Reverse dialectometry

Verb cluster variation in Dutch

Jeroen van Craenenbroeck

KU Leuven/CRISSP

Methods in Dialectology XV
August 11–15 4, 2014, Groningen, the Netherlands
This talk in one slide

- main topic: interaction between formal-theoretical and quantitative-statistical linguistics
- starting point: the massive amount of variation attested in Dutch verb clusters necessitates a collaboration between formal and quantitative approaches
- traditional dialectometry measures (dis)similarities between dialect locations based on their linguistic profile
- reverse dialectometry measures (dis)similarities between linguistic constructions based on their geographical spread and maps these results against formal-theoretical microparameters
- result: a method that can detect and identify grammatical (micro)parameters in a large and highly varied data set
This talk in one slide

- **main topic:** interaction between formal-theoretical and quantitative-statistical linguistics
This talk in one slide

- **main topic:** interaction between formal-theoretical and quantitative-statistical linguistics
- **starting point:** the massive amount of variation attested in Dutch verb clusters necessitates a collaboration between formal and quantitative approaches
This talk in one slide

- **main topic:** interaction between formal-theoretical and quantitative-statistical linguistics

- **starting point:** the massive amount of variation attested in Dutch verb clusters necessitates a collaboration between formal and quantitative approaches

- **traditional dialectometry** measures (dis)similarities between dialect locations based on their linguistic profile
This talk in one slide

• **main topic:** interaction between formal-theoretical and quantitative-statistical linguistics

• **starting point:** the massive amount of variation attested in Dutch verb clusters necessitates a collaboration between formal and quantitative approaches

• **traditional dialectometry** measures (dis)similarities between dialect locations based on their linguistic profile

• **reverse dialectometry** measures (dis)similarities between *linguistic constructions* based on their geographical spread and maps these results against formal-theoretical microparameters
This talk in one slide

- **main topic**: interaction between formal-theoretical and quantitative-statistical linguistics
- **starting point**: the massive amount of variation attested in Dutch verb clusters necessitates a collaboration between formal and quantitative approaches
- **traditional dialectometry** measures (dis)similarities between dialect locations based on their linguistic profile
- **reverse dialectometry** measures (dis)similarities between *linguistic constructions* based on their geographical spread and maps these results against formal-theoretical microparameters
- **result**: a method that can detect and identify grammatical (micro)parameters in a large and highly varied data set
Introduction: verb clusters & parameters

- highly simplified view of parameters & verb clusters: a difference in cluster order corresponds to a different parameter setting
Introduction: verb clusters & parameters

- highly simplified view of parameters & verb clusters: a difference in cluster order corresponds to a different parameter setting

**dialect A**

(1) dat hij gelachen heeft
that he laughed has
‘that he has laughed.’

**dialect B**

(2) dat hij heeft gelachen
that he has laughed
‘that he has laughed.’
Introduction: verb clusters & parameters

- highly simplified view of parameters & verb clusters: a difference in cluster order corresponds to a different parameter setting

**dialect A**

(1) dat hij gelachen heeft
that he laughed has
‘that he has laughed.’

**dialect B**

(2) dat hij heeft gelachen
that he has laughed
‘that he has laughed.’

- possible parameter setting:
  - dialect A: [+MoveParticipleAcrossAux]
  - dialect B: [−MoveParticipleAcrossAux]
Introduction: verb clusters & parameters

- highly simplified view of parameters & verb clusters: a difference in cluster order corresponds to a different parameter setting

**dialect A**

(1) dat hij *gelachen heeft*  
that he laughed  
‘that he has laughed.’

**dialect B**

(2) dat hij *heeft gelachen*  
that he has  
‘that he has laughed.’

- possible parameter setting:
  - dialect A: [+MoveParticipleAcrossAux]
  - dialect B: [−MoveParticipleAcrossAux]

- however, when faced with actual verb cluster variation data, things get *much* more complicated very quickly
e.g. the SAND-project:
• e.g. the SAND-project:
• e.g. the SAND-project:
  • dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands
e.g. the SAND-project:

- dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands

the SAND-questionnaire contained eight questions on word order in verb clusters:
- e.g. the SAND-project:
  - dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands
- the SAND-questionnaire contained eight questions on word order in verb clusters:
  - three two-verb clusters of the form AUXILIARY-PARTICIPLE
• e.g. the SAND-project:
  • dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands

• the SAND-questionnaire contained eight questions on word order in verb clusters:
  • three two-verb clusters of the form AUXILIARY-PARTICIPLE
  • one two-verb cluster of the form MODAL-INFinitIVE
• e.g. the SAND-project:
  • dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands
• the SAND-questionnaire contained eight questions on word order in verb clusters:
  • three two-verb clusters of the form AUXILIARY-PARTICIPLE
  • one two-verb cluster of the form MODAL-INFINITIVE
  • four three-verb clusters:
• e.g. the SAND-project:
  • dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands

