute (KNAW) & Utrecht University



Coordination

- Hans Bennis (Meertens Institute)
- Hans den Besten (University of Amsterdam)
- Magda Devos (University of Gent)
- Johan Rooryck (Leiden University)
- Johan Van der Auwera (University of Antwerp)



SAND team

project leader

Sjef Barbiers (Meertens Institute)

research

Jeroen van Craenenbroeck (Leiden University)

Marjo van Koppen (Leiden University / Meertens Institute)

Annemie Neuckermans (University Antwerpen)

Guido Vanden Wyngaerd (University Antwerpen)

Gunther de Vogelaer (University Ghent)

Henk Wolf (Fryske Akademy)

Hedde Zeijlstra (University of Amsterdam)

methodology

Leonie Cornips (Meertens Institute)

Willy Jongenburger (Meertens Institute)



SAND team

general assistance

Margreet van der Ham (Meertens Institute)

Irene Haslinger (Meertens Institute)

Vicky van den Heede (University Ghent)

Mathilde Jansen (Meertens Institute)

Alies MacLean (Meertens Institute)

Vivien Waszink (Meertens Institute)

technical assistance

Ilse van Gemert (Meertens Institute)

Jan Pieter Kunst (Meertens Institute)



Financial support

- VNC (Flemish-Dutch committee)
- KNAW (Royal Dutch Academy of Arts and Sciences)
- NWO (Dutch Science Foundation)
- FWO (Flemish Science Foundation)
- Meertens Institute (KNAW)
- Universities of Amsterdam, Antwerpen, Ghent & Leiden



Perspectives

theoretical perspective

- generative syntax
- language typology
- dialectology

empirical perspective

- dialect geography: Dutch speaking area in the Netherlands and Belgium



Different theoretical perspectives

Advantages

- Variation predicted by (generative) theory
- Variation to be expected on the basis of (typological) cross-linguistic generalizations
- Variation known to exist on the basis of dialectological and diachronic research



Different theoretical perspectives

Disadvantages

- Different views on the importance of a particular phenomenon
- Different views on which aspects should have priority in the description of a phenomenon
- Different views on what counts as an adequate “theory-neutral” description of the facts
- Manuals, papers, talks etc should be made accessible to a general linguistic audience



Empirical domains

- *Left-periphery of the clause*
e.g., complementizer agreement, subject doubling, relative clauses, Wh-clauses
- *Right-periphery of the clause*
e.g., verbal clusters, verb cluster interruption, IPP
- *Negation and quantification*
e.g., negative particle, negative concord, scope, negative quantifiers
- *Pronominal reference*
e.g., reflexives, reciprocals, weak and strong pronouns



Advantages

- Too much variation to cover everything
- In-depth investigation instead of broad and shallow
- Choice based on existing knowledge about variation, thus reduced risk of choosing a domain that shows no variation
- Domains that are theoretically interesting

~ [SAND] ... *Restricted empirical domains*

Disadvantages

- Risk to miss interesting variation in another domain
- (If you believe that dialectal variation is rapidly disappearing:) Missing the last chance to record dialectal variation



Project stages

1. Inventorization (2000)
2. Postal pilot (2000)
3. Fieldwork (2001-2002)
4. Telephone interviews (2003)
5. Data processing, storage, retrieval and visualization (2003-2004)



Stage 1: Inventorization

- SAND bibliography: 1300 titles
- Interviews with (dialect speaking)
linguists

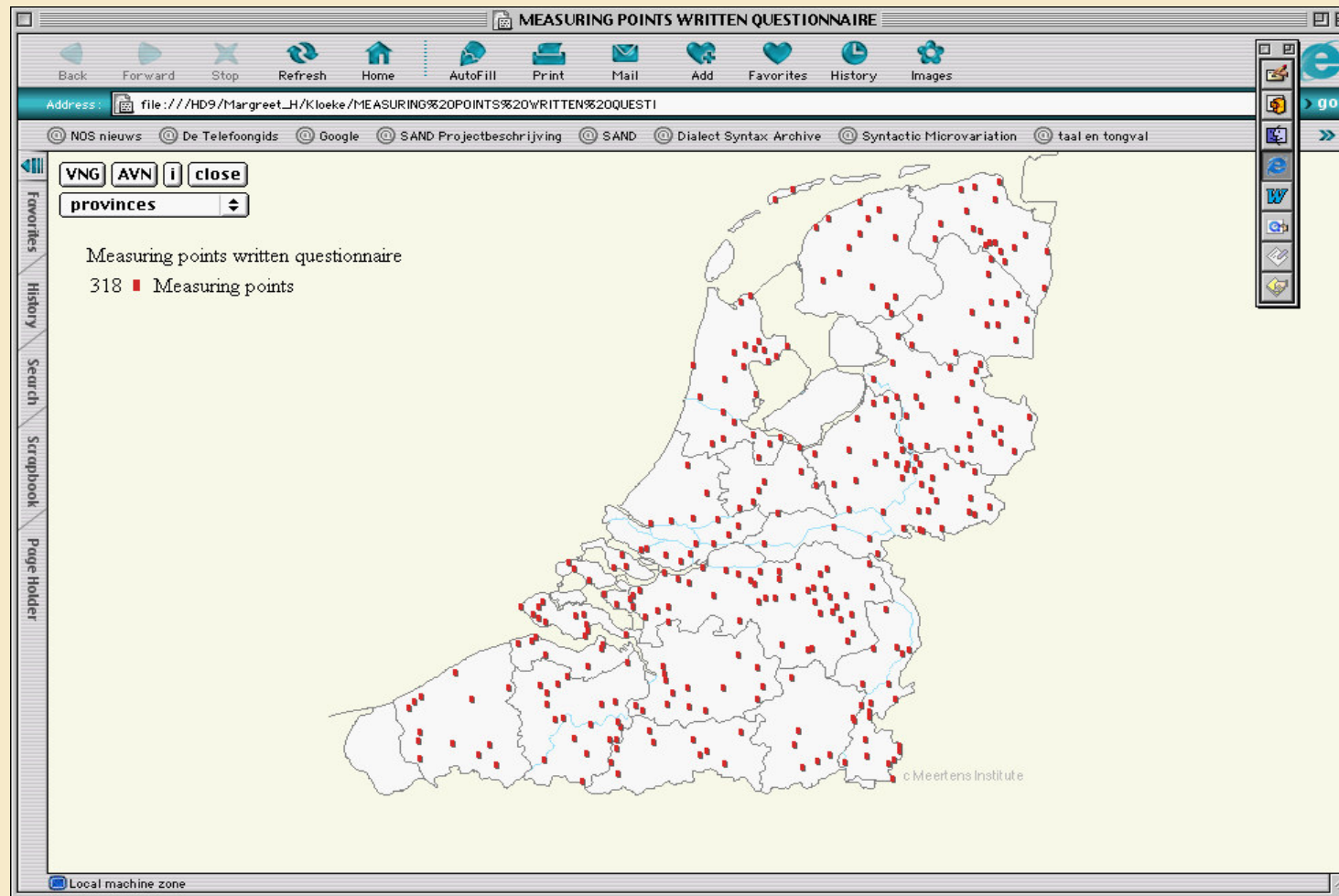


Stage II: Postal pilot (1)

- 321 measuring points (368 informants), reasonably even distribution over language area



Stage II: Postal pilot (2)



Map of measuring points / written questionnaire



Stage II: Postal pilot (3)

- Not controlled for social variables
- 393 test sentences
- Goals:
 - inventory of syntactic variables
 - first impression of distribution of variables
 - try out of different question types

EXAMPLE:



Stage II: Postal pilot (5)

Advantages

- Number of sentences that can be tested is larger
- Allows testing of sentences with a higher complexity
- Opportunity to test different methods of elicitation
- Yields data that the data from the oral interviews can be compared with
- Makes oral interviews more efficient: regionalized and multi-stage questions
- Less time and people consuming than oral interviews



Stage II: Postal pilot (6)

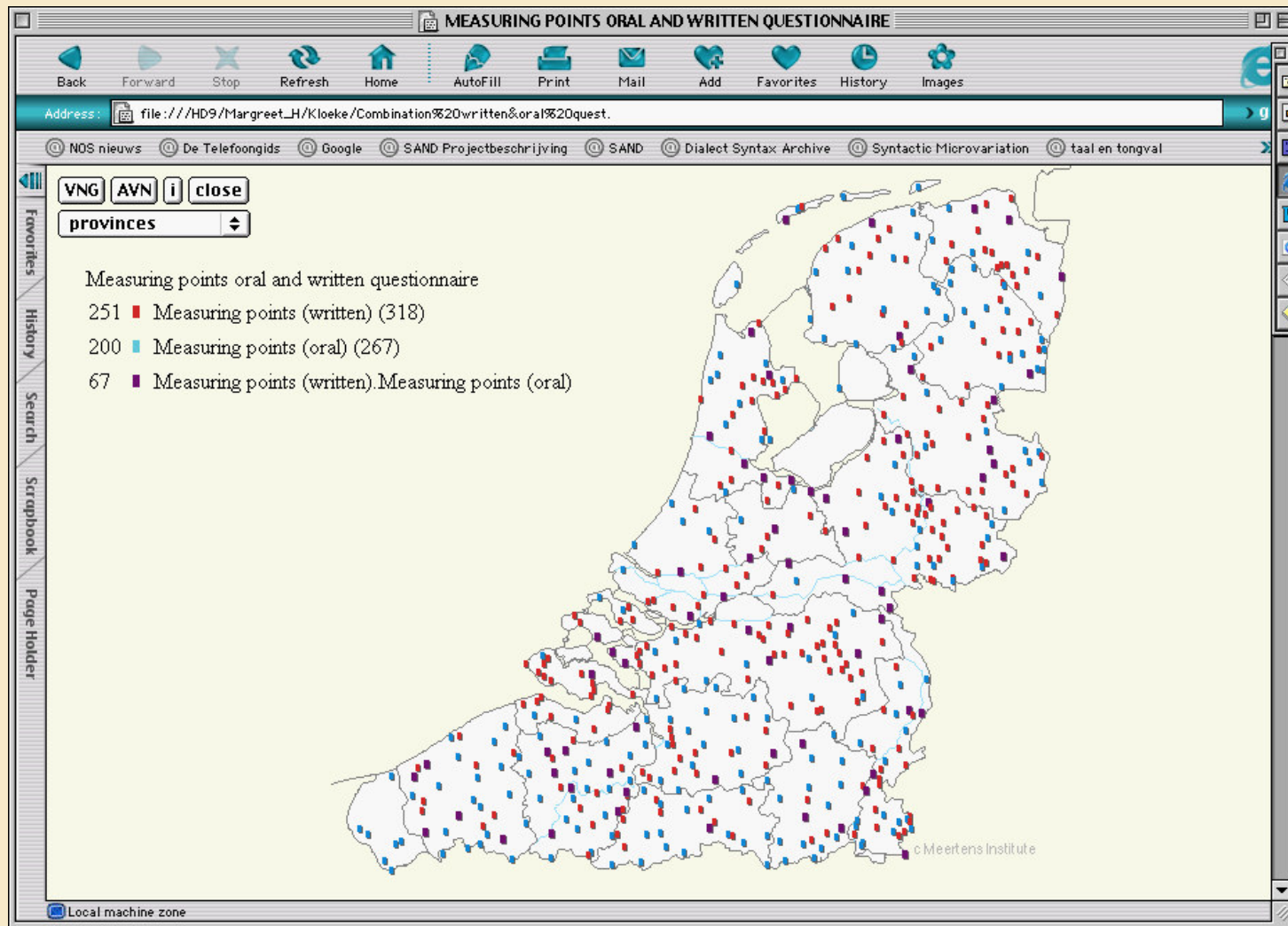
Disadvantages

- Validity of the data not always clear, e.g. a sentence may be rejected on phonological or lexical grounds
- Informants have to invent their own orthography. Syntactically relevant sounds may be omitted
- Written mode may trigger more formal, hence less dialectal behavior
- Impossible to observe and immediately respond to answers and reactions of informants



Stage III: Fieldwork (1)

- 267 oral interviews
(average length 1h 45 min)



measuring points oral + written questionnaire (2)



Stage III: Fieldwork (3)

Methodology

- Informants controlled for social variables
 - » 55-70 years old
 - » born and raised in location of interview
 - » parents born and raised in location of interview
 - » no longer than 7 years from home
 - » no higher education
 - » active dialect user in at least one social domain
- Both interviewer and informant are native speakers of the relevant dialect (to avoid accommodation); field worker as supervisor
- Regionalized and multi-stage questions



Stage III: Fieldwork (4)

Methodological problem

- Uniformity of methodology, given different traditions in different frameworks



Stage III: Elicitation tasks

⑩ Masked (relative) grammaticality judgements

- (i) Ik vind dat iedereen moet kunnen zwemmen.
I think that everyone must can.INF swim.INF
'I think that everybody should be able to swim.'
- (ii) Ik vind dat iedereen moet zwemmen kunnen.
- (iii) Ik vind dat iedereen zwemmen kunnen moet.

