Morphosyntactic microvariation in Dutch: a quantitative-parametric approach

Jeroen van Craenenbroeck & Marjo van Koppen
KU Leuven/Meertens Institute/Utrecht U

Goethe Universität Frankfurt am Main
8 February 2022

Outline

Main goals for today

Introduction: Kayne's dream

Quantitative analysis

- Correspondence Analysis
- Cluster Analysis
- Cluster Description
- Conclusion

Qualitative analysis

- Case study #1: PolP
- Case study #2: split DP
- Case study #3: split Force/Fin
- Combining the case studies: 7 parameters

The bigger picture: determinants of variation

Main goals for today

1. Develop a parametric analysis for a large data set of morphosyntactic variation in Dutch dialects.

Main goals for today

1. Develop a parametric analysis for a large data set of morphosyntactic variation in Dutch dialects.
2. Advocate for the combined use of quantitative (statistical) and qualitative (formal-theoretical) methods as a way towards achieving such an analysis.

Main goals for today

1. Develop a parametric analysis for a large data set of morphosyntactic variation in Dutch dialects.
2. Advocate for the combined use of quantitative (statistical) and qualitative (formal-theoretical) methods as a way towards achieving such an analysis.
3. Consider the bigger implications of this one case study for understanding the properties of and mechanisms behind variation in natural language.

Outline

Main goals for today

Introduction: Kayne's dream

Quantitative analysis

- Correspondence Analysis
- Cluster Analysis
- Cluster Description
- Conclusion

Qualitative analysis

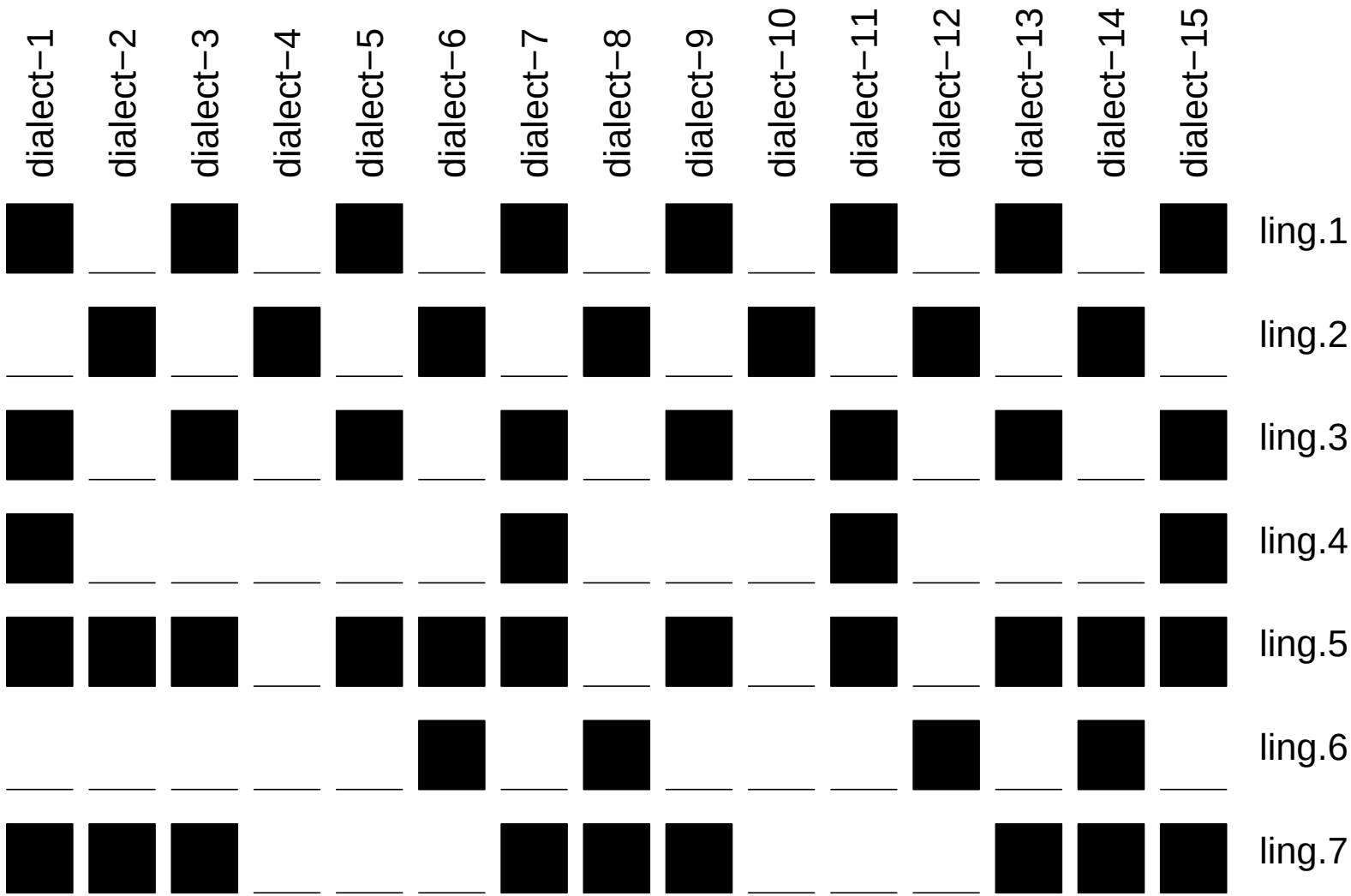
- Case study #1: PolP
- Case study #2: split DP
- Case study #3: split Force/Fin
- Combining the case studies: 7 parameters

The bigger picture: determinants of variation

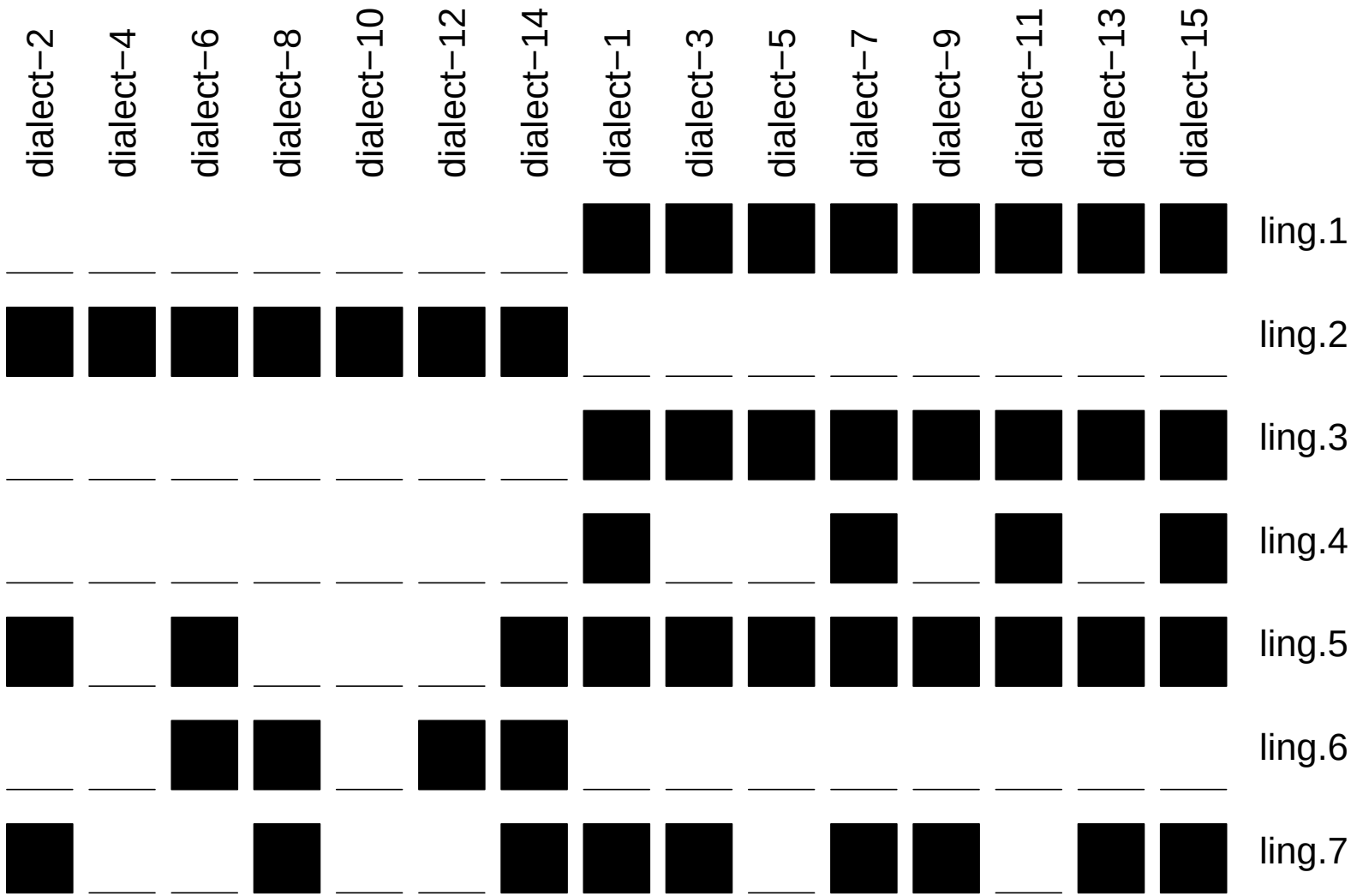
Introduction: Kayne's dream

"*If it were possible to experiment on languages, a syntactician would construct an experiment of the following type: take a language, alter a single one of its observable syntactic properties, examine the result to see what, if any, other property has changed as a consequence of the original manipulation.*" (Kayne 1996:xii)

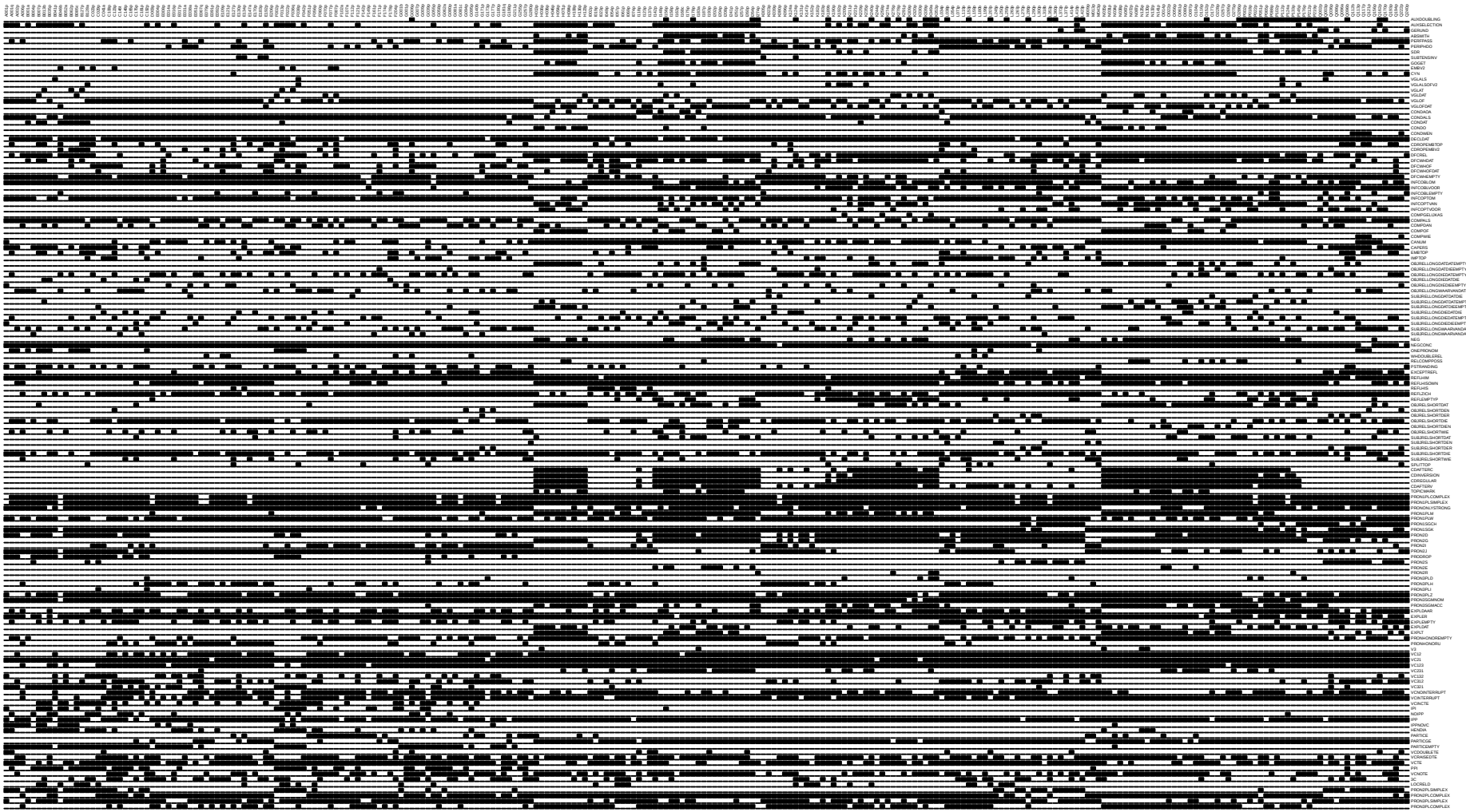
Introduction: Kayne's dream



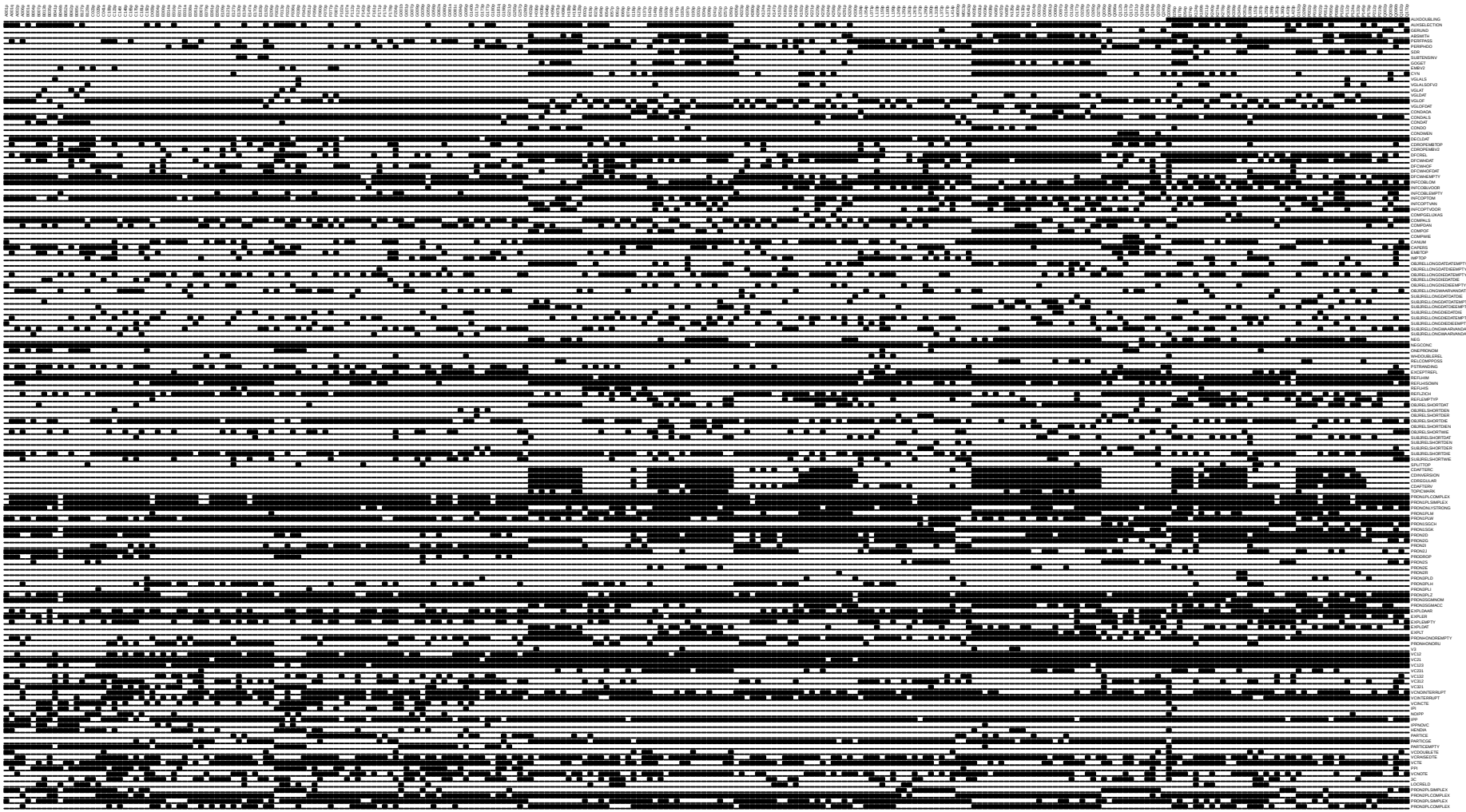
Introduction: Kayne's dream



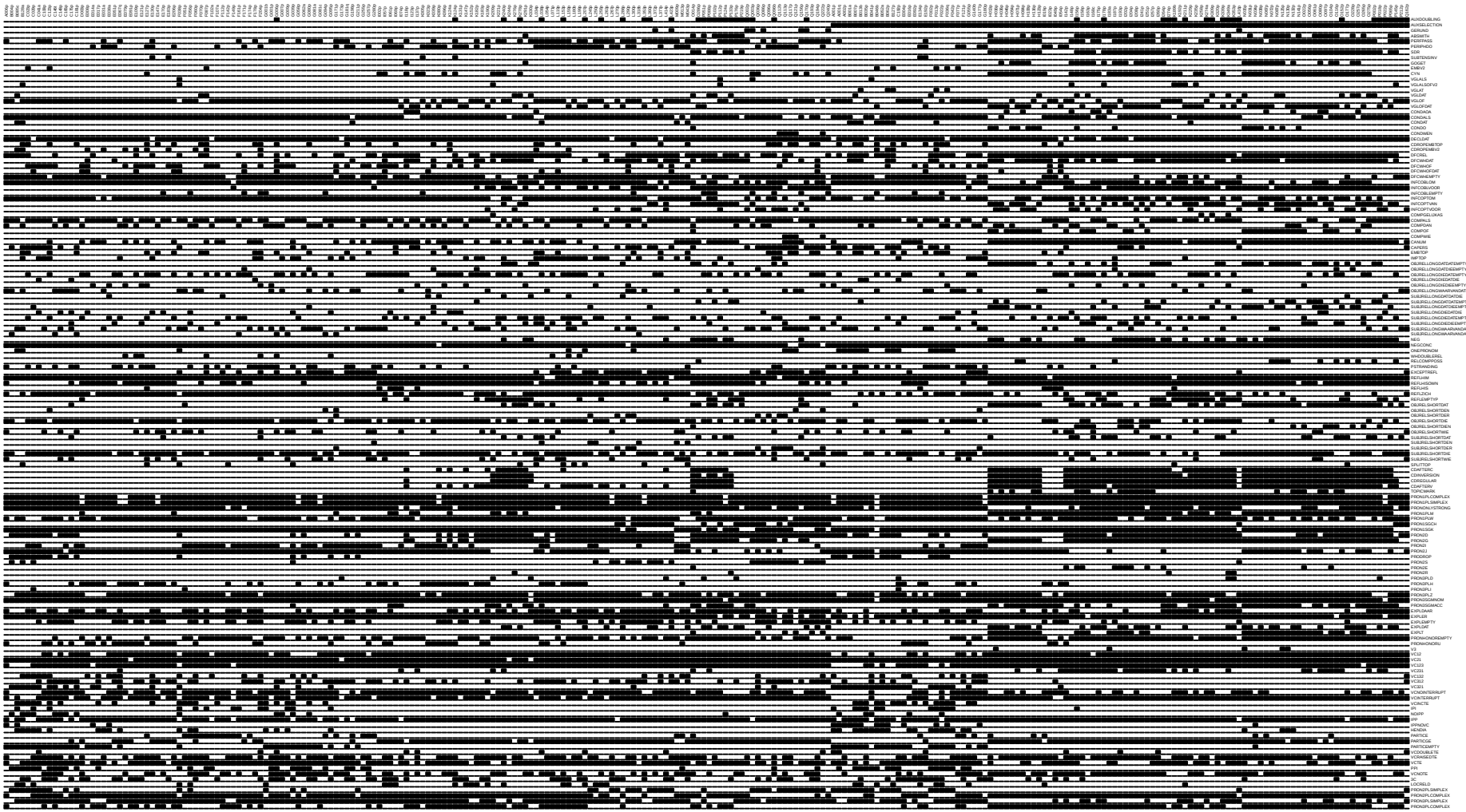
Introduction: Kayne's dream



Introduction: Kayne's dream



Introduction: Kayne's dream



Introduction: Kayne's dream

→ **The goal of the current research:**
To bring together two traditions of dialect research:

Introduction: Kayne's dream

→ **The goal of the current research:**
To bring together two traditions of dialect research:

1. **quantitative work** (e.g. Heeringa (2004), Spruit (2008), Heeringa and Nerbonne (2013), Wieling and Nerbonne (2015))

Introduction: Kayne's dream

→ **The goal of the current research:**
To bring together two traditions of dialect research:

1. **quantitative work** (e.g. Heeringa (2004), Spruit (2008), Heeringa and Nerbonne (2013), Wieling and Nerbonne (2015))
2. **formal-theoretical work** (e.g. Bayer (1984), Haegeman (1992), Hoekstra (1993), Penner (1994), Poletto (2000), Benincà and Poletto (2004))

Introduction: Kayne's dream

→ **The goal of the current research:**
To bring together two traditions of dialect research:

1. **quantitative work** (e.g. Heeringa (2004), Spruit (2008), Heeringa and Nerbonne (2013), Wieling and Nerbonne (2015))
2. **formal-theoretical work** (e.g. Bayer (1984), Haegeman (1992), Hoekstra (1993), Penner (1994), Poletto (2000), Benincà and Poletto (2004))

- **more specifically:**

Introduction: Kayne's dream

→ **The goal of the current research:**
To bring together two traditions of dialect research:

1. **quantitative work** (e.g. Heeringa (2004), Spruit (2008), Heeringa and Nerbonne (2013), Wieling and Nerbonne (2015))
2. **formal-theoretical work** (e.g. Bayer (1984), Haegeman (1992), Hoekstra (1993), Penner (1994), Poletto (2000), Benincà and Poletto (2004))

- **more specifically:**
  - use quantitative-statistical means to identify patterns in the data

Introduction: Kayne's dream

→ **The goal of the current research:**
To bring together two traditions of dialect research:

1. **quantitative work** (e.g. Heeringa (2004), Spruit (2008), Heeringa and Nerbonne (2013), Wieling and Nerbonne (2015))
2. **formal-theoretical work** (e.g. Bayer (1984), Haegeman (1992), Hoekstra (1993), Penner (1994), Poletto (2000), Benincà and Poletto (2004))

- **more specifically:**
  - use quantitative-statistical means to identify patterns in the data
  - use qualitative-theoretical means to interpret those patterns

Outline

Main goals for today

Introduction: Kayne's dream

Quantitative analysis

- Correspondence Analysis
- Cluster Analysis
- Cluster Description
- Conclusion

Qualitative analysis

- Case study #1: PolP
- Case study #2: split DP
- Case study #3: split Force/Fin
- Combining the case studies: 7 parameters

The bigger picture: determinants of variation

Quantitative analysis

Our quantitative analysis involves three steps:

Quantitative analysis

Our quantitative analysis involves three steps:

1. **Correspondence Analysis:** identifying the main tendencies

Quantitative analysis

Our quantitative analysis involves three steps:

1. **Correspondence Analysis:** identifying the main tendencies
2. **Cluster Analysis:** cluster the dialects into groups based on those tendencies

Quantitative analysis

Our quantitative analysis involves three steps:

1. **Correspondence Analysis:** identifying the main tendencies
2. **Cluster Analysis:** cluster the dialects into groups based on those tendencies
3. **Cluster Description:** identify the linguistic phenomena that are characteristic for those clusters

Quantitative analysis: Correspondence Analysis

The Correspondence Analysis proceeds in three steps:

1. raw data table:

	AUXDOUBL	AUXSEL	GERUND	ABSWITH	PERPASS	…
Midsland	0	1	0	0	0	…
Lies	0	1	0	0	1	…
West-Terschelling	0	1	0	0	0	…
Oosterend	0	0	0	0	1	…
Hollum	0	1	0	0	0	…
Schiermonnikoog	0	0	0	0	0	…
Ferwerd	0	1	0	0	0	…
Anjum	0	1	0	0	0	…
Kollum	0	1	0	0	0	…
Visvliet	0	1	0	0	0	…
…	…	…	…	…	…	…

Quantitative analysis: Correspondence Analysis

The Correspondence Analysis proceeds in three steps:

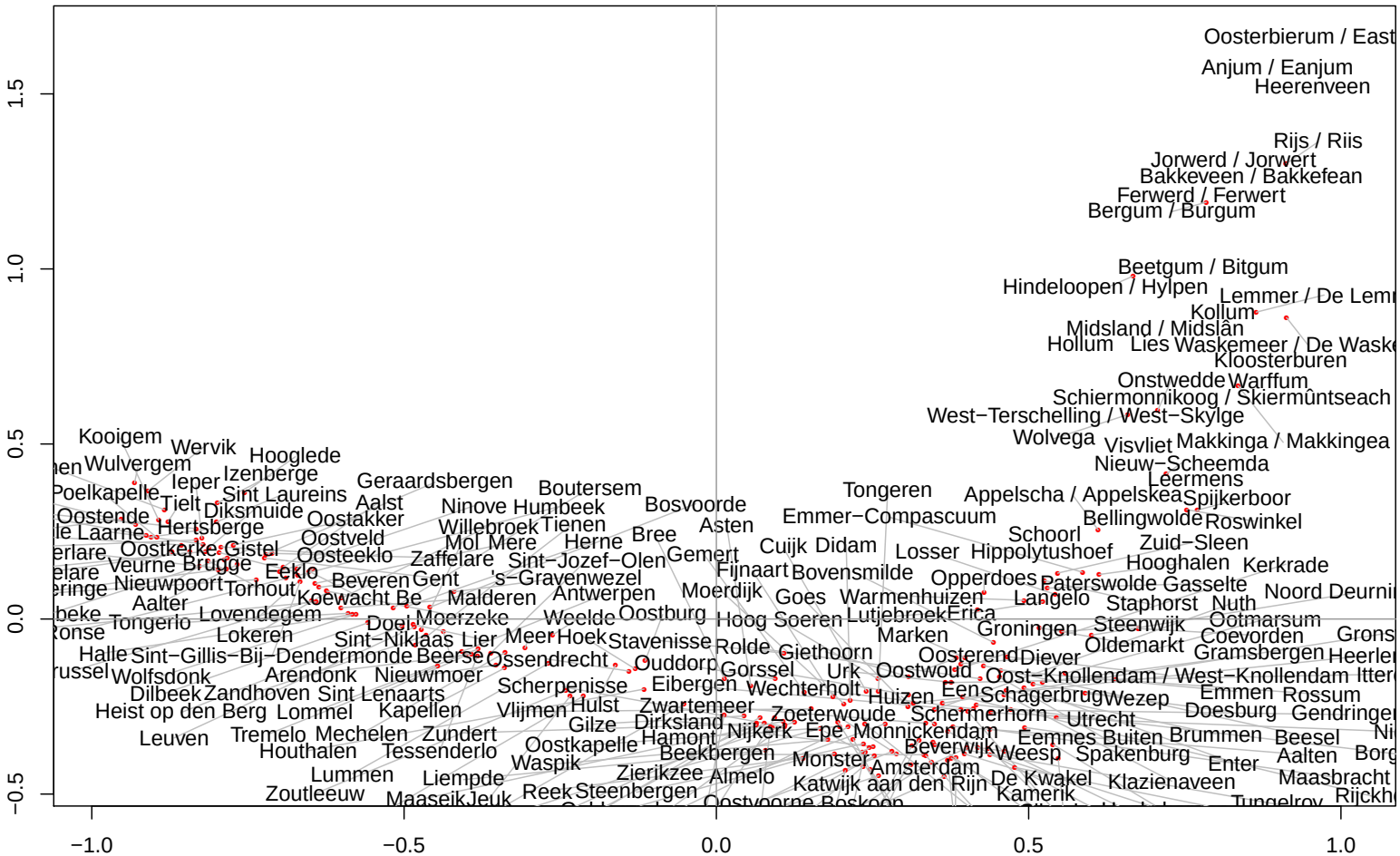
2. which is converted into a distance matrix

	Midsland	Lies	West-Terschelling	Oosterend	Hollum	..
Midsland	0					
Lies	5.099	0				
West-Terschelling	4.795	4.795	0			
Oosterend	6.000	5.656	5.196	0		
Hollum	4.898	5.099	4.358	5.477	0	
Schiermonnikoog	5.000	4.358	4.242	5.385	4.582	..
Ferwerd	5.099	5.099	5.385	6.633	4.690	..
Anjum	5.385	5.567	5.830	6.557	4.795	..
Kollum	4.795	4.358	4.898	5.385	5.196	..
Visvliet	5.169	5.567	5.477	5.744	5.385	..
...	...	...	...	...	...	..

Quantitative analysis: Correspondence Analysis

The Correspondence Analysis proceeds in three steps:

3. which undergoes dimension reduction:



Quantitative analysis: Cluster Analysis

- Cluster Analysis is a technique for combining observations into groups (clusters)

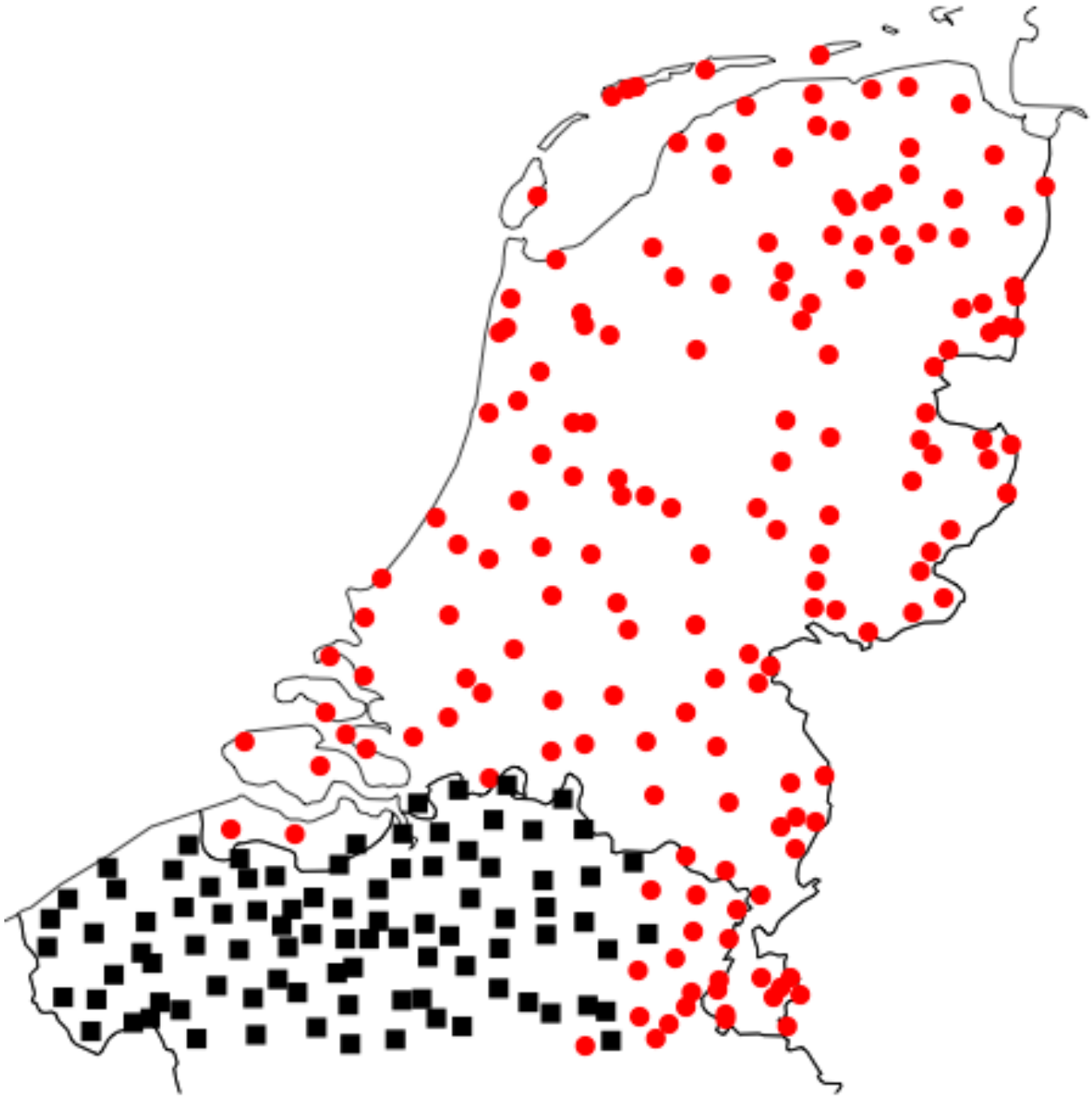
Quantitative analysis: Cluster Analysis

- Cluster Analysis is a technique for combining observations into groups (clusters)
- we are performing the Cluster Analysis based on the results of the Correspondence Analysis

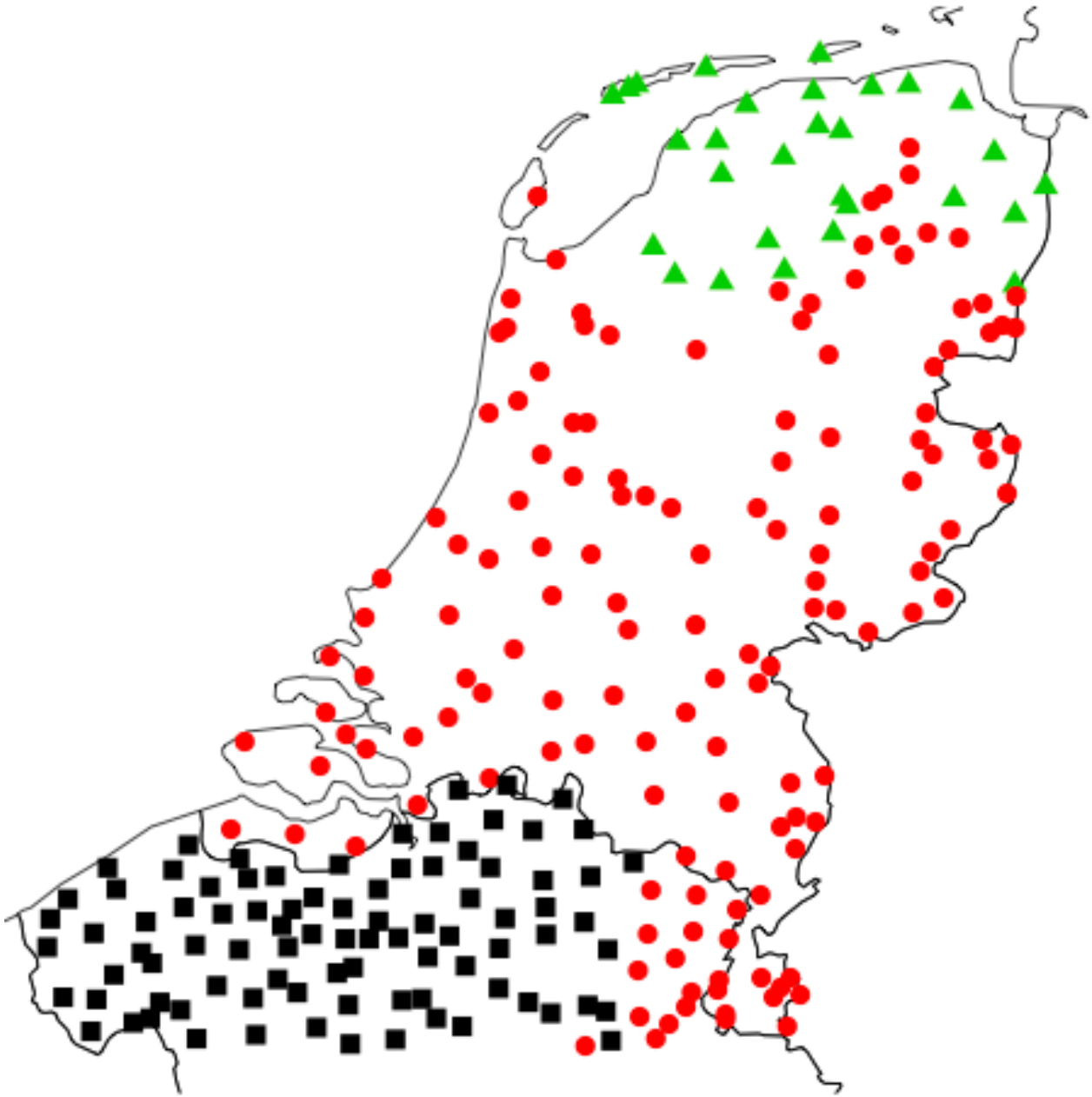
Quantitative analysis: Cluster Analysis

- Cluster Analysis is a technique for combining observations into groups (clusters)
- we are performing the Cluster Analysis based on the results of the Correspondence Analysis
- varying the number of clusters is a way of varying the granularity of the morphosyntactic variation patterns we are looking at