• the SAND-questionnaire contained eight questions on word order in verb clusters:
  • three two-verb clusters of the form AUXILIARY-PARTICIPLE
  • one two-verb cluster of the form MODAL-INFINITIVE
  • four three-verb clusters:
    • MODAL-MODAL-INFINITIVE
• e.g. the SAND-project:
  • dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands
• the SAND-questionnaire contained eight questions on word order in verb clusters:
  • three two-verb clusters of the form AUXILIARY-PARTICLE
  • one two-verb cluster of the form MODAL-INFINITIVE
  • four three-verb clusters:
    • MODAL-MODAL-INFINITIVE
    • MODAL-AUXILIARY-PARTICLE
• e.g. the SAND-project:
  • dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands

• the SAND-questionnaire contained eight questions on word order in verb clusters:
  • three two-verb clusters of the form AUXILIARY-PARTICIPLE
  • one two-verb cluster of the form MODAL-INFINITIVE
  • four three-verb clusters:
    • MODAL-MODAL-INFINITIVE
    • MODAL-AUXILIARY-PARTICIPLE
    • AUXILIARY-AUXILIARY-INFINITIVE
• e.g. the SAND-project:
  • dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands
• the SAND-questionnaire contained eight questions on word order in verb clusters:
  • three two-verb clusters of the form AUXILIARY-PARTICIPLE
  • one two-verb cluster of the form MODAL-INFINITIVE
  • four three-verb clusters:
    • MODAL-MODAL-INFINITIVE
    • MODAL-AUXILIARY-PARTICIPLE
    • AUXILIARY-AUXILIARY-INFINITIVE
    • AUXILIARY-MODAL-INFINITIVE
• e.g. the SAND-project:
  • dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands

• the SAND-questionnaire contained eight questions on word order in verb clusters:
  • three two-verb clusters of the form AUXILIARY-PARTICIPLE
  • one two-verb cluster of the form MODAL-INFINITIVE
  • four three-verb clusters:
    • MODAL-MODAL-INFINITIVE
    • MODAL-AUXILIARY-PARTICIPLE
    • AUXILIARY-AUXILIARY-INFINITIVE
    • AUXILIARY-MODAL-INFINITIVE

• for a total of 31 cluster orders
e.g. the SAND-project:
- dialect interviews in 267 dialect locations in Belgium, France, and the Netherlands

the SAND-questionnaire contained eight questions on word order in verb clusters:
- three two-verb clusters of the form AUXILIARY-PARTICIPLE
- one two-verb cluster of the form MODAL-INFINITIVE
- four three-verb clusters:
  - MODAL-MODAL-INFINITIVE
  - MODAL-AUXILIARY-PARTICIPLE
  - AUXILIARY-AUXILIARY-INFINITIVE
  - AUXILIARY-MODAL-INFINITIVE

for a total of 31 cluster orders

if we map, for each of the 267 SAND-dialects, which dialect has which combination of cluster orders, we find 137 different combinations of verb cluster orders
• this state of affairs raises fundamental questions for parameter theory:
• this state of affairs raises fundamental questions for parameter theory:
  • are there really (grammatical) microparameters distinguishing between all of these 137 dialect types?
• this state of affairs raises fundamental questions for parameter theory:
  • are there really (grammatical) microparameters distinguishing between all of these 137 dialect types?
  • if there are, what are they and how can we detect them?
• this state of affairs raises fundamental questions for parameter theory:
  • are there really (grammatical) microparameters distinguishing between all of these 137 dialect types?
  • if there are, what are they and how can we detect them?
  • more generally, how can we distinguish between the signal and the noise in such large and highly variable datasets?
• this state of affairs raises fundamental questions for parameter theory:
  • are there really (grammatical) microparameters distinguishing between all of these 137 dialect types?
  • if there are, what are they and how can we detect them?
  • more generally, how can we distinguish between the signal and the noise in such large and highly variable datasets?
• **in this talk** I use statistical methods to detect and identify grammatical microparameters regulating (part of) the variation found in Dutch verb clusters
A dialectometric analysis

- **dialectometry** is a subdiscipline of linguistics that uses computational and quantitative techniques in dialectology (Nerbonne and Kretzschmar Jr., 2013)
A dialectometric analysis

- **dialectometry** is a subdiscipline of linguistics that uses computational and quantitative techniques in dialectology (Nerbonne and Kretzschmar Jr., 2013)
- in a typical dialectometric analysis locations are used as individuals and linguistic phenomena as variables → we’re measuring similarities and differences between dialect locations based on their linguistic profile
A dialectometric analysis