⑩ Translation tasks

- (iv) Als we sober leven, leven we gelukkig.
If we austerely live, live we happily
'If we live austerely, we will be living happily.'

⑩ Cloze tests

- (v) Jan wast
- John washes



Stage III: Fieldwork (6)

Mondelinge Vragenlijst.pdf

512 z15a *Zo'n ding een(e) heb ik nog nooit gezien!*

1. Prioriteitsklasse
Hoog

2. Trefwoorden

388	algemeen	DP ev o	
	grammaticaal		
337	Structuur DP	lidwoord	een-dubbeling

3. Waar opvragen?
K, L, Q

4. Trap
Als voorkomt, dan 513, 514 (z15b-c) opvragen

5. Instructies
Als voorkomt, controleren of 'een(e)' na 'ding' ook weggelaten kan worden

6. Vraagtype

- Komt deze zin voor in uw dialect?
- Hoe gebruikelijk is deze zin in uw dialect?
- Vertaal
- Vertaal op zoveel mogelijk manieren
- Plaatje

513 z15b *Zo een vrouw een(e) kun je maar beter niet tegenspreken*

1. Prioriteitsklasse

106 of 166 8.26 x 11.69 in



Stage III: Fieldwork (7)

Advantages

- More natural communicative setting
- Spoken data
- Possibility to observe and respond to answers and reactions of informants
- Greater validity of the data; less accommodation



Stage III: Fieldwork (8)

Disadvantages

- More time and people and money consuming than written interviews
- Sentences that require longer reflection are more difficult to test



Stage IV: Telephone interviews

- Testing sentences that did not yield a (relevant) answer
- Checking the reliability of certain answers
- Additional test sentences to get complete paradigms



Stage V: Data management

- i. Data processing and enrichment
 - Digitalization
 - Transcription
 - POS Tagging
- ii. Data storage
- iii. Data retrieval
- iv. Data visualization

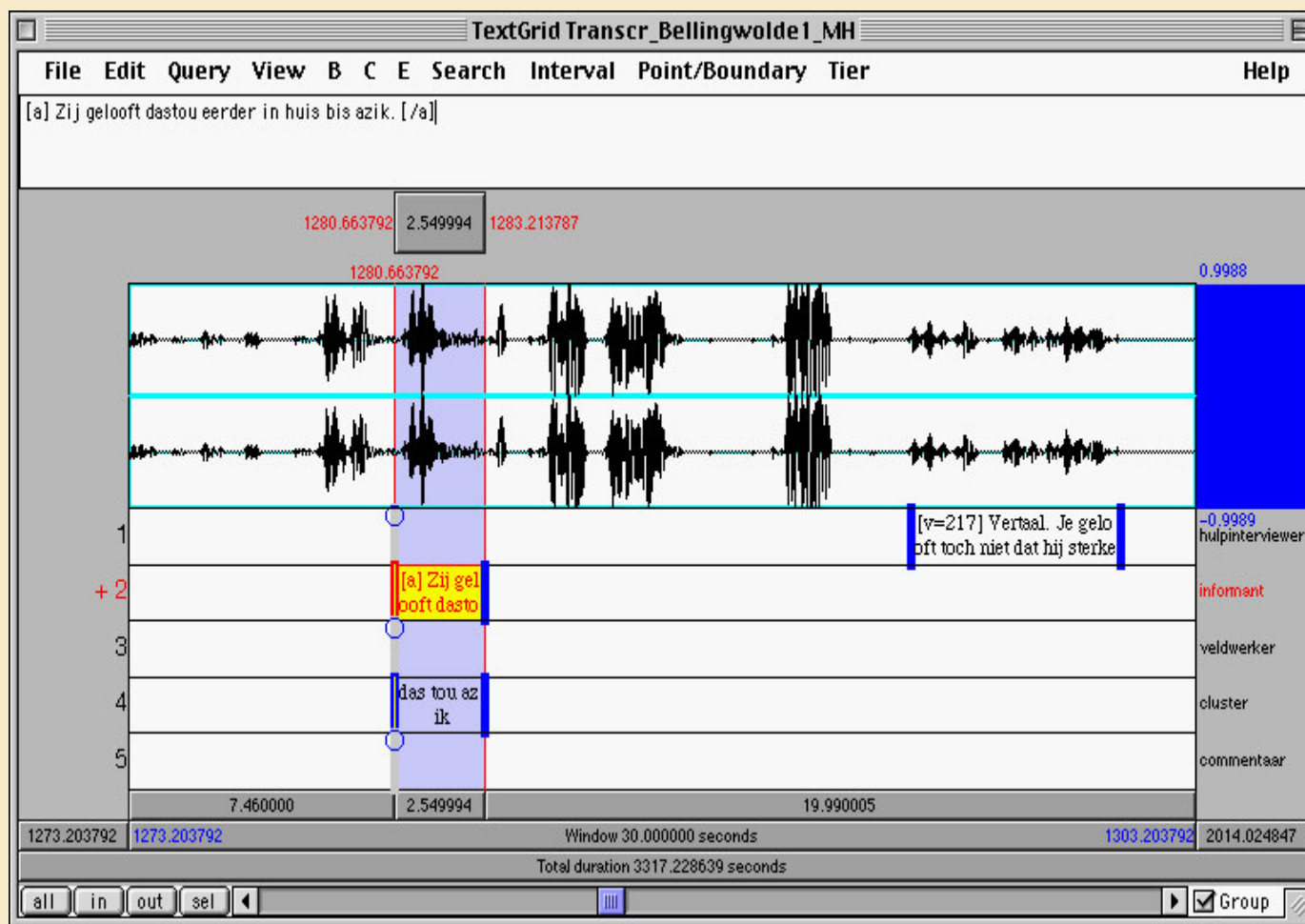


Data processing

- Recordings read into computer without conversion (system used: Sadie DAW. Sample frequency: 44.1 kHz, 16 bits)
- Digital recordings (DAT recorder /minidisk)
- Transcription with PRAAT (Boersma, University of Amsterdam)



Transcription (1)



Transcription in PRAAT (Boersma)



Transcription (2)

Issues

- Phonetic transcription requires too many resources
- Orthographic transcription obscures all kinds of relevant (morpho-)syntactic properties
- There are no orthographic conventions for most of the dialects
- Morpheme and word boundaries are often unclear (e.g. in clitic clusters)



Transcription (3)

Choices

- Compromise between orthographic and literal transcription
- Literal transcription of function words including (potential) inflection
- Standardization of lexical dialect words (advantageous for automatic lemmatization and POS tagging)
- Clusters are transcribed unsegmented and segmented



POS Tagging (1)

- Automatic pretagging
 - memory-based probabilistic tagger (Van den Bosch, University of Tilburg)
- Manual correction and completion
- Extended tag set (EAGLES standard)



POS Tagging (2)

Interval 10 uit tier 'informant' van file Tra...

Address: http://www.p.jmi.knaw.nl/sand/tagging/tag_interval.php?interval_id=17492

Interval 10 uit tier 'informant' van file Transcr_Bellingwolde1_MH.TextGr

[a] Vertel maar niet wel ze roep kend ha. [a]

	CGN-tags	SAND-woordsoort	toegekende SAND-tags
vertel	<input checked="" type="radio"/> WW(pv,tgw,ev) vertellen 100.00 <input type="radio"/> anders	<input type="text" value=""/>	<input type="text" value="V(inv,tt,imp,2,s,trans)"/>
maar	<input checked="" type="radio"/> BW() maar 95.84 <input type="radio"/> SPEC(meta) _ 0.00 <input type="radio"/> N(soort,ev,basis,onz,stan) maar 0.00 <input type="radio"/> N(soort,ev,basis,zijd,stan) maar 0.00 <input type="radio"/> VG(neven) maar 0.00 <input type="radio"/> anders	<input type="text" value="ADV"/>	<input type="text" value=""/>
niet	<input checked="" type="radio"/> BW() niet 99.98 <input type="radio"/> SPEC(meta) _ 0.00 <input type="radio"/> N(soort,ev,basis,onz,stan) niet 0.00 <input type="radio"/> N(soort,ev,basis,zijd,stan) niet 0.00 <input type="radio"/> WW(pv,tgw,ev) nieten 0.00 <input type="radio"/> anders	<input type="text" value="ADV"/>	<input type="text" value=""/>
wel	<input checked="" type="radio"/> BW() wel 99.98 <input type="radio"/> N(soort,ev,basis,onz,stan) wel 0.00 <input type="radio"/> N(soort,ev,basis,zijd,stan) wel 0.00 <input type="radio"/> TSW() wel 0.00 <input type="radio"/> WW(pv,tgw,ev) wellen 0.00 <input type="radio"/> anders	<input type="text" value="ADV"/>	<input type="text" value=""/>
ze	<input checked="" type="radio"/> VNW(pers,pron,stan,red,3,ev,fem) ze 93.52 <input type="radio"/> VNW(pers,pron,stan,red,3,mv) ze 6.12 <input type="radio"/> SPEC(meta) _ 0.00 <input type="radio"/> TSW(dial) ze 0.00 <input type="radio"/> anders	<input type="text" value="PRON"/>	<input type="text" value=""/>
roep	<input checked="" type="radio"/> N(soort,ev,basis,zijd,stan) roep 100.00 <input type="radio"/> WW(pv,tgw,ev) roepen 0.00 <input type="radio"/> anders	<input type="text" value="N"/>	<input type="text" value=""/>

Internet zone

Tagging (Kunst)



POS Tagging (3)

Interval 10 uit tier 'informant' van file Tra...