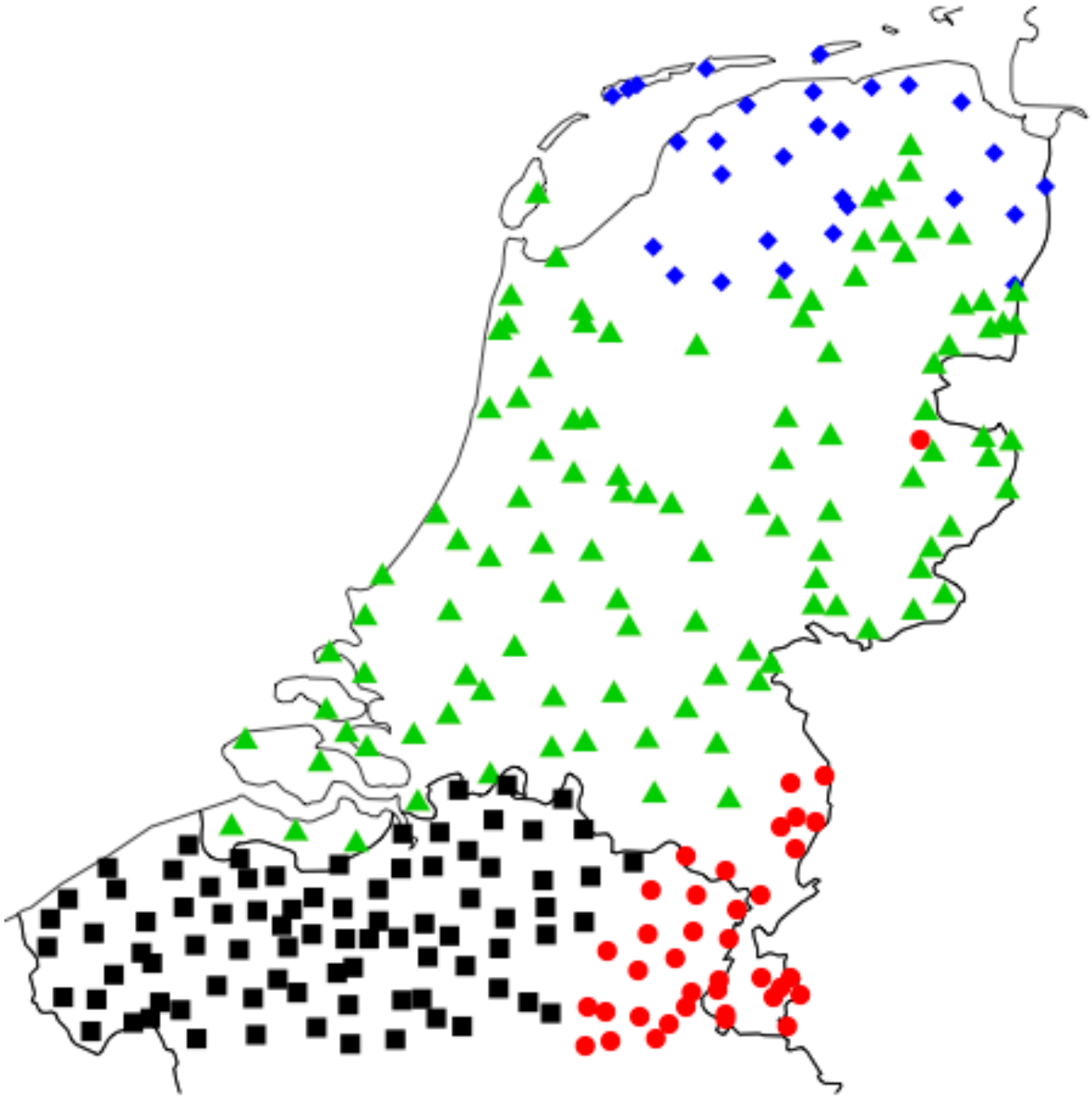
Quantitative analysis: Cluster Analysis



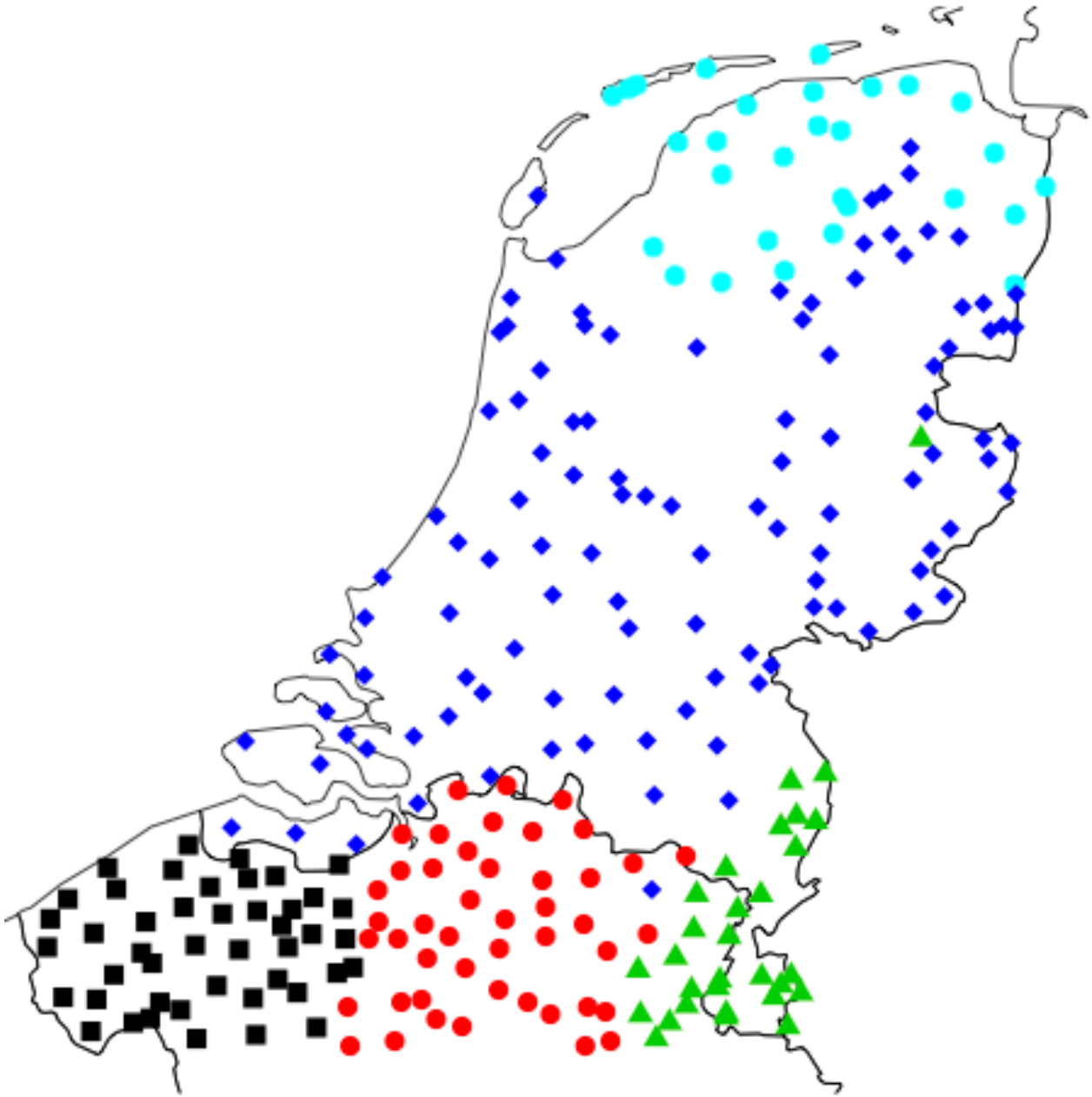
Quantitative analysis: Cluster Analysis



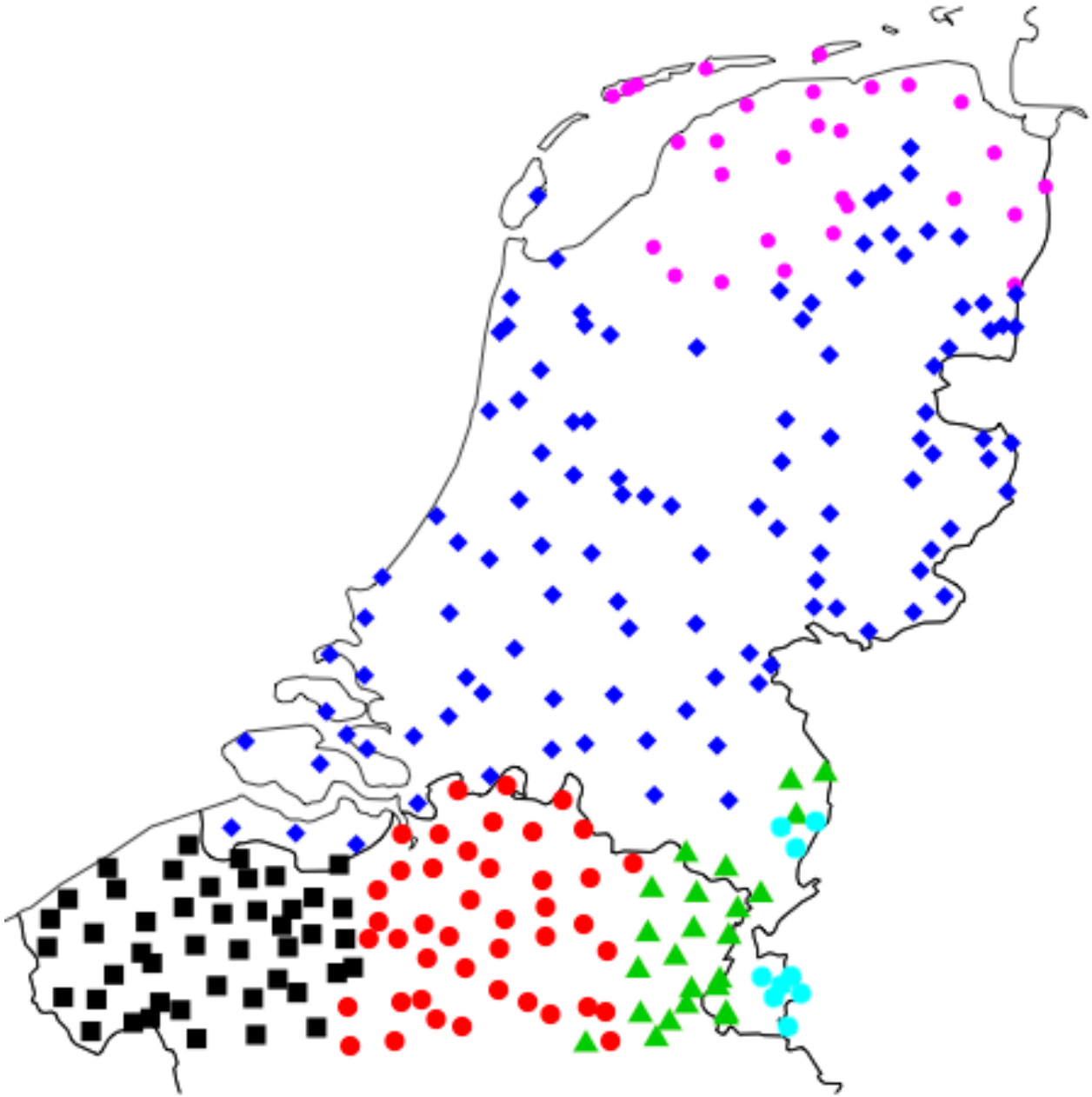
Quantitative analysis: Cluster Analysis



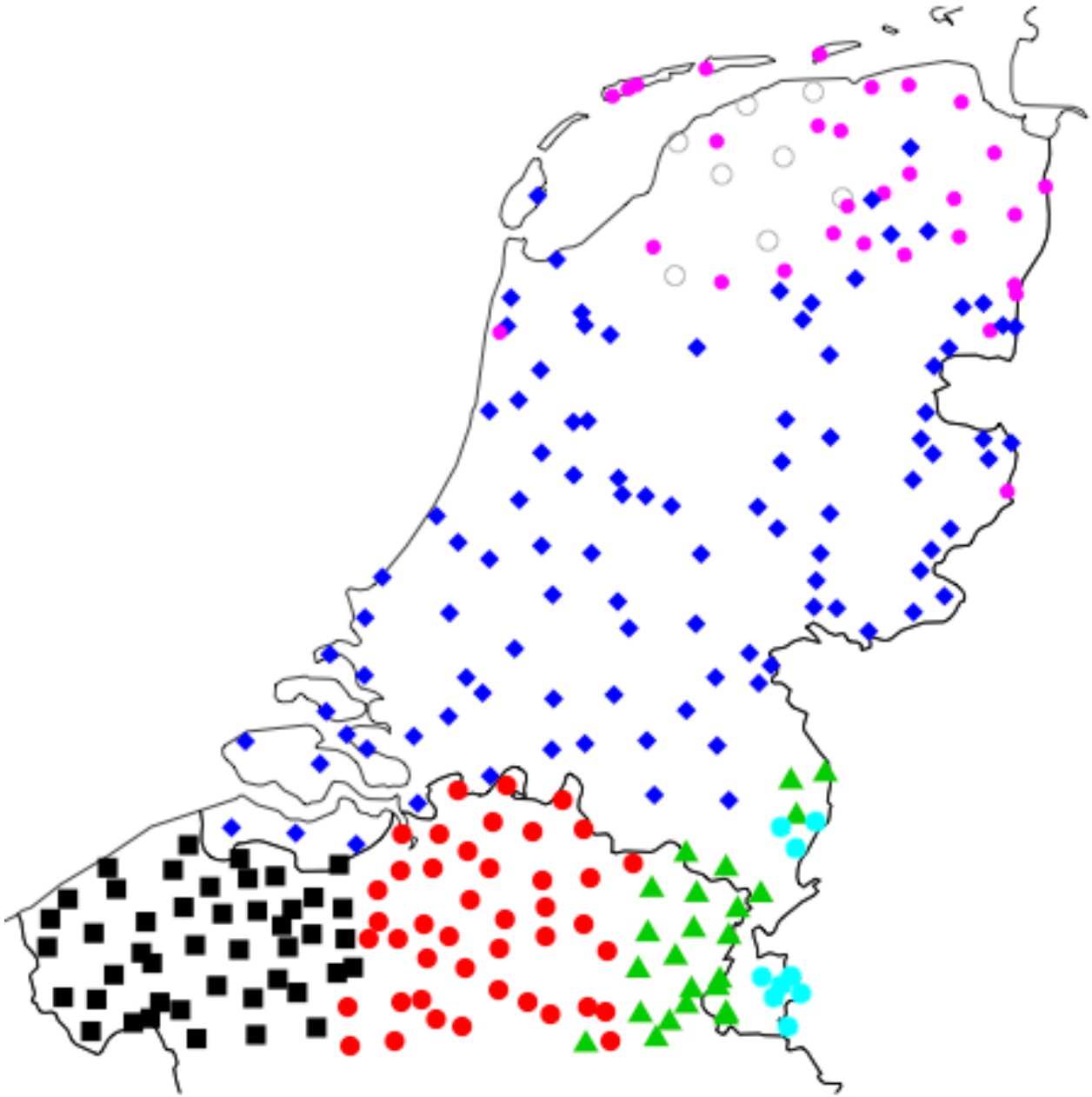
Quantitative analysis: Cluster Analysis



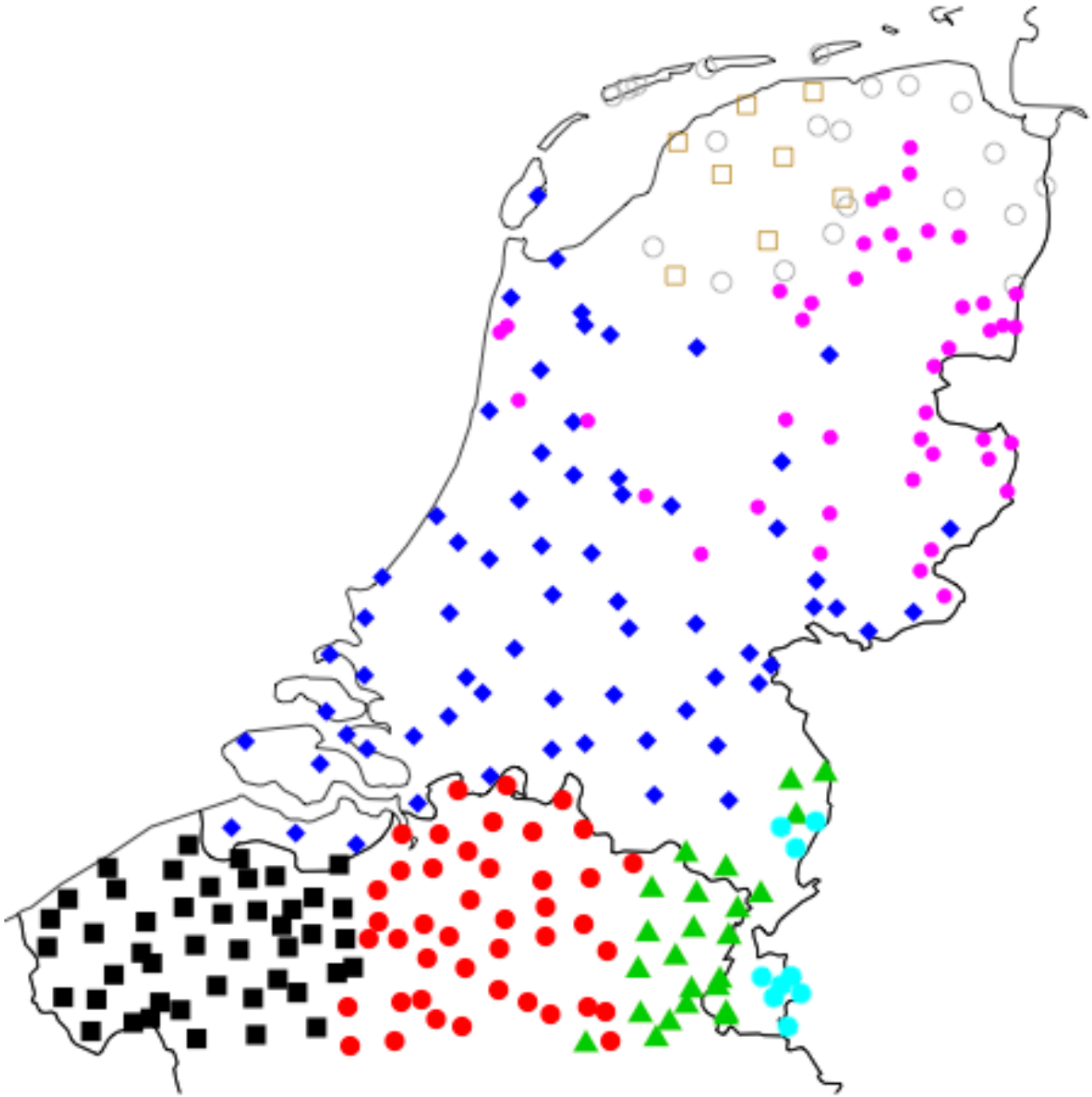
Quantitative analysis: Cluster Analysis



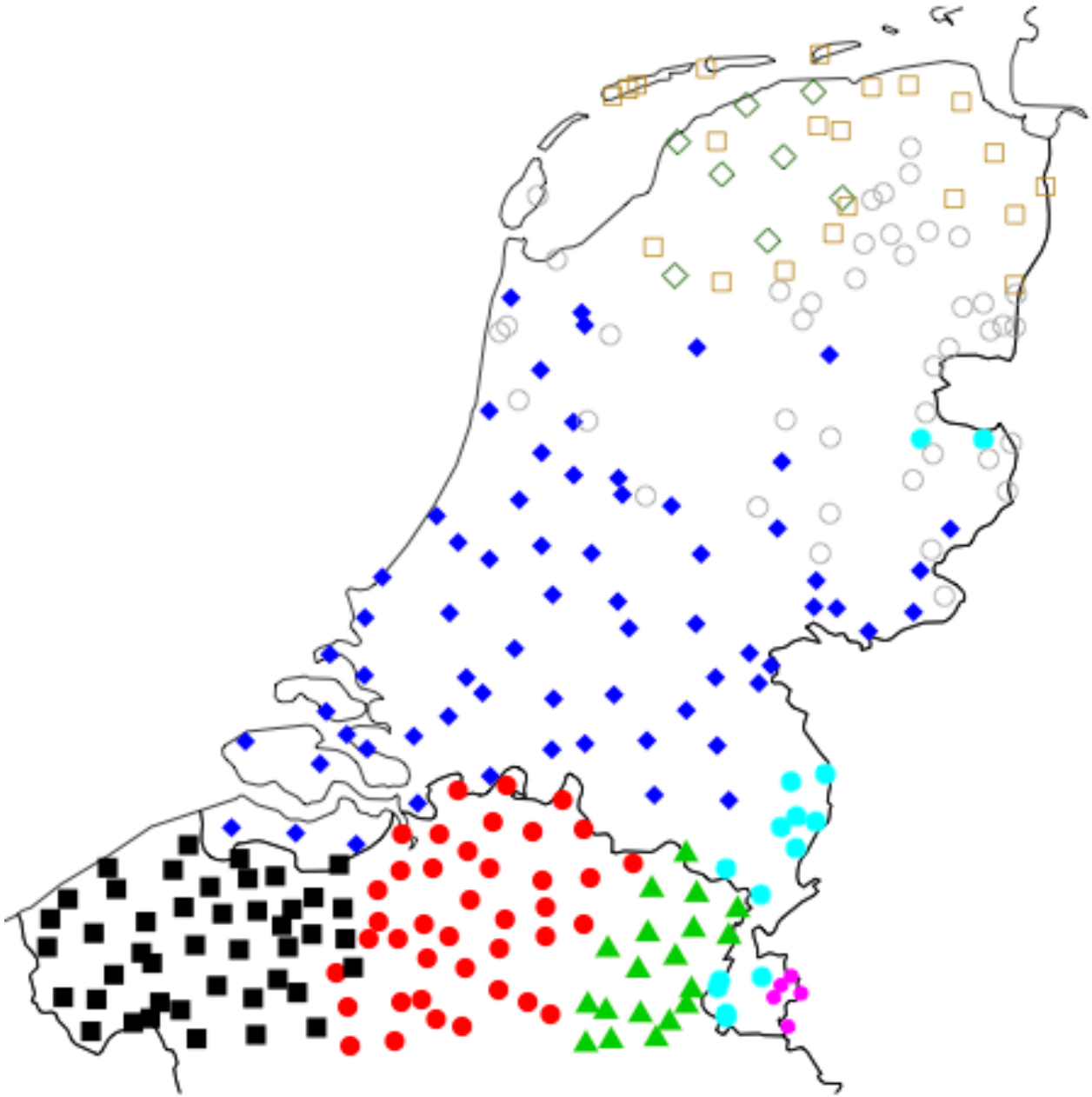
Quantitative analysis: Cluster Analysis



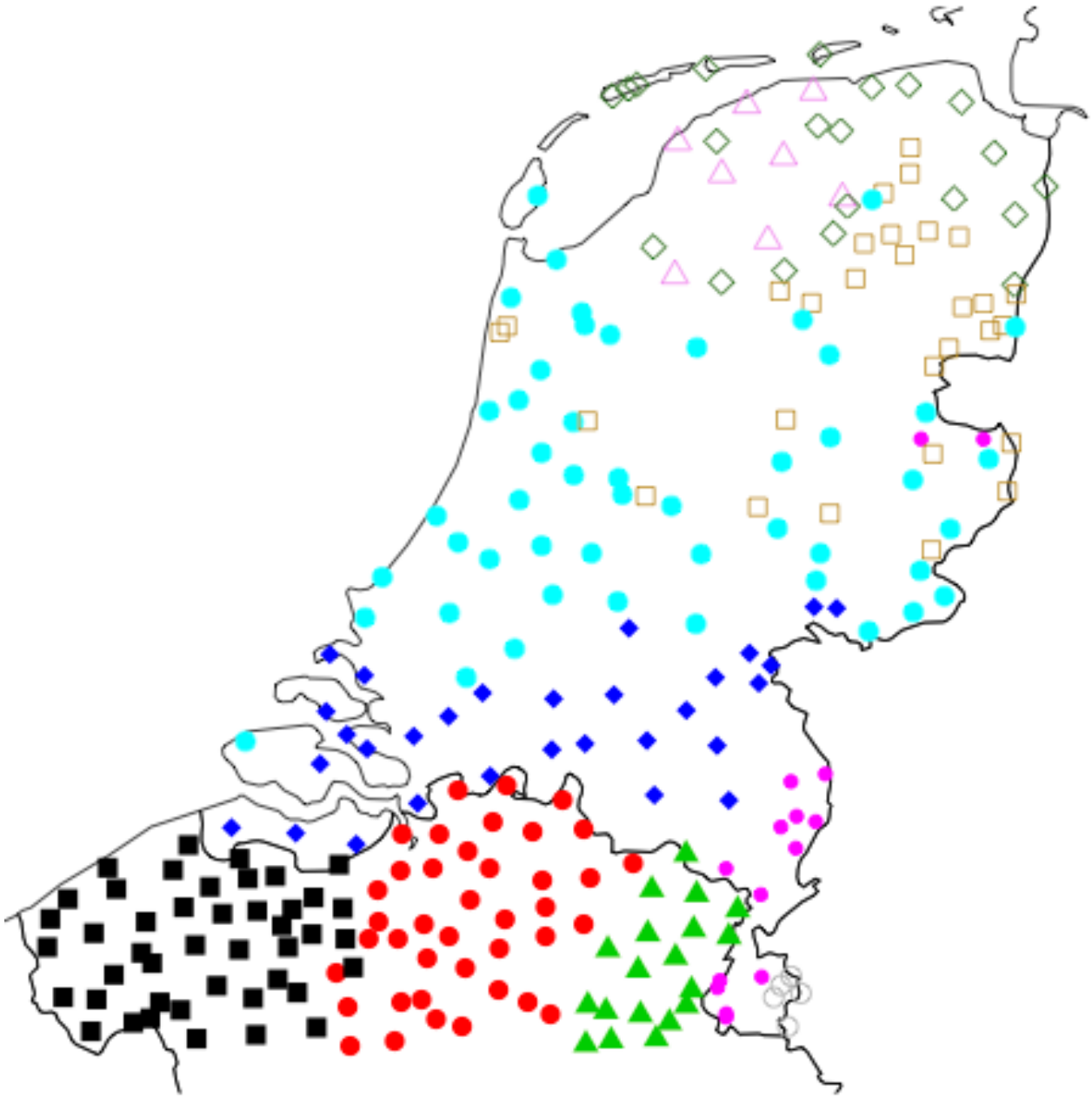
Quantitative analysis: Cluster Analysis



Quantitative analysis: Cluster Analysis



Quantitative analysis: Cluster Analysis



Quantitative analysis: Cluster Analysis

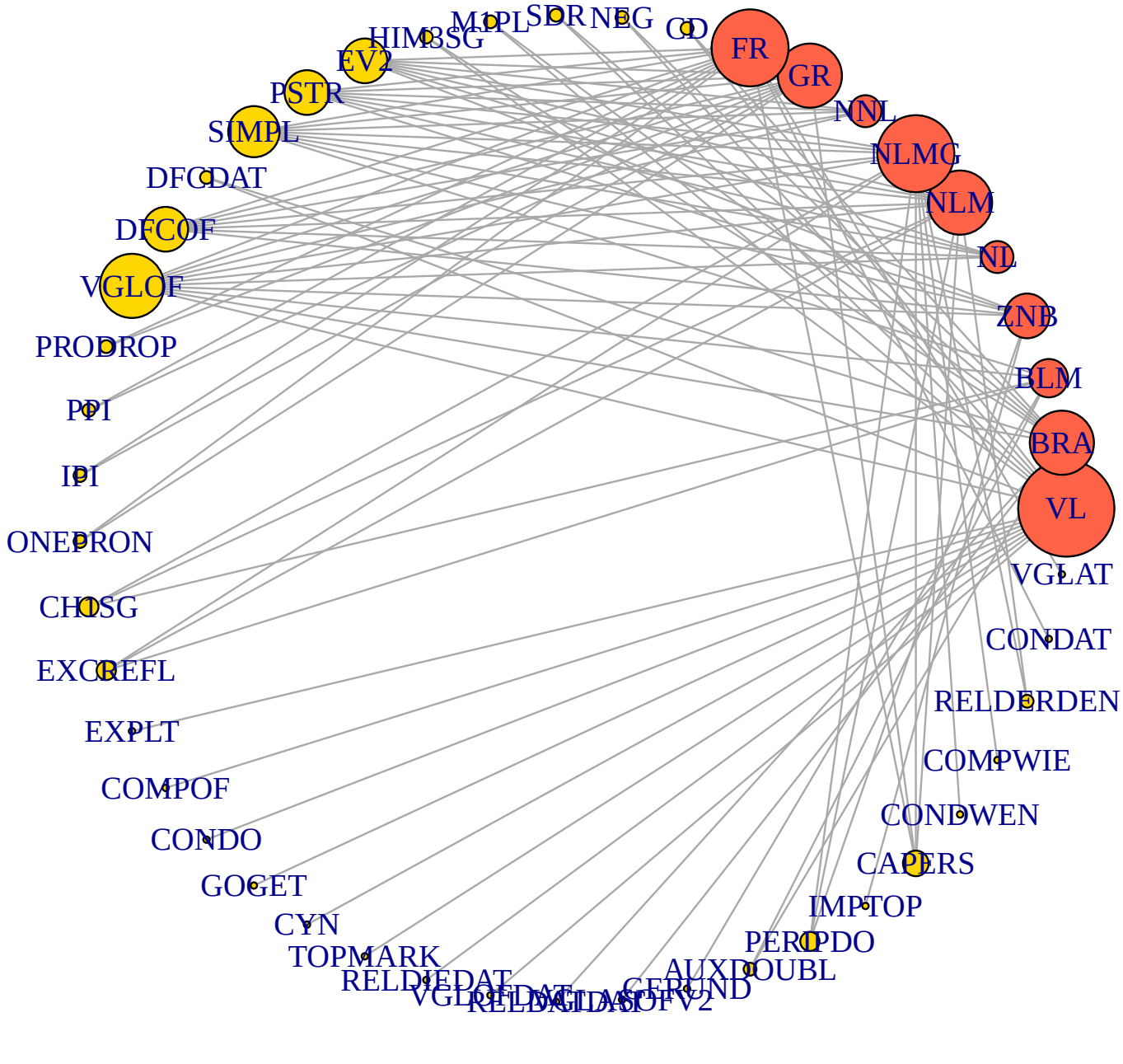
Quantitative analysis: Cluster Description

- we can now list, for any cluster (of any granularity) which linguistic phenomena are significantly more present in that cluster than would be expected by chance

Quantitative analysis: Cluster Description

- we can now list, for any cluster (of any granularity) which linguistic phenomena are significantly more present in that cluster than would be expected by chance
- in other words, which linguistic features are characteristic for which dialect area?

Quantitative analysis: Cluster Description



Quantitative analysis: Conclusion

- the quantitative analysis has allowed us to reduce the dataset from 260 dialect locations and 146 linguistic phenomena to (a maximum of) 10 dialect areas and 37 linguistic phenomena

Quantitative analysis: Conclusion

- the quantitative analysis has allowed us to reduce the dataset from 260 dialect locations and 146 linguistic phenomena to (a maximum of) 10 dialect areas and 37 linguistic phenomena
- those 37 phenomena do the bulk of the work towards explaining the variance in the data set

Quantitative analysis: Conclusion

- the quantitative analysis has allowed us to reduce the dataset from 260 dialect locations and 146 linguistic phenomena to (a maximum of) 10 dialect areas and 37 linguistic phenomena
- those 37 phenomena do the bulk of the work towards explaining the variance in the data set

$\rightarrow$ they will serve as input for the qualitative analysis

Outline

Main goals for today

Introduction: Kayne's dream

Quantitative analysis

- Correspondence Analysis
- Cluster Analysis
- Cluster Description
- Conclusion

Qualitative analysis

- Case study #1: PolP
- Case study #2: split DP
- Case study #3: split Force/Fin
- Combining the case studies: 7 parameters

The bigger picture: determinants of variation