- **dialectometry** is a subdiscipline of linguistics that uses computational and quantitative techniques in dialectology (Nerbonne and Kretzschmar Jr., 2013)
- in a typical dialectometric analysis locations are used as individuals and linguistic phenomena as variables → we’re measuring similarities and differences between dialect locations based on their linguistic profile
- often used method: Multidimensional Scaling (MDS)
A dialectometric analysis

• **dialectometry** is a subdiscipline of linguistics that uses computational and quantitative techniques in dialectology (Nerbonne and Kretzschmar Jr., 2013)

• in a typical dialectometric analysis locations are used as individuals and linguistic phenomena as variables → we’re measuring similarities and differences between dialect locations based on their linguistic profile

• often used method: Multidimensional Scaling (MDS)

• starting point: data table with dialects in rows and cluster orders in columns
<table>
<thead>
<tr>
<th>Location</th>
<th>AUX1(be.sg)-PART2</th>
<th>PART2-AUX1(be.sg)</th>
<th>AUX1(have.sg)-PART2</th>
<th>PART2-AUX1(have.sg)</th>
<th>AUX1(have.pl)-PART2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midsland / Midslân</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Lies</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>West-Terschelling</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Oosterend</td>
<td>NA</td>
<td>NA</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Hollum</td>
<td>no</td>
<td>yes</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Schiermonnikoog</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Ferwerd / Ferwert</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Anjum / Eanjum</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Kollum</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Visvliet</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Oosterbierum / Ee</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Beetgum / Bitgum</td>
<td>no</td>
<td>yes</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Bergum / Burgum</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Jorwerd / Jorwert</td>
<td>no</td>
<td>yes</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Bakkeveen / Bakke</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Waskemeer / De V</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Kloosterburen</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Warffum</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Leermens</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Groningen</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Nieuw-Scheemda</td>
<td>NA</td>
<td>NA</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Langelo</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
</tbody>
</table>
- step 1: convert the data table into a $267 \times 267$ (symmetric) distance matrix, whereby for each pair of locations a distance between them is calculated based on the linguistic features they share
### A dialectometric analysis