Address: http://www.p.jmi.knaw.nl/sand/tagging/tag_interval.php

NOS nieuws De Telefoongids Google SAND Projectbeschrijving SAND Dialect Syntax Archive Syntactic Microvariation taal en tongval

Interval 10 uit tier 'informant' van file Transcr_Bellingwolde1_MH.TextGr

[a] Vertel maar niet wel ze roep kend ha. [a]

	SAND-tag	lemma
vertel	V(inv,tt,imp,2,s,trans)	vertellen
maar	ADV INFL vrij TYPE VORM	maar
niet	ADV INFL vrij kwant-neg VORM	niet
wel	PRON INFL N CASUS 3 s GENUS d-obj wh VORM sterk simpl DUB	wie
ze	PRON INFL vrij CASUS 3 s v subj pers VORM zwak simpl DUB	ze
roep	V INFL eind3 PART WIJZE PERS GETAL AUX TYPE	roepen
kend	V INFL eind2 PART WIJZE PERS GETAL AUX TYPE	kennen
ha	V INFL POS FIN WIJZE PERS GETAL AUX TYPE	ha

Internet zone

- stage 4.3 Tagging-2 (Kunst)



POS Tagging (4)

Issues

- EAGLES Standard does not make sufficient distinctions
- Automatic taggers require sufficiently large training set for each dialect
- Fine grained POS tagging desirable but very time (people, money) consuming and tedious (if done manually)
- Incorporation of function and word order information necessary if syntactic annotation is not feasible



*Data storage, retrieval
and visualization*

Dynamic Atlas ([DynaSAND](#))

- database
- search engine
- cartographic tool



DynaSAND

Issues

- Freeware / open source
- Flexibility: relational database
Database: MySQL
- Accessibility: on-line
Web interface: PHP
- Compatibility with other databases (common search engine)
- Cartographic tool: SVG



SAND Results



Results

- [DynaSAND](#): On-line Dynamic Atlas (2006)
- Syntactic Atlas of the Dutch Dialects,
Volume 1; 2005
- Syntactic Atlas of the Dutch Dialects,
Volume 2; 2008



DynaSAND

On-line research tool

- ⑩ Database
- ⑩ Search engine
- ⑩ Cartographic tool



DynaSAND

Database

Postal pilot

156,000 Q-A-pairs
368 informants
393 testzinnen

Fieldwork

45.000 Q-A-pairs (= 425 hours of speech
= 1.200.000 words in transcription)
607 informants in 267 locations in The
Netherlands, Belgium and France

Telephone interviews

26.000 Q-A-pairs
252 informants/locations
105 test sentences



DynaSAND

Search engine

Searching with locations, test sentences, (combinations of) words, strings, tags

Cartographic software

Generating maps to visualize geographic distribution of and potential correlations between syntactic variables



SAND Vol. I

Barbiers, S., H. Bennis, G. De Vogelaer, M. Devos, M. van der Ham
SAND Vol. 1 (Amsterdam University Press, 2005)
145 maps + book (80 p.); English & Dutch

Ch. 1: Complementizers and complementizer agreement

Ch. 2: Subject pronouns

Ch. 3: Subject doubling and clitisation to *yes/no*

Ch. 4: Reflexive and reciprocal pronouns

Ch. 5: Fronting



SAND Vol. II

Barbiers, S., J. Van der Auwera, H. Bennis, Eefje Boef, G. De Vogelaer., M. van der Ham

SAND Vol. 2 (Amsterdam University Press, 2008;
87 maps, 78 pages)

Ch. 1: Verb clusters

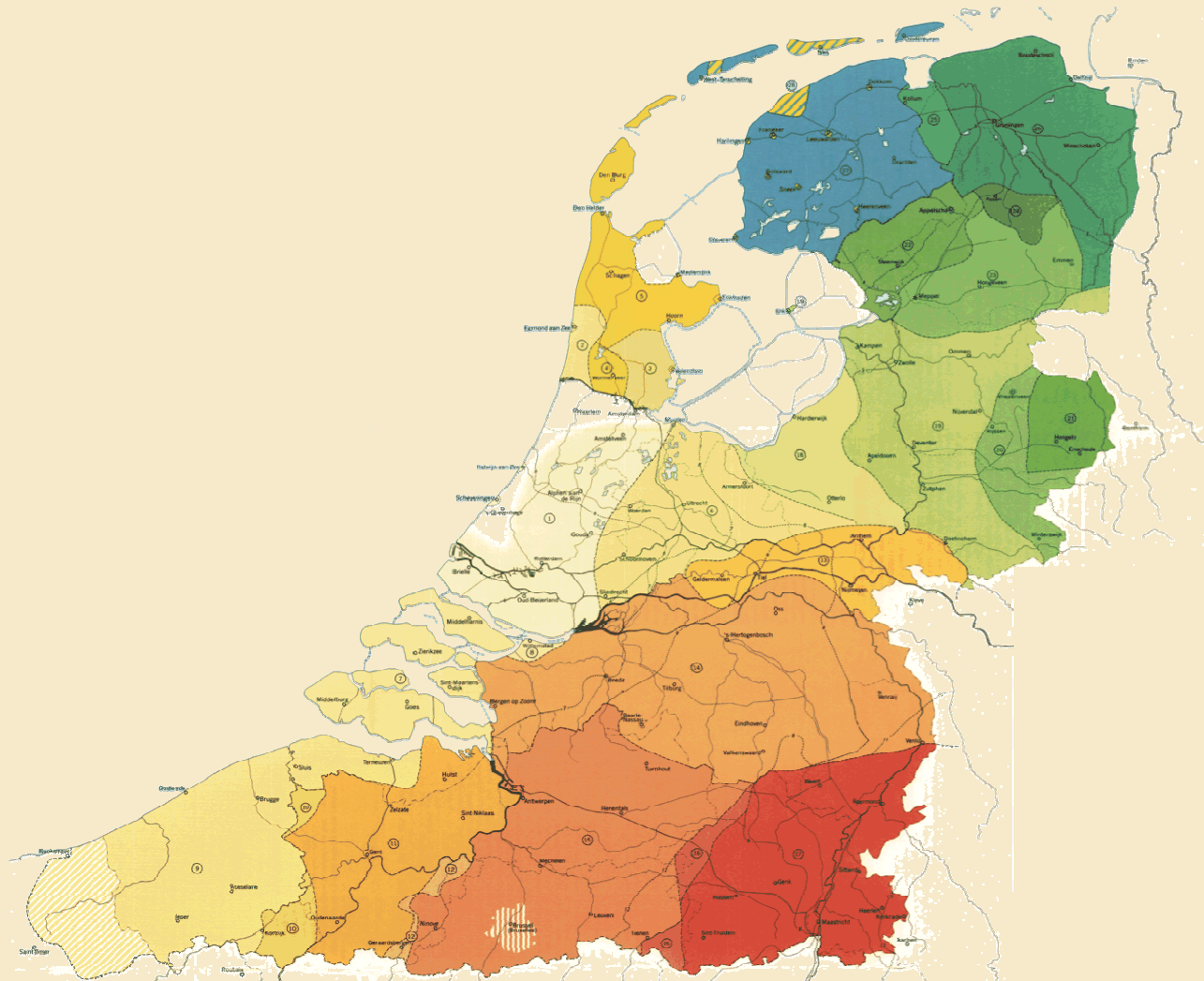
Ch. 2: Verb cluster: interruption and morphosyntax

Ch. 3: Auxiliaries and verb positions

Ch. 4: Negation and quantification



Psychological dialect map





Complementizer agreement

- (1) Ik vind da-**st** tu het niet zien mag-**st**
 I think that.2SING you it not see may.2SING
 ‘In my opinion you are not allowed to see it.’ (Gronings)
- (2) As-**e** me sober leev-**e** [...]
 if.1PL we austerely live.1PL [...]
 ‘If we live austerely, [...]’ (South-Hollandic)
- (3) We peize me da-**n** ze niet zo slim zij-**n**.
 we think we that-3PL they not so smart are
 ‘We think that they are not so smart.’ (East-Flemish)

Koppen, M. van (2005). *One probe, two goals. Aspects of Agreement in Dutch Dialects*. Leiden, LOT Dissertation 105.

De Vogelaer, G. (2005). *Persoonsmarkering in de dialecten in het Nederlandse taalgebied*. Dissertation Ghent.



Complementizer agreement

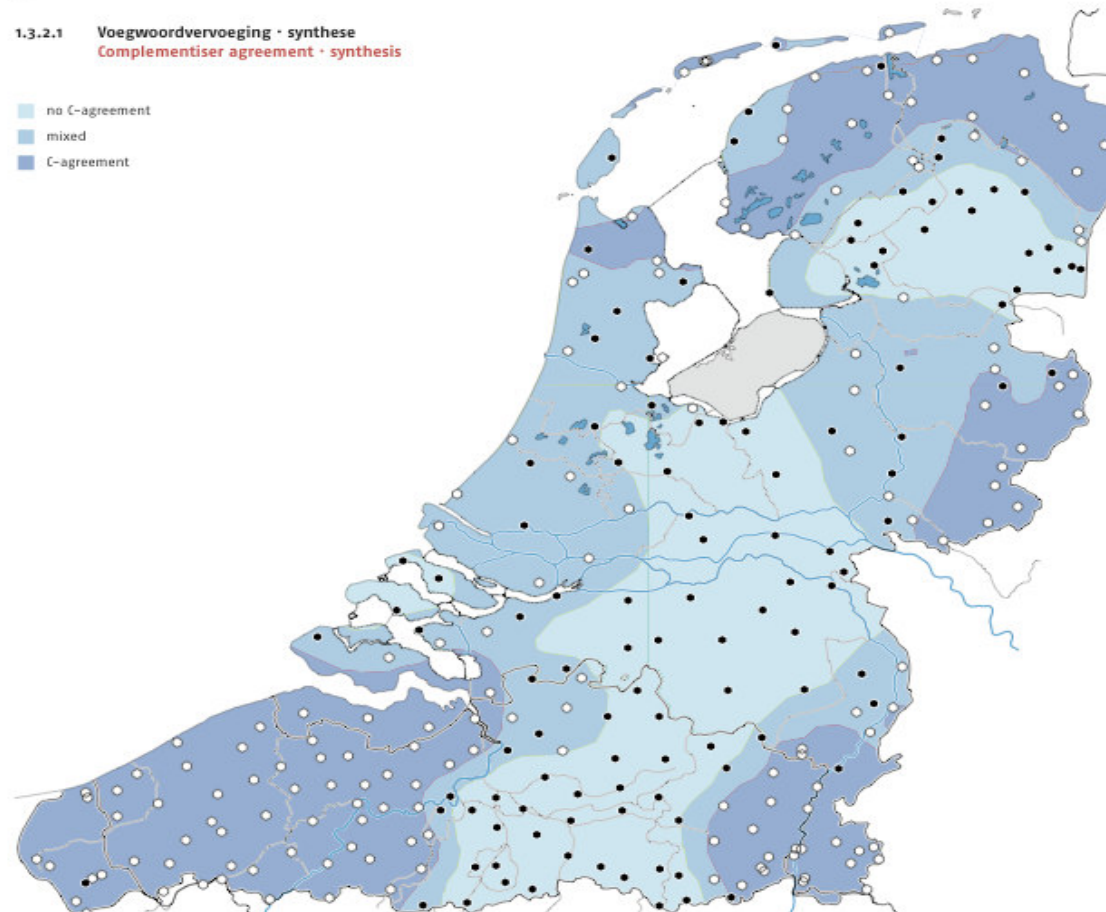
Voegwoorden **Complementisers**

19

a

1.3.2.1 Voegwoordvervoeging · *synthese* **Complementiser agreement · *synthesis***

- no C-agreement
- mixed
- C-agreement





Subject doubling

(4) **Ge** gelooft **gulder** niet dat hij sterker is als **ge gulder**.
you.W believe you.S not that he stronger is than you.W you.S
'You don't believe that he is stronger than you.' (West- & East Flemish)

(5) **'K** zal **'k** 't **ik** krijgen.
I.W will I.W it I.S get
'I will get it' (West- & East-Flemish)

De Vogelaer, G. (2005). *Persoonsmarkering in de dialecten in het Nederlandse taalgebied*. Dissertation Ghent.