Qualitative analysis

- **Central question:** to what extent can we make sense of the 37 phenomena retained in the quantitative analysis from a formal-theoretical point of view?

Qualitative analysis

- **Central question:** to what extent can we make sense of the 37 phenomena retained in the quantitative analysis from a formal-theoretical point of view?
- Three case studies characterizing the first split, i.e. the SOUTH (Flanders and Brabant/Antwerp in Belgium) vs. the NORTH (the Netherlands + part of Belgian Limburg):

Qualitative analysis

- **Central question:** to what extent can we make sense of the 37 phenomena retained in the quantitative analysis from a formal-theoretical point of view?
- Three case studies characterizing the first split, i.e. the SOUTH (Flanders and Brabant/Antwerp in Belgium) vs. the NORTH (the Netherlands + part of Belgian Limburg):
  1. a separate polarity phrase

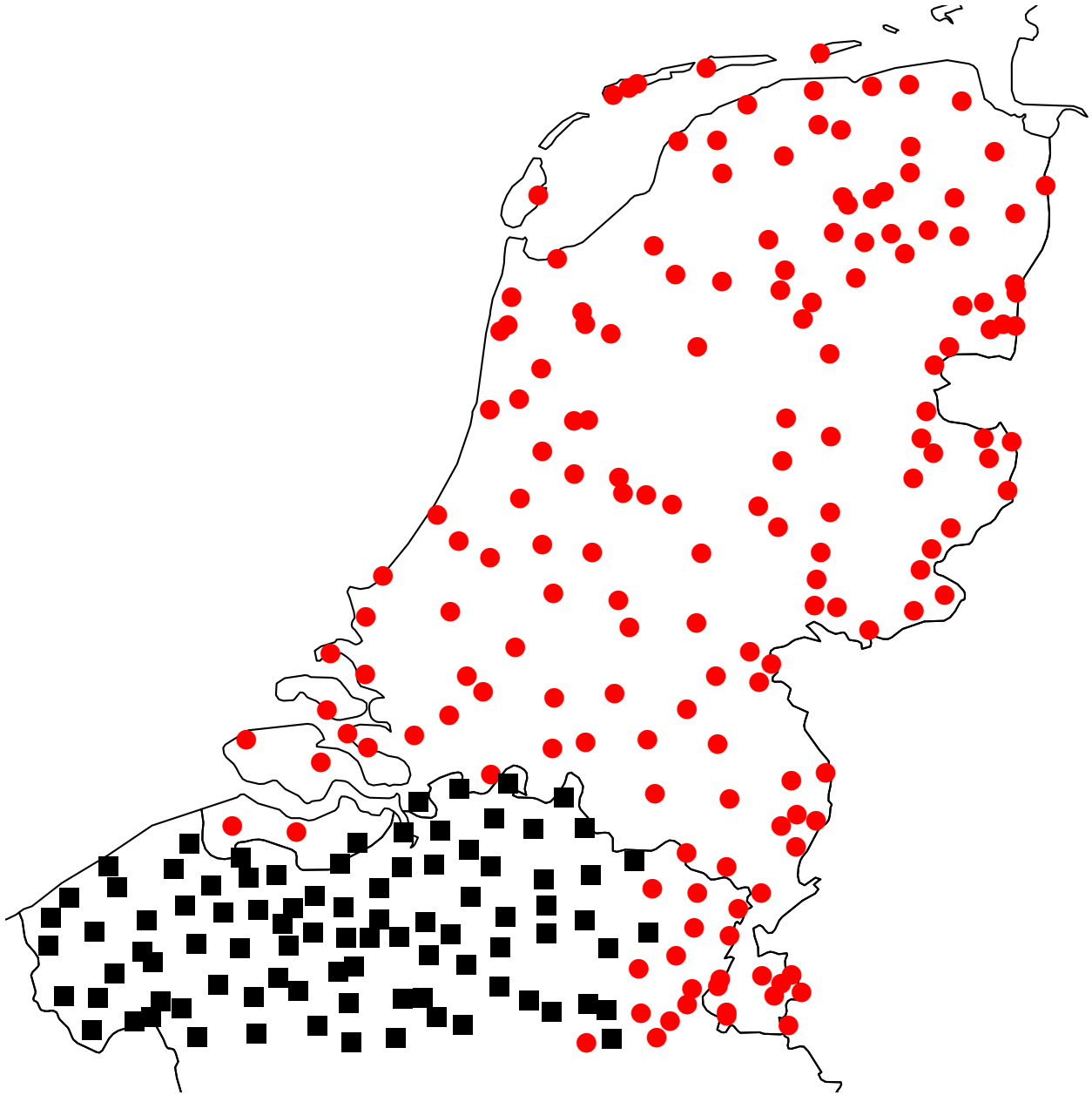
Qualitative analysis

- **Central question:** to what extent can we make sense of the 37 phenomena retained in the quantitative analysis from a formal-theoretical point of view?
- Three case studies characterizing the first split, i.e. the SOUTH (Flanders and Brabant/Antwerp in Belgium) vs. the NORTH (the Netherlands + part of Belgian Limburg):
  1. a separate polarity phrase
  2. a split DP-layer

Qualitative analysis

- **Central question:** to what extent can we make sense of the 37 phenomena retained in the quantitative analysis from a formal-theoretical point of view?
- Three case studies characterizing the first split, i.e. the SOUTH (Flanders and Brabant/Antwerp in Belgium) vs. the NORTH (the Netherlands + part of Belgian Limburg):
  1. a separate polarity phrase
  2. a split DP-layer
  3. a split Force/Fin-layer

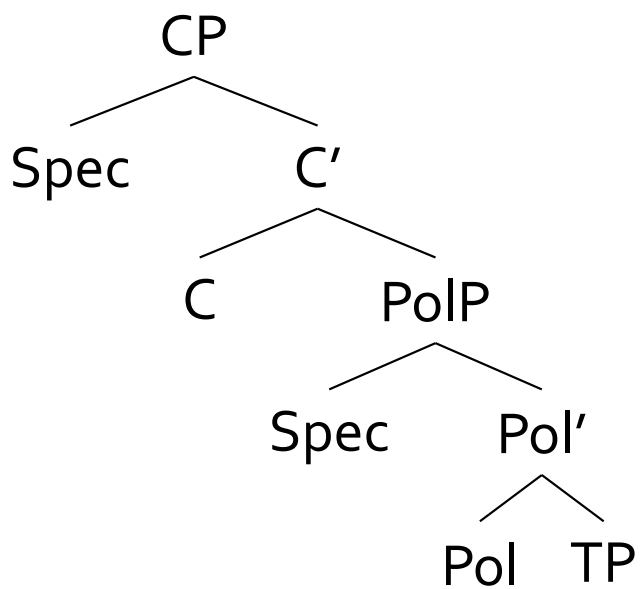
Qualitative analysis



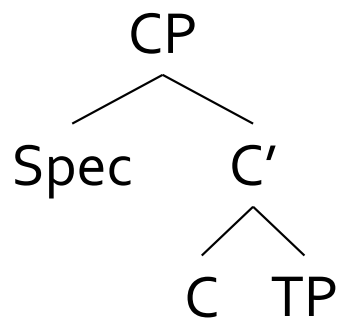
Case study #1: PolP

the SPLIT C-POL Parameter
The CP-domain {does/does not} project a separate PolP.

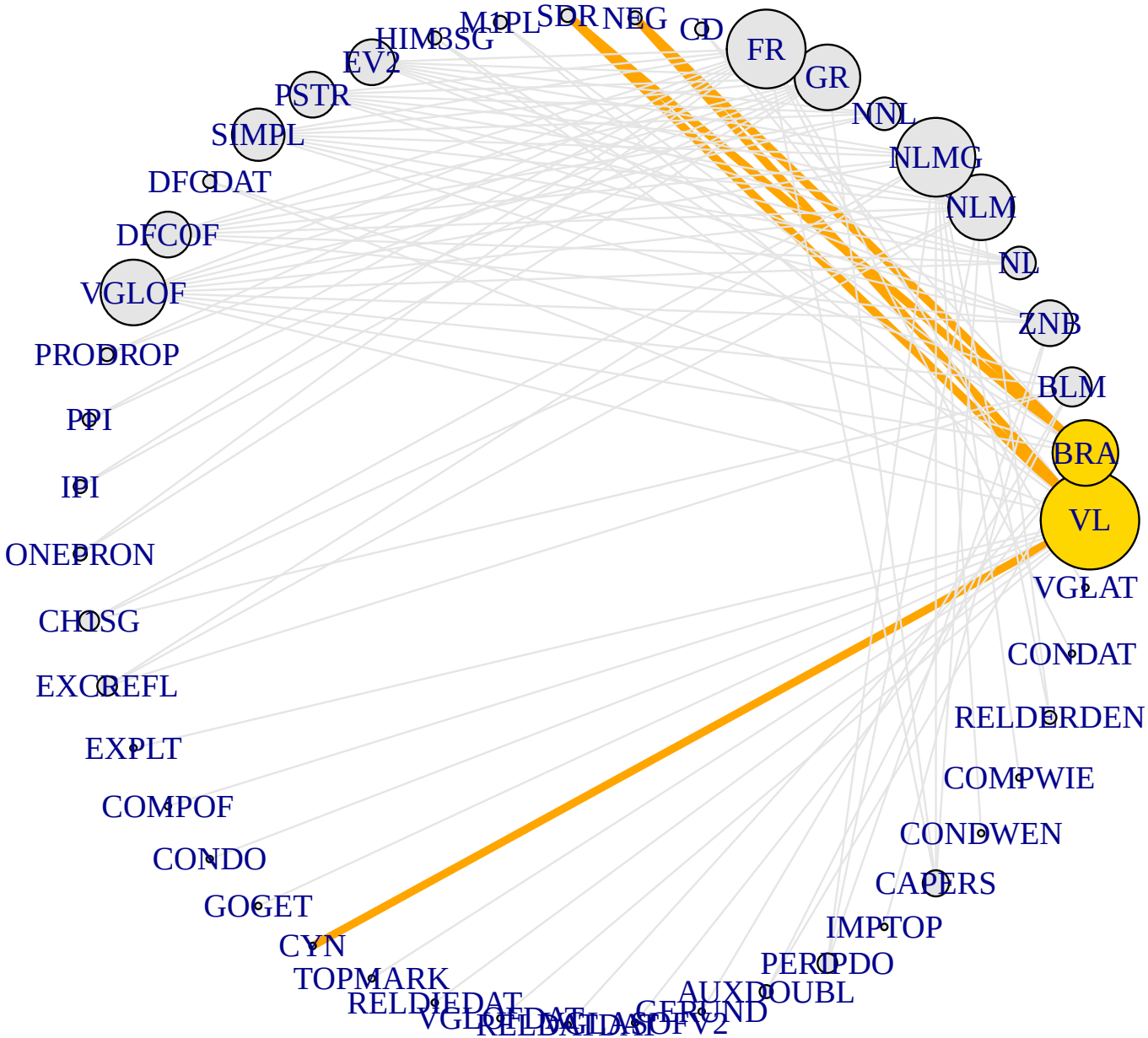
+ **Split C-Pol-parameter**



– **Split C-Pol-parameter**



Case study #1: PolP



Case study #1: PolP

- The following phenomena are characteristic of the South:

short *do* replies

(1) A: IJ zal nie komen. B: **IJ doet**.
he will not come he does
'A: He won't come. B: Yes, he will.'

negative clitic

(2) K **en** goa nie noar schole.
I NEG go not to school
'I'm not going to school.'

clitics on *yes* and *no*

(3) A: Wilde nog koffie, Jan? B: **Ja-k**.
want.you PART coffee Jan Yes-I
'A: Do you want some more coffee, Jan? B: Yes.'