<table>
<thead>
<tr>
<th></th>
<th>Midsland</th>
<th>Lies</th>
<th>West-Tersch</th>
<th>Oosteren</th>
<th>Holumm</th>
<th>Schiermond</th>
<th>Ferwerd</th>
<th>Anjum / Kollum</th>
<th>Visvliet</th>
<th>Oosterblok</th>
<th>Beetgum</th>
<th>Bergum / Jorwerd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midsland / Mids</td>
<td>0,000</td>
<td>0,500</td>
<td>0,333</td>
<td>0,706</td>
<td>0,250</td>
<td>0,647</td>
<td>0,357</td>
<td>0,250</td>
<td>0,611</td>
<td>0,650</td>
<td>0,533</td>
<td>0,545</td>
</tr>
<tr>
<td>Lies</td>
<td>0,500</td>
<td>0,000</td>
<td>0,444</td>
<td>0,750</td>
<td>0,588</td>
<td>0,375</td>
<td>0,471</td>
<td>0,563</td>
<td>0,444</td>
<td>0,444</td>
<td>0,632</td>
<td>0,714</td>
</tr>
<tr>
<td>West-Terschell</td>
<td>0,333</td>
<td>0,444</td>
<td>0,000</td>
<td>0,789</td>
<td>0,429</td>
<td>0,667</td>
<td>0,286</td>
<td>0,429</td>
<td>0,632</td>
<td>0,600</td>
<td>0,500</td>
<td>0,429</td>
</tr>
<tr>
<td>Oosterend</td>
<td>0,706</td>
<td>0,750</td>
<td>0,789</td>
<td>0,000</td>
<td>0,706</td>
<td>0,765</td>
<td>0,737</td>
<td>0,538</td>
<td>0,563</td>
<td>0,600</td>
<td>0,600</td>
<td>0,727</td>
</tr>
<tr>
<td>Holumm</td>
<td>0,250</td>
<td>0,588</td>
<td>0,429</td>
<td>0,706</td>
<td>0,000</td>
<td>0,667</td>
<td>0,167</td>
<td>0,000</td>
<td>0,625</td>
<td>0,400</td>
<td>0,714</td>
<td>0,462</td>
</tr>
<tr>
<td>Schiermonniken</td>
<td>0,647</td>
<td>0,375</td>
<td>0,667</td>
<td>0,765</td>
<td>0,667</td>
<td>0,000</td>
<td>0,625</td>
<td>0,353</td>
<td>0,682</td>
<td>0,308</td>
<td>0,706</td>
<td>0,571</td>
</tr>
<tr>
<td>Ferwerd</td>
<td>0,357</td>
<td>0,471</td>
<td>0,286</td>
<td>0,737</td>
<td>0,167</td>
<td>0,625</td>
<td>0,000</td>
<td>0,182</td>
<td>0,588</td>
<td>0,308</td>
<td>0,000</td>
<td>0,333</td>
</tr>
<tr>
<td>Anjum / Eeanjum</td>
<td>0,250</td>
<td>0,563</td>
<td>0,429</td>
<td>0,538</td>
<td>0,000</td>
<td>0,667</td>
<td>0,182</td>
<td>0,000</td>
<td>0,571</td>
<td>0,625</td>
<td>0,417</td>
<td>0,556</td>
</tr>
<tr>
<td>Kollum</td>
<td>0,611</td>
<td>0,444</td>
<td>0,632</td>
<td>0,563</td>
<td>0,625</td>
<td>0,400</td>
<td>0,588</td>
<td>0,571</td>
<td>0,353</td>
<td>0,625</td>
<td>0,643</td>
<td>0,429</td>
</tr>
<tr>
<td>Visvliet</td>
<td>0,650</td>
<td>0,444</td>
<td>0,600</td>
<td>0,714</td>
<td>0,556</td>
<td>0,682</td>
<td>0,625</td>
<td>0,353</td>
<td>0,000</td>
<td>0,588</td>
<td>0,500</td>
<td>0,667</td>
</tr>
<tr>
<td>Oosterbierum</td>
<td>0,533</td>
<td>0,632</td>
<td>0,500</td>
<td>0,600</td>
<td>0,462</td>
<td>0,706</td>
<td>0,308</td>
<td>0,417</td>
<td>0,625</td>
<td>0,588</td>
<td>0,167</td>
<td>0,571</td>
</tr>
<tr>
<td>Beetgum / Bintum</td>
<td>0,545</td>
<td>0,714</td>
<td>0,500</td>
<td>0,727</td>
<td>0,500</td>
<td>0,750</td>
<td>0,333</td>
<td>0,556</td>
<td>0,643</td>
<td>0,500</td>
<td>0,167</td>
<td>0,000</td>
</tr>
<tr>
<td>Bergum / Buntum</td>
<td>0,500</td>
<td>0,500</td>
<td>0,429</td>
<td>0,813</td>
<td>0,500</td>
<td>0,571</td>
<td>0,333</td>
<td>0,500</td>
<td>0,429</td>
<td>0,667</td>
<td>0,571</td>
<td>0,500</td>
</tr>
<tr>
<td>Jorwerd / Jorwerd</td>
<td>0,692</td>
<td>0,667</td>
<td>0,583</td>
<td>0,846</td>
<td>0,545</td>
<td>0,667</td>
<td>0,400</td>
<td>0,600</td>
<td>0,571</td>
<td>0,692</td>
<td>0,500</td>
<td>0,455</td>
</tr>
<tr>
<td>Bakkeveen / Bakk</td>
<td>0,400</td>
<td>0,500</td>
<td>0,438</td>
<td>0,706</td>
<td>0,385</td>
<td>0,563</td>
<td>0,357</td>
<td>0,385</td>
<td>0,579</td>
<td>0,533</td>
<td>0,545</td>
<td>0,385</td>
</tr>
<tr>
<td>Waskemeer / Was</td>
<td>0,438</td>
<td>0,526</td>
<td>0,556</td>
<td>0,818</td>
<td>0,500</td>
<td>0,588</td>
<td>0,471</td>
<td>0,533</td>
<td>0,471</td>
<td>0,652</td>
<td>0,588</td>
<td>0,667</td>
</tr>
<tr>
<td>Kloosterburen</td>
<td>0,500</td>
<td>0,412</td>
<td>0,611</td>
<td>0,810</td>
<td>0,563</td>
<td>0,357</td>
<td>0,529</td>
<td>0,600</td>
<td>0,333</td>
<td>0,636</td>
<td>0,706</td>
<td>0,667</td>
</tr>
<tr>
<td>Warfum</td>
<td>0,563</td>
<td>0,438</td>
<td>0,667</td>
<td>0,737</td>
<td>0,625</td>
<td>0,429</td>
<td>0,588</td>
<td>0,643</td>
<td>0,400</td>
<td>0,652</td>
<td>0,600</td>
<td>0,636</td>
</tr>
<tr>
<td>Leermens</td>
<td>0,667</td>
<td>0,652</td>
<td>0,739</td>
<td>0,550</td>
<td>0,773</td>
<td>0,650</td>
<td>0,739</td>
<td>0,722</td>
<td>0,389</td>
<td>0,455</td>
<td>0,667</td>
<td>0,571</td>
</tr>
<tr>
<td>Groningen</td>
<td>0,714</td>
<td>0,682</td>
<td>0,714</td>
<td>0,636</td>
<td>0,783</td>
<td>0,762</td>
<td>0,800</td>
<td>0,778</td>
<td>0,471</td>
<td>0,476</td>
<td>0,684</td>
<td>0,737</td>
</tr>
<tr>
<td>Nieuw-Scheerden</td>
<td>0,650</td>
<td>0,682</td>
<td>0,650</td>
<td>0,652</td>
<td>0,773</td>
<td>0,762</td>
<td>0,739</td>
<td>0,722</td>
<td>0,556</td>
<td>0,358</td>
<td>0,647</td>
<td>0,615</td>
</tr>
<tr>
<td>Langeloo</td>
<td>0,727</td>
<td>0,524</td>
<td>0,739</td>
<td>0,652</td>
<td>0,792</td>
<td>0,650</td>
<td>0,760</td>
<td>0,647</td>
<td>0,550</td>
<td>0,500</td>
<td>0,700</td>
<td>0,824</td>
</tr>
</tbody>
</table>