Subject doubling

Subjectverdubbeling Subject doubling

55

a

3.1.3.4.1 Subjectverdubbeling 2 meervoud Subject doubling 2 plural

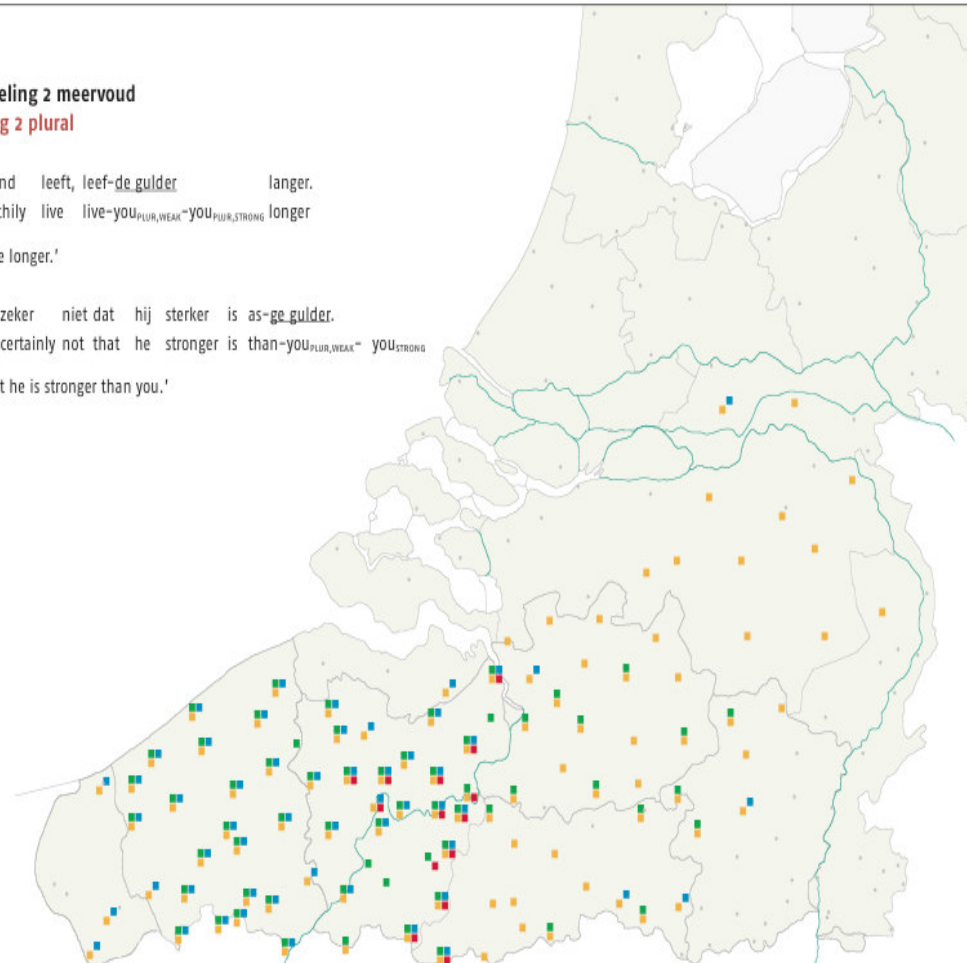
As-ge gulder gezond leeft, leef-de gulder langer.
if-you_{PLUR,WEAK}-you_{PLUR,STRONG} healthy live live-you_{PLUR,WEAK}-you_{PLUR,STRONG} longer

'If you live healthily you will live longer.'

Ge gelooft gulder zeker niet dat hij sterker is as-ge gulder.
you_{PLUR,WEAK} believe you_{PLUR,STRONG} certainly not that he stronger is than-you_{PLUR,WEAK}-you_{STRONG}

'You do not seem to believe that he is stronger than you.'

■ V _{FIN} —	99
■ — V _{FIN} —	62
■ C —	54
■ C _{COMPAR} —	14





Subject clitisation to yes/no

- (6) Q: Hebben ze al gegeten?
have they already eaten
A: Jaa(-n)-ze
yes(.PLUR).they

Van Craenenbroeck, J. (2004). *Ellipsis in Dutch dialects*. Leiden, LOT
Dissertation 96.



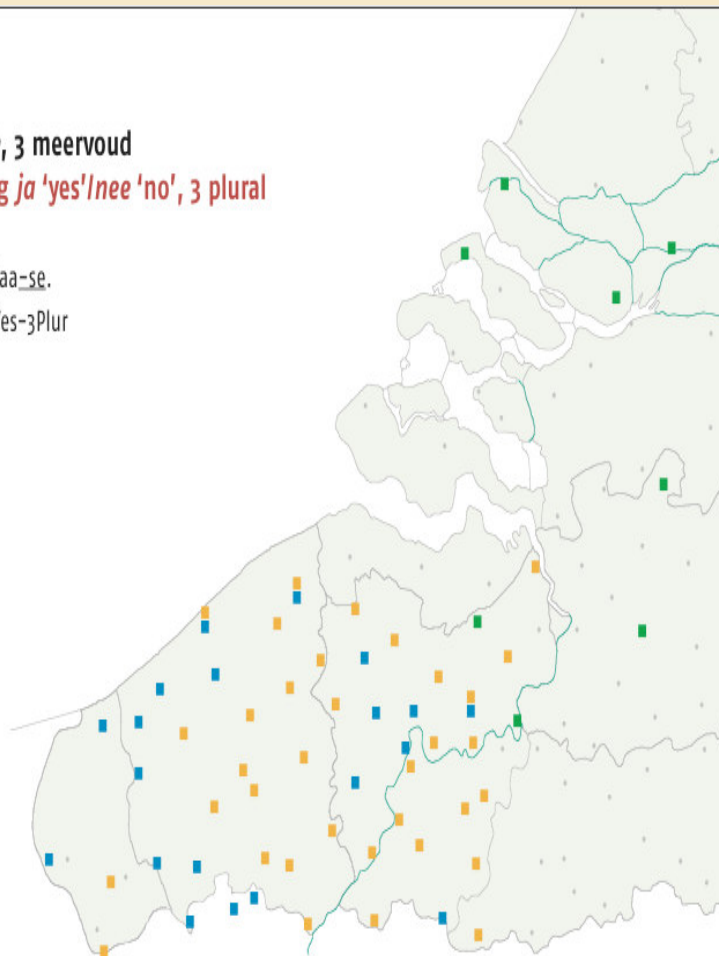
Clitics on yes/no

a

3.2.3.7.1 Clitica en vervoeging na *ja/nee*, 3 meervoud *Clitics and agreement following ja 'yes' / nee 'no', 3 plural*

Vraag: Hebben ze al gegeten? Antwoord: Jaa-se.
Question: Did they already eat? Answer: Yes-3Plur

■ jaa-s(e)	35
■ jaa-ns	20
■ jaa-t	9





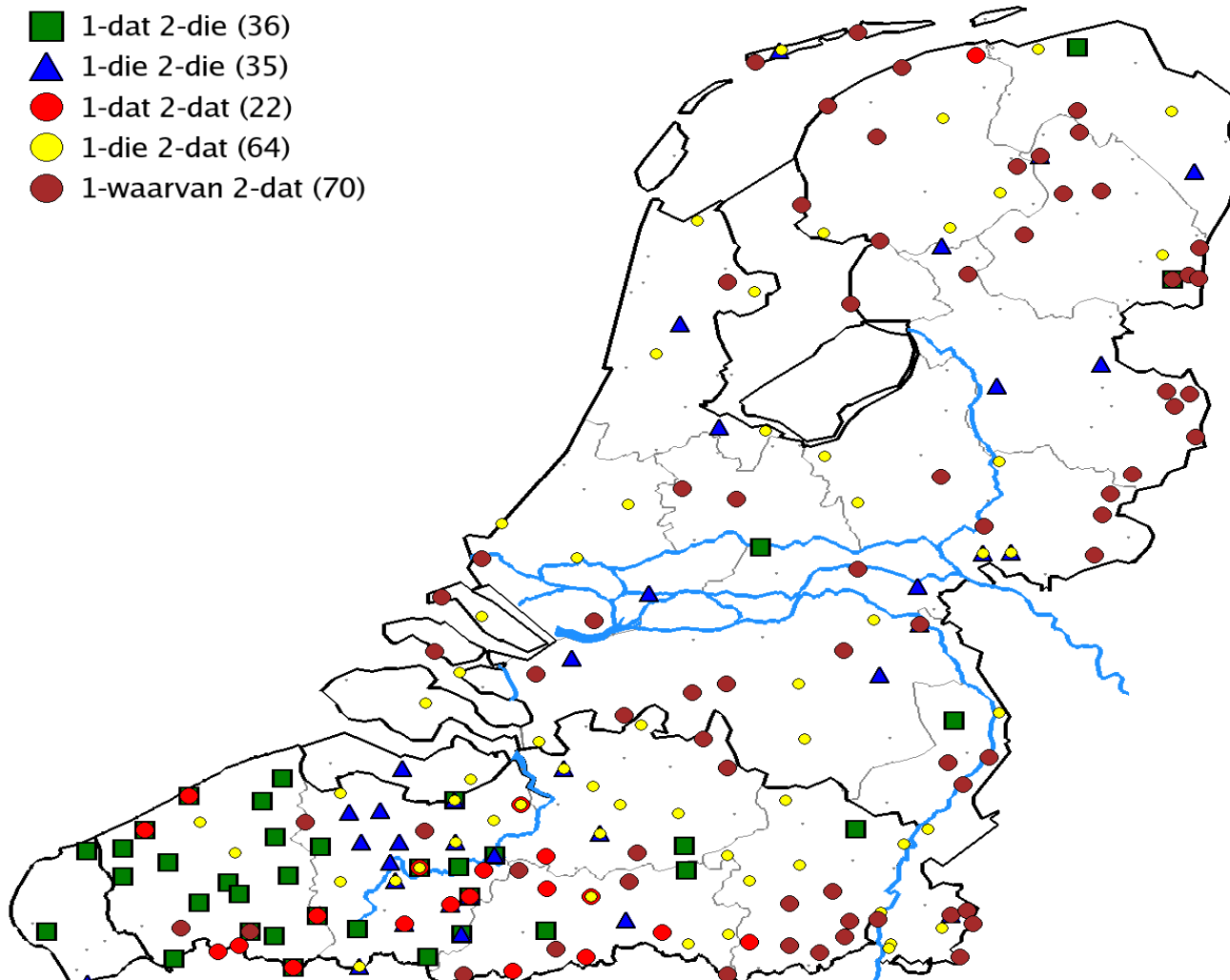
Fronting

(7) Dit is de man 1-die ik denk 2-dat het verhaal verteld heeft
this is the man 1-who I think 2-that the story told has
'This the man who I think has told the story.'