Case study #1: PolP

(4) A: IJ zal nie komen. B: **IJ doet**.
he will not come he does
'A: He won't come. B: Yes, he will.

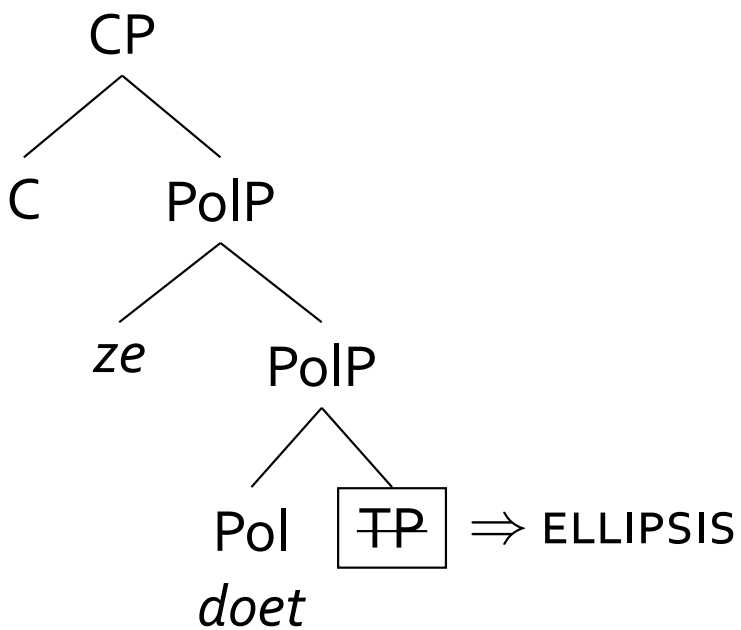
- van Craenenbroeck (2010): short *do* replies only occur in non-embedded contradictory polar replies to declarative clauses $\rightarrow$ TP-ellipsis licensed by a left peripheral polarity head:

Case study #1: PolP

(4) A: IJ zal nie komen. B: **IJ doet**.
he will not come he does
'A: He won't come. B: Yes, he will.

- van Craenenbroeck (2010): short *do* replies only occur in non-embedded contradictory polar replies to declarative clauses → TP-ellipsis licensed by a left peripheral polarity head:

(5)



Case study #1: PolP

- **supporting evidence:** short *do* replies are only compatible with high left-peripheral adverbs:

(6) A: Jef zeit da gou veel geldj etj. B: K'en duu { pertang
Jef says that you much money have I.NEG doe however
/ * nie mieje. }
not anymore
'A: Jef says you have a lot of money. B: I don't, however/*anymore.'

Case study #1: PolP

- **supporting evidence:** short *do* replies are only compatible with high left-peripheral adverbs:

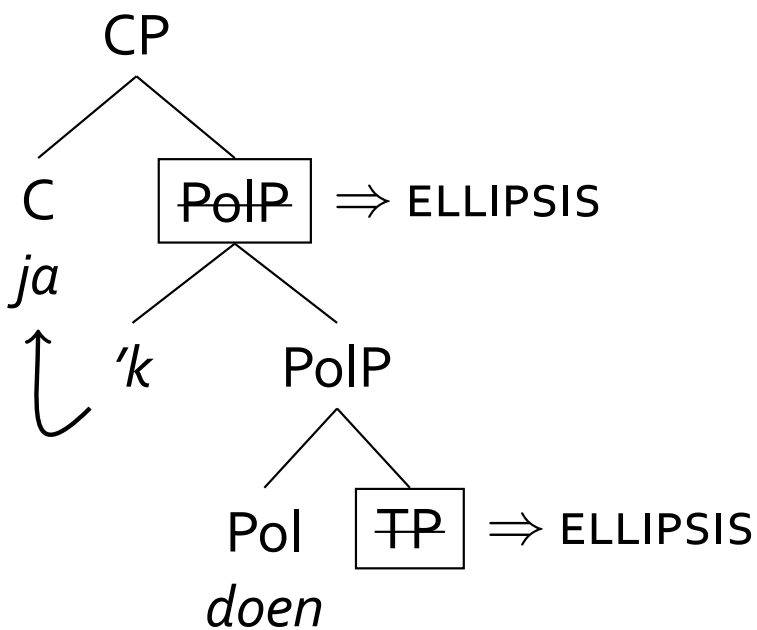
(6) A: Jef zeit da gou veel geldj etj. B: K'en duu { pertang
Jef says that you much money have I.NEG doe however
/ * nie mieje. }
not anymore
'A: Jef says you have a lot of money. B: I don't, however/*anymore.'

- the negative clitic *en* also fits this pattern: it too occupies a high Pol-head in the left periphery (van Craenenbroeck 2010).

Case study #1: PolP

- the occurrence of clitics on 'yes' and 'no' are derived from short *do* replies: they involve further (higher) ellipsis of an already truncated structure (van Craenenbroeck 2010)

(7)



Case study #1: PolP

the SPLIT C-POL Parameter
The CP-domain {does/does not} project a separate PolP.

- **SOUTH**: the CP-domain DOES project a separate PolP
- **NORTH**: the CP-domain DOES NOT project a separate PolP.

Case study #1: PolP

the SPLIT C-POL Parameter
The CP-domain {does/does not} project a separate PolP.

- **SOUTH**: the CP-domain DOES project a separate PolP
- **NORTH**: the CP-domain DOES NOT project a separate PolP.

	NEG	SDR	CYN
SOUTH (FL)	+	+	+
SOUTH (BRA)	+	+	-
NORTH	-	-	-

Case study #1: PolP

the SPLIT C-POL Parameter
The CP-domain {does/does not} project a separate PolP.

- **SOUTH**: the CP-domain DOES project a separate PolP
- **NORTH**: the CP-domain DOES NOT project a separate PolP.

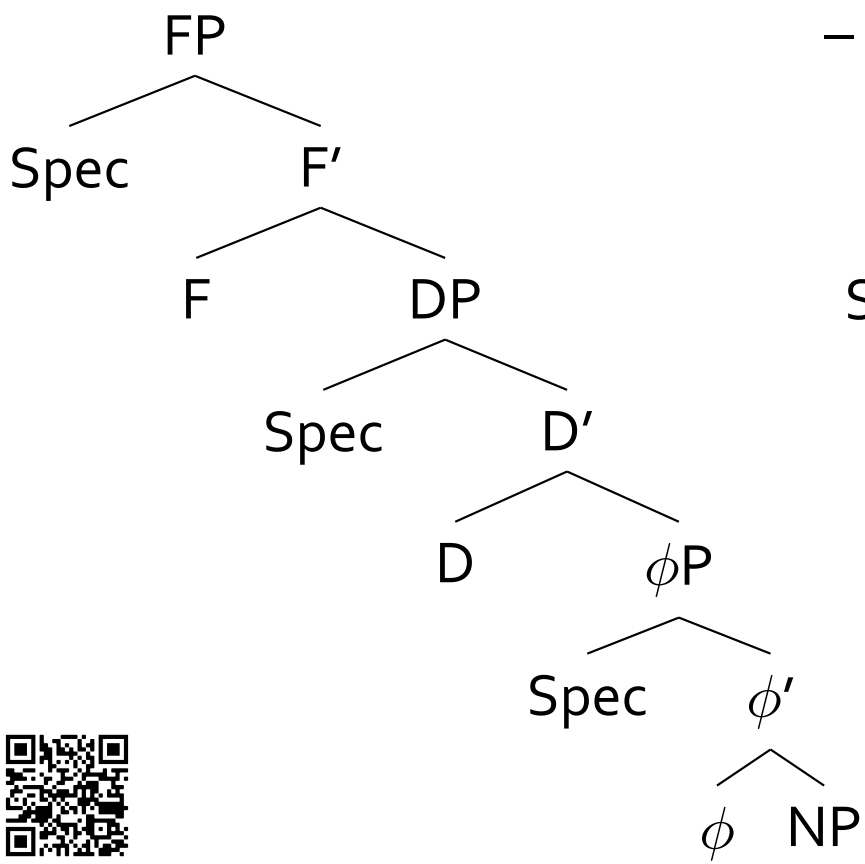
	NEG	SDR	CYN
SOUTH (FL)	+	+	+
SOUTH (BRA)	+	+	-
NORTH	-	-	-

note: For CYN a SPLIT C-POL parameter is a necessary but not a sufficient condition. A further parameter is necessary to license CYN. This parameter is set to + in FL but not in BRA.

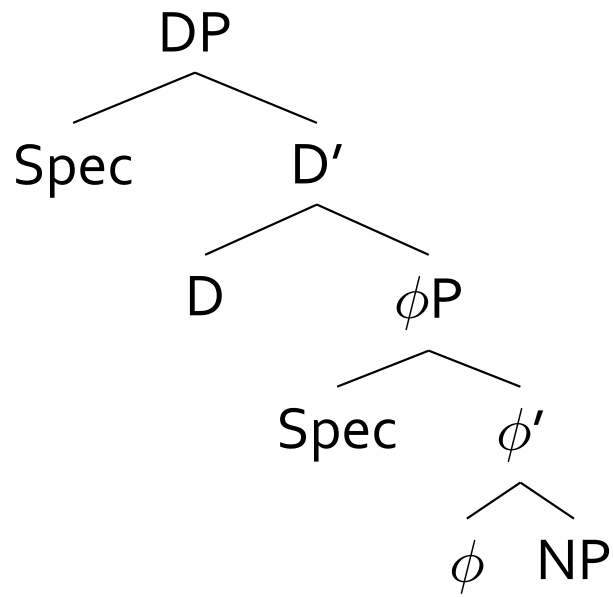
Case study #2: split DP

the SPLIT-D Parameter
DP {does/does not} have an extended left periphery.

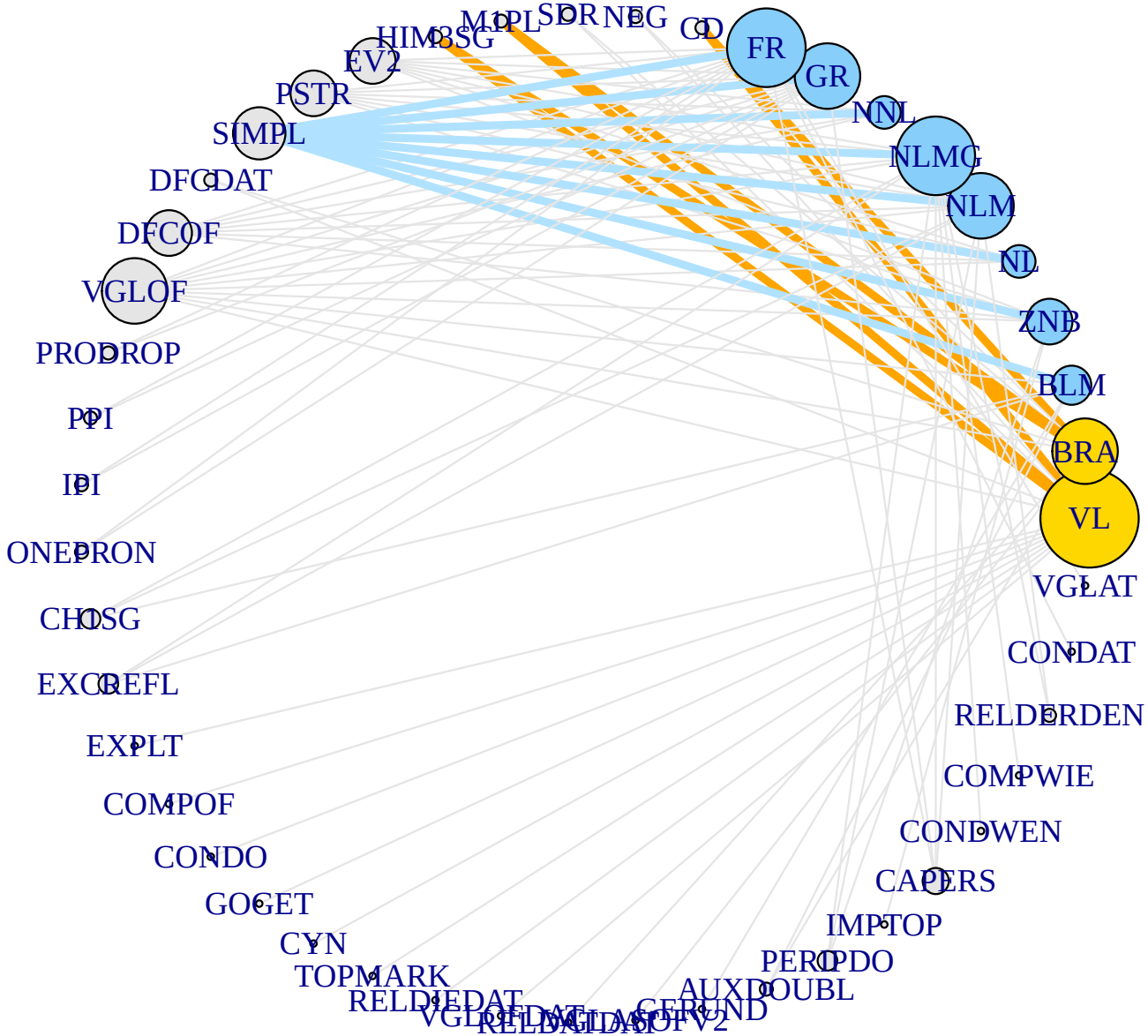
+ **Split D-parameter**



– **Split D-parameter**



Case study #2: split DP



Case study #2: split DP

- The following phenomena are characteristic of the South:

clitic doubling

(8) da-**ze** **zaaile** lachen.
that-they$_{\text{CLITIC}}$ they$_{\text{STRONG}}$ laugh
'that they are laughing.'

m-form of 1.pl subject pronoun

(9) **Me** zijn doa nooit geweest.
we are there never been
'We have never been there.'

accusative 3.sg.masc pronoun in subject position

(10) **Em** is dood.
him is dead
'He is dead.'

Case study #2: split DP

- In addition: complex plural pronouns in the South (11) and simplex plural pronouns in the North (12):

complex plural pronouns

(11) **Gu-lder** gelooft toch nie da **zu-lder** armer zijn
you-people believe PART not that they-people poorer are
dan **wu-lder**.
than we-people
'You won't believe that they are poorer than us.'

simplex plural pronouns

(12) **Jim** gelove jammer genoeg net dat
You_{pl}-SIMPLEX believe unfortunately enough not that
sij it minder ha dan **wij**
they-SIMPLEX it less have than we-SIMPLEX.
'Unfortunately you do not believe that they are less well off than we are.'

Case study #2: split DP

(13) da-**ze** **zaaile** lachen.
that-they$_{\text{CLITIC}}$ they$_{\text{STRONG}}$ laugh
'that they are laughing.'

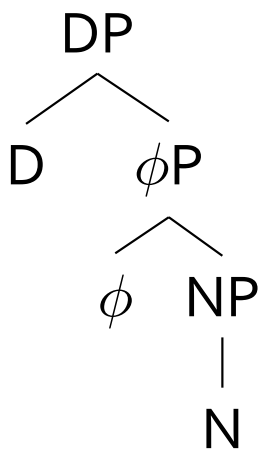
- **starting point:** van Craenenbroeck and van Koppen (2008)'s analysis of clitic doubling:

Case study #2: split DP

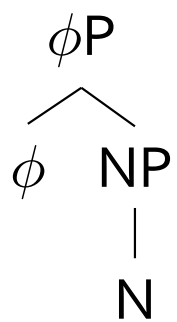
(13) da-**ze** **zaaile** lachen.
that-they$_{\text{CLITIC}}$ they$_{\text{STRONG}}$ laugh
'that they are laughing.'

- **starting point:** van Craenenbroeck and van Koppen (2008)'s analysis of clitic doubling:
- **step one:** strong pronouns in doubling dialects are pro-DPs, while subject clitics are pro-ϕPs (Déchaine and Wiltschko 2002)

(14) **strong subject pronoun**



(15) **subject clitic**

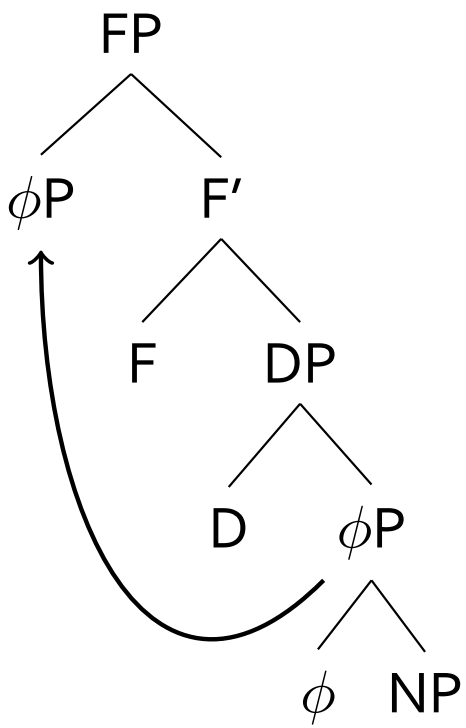


Case study #2: split DP

- **step two:** a clitic-doubled subject is base-generated as a big DP; clitics are the result of ϕP-movement into the extended left periphery of the DP

$\Rightarrow$ there has to be an additional layer above DP to host the movement of the clitic (FP) in order to avoid an anti-locality violation (Abels (2003)):

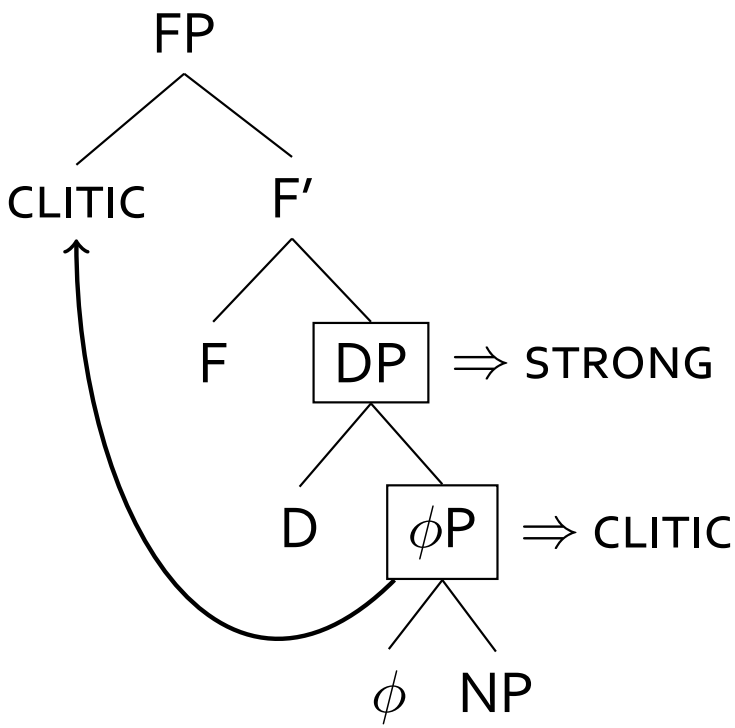
(16)



Case study #2: split DP

- **step three:** when the resulting structure is handed over to PF, the moved ϕP is spelled out as a subject clitic, and the DP as a strong pronoun

(17)



Case study #2: split DP

the SPLIT-D Parameter
DP {does/does not} have an extended left periphery.