*This is a table showing dialectometric analysis results for various locations.*
• step 2: reduce this 267-dimensional matrix to a two- or three-dimensional one, so that it can easily be visualized
• step 3: project back onto a geographical map
This talk in one slide

Introduction

A dialectometric analysis

Reverse dialectometry

Conclusion

References
• shortcomings of this approach for my current purposes:
• shortcomings of this approach for my current purposes:
  1. the linguistic constructions themselves play only an indirect role in the outcome of the analysis: we can see when two dialects differ, but we don’t see which cluster orders are responsible for this difference or how they cluster or correlate
shortcomings of this approach for my current purposes:

1. the linguistic constructions themselves play only an indirect role in the outcome of the analysis: we can see when two dialects differ, but we don’t see which cluster orders are responsible for this difference or how they cluster or correlate

2. there is no link between the data that feed into the quantitative analysis and the formal theoretical literature on verb clusters
Reverse dialectometry

- **proposal:** two changes to the classical dialectometric setup:
Reverse dialectometry

- proposal: two changes to the classical dialectometric setup:
  1. cluster orders are *individuals* rather than variables, i.e. instead of calculating differences between dialect locations, we measure differences between linguistic constructions
Reverse dialectometry

• **proposal:** two changes to the classical dialectometric setup:
  
  1. cluster orders are *individuals* rather than variables, i.e. instead of calculating differences between dialect locations, we measure differences between linguistic constructions
  
  2. Multiple Correspondence Analysis (MCA) instead of Multidimensional Scaling (MDS): involves the same kind of dimension reduction, but applied simultaneously to individuals and variables → will allow for the inclusion of formal theoretical variables alongside geographical ones
Reverse dialectometry

• proposal: two changes to the classical dialectometric setup:
  1. cluster orders are *individuals* rather than variables, i.e. instead of calculating differences between dialect locations, we measure differences between linguistic constructions
  2. Multiple Correspondence Analysis (MCA) instead of Multidimensional Scaling (MDS): involves the same kind of dimension reduction, but applied simultaneously to individuals and variables → will allow for the inclusion of formal theoretical variables alongside geographical ones

• starting point: a data table with cluster orders as rows and dialect locations as columns
<table>
<thead>
<tr>
<th></th>
<th>Midland</th>
<th>Lies</th>
<th>West.Tersch</th>
<th>Oosterend</th>
<th>Hollum</th>
<th>Schiermonnl</th>
<th>Ferwerd</th>
<th>Anjum</th>
<th>Kollum</th>
<th>Visvliet</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUX1(be.sg)-PART2</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>PART2-AUX1(be.sg)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>NA</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>AUX1(have.sg)-PART2</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>PART2-AUX1(have.sg)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>NA</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>AUX1(have.pl)-PART2</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>PART2-AUX1(have.pl)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>MOD1(sg)-INF2</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>INF2-MOD1(sg)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>MOD2-INF3-MOD1(sg)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>MOD1(sg)-MOD2-INF3</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>MOD1(sg)-INF3-MOD2</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>INF3-MOD2-MOD1(sg)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>INF3-MOD1(sg)-MOD2</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>MOD1(sg)-AUX2(have)-PART3</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>MOD1(sg)-PART3-AUX2(have)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>PART3-MOD1(sg)-AUX2(have)</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>PART3-AUX2(have)-MOD1(sg)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>NA</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>AUX1(be.sg)-AUX2(go)-INF3</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>AUX1(be.sg)-INF3-AUX2(go)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>AUX2(go)-AUX1(be.sg)-INF3</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>no</td>
</tr>
<tr>
<td>AUX2(go)-INF3-AUX1(be.sg)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>INF3-AUX1(be.sg)-AUX2(go)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>NA</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>INF3-AUX2(go)-AUX1(be.sg)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>NA</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>AUX1(have.sg)-MOD2(inf)-INF3</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>AUX1(have.sg)-INF3-MOD2(part)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>AUX1(have.sg)-INF3-MOD2(inf)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>MOD2(inf)-INF3-AUX1(have.sg)</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>INF3-AUX1(have.sg)-MOD2(inf)</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>INF3-AUX1(have.sg)-MOD2(part)</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>INF3-MOD2(part)-AUX1(have.sg)</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>INF3-MOD2(inf)-AUX1(have.sg)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
• transform to a distance matrix and reduce its dimensionality
Two-dimensional representation of the 31 SAND-verb cluster orders
- **note:** each point now represents a particular cluster order and closeness of points indicates how alike two verb cluster orders are based on their geographical spread
• **note:** each point now represents a particular cluster order and closeness of points indicates how alike two verb cluster orders are based on their geographical spread