Long subject relatives



- 1-dat 2-die (36)
- ▲ 1-die 2-die (35)
- 1-dat 2-dat (22)
- 1-die 2-dat (64)
- 1-waarvan 2-dat (70)





Reflexive microvariation

Strong reflexives: forms

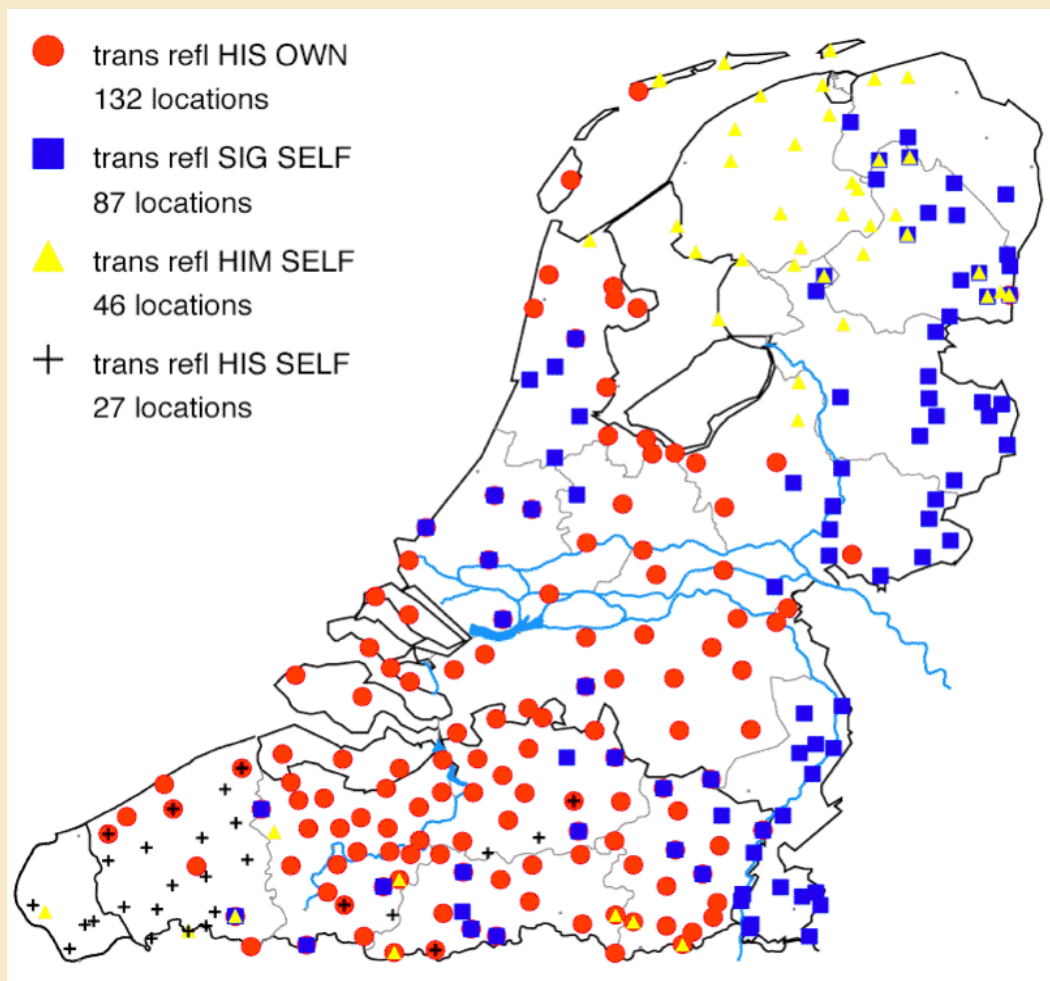
- a. Toon bekeek REFL
Toon looked-at REFL
- b. Eduard kent REFL
Edward knows REFL

forms:

zichzelf	(sig-self)
hemzelf	(him-self)
zijnzelf	(his-self)
zijn eigen	(his own)



Strong reflexives: distribution





Weak reflexives: forms

Jan herinnert REFL dat verhaal

John remembers REFL that story

forms:

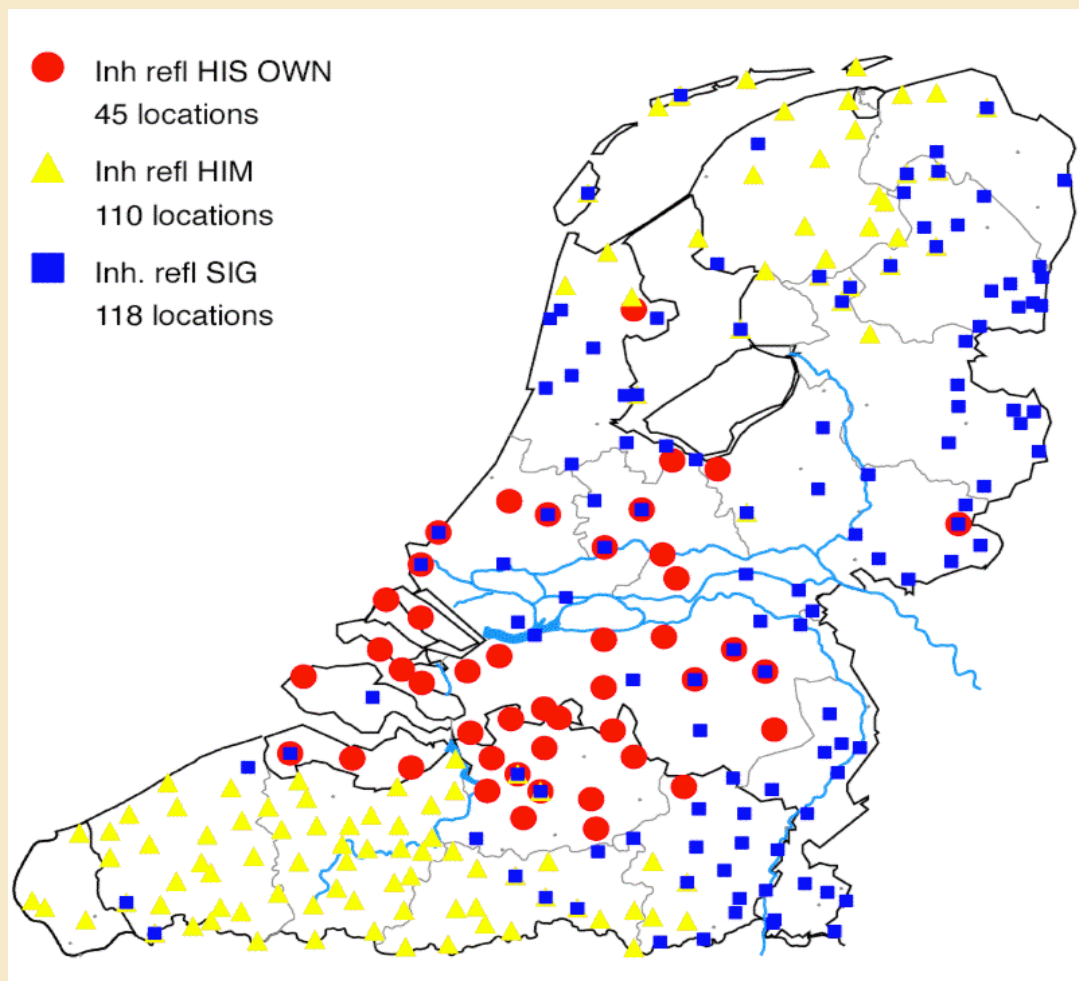
zich (sig)

hem (him)

zijn eigen (his own)



Weak reflexives: distribution

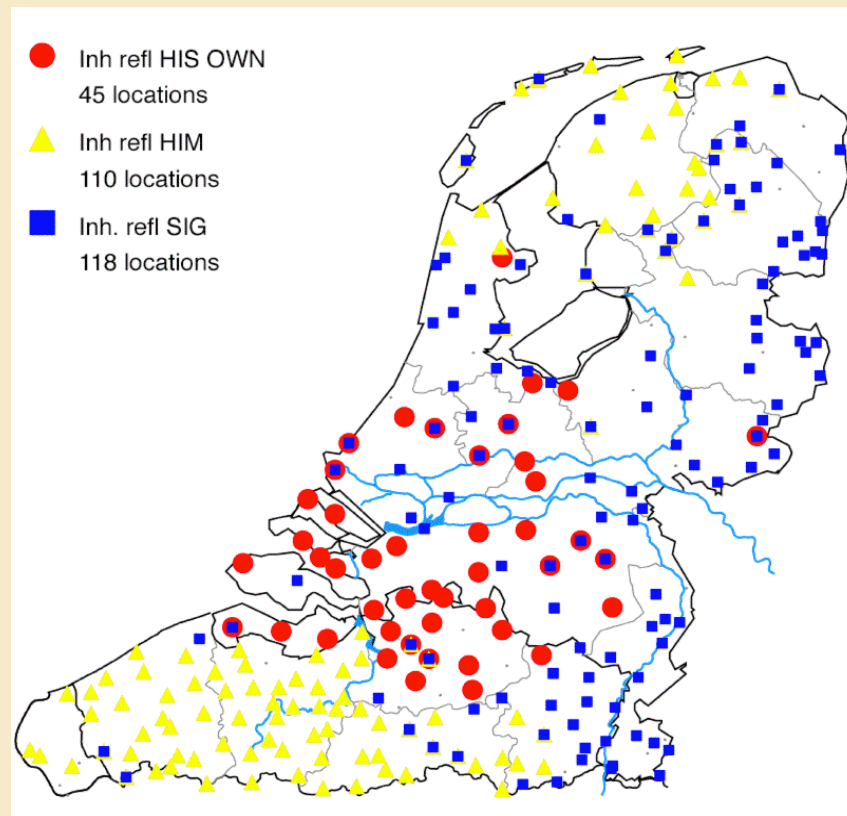
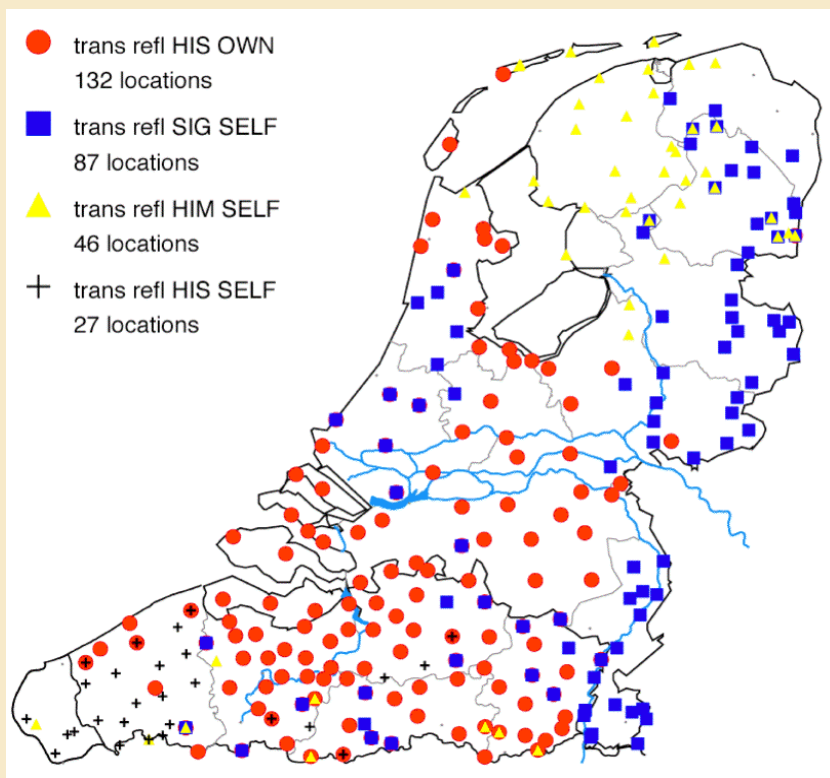




Strong and weak reflexives

STRONG

WEAK





Verb clusters

(8) *modal - modal - main verb*

Ik vind dat iedereen **moet kunnen** zwemmen.

I think that everyone **must can.INF** swim.INF

‘I think that everybody should be able to swim.’