- **SOUTH**: the DP-domain DOES have an extended left periphery
- **NORTH**: the DP-domain DOES NOT have an extended left periphery

Case study #2: split DP

the SPLIT-D Parameter
DP {does/does not} have an extended left periphery.

- **SOUTH**: the DP-domain DOES have an extended left periphery
- **NORTH**: the DP-domain DOES NOT have an extended left periphery

	CD
SOUTH	+
NORTH	-

Case study #2: split DP

- **supporting evidence:** Barbiers et al. (2016) argue for a similar big DP+movement-analysis for another linguistic phenomenon that is characteristic of the South: demonstrative doubling.

(18) **De die** zou k ik wiln op eetn.
the those would I_{CLITIC} I_{STRONG} want up eat
'I would like to eat those.'

Case study #2: split DP

- **step one**: the definite article in demonstrative doubling pronominalizes ϕP, i.e. the part of the DP-structure hosting the noun, numerals, and adjectives:

(19)
a. de dien
the that
'that one'

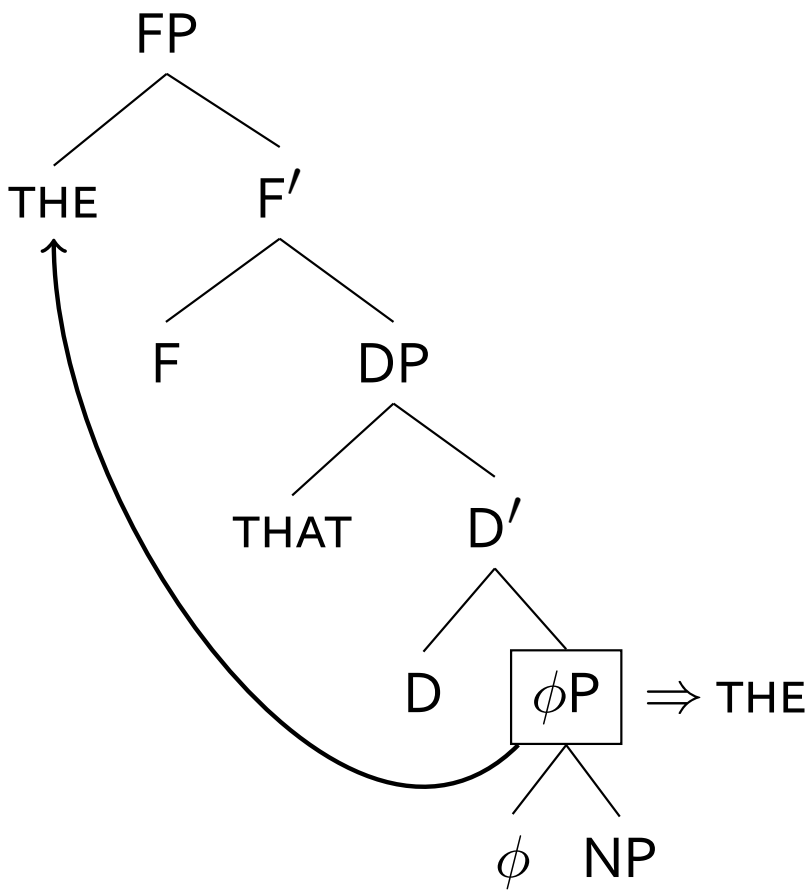
b. (* de) dien opa
the that grandfather
'that grandfather'

c. De dieje (* twee) (* rode) liggen op de tafel.
the those two red are on the table
'Those are on the table.'

Case study #2: split DP

- **step two**: ϕP moves into the left periphery of the DP; anti-locality again requires that the left periphery of DP be complex.

(20)



Case study #2: split DP

Further supporting evidence from possessive structures:

1. dialects with a negative setting for the D-parameter lack demonstrative doubling because they lack the additional DP-layer (no landing site for the definite article)

Case study #2: split DP

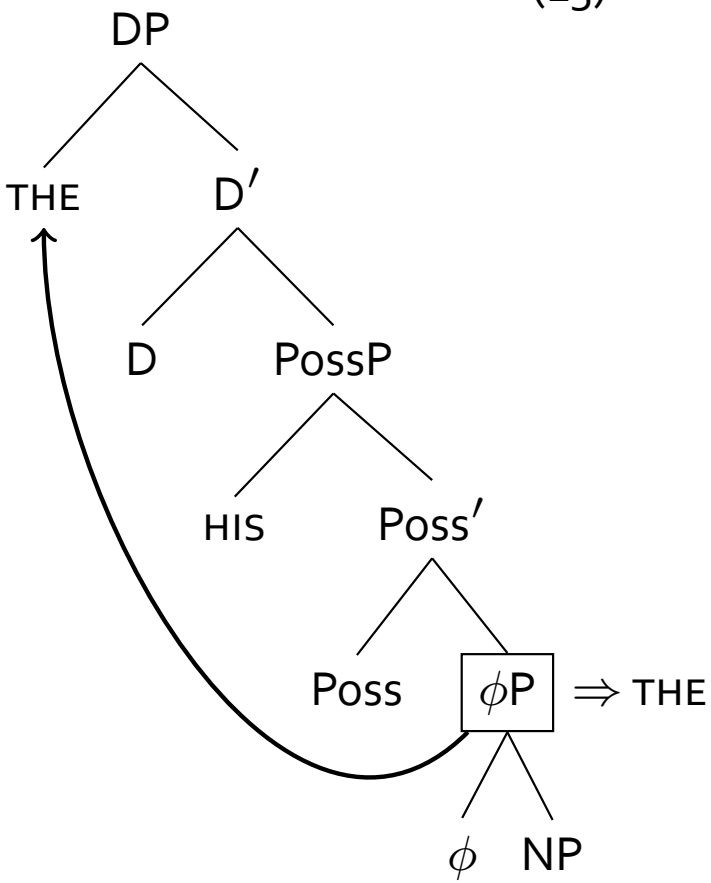
Further supporting evidence from possessive structures:

1. dialects with a negative setting for the D-parameter lack demonstrative doubling because they lack the additional DP-layer (no landing site for the definite article)
2. these dialects (as well as the dialects with a positive setting for the D-parameter) do have THE+possessive pronoun:

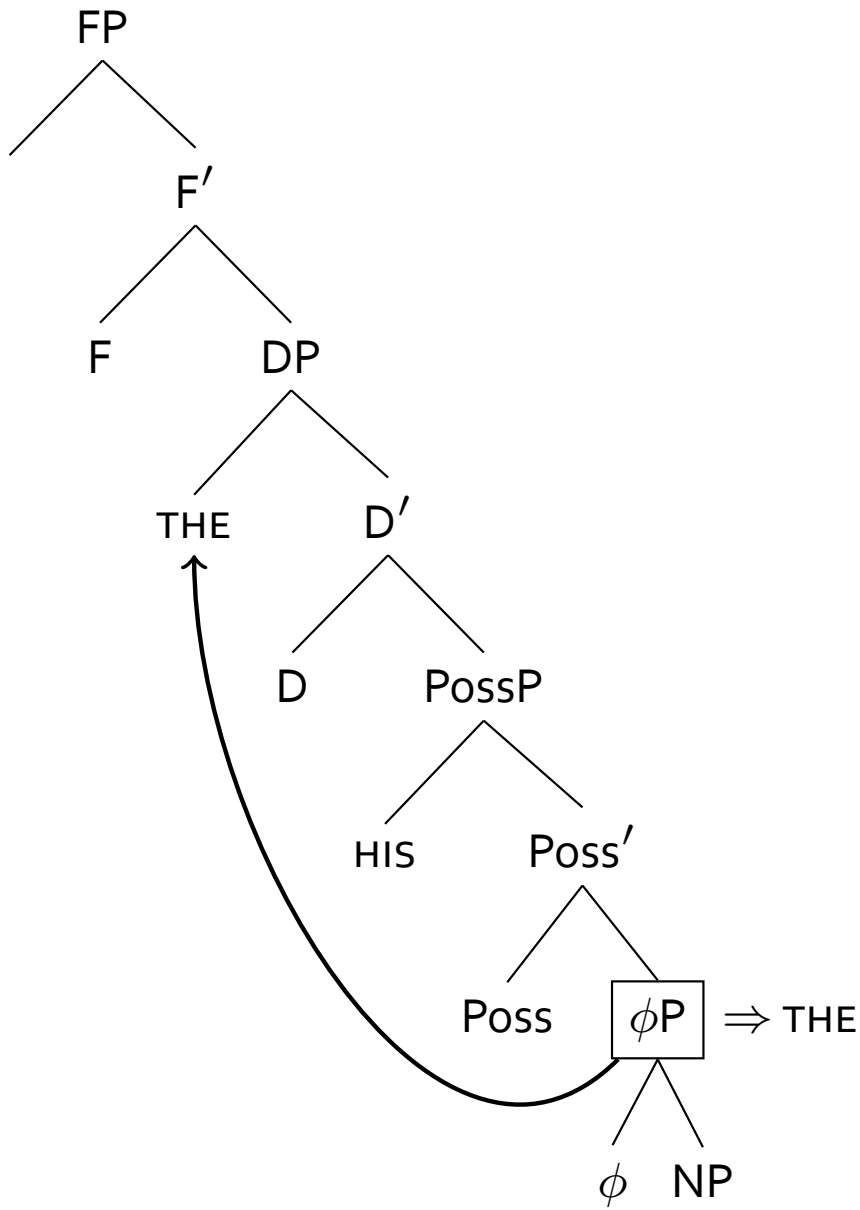
(21) Ik vin **de zaine** ech geweldig.
I find the his really great
'I find his really great.' (–split DP parameter)

Case study #2: split DP

(22)



(23)



Case study #2: split DP

3. however, only dialects with a positive setting of the D-parameter allow doubling in THE+possessive pronoun:

Case study #2: split DP

3. however, only dialects with a positive setting of the D-parameter allow doubling in THE+possessive pronoun:

(24) **Toin de zijnen** is geweldig.
Teun the his is great
'Teun's is great.' (+SPLIT DP-Parameter)

Case study #2: split DP

3. however, only dialects with a positive setting of the D-parameter allow doubling in THE+possessive pronoun:

(24) **Toin de zijnen** is geweldig.
Teun the his is great
'Teun's is great.' (+SPLIT DP-Parameter)

(25) Ik vin (* **Teun**) **de zaine** ech geweldig.
I find Teun the his really great
'I find his really great.' (–SPLIT DP-Parameter)

Case study #2: split DP

3. however, only dialects with a positive setting of the D-parameter allow doubling in THE+possessive pronoun:

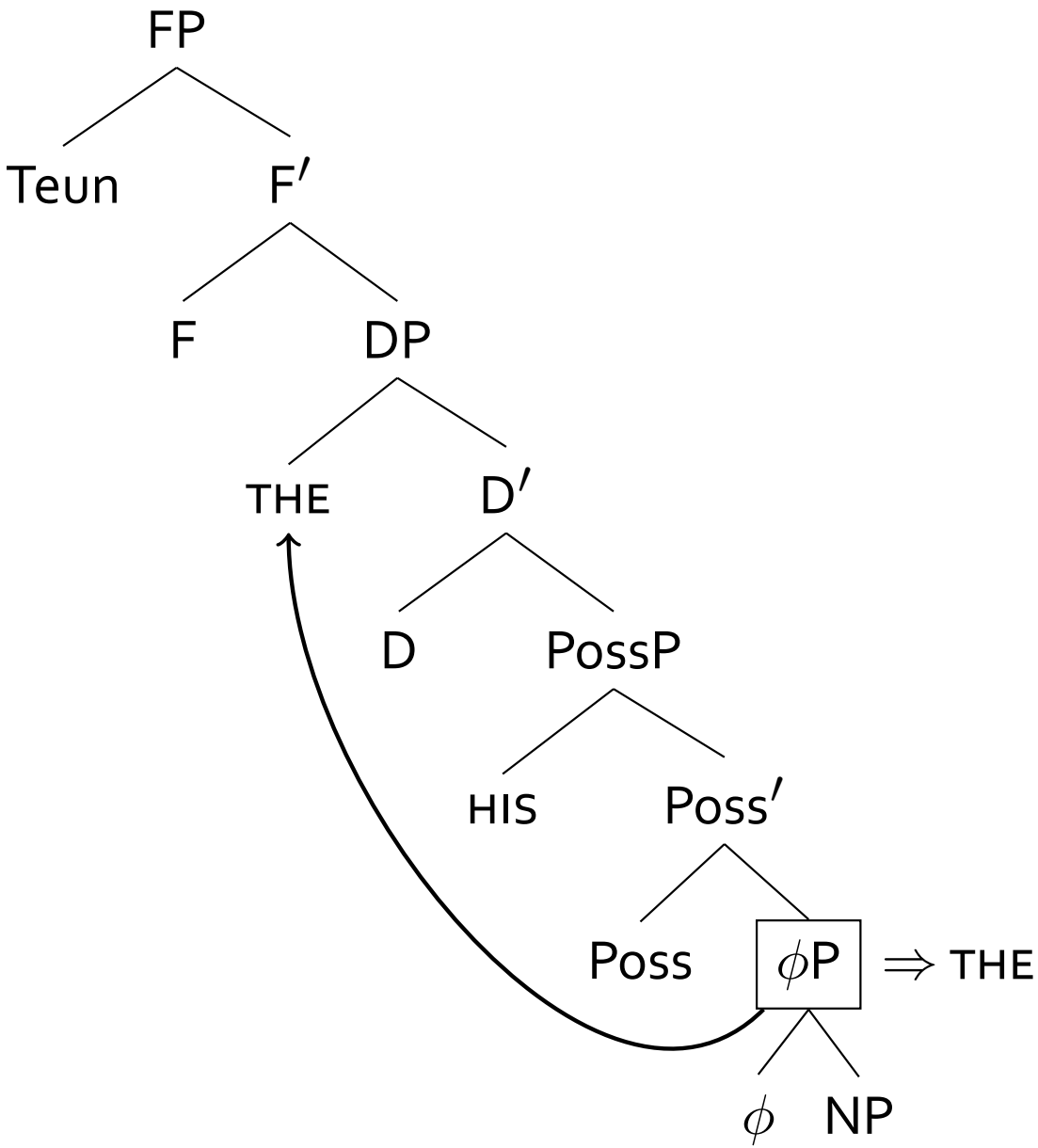
(24) **Toin de zijnen** is geweldig.
Teun the his is great
'Teun's is great.' (+SPLIT DP-Parameter)

(25) Ik vin (* **Teun**) **de zaine** ech geweldig.
I find Teun the his really great
'I find his really great.' (–SPLIT DP-Parameter)

→ this can be explained by the presence of an additional layer in the +Split D-dialects:

Case study #2: split DP

(26)



Case study #2: split DP

the SPLIT-D Parameter
DP {does/does not} have an extended left periphery.

- **SOUTH**: the DP-domain DOES have an extended left periphery
- **NORTH**: the DP-domain DOES NOT have an extended left periphery

Case study #2: split DP

the SPLIT-D Parameter
DP {does/does not} have an extended left periphery.

- **SOUTH**: the DP-domain DOES have an extended left periphery
- **NORTH**: the DP-domain DOES NOT have an extended left periphery

	CD	DD	THE POSS	POSS THE POSS
SOUTH	+	+	+	+
NORTH	-	-	+	-

Case study #2: split DP

- Can this analysis also give us a handle on the variation concerning pronouns?

m-form of 1.pl subject pronoun

(27) **Me** zijn doa nooit geweest.
we are there never been
'We have never been there.'

accusative 3.sg.masc pronoun in subject position

(28) **Em** is dood.
him is dead
'He is dead'

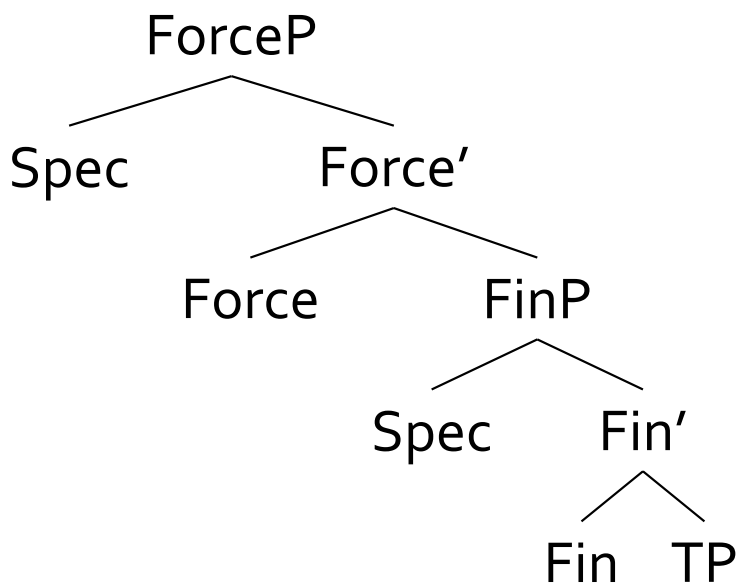
complex plural pronouns

(29) **Gu-lder** gelooft toch nie da **zu-lder** armer zijn dan
you-people believe PART not that they-people poorer are than
wu-lder.
we-people
'You won't believe that they are poorer than us.'

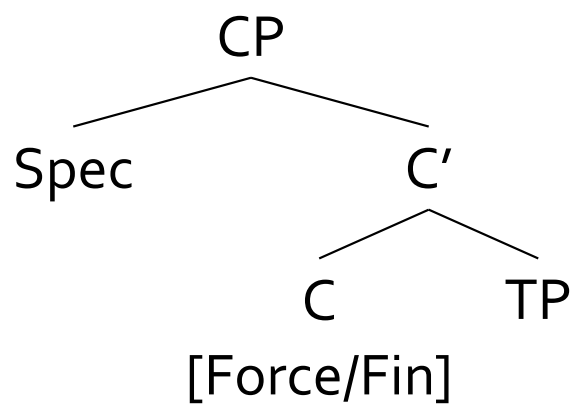
Case study #3: split Force/Fin

the Split Force/Fin-Parameter
the CP-domain {does/does not} have a split Force/Fin.