• if this likeness is the result of grammatical microparameters, then verb cluster orders that are ‘closeby’ should be the result of the same parameter setting, i.e. parameters create **natural classes** of verb cluster orders
• note: each point now represents a particular cluster order and closeness of points indicates how alike two verb cluster orders are based on their geographical spread

• if this likeness is the result of grammatical microparameters, then verb cluster orders that are ‘closeby’ should be the result of the same parameter setting, i.e. parameters create natural classes of verb cluster orders

• in order to find those parameters, we can also encode the cluster orders in terms of their theoretical linguistic analyses
• e.g. in Barbiers (2005)’s analysis cluster orders involve the following microparameters:
• e.g. in Barbiers (2005)'s analysis cluster orders involve the following microparameters:
  • [±base-generation]: can the order be base-generated?
• e.g. in Barbiers (2005)'s analysis cluster orders involve the following microparameters:
  • [±base-generation]: can the order be base-generated?
  • [±movement]: can the order be derived via movement?
e.g. in Barbiers (2005)'s analysis cluster orders involve the following microparameters:

- [±base-generation]: can the order be base-generated?
- [±movement]: can the order be derived via movement?
- [±pied-piping]: does the derivation involve pied-piping?
• e.g. in Barbiers (2005)'s analysis cluster orders involve the following microparameters:
  • [±base-generation]: can the order be base-generated?
  • [±movement]: can the order be derived via movement?
  • [±pied-piping]: does the derivation involve pied-piping?
  • [±feature-checking violation]: does the order involve a feature checking violation?
e.g. in Barbiers (2005)’s analysis cluster orders involve the following microparameters:

- [±base-generation]: can the order be base-generated?
- [±movement]: can the order be derived via movement?
- [±pied-piping]: does the derivation involve pied-piping?
- [±feature-checking violation]: does the order involve a feature checking violation?

and the SAND cluster orders can be encoded in terms of these microparameters
### This talk in one slide

#### Introduction

- A dialectometric analysis
- Reverse dialectometry

#### A dialectometric analysis

<table>
<thead>
<tr>
<th>AUX1(be.sg)-PART2</th>
<th>Barbiers-base-generation</th>
<th>Barbiers-movement</th>
<th>Barbiers-spec-pied-piping</th>
<th>Barbiers-feature.checking-failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>yesBase</td>
<td>noMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
<td></td>
</tr>
<tr>
<td>PART2-AUX1(be.sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX1(have.sg)-PART2</td>
<td>yesBase</td>
<td>noMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>PART2-AUX1(have.sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX1(have.pl)-PART2</td>
<td>yesBase</td>
<td>noMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>PART2-AUX1(have.pl)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>MOD1(sg)-INF2</td>
<td>yesBase</td>
<td>noMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>INF2-MOD1(sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>MOD2-INF3-MOD1(sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>MOD1(sg)-MOD2-INF3</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>MOD3-INF3-MOD2</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>INF3-MOD2-MOD1(sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>INF3-MOD1(sg)-MOD2</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>MOD1(sg)-AUX2(have)-PART3</td>
<td>yesBase</td>
<td>noMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>MOD1(sg)-PART3-AUX2(have)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>PART3-MOD1(sg)-AUX2(have)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>PART3-AUX2(have)-MOD1(sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>yesPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX1(be.sg)-AUX2(go)-INF3</td>
<td>yesBase</td>
<td>noMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX1(be.sg)-INF3-AUX2(go)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX2(go)-AUX1(be.sg)-INF3</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX2(go)-INF3-AUX1(be.sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>INF3-AUX1(be.sg)-AUX2(go)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>yesFeatCheckFail</td>
</tr>
<tr>
<td>INF3-AUX2(go)-AUX1(be.sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX1(have.sg)-MOD2(inf)-INF3</td>
<td>yesBase</td>
<td>noMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX1(have.sg)-INF3-MOD2(part)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>AUX1(have.sg)-INF3-MOD2(inf)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>MOD2(inf)-INF3-AUX1(have.sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>INF3-AUX1(have.sg)-MOD2(inf)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>yesFeatCheckFail</td>
</tr>
<tr>
<td>INF3-AUX1(have.sg)-MOD2(part)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>yesFeatCheckFail</td>
</tr>
<tr>
<td>INF3-MOD2(part)-AUX1(have.sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>INF3-MOD2(inf)-AUX1(have.sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
<tr>
<td>INF3-MOD2(inf)-AUX1(have.sg)</td>
<td>noBase</td>
<td>yesMvt</td>
<td>noPiedP</td>
<td>noFeatCheckFail</td>
</tr>
</tbody>
</table>
Barbiers's microparameters thus serve as supplementary variables in the data table.
• Barbiers’s microparameters thus serve as supplementary variables in the data table