(9) *modal - perf. auxiliary - main verb*

dat hij voor drie uur de wagen **moet hebben** gemaakt.

that he before three o'clock the car **must have.INF** repaired.PART

‘that he should have repaired the car before three o'clock’

(10) *perf. auxiliary - aspectual - main verb*

Vertel niet wie hij **had kunnen** roepen

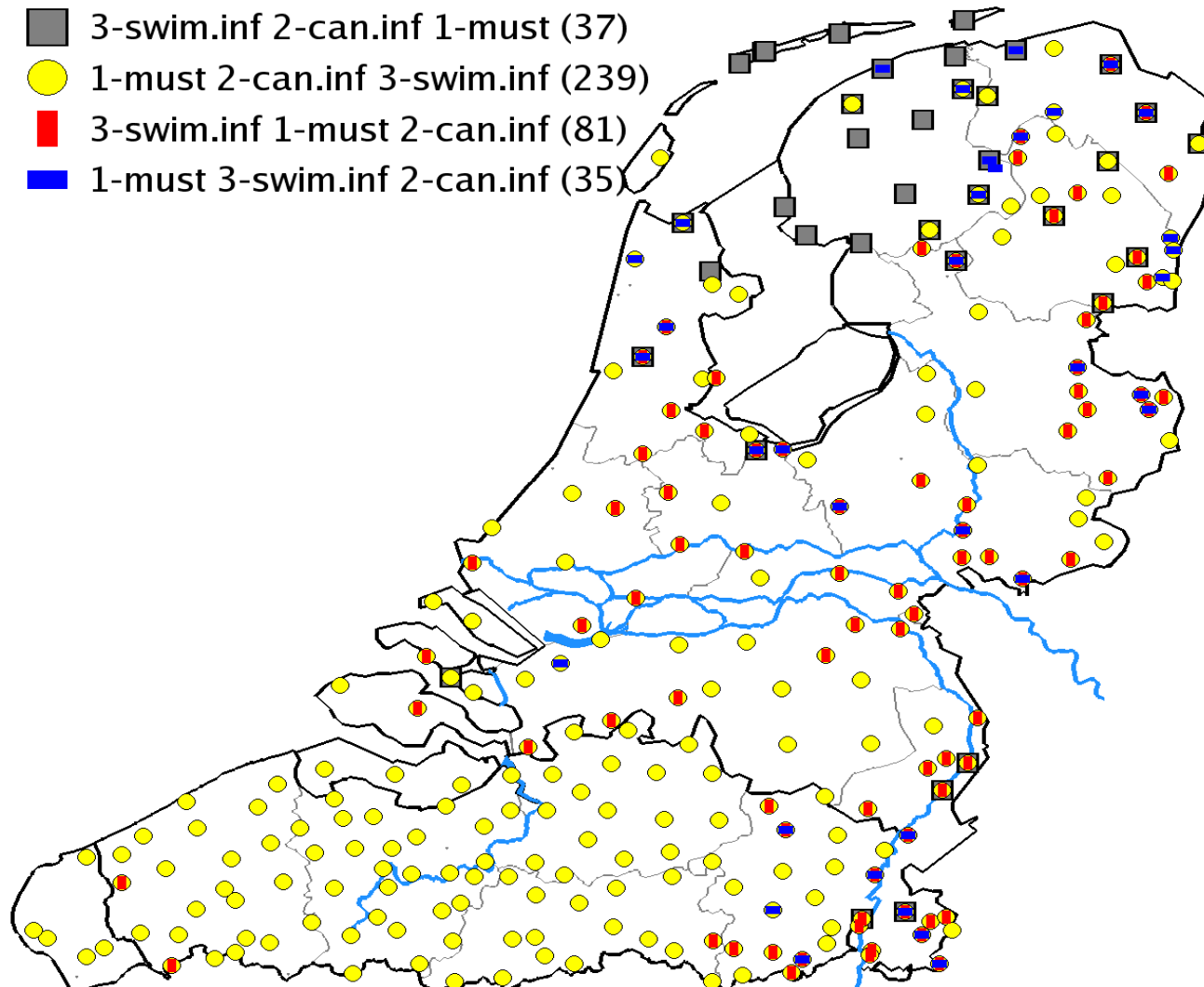
tell not who he **had can.INF** call.INF

‘Don’t say who he could have called.’



must can swim

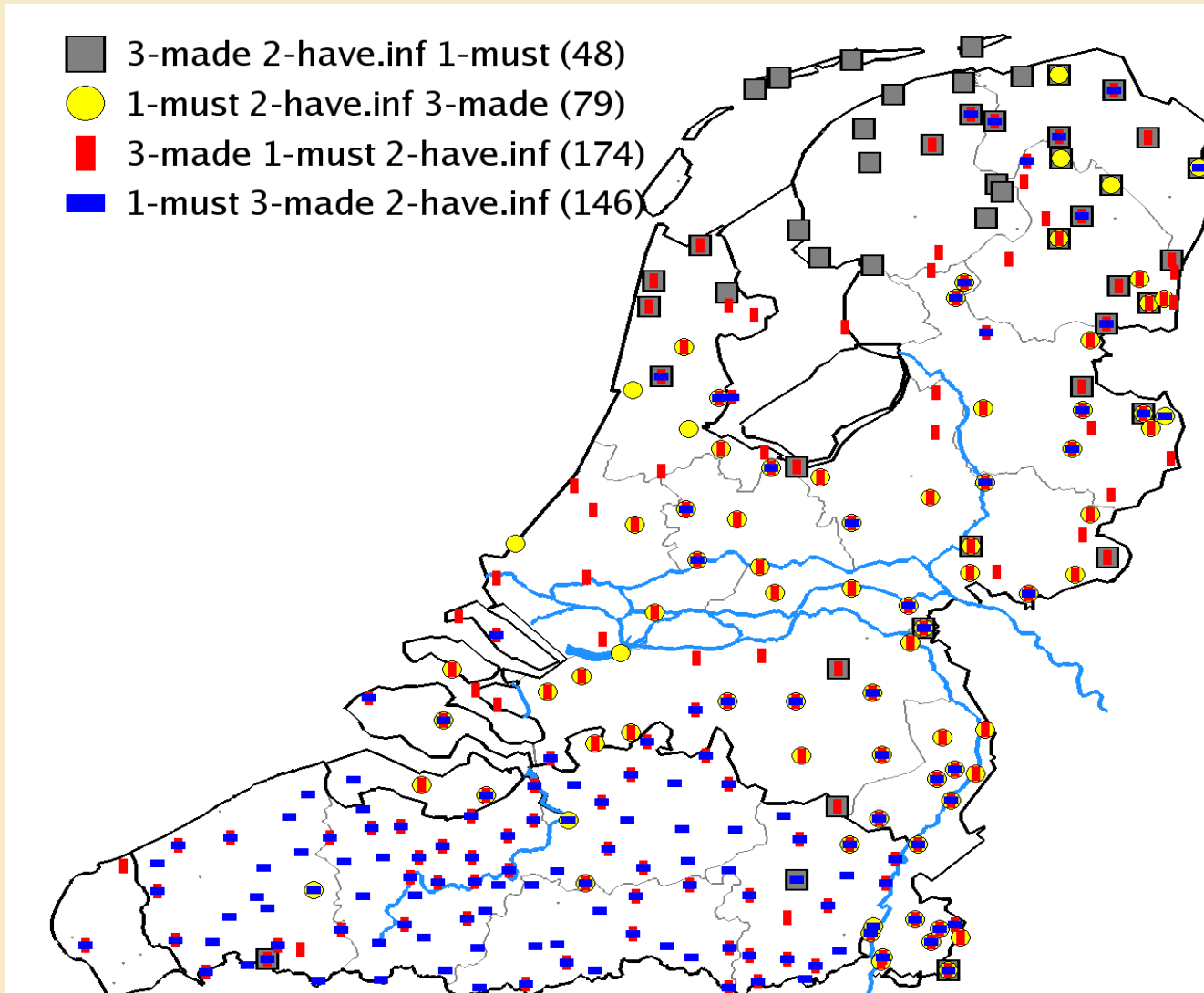
- 3-swim.inf 2-can.inf 1-must (37)
- 1-must 2-can.inf 3-swim.inf (239)
- 3-swim.inf 1-must 2-can.inf (81)
- 1-must 3-swim.inf 2-can.inf (35)





must have made

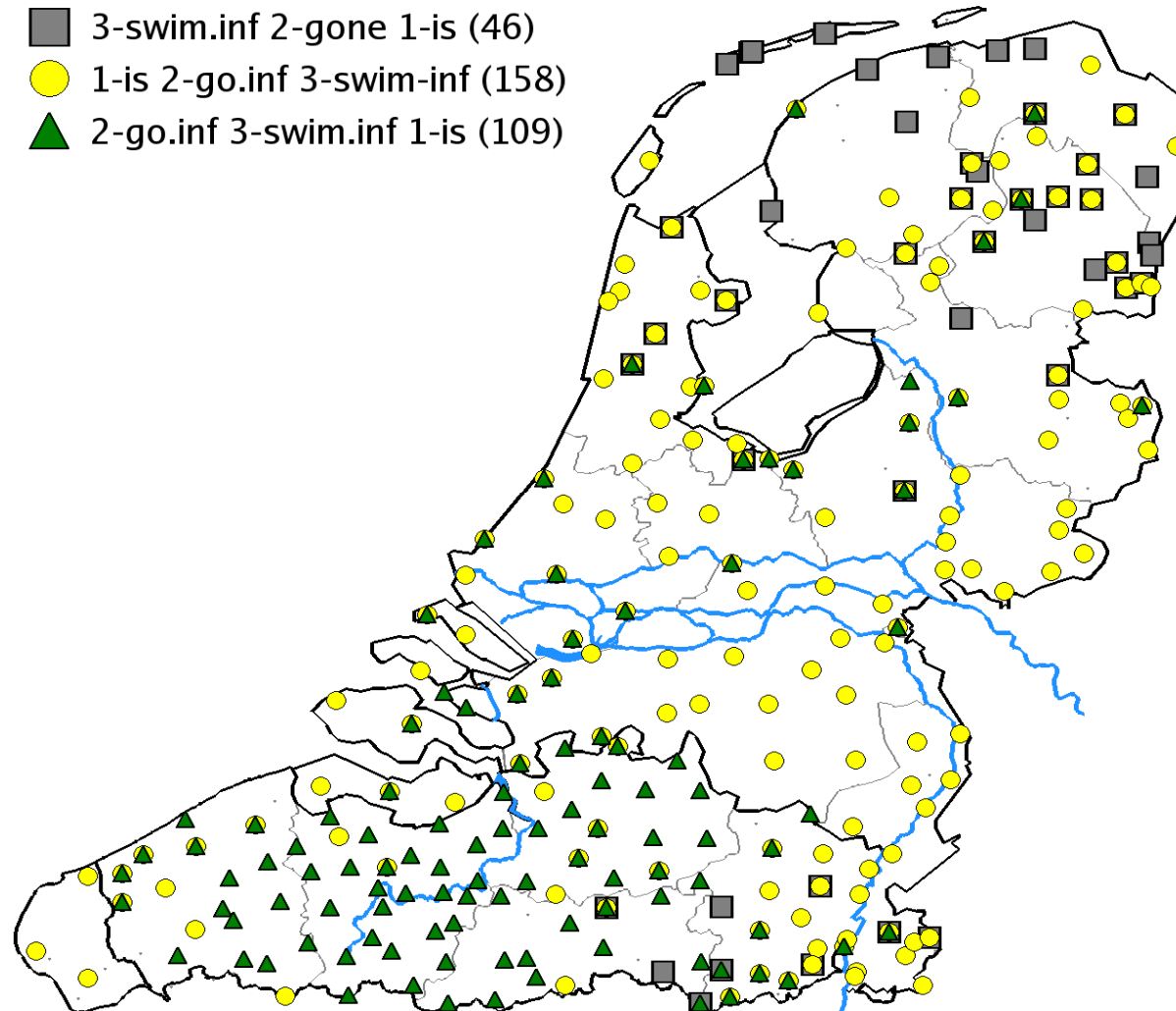
- 3-made 2-have.inf 1-must (48)
- 1-must 2-have.inf 3-made (79)
- 3-made 1-must 2-have.inf (174)
- 1-must 3-made 2-have.inf (146)