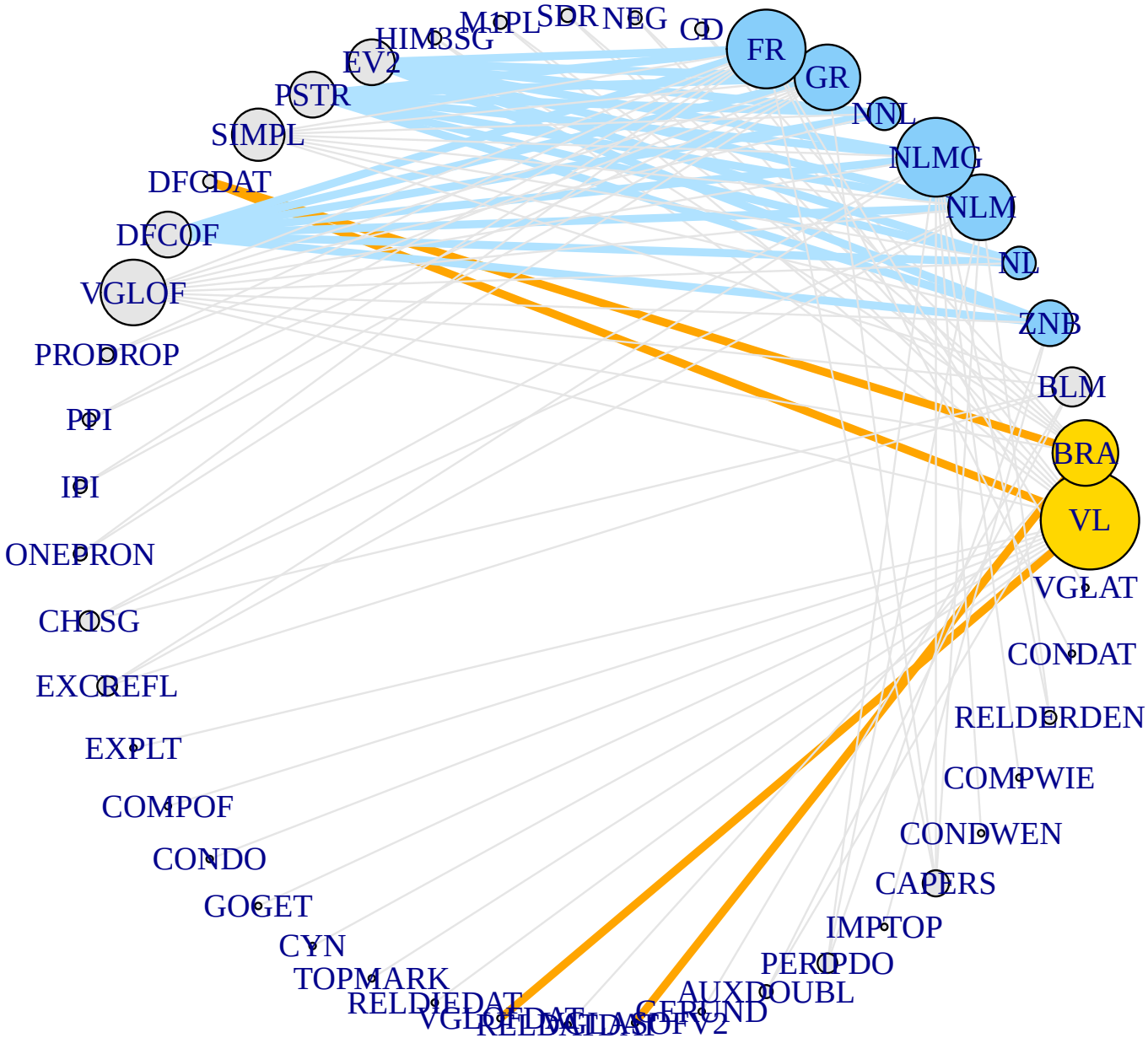
+ **Split Force/Fin-parameter**



– **Split Force/Fin-parameter**



Case study #3: split Force/Fin



Case study #3: split Force/Fin

- The following phenomena are characteristic of the South:

doubly filled COMP with *dat* 'that'

(30) Zeg ma nie **wien da**-se zie hadde wiln roepn.
tell but not who that-they$_{clitic}$ they$_{strong}$ had want call
'Don't tell me who they planned to call.'

***of* 'if' + *dat* 'that' in an 'as if'-clause**

(31) T is juist lijk **of dat**-er etwien in dn hof stoat.
it is just like if that-there someone in the garden stands
'It looks as if there is someone in the garden.'

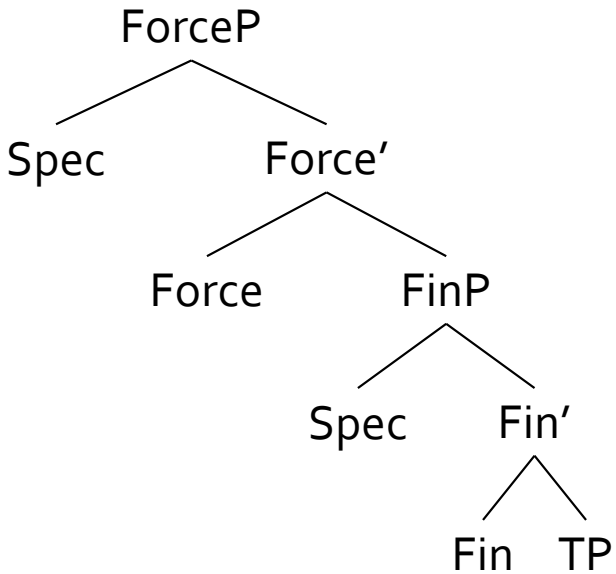
***of* 'if' + embedded V2 in an 'as if'-clause**

(32) T is precies **of** d'r **staat** d'r enen in den hof.
it is exactly if there stands there someone in the garden
'It looks as if there is someone in the garden.'

Case study #3: split Force/Fin

SOUTH: Split Force/Fin

(33)



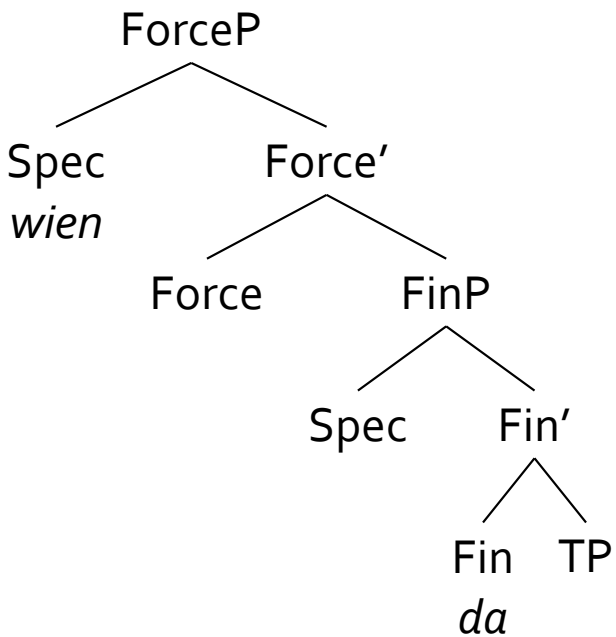
Assumptions about the left periphery:

1. FinP has to contain overt material (every sentence has to be marked as finite).
2. *wh*-phrases are merged in specForceP
3. Generalized Doubly Filled Comp Filter (GDFCF): A feature cannot be spelled out twice

Case study #3: split Force/Fin

(34) … **wien da**-se zie hadde wiln roepn.
who that-they$_{clitic}$ they$_{strong}$ had want call
'… who they planned to call.'

(35)

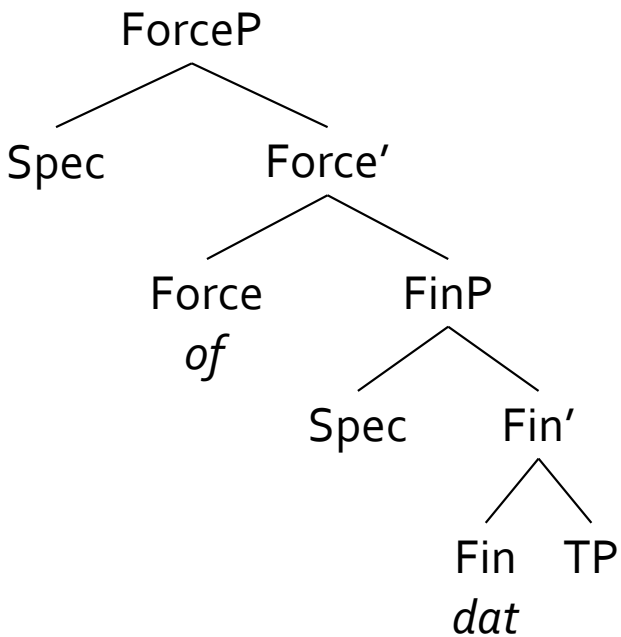


- feature specification of *dat*: +Fin
- *dat* has to be spelled out to realize FinP

Case study #3: split Force/Fin

(36) T is juist lijk **of dat**-er etwien in dn hof stoat.
it is just like if that-there someone in the garden stands
'It looks as if there is someone in the garden.'

(37)

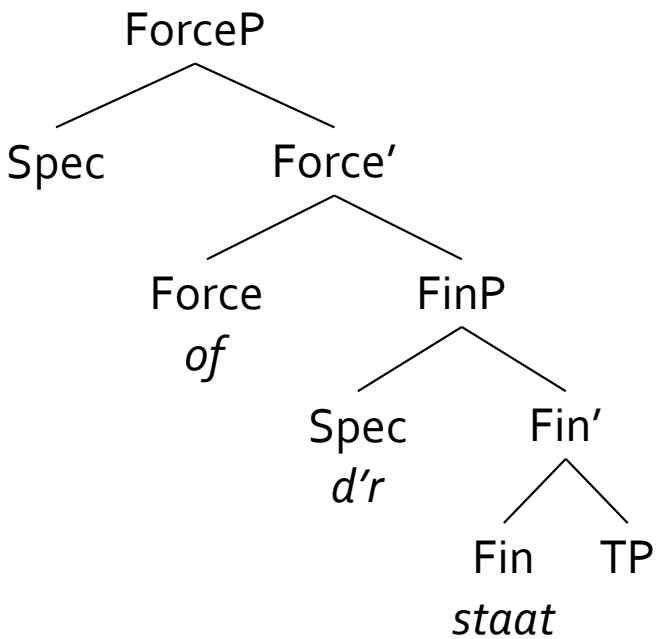


- feature specification of the complementizers: *dat*:+Fin, *of*:+Force
- *dat* has to be spelled out to realize FinP.

Case study #3: split Force/Fin

(38) T is precies **of** d'r **staat** d'r enen in den hof.
it is exactly if there stands there someone in the garden
'It looks as if there is someone in the garden.'

(39)



- feature specification of *of*: +Force
- the verb realizes FinP

Case study #3: split Force/Fin

Predictions:

1. Doubly filled COMP should be obligatory in embedded *wh*-clauses in the South $\rightarrow$ confirmed

Case study #3: split Force/Fin

Predictions:

1. Doubly filled COMP should be obligatory in embedded *wh*-clauses in the South $\rightarrow$ confirmed

(40) … **wien** *(**da**) se zie hadde wiln roepn.
who that $\text{they}_{\text{clitic}}$ $\text{they}_{\text{strong}}$ had want call
'… who they planned to call.'

Case study #3: split Force/Fin

Predictions:

1. Doubly filled COMP should be obligatory in embedded *wh*-clauses in the South $\rightarrow$ confirmed

   (40) … **wien** *(**da**) se zie hadde wiln roepn.
   who that $\text{they}_{\text{clitic}}$ $\text{they}_{\text{strong}}$ had want call
   '… who they planned to call.'

2. No doubly filled COMP with *of* 'if' in the South $\rightarrow$ confirmed

Case study #3: split Force/Fin

Predictions:

1. Doubly filled COMP should be obligatory in embedded *wh*-clauses in the South → confirmed

(40) … **wien** *(**da**) se zie hadde wiln roepn.
who that $they_{clitic}$ $they_{strong}$ had want call
'… who they planned to call.'

2. No doubly filled COMP with *of* 'if' in the South → confirmed

(41) *… **wien of** se zie hadde wiln roepn.
who if $they_{clitic}$ $they_{strong}$ had want call
'… who they planned to call.'

Case study #3: split Force/Fin

- The following phenomena are characteristic of the North:

doubly filled COMP with *of* 'if'

(42) Vertel mie eens **wel of** ze had kenn roepn.
Tell me PART who if she had can call
'Tell me who she could have been calling.'

embedded V2 with complementizer drop

(43) Ik geloof **deze jongens vindt** ze allemaal wel aardig.
I believe these guys finds she all PART nice
'I believe that she likes all of these guys.'

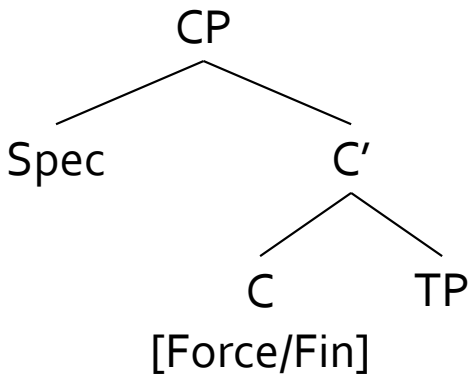
preposition stranding

(44) **Die rare jongen** ben ik **mee** naar de markt west.
that strange boy am I with to the market been
'With that strange boy I went to the market.'

Case study #3: split Force/Fin

NORTH: NO Split Force/Fin

(45)



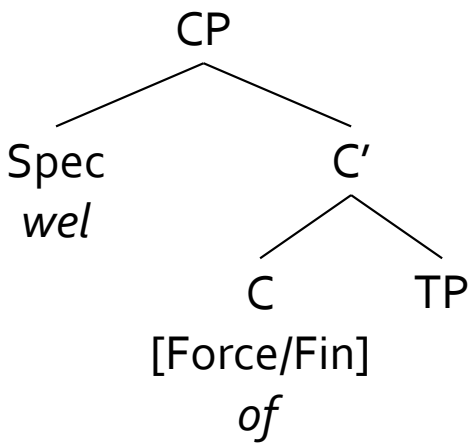
Assumptions about the left periphery:

1. FinP has to contain overt material (every sentence has to be marked as finite).
2. *wh*-phrases are merged in specForceP
3. Generalized Doubly Filled Comp Filter (GDFCF): A feature cannot be spelled out twice

Case study #3: split Force/Fin

(46) Vertel mie eens **wel** **of** ze had kenn roepn.
Tell me PART who if she had can call
'Tell me who she could have been calling.'

(47)

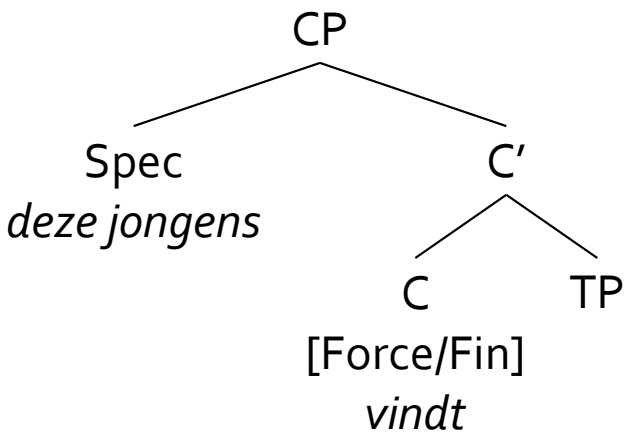


- feature specification of the complementizers: *dat*:[+Force,+Fin], *of*:[+Force,+Fin]
- doubly filled COMP: *dat* has the wrong value for Force, *of* is allowed if it spells out Fin (GDFCF).

Case study #3: split Force/Fin

(48) Ik geloof **deze jongens vindt** ze allemaal wel aardig.
I believe these guys finds she all PART nice
'I believe that she likes all of these guys.' (EV2)

(49)



- the finite verb realizes FinP

Case study #3: split Force/Fin

Predictions:

1. No doubly filled COMP with *dat* 'that' in the North $\rightarrow$ confirmed

Case study #3: split Force/Fin

Predictions:

1. No doubly filled COMP with *dat* 'that' in the North $\rightarrow$ confirmed

(50) *Vertel mie eens **wel dat** ze had kenn roepn.
Tell me PART who that she had can call
'Tell me who she could have been calling.'

Case study #3: split Force/Fin

Predictions:

1. No doubly filled COMP with *dat* 'that' in the North → confirmed

   (50) *Vertel mie eens **wel dat** ze had kenn roepn.
   Tell me PART who that she had can call
   'Tell me who she could have been calling.'

2. Doubly filled COMP should be optional in the North → confirmed

Case study #3: split Force/Fin

Predictions:

1. No doubly filled COMP with *dat* 'that' in the North → confirmed

   (50) *Vertel mie eens **wel dat** ze had kenn roepn.
   Tell me PART who that she had can call
   'Tell me who she could have been calling.'

2. Doubly filled COMP should be optional in the North → confirmed

   (51) Vertel mie eens **wel** (**of**) ze had kenn roepn.
   Tell me PART who if she had can call
   'Tell me who she could have been calling.'

Case study #3: split Force/Fin

the Split Force/Fin-Parameter
the CP-domain {does/does not} have a split Force/Fin.

- **SOUTH**: the CP-domain DOES have a split Force/Fin
- **NORTH**: the CP-domain DOES NOT have a split Force/Fin

Case study #3: split Force/Fin

the Split Force/Fin-Parameter
the CP-domain {does/does not} have a split Force/Fin.

- **SOUTH**: the CP-domain DOES have a split Force/Fin
- **NORTH**: the CP-domain DOES NOT have a split Force/Fin

	WH-DAT	WH-OF	WH-EMPTY	EV2	VGLOFV2	VGLOFDAT
SOUTH (FL)	+	-	-	-	-	+
SOUTH (BRA)	+	-	-	-	+	-
NORTH	-	+	+	+	-	-

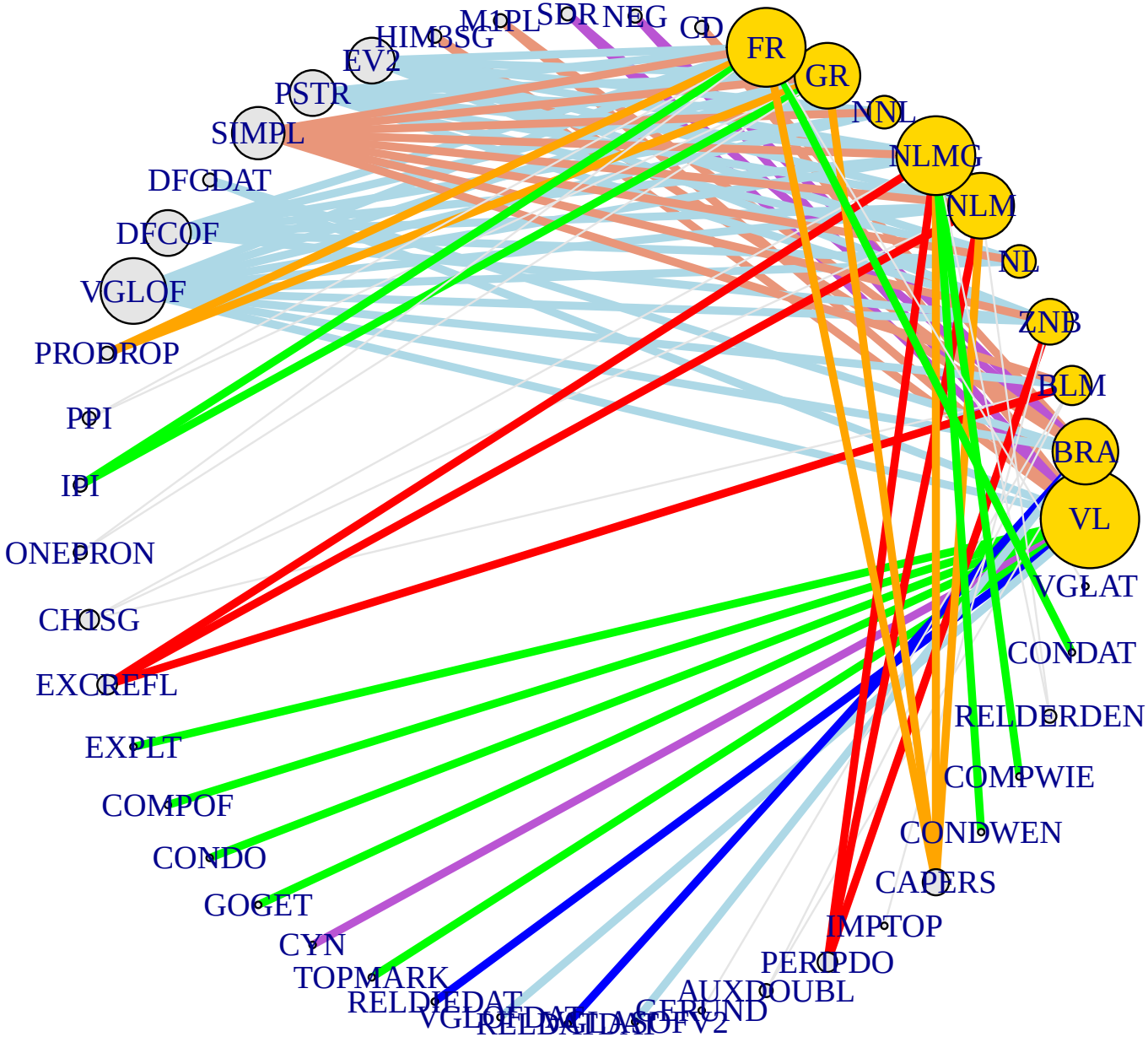
Case study #3: split Force/Fin

- Can this analysis also give us a handle on the variation concerning P-stranding?