• **in total:** 70 additional variables distilled from the theoretical literature on verb clusters:
• Barbiers’s microparameters thus serve as supplementary variables in the data table
• **in total:** 70 additional variables distilled from the theoretical literature on verb clusters:
  • the analyses of Barbiers (2005), Barbiers and Bennis (2010), Abels (2011), Haegeman and Riemsdijk (1986), Bader (2012), and Schmid and Vogel (2004)
- Barbiers's microparameters thus serve as supplementary variables in the data table
- **in total**: 70 additional variables distilled from the theoretical literature on verb clusters:
  - a head-initial head movement analysis, a head-final head movement analysis, a head-initial XP-movement analysis, a head-final XP-movement analysis (all based on Wurmbrand (2005))
• Barbiers's microparameters thus serve as supplementary variables in the data table

• **in total:** 70 additional variables distilled from the theoretical literature on verb clusters:
  
  • the analyses of Barbiers (2005), Barbiers and Bennis (2010), Abels (2011), Haegeman and Riemsdijk (1986), Bader (2012), and Schmid and Vogel (2004)
  
  • a head-initial head movement analysis, a head-final head movement analysis, a head-initial XP-movement analysis, a head-final XP-movement analysis (all based on Wurmbrand (2005))
  
  • 17 additional variables based on the theoretical literature, but not linked to a specific analysis
• proposal (I): the number of microparameters responsible for the verb cluster variation = the number of dimensions we reduce our data set to
• **note:** there seems to be a clear cut-off point after the third dimension
• note: there seems to be a clear cut-off point after the third dimension

• together, the first three dimensions account for 78.46% of the variation in the SAND verb cluster data
• **note:** there seems to be a clear cut-off point after the third dimension

• together, the first three dimensions account for 78.46% of the variation in the SAND verb cluster data

• this means that roughly 80% of the variation in verb cluster ordering in SAND can be reduced to three microparameters
• **note:** there seems to be a clear cut-off point after the third dimension

• together, the first three dimensions account for 78.46% of the variation in the SAND verb cluster data

• this means that roughly 80% of the variation in verb cluster ordering in SAND can be reduced to three microparameters

• in order to know what those microparameters are, we need to *interpret* the first three dimensions
- **proposal (I):** the number of microparameters responsible for the verb cluster variation = the number of dimensions we reduce our data set to

- **proposal (II):** the identity of those microparameters = the interpretation of the dimensions
• **proposal (I):** the number of microparameters responsible for the verb cluster variation = the number of dimensions we reduce our data set to

• **proposal (II):** the identity of those microparameters = the interpretation of the dimensions

• the degree of similarity/correlation between a dimension and a linguistic variable can be determined by:
  1. visual inspection of a color-coded map
  2. calculating the squared correlation ratio ($\eta^2$): value between 0 and 1 indicating the strength of the link between a dimension and a particular categorical variable; can be interpreted as the percentage of variation on the dimension that can be explained by that categorical variable
Dimension 1

- is related to the position of infinitives and participles *vis-à-vis* their selecting verbs (modals and auxiliaries respectively)
Dimension 1

- is related to the position of infinitives and participles *vis-à-vis* their selecting verbs (modals and auxiliaries respectively)
- more specifically, the variable `InfMod.AuxPart`
Dimension 1

- is related to the position of infinitives and participles *vis-à-vis* their selecting verbs (modals and auxiliaries respectively)
- more specifically, the variable $\text{InfMod.AuxPart}$:
  - set to ‘no’ when the modal precedes the infinitive (when present) and the participle precedes the auxiliary (when present)
Dimension 1

- is related to the position of infinitives and participles *vis-à-vis* their selecting verbs (modals and auxiliaries respectively)
- more specifically, the variable `InfMod.AuxPart`:
  - set to ‘no’ when the modal precedes the infinitive (when present) and the participle precedes the auxiliary (when present)
  - set to ‘yes’ when at least one of these conditions is not met
Dimension 1