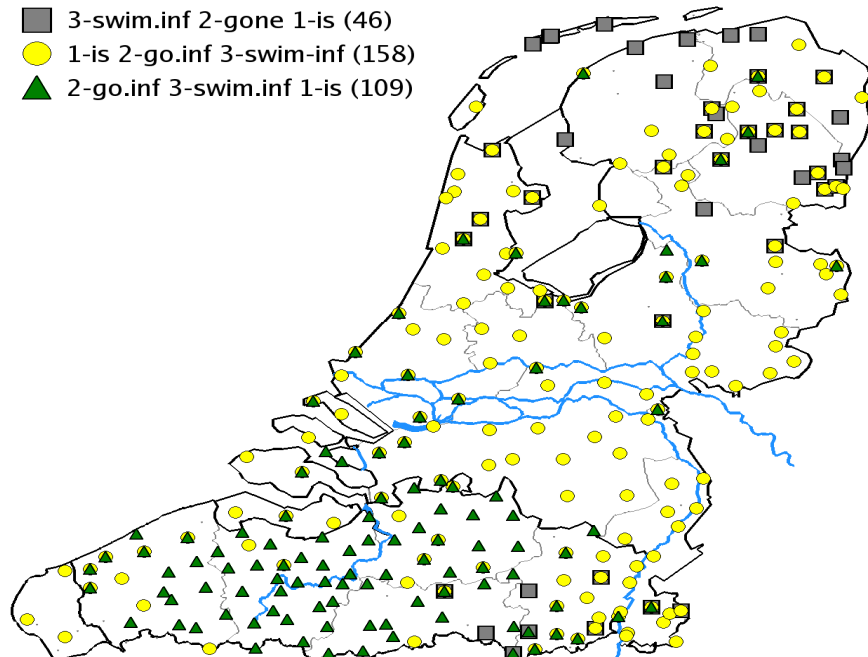
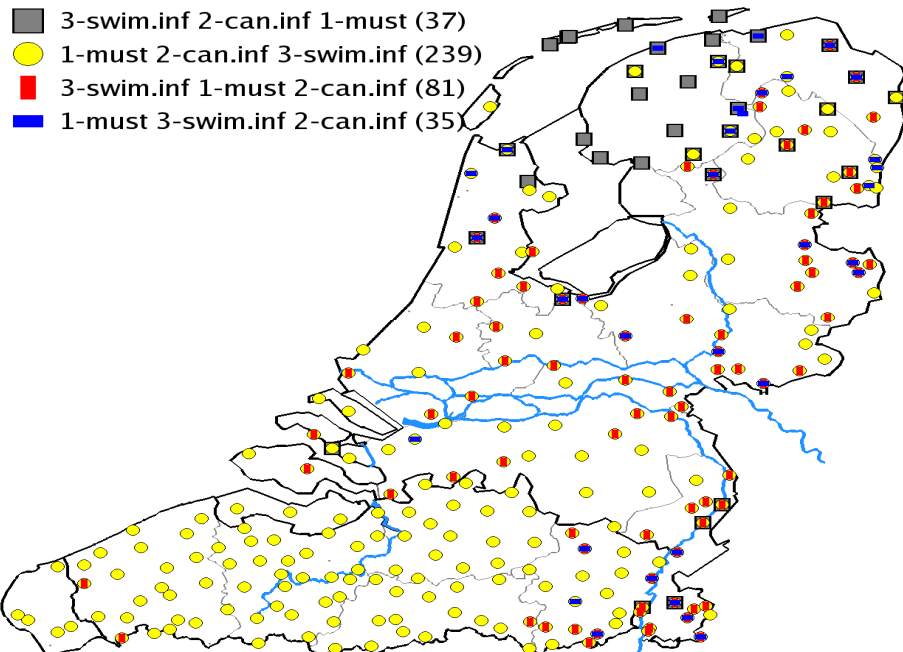
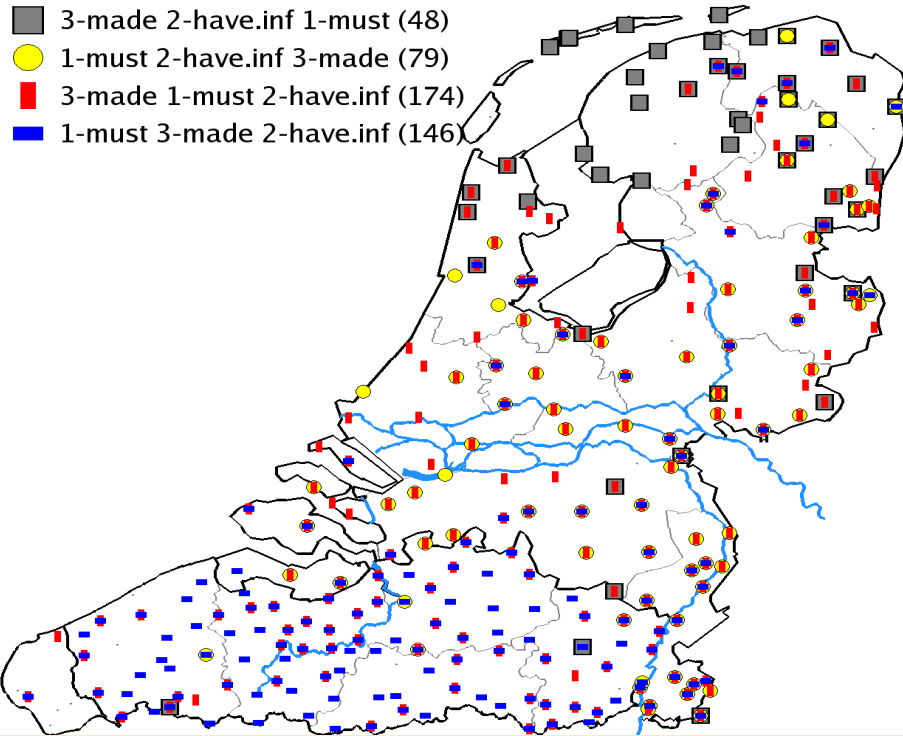


is go swim

- 3-swim.inf 2-gone 1-is (46)
- 1-is 2-go.inf 3-swim-inf (158)
- ▲ 2-go.inf 3-swim.inf 1-is (109)



three cluster types





Verb clusters

Missing orders

- * can.INF - must.FIN - swim.INF (2-1-3)
 - * have.INF - must.FIN - repaired.PART (2-1-3)
 - * go.INF - is.FIN - swim.INF (2-1-3)
- * can.INF - swim.INF - must.FIN (2-3-1)
 - * have.INF - repaired.PART - must.FIN (2-3-1)
 - √ go.INF - swim.INF - is.FIN (2-3-1)
- √ swim.INF - must.FIN - can.INF (3-1-2)
 - √ repaired.PART - must.FIN - have.INF (3-1-2)
 - * swim.INF - is.FIN - go.INF (3-1-2)



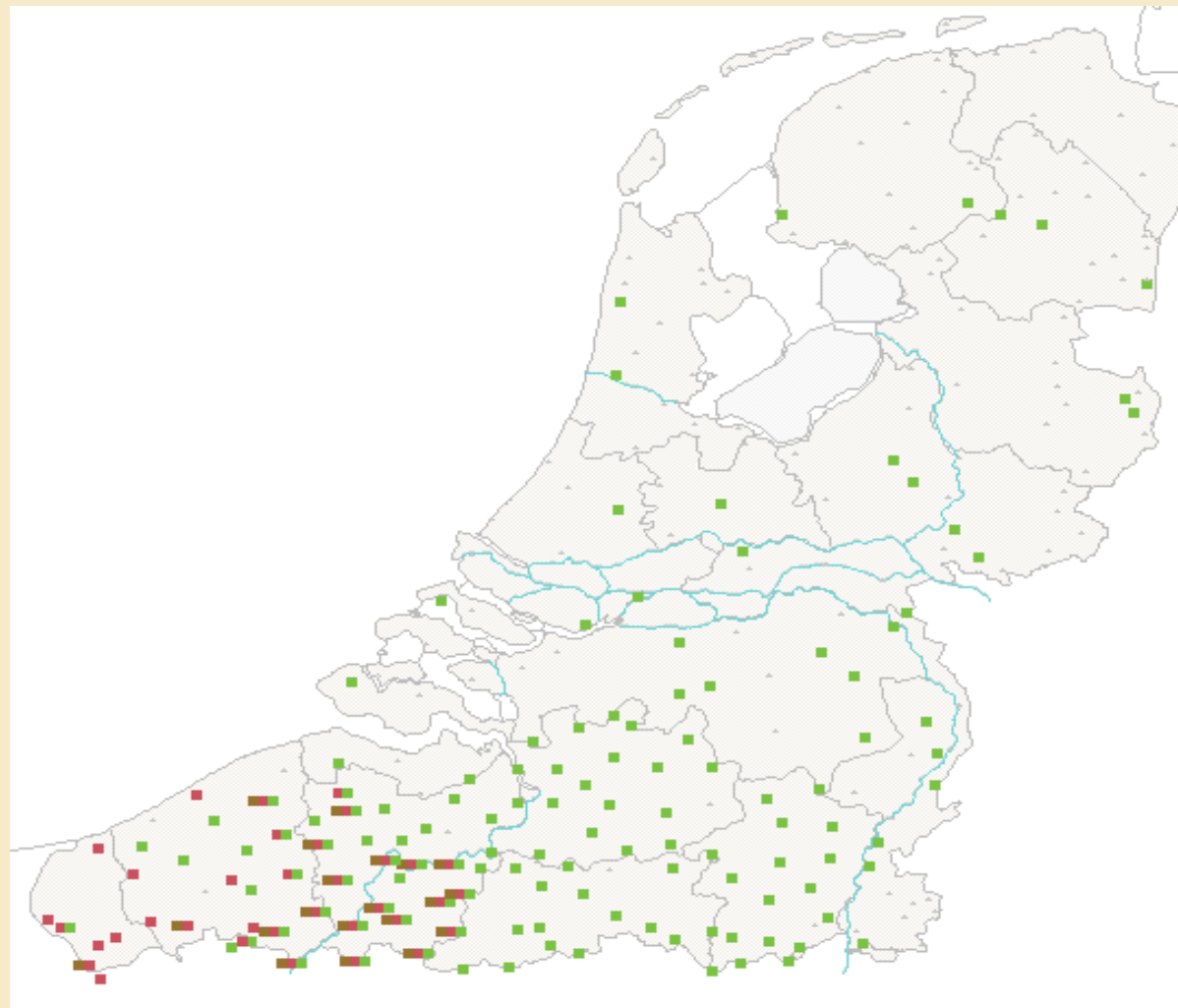
*Negative particles
and negative concord*

- (11) 'K en e niemand nie gezien
I NEG have nobody not seen
'I have not seen anybody.'