(52) **Die rare jongen** ben ik **mee** naar de markt west.
that strange boy am I with to the market been
'With that strange boy I went to the market.' NORTH

(53) *__Die rare jongen__ ben ik __mee__ naar de markt west.
that strange boy am I with to the market been
'With that strange boy I went to the market.' SOUTH

Combining the case studies: 7 parameters



Combining the case studies: 7 parameters

- We can bring back these 37 linguistic phenomena to 7 parameters:

	VL	BRA	BLM	ZNB	NL	NLM	NLMG	NNL	GR	FR
SPLIT C-POL	+	+	-	-	-	-	-	-	-	-
SPLIT D	+	+	-	-	-	-	-	-	-	-
SPLIT Force/FIN	+	+	-	-	-	-	-	-	-	-
SPLIT TP	-	-	+	+	-	+	+	-	-	-
SPLIT C_3	+	-	-	-	-	-	+	-	-	+
AGR C-num	+	-	-	-	-	-	-	-	-	-
AGR C-pers	-	-	-	-	-	+	+	-	+	+

Combining the case studies: 7 parameters

- We can bring back these 37 linguistic phenomena to 7 parameters:

	VL	BRA	BLM	ZNB	NL	NLM	NLMG	NNL	GR	FR
SPLIT C-POL	+	+	-	-	-	-	-	-	-	-
SPLIT D	+	+	-	-	-	-	-	-	-	-
SPLIT Force/FIN	+	+	-	-	-	-	-	-	-	-
SPLIT TP	-	-	+	+	-	+	+	-	-	-
SPLIT C_3	+	-	-	-	-	-	+	-	-	+
AGR C-num	+	-	-	-	-	-	-	-	-	-
AGR C-pers	-	-	-	-	-	+	+	-	+	+

- **Split TP-parameter:** The TP-domain {is/is not} split.

Combining the case studies: 7 parameters

- We can bring back these 37 linguistic phenomena to 7 parameters:

	VL	BRA	BLM	ZNB	NL	NLM	NLMG	NNL	GR	FR
SPLIT C-POL	+	+	-	-	-	-	-	-	-	-
SPLIT D	+	+	-	-	-	-	-	-	-	-
SPLIT Force/FIN	+	+	-	-	-	-	-	-	-	-
SPLIT TP	-	-	+	+	-	+	+	-	-	-
SPLIT C_3	+	-	-	-	-	-	+	-	-	+
AGR C-num	+	-	-	-	-	-	-	-	-	-
AGR C-pers	-	-	-	-	-	+	+	-	+	+

- **Split TP-parameter:** The TP-domain {is/is not} split.
- **Split C3-parameter:** The CP-domain {does/does not} have separate projections for comparatives and conditionals.

Combining the case studies: 7 parameters

- We can bring back these 37 linguistic phenomena to 7 parameters:

	VL	BRA	BLM	ZNB	NL	NLM	NLMG	NNL	GR	FR
SPLIT C-POL	+	+	-	-	-	-	-	-	-	-
SPLIT D	+	+	-	-	-	-	-	-	-	-
SPLIT Force/FIN	+	+	-	-	-	-	-	-	-	-
SPLIT TP	-	-	+	+	-	+	+	-	-	-
SPLIT C_3	+	-	-	-	-	-	+	-	-	+
AGR C-num	+	-	-	-	-	-	-	-	-	-
AGR C-pers	-	-	-	-	-	+	+	-	+	+

- **Split TP-parameter:** The TP-domain {is/is not} split.
- **Split C3-parameter:** The CP-domain {does/does not} have separate projections for comparatives and conditionals.
- **AGR C-num-parameter:** C {does/does not} bear an unvalued number feature.

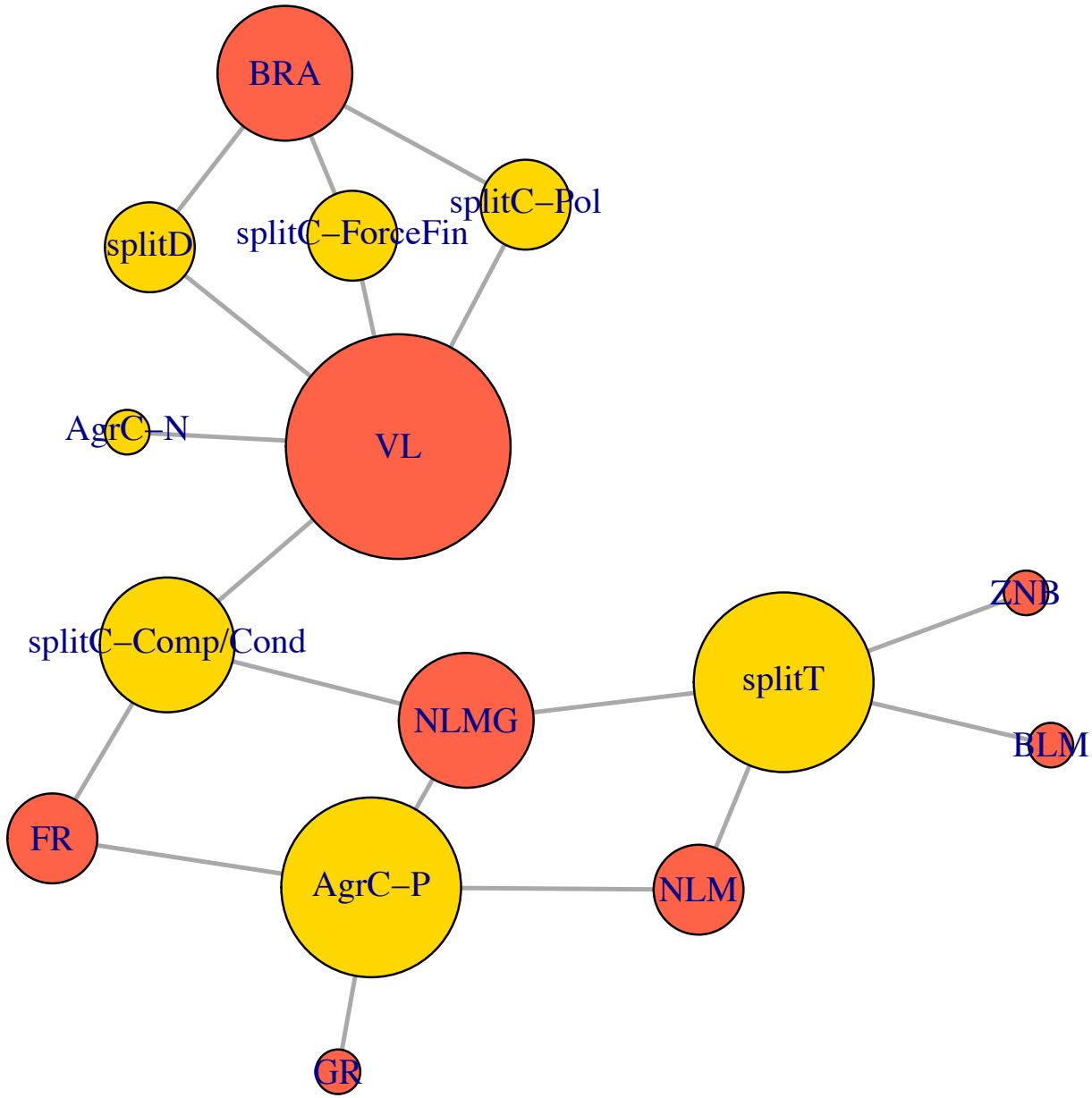
Combining the case studies: 7 parameters

- We can bring back these 37 linguistic phenomena to 7 parameters:

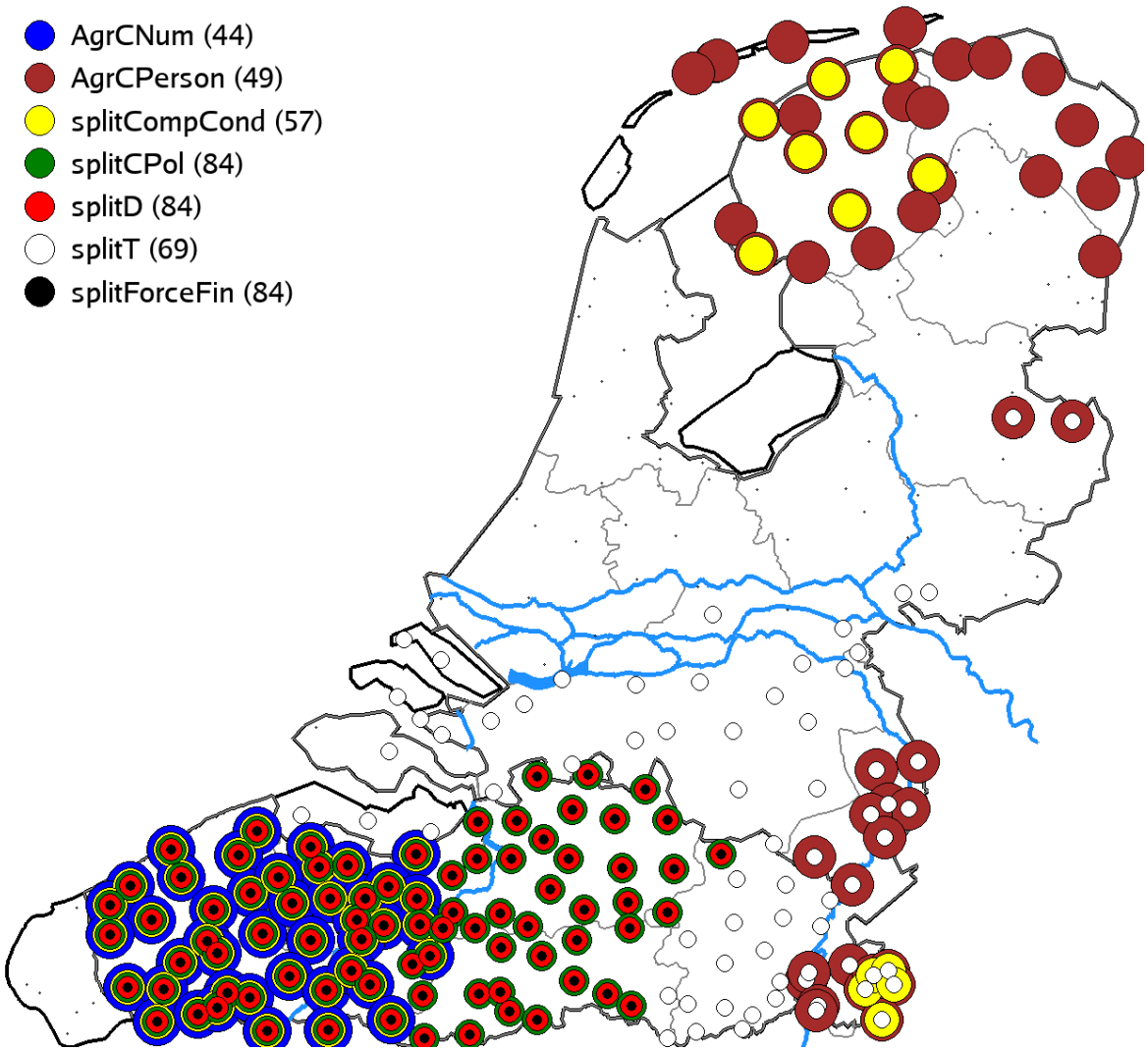
	VL	BRA	BLM	ZNB	NL	NLM	NLMG	NNL	GR	FR
SPLIT C-POL	+	+	-	-	-	-	-	-	-	-
SPLIT D	+	+	-	-	-	-	-	-	-	-
SPLIT Force/FIN	+	+	-	-	-	-	-	-	-	-
SPLIT TP	-	-	+	+	-	+	+	-	-	-
SPLIT C_3	+	-	-	-	-	-	+	-	-	+
AGR C-num	+	-	-	-	-	-	-	-	-	-
AGR C-pers	-	-	-	-	-	+	+	-	+	+

- **Split TP-parameter:** The TP-domain {is/is not} split.
- **Split C3-parameter:** The CP-domain {does/does not} have separate projections for comparatives and conditionals.
- **AGR C-num-parameter:** C {does/does not} bear an unvalued number feature.
- **AGR C-pers-parameter:** C {does/does not} bear an unvalued person feature.

Combining the case studies: 7 parameters



Combining the case studies: 7 parameters



Outline

Main goals for today

Introduction: Kayne's dream

Quantitative analysis

- Correspondence Analysis
- Cluster Analysis
- Cluster Description
- Conclusion

Qualitative analysis

- Case study #1: PolP
- Case study #2: split DP
- Case study #3: split Force/Fin
- Combining the case studies: 7 parameters

The bigger picture: determinants of variation

The bigger picture: determinants of variation

- our ten dialect groups differ:

The bigger picture: determinants of variation

- our ten dialect groups differ:
  1. in whether or not a morphosyntactic feature heads its own projection (SPLIT)

The bigger picture: determinants of variation

- our ten dialect groups differ:
  1. in whether or not a morphosyntactic feature heads its own projection (SPLIT)
  2. in the extent to which this happens

The bigger picture: determinants of variation

- our ten dialect groups differ:
  1. in whether or not a morphosyntactic feature heads its own projection (SPLIT)
  2. in the extent to which this happens
  3. in whether or not a morphosyntactic feature triggers Agree (AGR)

The bigger picture: determinants of variation

- our ten dialect groups differ:
  1. in whether or not a morphosyntactic feature heads its own projection (SPLIT)
  2. in the extent to which this happens
  3. in whether or not a morphosyntactic feature triggers Agree (AGR)

→ reminiscent of Longobardi (2005)'s Principles & Schemata:

(54) **Parameter Schema:**

a. Is F, F a functional feature, grammaticalized?
b. Is F, F a grammaticalized feature, checked by X, X a lexical category?
c. Is F, F a grammaticalized feature, spread on Y, Y a lexical category?
d. Is F, F a grammaticalized feature checked by X, strong (i.e. overtly attracts X)?

The bigger picture: determinants of variation

- our ten dialect groups differ:
  1. in whether or not a morphosyntactic feature heads its own projection (SPLIT)
  2. in the extent to which this happens
  3. in whether or not a morphosyntactic feature triggers Agree (AGR)

→ and of Biberauer and Roberts (2013)'s parameter hierarchies:

Parameter Hierarchy
For a given value v_i of a parametrically variant feature F:

- **Macroparameters:** all heads of the relevant type share v_i
- **Mesoparameters:** all heads of a given naturally definable class, a subset of the full class of heads of the relevant type, e.g. [+V], share v_i
- **Microparameters:** a small subclass of functional heads (e.g. modal auxiliaries, pronouns) shows v_i
- **Nanoparameters:** one or more individual lexical items is/are specified for v_i

The bigger picture: determinants of variation

Are A′-features grammaticalized?

NO
Consistently poor left periphery

Yes
Are **ALL** A′-features grammaticalized?

Yes
Consistently rich left periphery

No
Are **SOME** A′-features grammaticalized?

C *v* D
Mixed effects of left-peripheral richness

To sum up

To sum up

1. We have developed a parametric analysis for a large data set of morphosyntactic variation in Dutch dialects and have reduced the core tendencies in that variation to seven grammatical parameters.

To sum up

1. We have developed a parametric analysis for a large data set of morphosyntactic variation in Dutch dialects and have reduced the core tendencies in that variation to seven grammatical parameters.
2. In identifying those core tendencies we have crucially relied on quantitative-statistical means, but in identifying the grammatical parameters we started from formal-theoretical analyses.

To sum up

1. We have developed a parametric analysis for a large data set of morphosyntactic variation in Dutch dialects and have reduced the core tendencies in that variation to seven grammatical parameters.
2. In identifying those core tendencies we have crucially relied on quantitative-statistical means, but in identifying the grammatical parameters we started from formal-theoretical analyses.
3. At a more general level, these dialects seem to differ from one another in the choice of the morphosyntactic features that are grammaticalized and the degree to which they are.

References I

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral Dissertation, University of Connecticut at Storrs.

Barbiers, Sjef, Marjo van Koppen, Hans Bennis, and Norbert Corver. 2016. Microcomparative MOrphosyntactic REsearch (MIMORE): Mapping partial grammars of Flemish, Brabantish and Dutch. *Lingua* 178:5–31.

Bayer, Josef. 1984. COMP in Bavarian syntax. *The Linguistic Review* 3:209–274.

Benincà, Paola, and Cecilia Poletto. 2004. Topic, focus, and V2: Defining the CP sublayers. In *The structure of CP and IP*, ed. Luigi Rizzi, 52–75. Oxford: Oxford University Press.

Biberauer, Theresa, and Ian Roberts. 2013. *Challenges to linearization*. Berlin: Mouton de Gruyter.

van Craenenbroeck, Jeroen. 2010. *The syntax of ellipsis. Evidence from Dutch dialects*. New York: OUP.

van Craenenbroeck, Jeroen, and Marjo van Koppen. 2008. Pronominal doubling in Dutch dialects: big DPs and coordinations. In *Microvariation in syntactic doubling.*, ed. Sjef Barbiers, Olaf Koeneman, Marika Lekakou, and Margreet van der Ham, volume 36 of *Syntax and Semantics*, 207–249. Bingley: Emerald.

Déchaine, Rose-Marie, and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33:409–442.

Haegeman, Liliane. 1992. *Theory and description in generative syntax*. Cambridge: Cambridge University Press.

Heeringa, Wilbert. 2004. Measuring dialect pronunciation differences using Levenshtein distance. Doctoral Dissertation, University of Groningen.

Heeringa, Wilbert, and John Nerbonne. 2013. Dialectometry. In *Language and Space. An International Handbook of Linguistic Variation. Volume 3: Dutch*, ed. Frans Hinskens and Johan Taeldeman, volume 30 of *Handbooks of Linguistics and Communication Science*, 624–645. Berlin/Boston: De Gruyter.

Hoekstra, Eric. 1993. Dialectal variation inside CP as parametric variation. In *Dialektsyntax*, ed. Werner Abraham and Josef Bayer, volume 5 of *Linguistische Berichte/Sonderheft*, 161–179. Opladen: Westdeutscher Verlag.

Kayne, Richard. 1996. Microparametric syntax: some introductory remarks. In *Microparametric syntax and dialect variation*, ed. J.R. Black and Virginia Motapanyane, ix–xviii. Amsterdam: John Benjamins.

References II

Longobardi, Giuseppe. 2005. A minimalist program for parametric linguistics? In *Organizing grammar. studies in honor of Henk van Riemsdijk*, ed. Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster, 407–414. Berlin: Mouton de Gruyter.

Penner, Zvi. 1994. Asking questions without CPs? On the acquisition of root wh-questions in Bernese Swiss German and Standard German. In *Language acquisition studies in generative grammar*, ed. Teun Hoekstra and Bonnie D. Schwartz, 177–214. Amsterdam: John Benjamins Publishing Company.

Poletto, Cecilia. 2000. *The higher functional field: Evidence from Northern Italian dialects*. Oxford University Press.

Spruit, Marco René. 2008. Quantitative perspectives on syntactic variation in Dutch dialects. Doctoral Dissertation, Universiteit van Amsterdam.

Wieling, Martijn, and John Nerbonne. 2015. Advances in dialectometry. *Annual Review of Linguistics* 1:243–264.