- is related to the position of infinitives and participles *vis-à-vis* their selecting verbs (modals and auxiliaries respectively)
- more specifically, the variable `InfMod.AuxPart`:
  - set to ‘no’ when the modal precedes the infinitive (when present) and the participle precedes the auxiliary (when present)
  - set to ‘yes’ when at least one of these conditions is not met
- this variable has a $\eta^2$ of 0.6142
Dimension 2

- is related to the ‘slope’ of the cluster: ascending or descending
Dimension 2

- is related to the ‘slope’ of the cluster: ascending or descending
- more specifically, the variable $\text{FINALDescent}$:
Dimension 2

- is related to the ‘slope’ of the cluster: ascending or descending
- more specifically, the variable \texttt{FINAL_DESCENT}:
  - set to ‘yes’ if the cluster ends in a descending order

\begin{center}
\begin{tabular}{c c c c}
  21 & 12 & 132 & 123 \\
  321 & 312 & 231 & 213 \\
\end{tabular}
\end{center}
Dimension 2

- is related to the ‘slope’ of the cluster: ascending or descending
- more specifically, the variable `FINAL_DESCENT`:
  - set to ‘yes’ if the cluster ends in a descending order
  - set to ‘no’ if it ends in an ascending order
Dimension 2

- is related to the ‘slope’ of the cluster: ascending or descending
- more specifically, the variable $\text{FinalDescent}$:
  - set to ‘yes’ if the cluster ends in a descending order
  - set to ‘no’ if it ends in an ascending order

<table>
<thead>
<tr>
<th>FinalDescent_yes</th>
<th>FinalDescent_no</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td>132</td>
<td>123</td>
</tr>
<tr>
<td>321</td>
<td>312</td>
</tr>
<tr>
<td>231</td>
<td>213</td>
</tr>
</tbody>
</table>
Dimension 2

- is related to the ‘slope’ of the cluster: ascending or descending
- more specifically, the variable $\text{FINALDescent}$:
  - set to ‘yes’ if the cluster ends in a descending order
  - set to ‘no’ if it ends in an ascending order

<table>
<thead>
<tr>
<th>FinalDescent_yes</th>
<th>FinalDescent_no</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td>132</td>
<td>123</td>
</tr>
<tr>
<td>321</td>
<td>312</td>
</tr>
<tr>
<td>231</td>
<td>213</td>
</tr>
</tbody>
</table>

- this variable has a $\eta^2$ of 0.382
Dimension 3

- is strongly correlated with head-finality
Dimension 3

- is strongly correlated with head-finality
- a variable like `HeadFinalBaseOrder` that separates strictly head-final orders from all others has a \( \eta^2 \) of 0.686
This talk in one slide

Introduction

A dialectometric analysis

Reverse dialectometry

Conclusion

References

Dimension 3

Dimension 3 vs. Bader's (2012) base-generated order
Conclusion

- from these three parameters, a rough, parametrized analysis of verb clusters arises:
• from these three parameters, a rough, parametrized analysis of verb clusters arises:

  1. a head-final base order
from these three parameters, a rough, parametrized analysis of verb clusters arises:

1. a head-final base order
2. which dialects can diverge from or not: $\pm$Movement (dimension 3)
Conclusion

- from these three parameters, a rough, parametrized analysis of verb clusters arises:
  1. a head-final base order
  2. which dialects can diverge from or not: [±Movement] (dimension 3)
  3. those that diverge can diverge strongly or not: Economy of Movement (dimension 2)
from these three parameters, a rough, parametrized analysis of verb clusters arises:

1. a head-final base order
2. which dialects can diverge from or not: [±Movement] (dimension 3)
3. those that diverge can diverge strongly or not: Economy of Movement (dimension 2)
4. above and beyond all this, a headedness parameter regulates the order of infinitives and participles vis-à-vis their selecting verbs: [±ModInf&PartAux] (dimension 1)
Conclusion

- roughly 80% of the variation found in Dutch verb cluster orders can be reduced to three grammatical microparameters by applying a statistical analysis to the data
Conclusion

• roughly 80% of the variation found in Dutch verb cluster orders can be reduced to three grammatical microparameters by applying a statistical analysis to the data
• more generally, there is room for fruitful collaboration between formal-theoretical and quantitative-statistical linguistics:
Conclusion

- roughly 80% of the variation found in Dutch verb cluster orders can be reduced to three grammatical microparameters by applying a statistical analysis to the data.

- more generally, there is room for fruitful collaboration between formal-theoretical and quantitative-statistical linguistics:
  - the former can guide the interpretation of results from the latter.
Conclusion

- roughly 80% of the variation found in Dutch verb cluster orders can be reduced to three grammatical microparameters by applying a statistical analysis to the data
- more generally, there is room for fruitful collaboration between formal-theoretical and quantitative-statistical linguistics:
  - the former can guide the interpretation of results from the latter
  - the latter can help evaluate and test hypotheses of the former
References I


References II