Zeijlstra, H. (2004). *Sentential negation and negative concord*.
Amsterdam, LOT Dissertation 101



*Negative particle
and negative concord*





Prospects

- Variation in Inflection (Variflex)
- Determinants of Dialectal Variation
- European Dialect Syntax (EDISYN)
- International Cooperation



European cooperation

Issues

- Standardization of methodology
- Standardization of POS tagging
- Standardization of data storage, retrieval and visualization
- Empirical coherence of dialect syntax research

EDISYN: European Dialect Syntax

EDISYN

European Dialect Syntax

Sjef Barbiers

Meertens Institute and Utrecht University

- Syntactic Microvariation
(Amsterdam, August 2000; SAND)
- Dialect Syntax in the West Germania
(Freiburg, November 2001; FRED)
- European Dialect Syntax (ESF)
(Padova, September 2003; ASIS)
- Grand Meeting Scandiasyn
(Leikanger, August 2005; Scandiasyn/NORMS)
- Syntactic Doubling in European Dialects
(Amsterdam, March 2006)

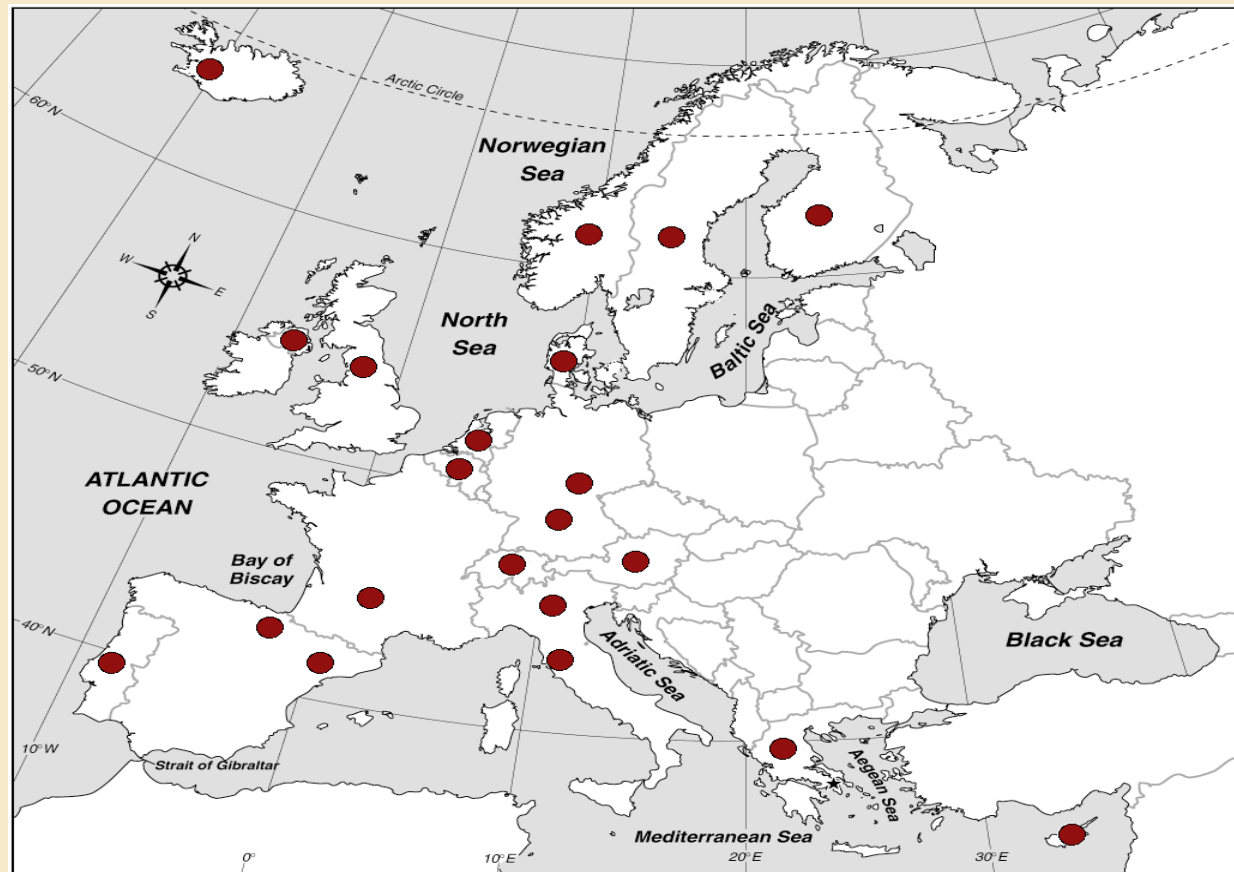
EDISYN

EDISYN Goals

- Network of European dialect syntacticians that use similar standards w.r.t
 - methodology of data collection
 - data storage and annotation
 - data retrieval
 - cartography
- Description and analysis of syntactic doubling phenomena

EDISYN

Current network



EDISYN

Building a European network

Network participants

- ASIS (Poletto et al.; Manzini et al.; Italian dialects)
- SCANDIASYN (Vangsnes et al., Scandinavian dialects)
- NORMS (Svenonius et al., Scandinavian dialects)
- FRED (Kortmann, English dialects)
- CORDIASYN (Martins et al., Portuguese dialects)
- Alemannic (Bayer, Brandnerl), Bavarian (Weiß), Catalan dialects (Rigau), Spanish dialects (Fernandez), French dialect (Sauzet, Sportiche, Dagnac), Balkan Slavic (Tomic), Swiss German (Glaser), Basque dialects (Fernandez et al., Etxeparre et al.), Greek dialects (Ralli), English dialects (Buchstaller, Corrigan & Holmberg), Russian and Balkan dialects (Sobolev, Rusakova), Armenian (Khurshudian), Estonian (Lindstrom)

EDISYN

EDISYN team

Sjef Barbiers (project leader)

Marika Lekakou (postdoc)

Eefje Boef (PhD)

Franca Wesseling (research assistant)

Jan Pieter Kunst (software developer)

(all Meertens Institute)

Olaf Koeneman (U of Amsterdam)

EDISYN

www.dialectsyntax.org

- Dialect syntax research manual
 - Cornips & Poletto (2005, to appear) on Methodology
 - Barbiers, Cornips & Kunst (to appear) on the SAND corpus
- Distributed network of dialect syntax databases with common search engine and cartography
- Microvariation in Syntactic Doubling (Bingley: Emerald; 2008)
- Monography on Syntactic Doubling
- Workshops:
 - Spring 2010 (Basque Country)

EDISYN

Syntactic doubling

- (1) Subject pronoun doubling and subject agreement doubling

Ze peiz-n da-n **ze** **ziender** rijker zij-n *West*
Flemish

they think-3pl that-3pl they.W they.S richer are-3pl
'They think that they are richer.'

- (2) Wh-word doubling

Wel denkst **wel** ik in de stad ontmoet heb. *Drents*
who think-2pl who I in the city met have

'Who do you think I met in the city?'

- (3) Auxiliary doubling

K-em da gezegd **gehad** *Brabants*
I-have that said-PART had-PART

'I have said that.'

EDISYN

Syntactic doubling

(4) Verb doubling / “do-support”

Doe het brood eve **snije.**
do the bread particle cut

Zeeuws

‘Please cut the bread.’

(5) Participial morphology doubling (PPI)

Zol hee dat **edane** hemmn **ekund**
would he that done-PART have could-PART

Gelders

‘Could he have done that?’

(6) Infinitival morphology doubling (IPP)

Hij had moet-**en** werk-**en**
majority of dialects
he had must-INF work-INF

Dutch &

‘He should have worked’

EDISYN

Syntactic doubling

(7) Negative concord

‘t **En** danst-ij **niemand nie**
Flemish

East-

it NEGdances-it nobody not

‘Nobody is dancing.’

(8) Focus particle doubling (PPI)

Maar één wedstrijd heb ik **maar** gezien *Colloquial*
Dutch

only one game have I only seen

‘I have seen only one game.’

(9) Indefinite determiner doubling

Zo ‘n ding **één** ha ik ze leve nie gezie
Limburgian

EDISYN *Doubling in Europe: pronouns*

- Subject doubling
- Clitic doubling/duplication
- Wh word doubling: short & long
- Relative pronoun doubling
- Possessor doubling
- Locative pronoun doubling
- Resumption

EDISYN

Doubling in Europe: extended VP
(i)

- Agreement
- TMA doubling (including English double modals)
- Infinitival marker doubling
- Participial morphology doubling
- Infinitival morphology doubling (IPP)

EDISYN *Doubling in Europe: extended VP*

- Auxiliary doubling: BE, HAVE, GO, LET
- Periphrastic DO
- Lexical verb doubling
- Double complementizers
- Complementizer agreement
- Doubly filled Comp

EDISYN *Doubling in Europe: quantifiers*

- Focus particle doubling
- Negative Concord
- Floating quantifier doubling
- Determiner doubling ((in-)definite)

EDISYN

Doubling in Europe: other

- Preposition doubling
- Comparative and superlative doubling
- Double conjunctions
- Demonstrative reinforcement

EDISYN

Main Issues

1. Difference between standard languages and dialects (?)
2. Theory of (micro)variation
3. Syntactic structure
4. Semantic / pragmatic functions of doubling
5. Geography

EDISYN *Theoretical interest of doubling*

- More doubling in dialects than in standard languages?

If so, why?

- Only functional material?
- Economy (Chomsky 1995)
- Compositionality (Frege 1923)
- Theta-criterion (Chomsky 1981)
- Semantic import?

- Multiple spell out of chain links (e.g. Nunes 2004)
e.g. Wh-doubling, periphrastic DO, subject doubling
 - (i) Who do you think who I met who
 - (Almost) missing: lexical doubling
 - Theory of morphological reanalysis?
 - What counts as identical and what not?
 - With distinct doubles: why is the first always underspecified? (cf. Barbiers 2006)

EDISYN

Multiple spell out

1. Do these spell-out options correlate with other properties/features of the language systems involved?
2. If not, how to account for cross-linguistic differences? Distinction between ungrammatical structures and unrealized structures?
3. (Why) do dialects show a tendency to spell out more than one position?
 - Economy
 - Normative pressure (avoidance of redundancy)

- One constituent hypothesis (e.g. Uriagareka 1995, Belletti 2003, Kayne 1994) e.g. subject doubling, resumptive pronouns
 - (i) In non-splitting varieties: Do we find these big XPs?
 - (ii) Theory of subextraction
 - (iii) Should absolute identity of doubles not be excluded, given that every part checks a different feature?
 - (iv) Economy of split checking?

EDISYN

Analyses of doubling- III

- Spell out of Spec Head checking configuration (e.g. Rizzi 1991, Haegeman 1995; Wh/Foc/Neg criterion)
e.g. doubly filled Comp, negative concord
(iv) I don't know **who if** John met.
- Spell out of agreement at a distance / covert checking (e.g. Sportiche 1998); split constituent hypothesis
e.g. verb subject agreement, clitic doubling

- Morphological visibility (Baker 1996): Arguments must be morphologically realized on predicate
e.g. agreement on V + empty pronoun + full DP
- Semantic/pragmatic import
 - focus
 - affirmation / negation
 - emphasis
 - specificity (dislocation configurations)
 - empathy (Nuyts 1995; subject doubling)

EDISYN

Analyses of doubling-V

- Doubling to provide a second argument
e.g. focus particle doubling (Bayer 1996, Kayne 2000, Barbiers 2003)
dat hij [maar [_{PP} op één man]] maar [_{TP} boos [~~maar [_{PP} op één man]]] is]
that he [only at one man] only angry [only at one man]
is~~
- Doubling to reduce dyadic to monadic predicate
(Barbiers 1995, 2000)
Atom of interpretation: pred (X,Y)
Monadic predication if X = Y,
i.e. if Y agrees with X; Y is a copy of X; Y is a pronoun coreferential with X.

EDISYN

CONCLUSION