



# Groene vs. rode woordvolgorde: een poging tot verklaring

Gert De Sutter



University of Leuven  
RU Quantitative Lexicology and Variational Linguistics

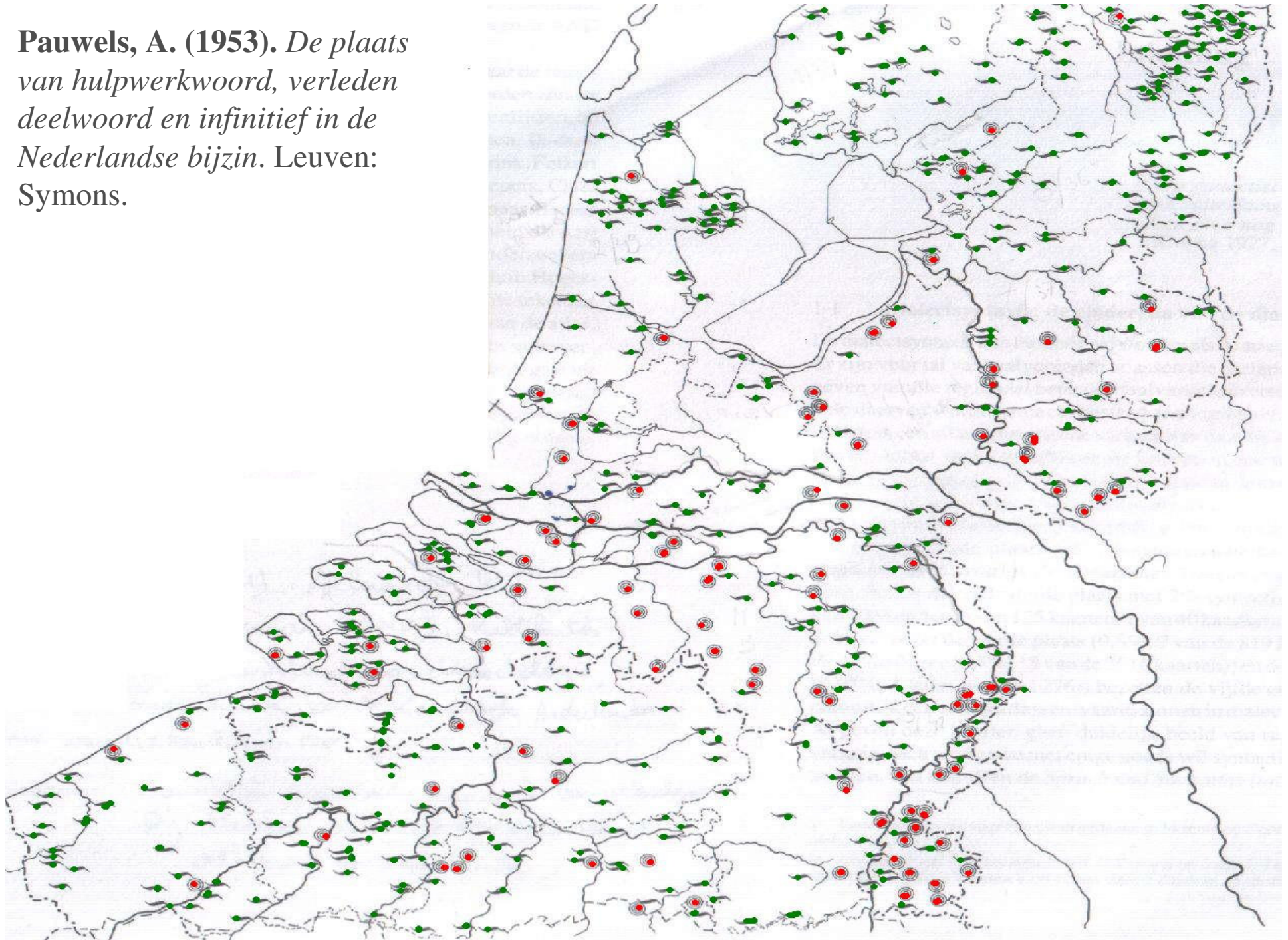
# Onderzoeksobject

- Woordvolgordevariatie in werkwoordelijke eindgroepen  
[...] dat Jan dat boek **gestolen**<sub>PART</sub> **heeft**<sub>AUX</sub>.  
[...] dat Jan dat boek **heeft**<sub>AUX</sub> **gestolen**<sub>PART</sub>.
- Beide varianten zijn grammaticaal en semantisch-pragmatisch equivalent
- Zogenaamde groene en rode woordvolgorde

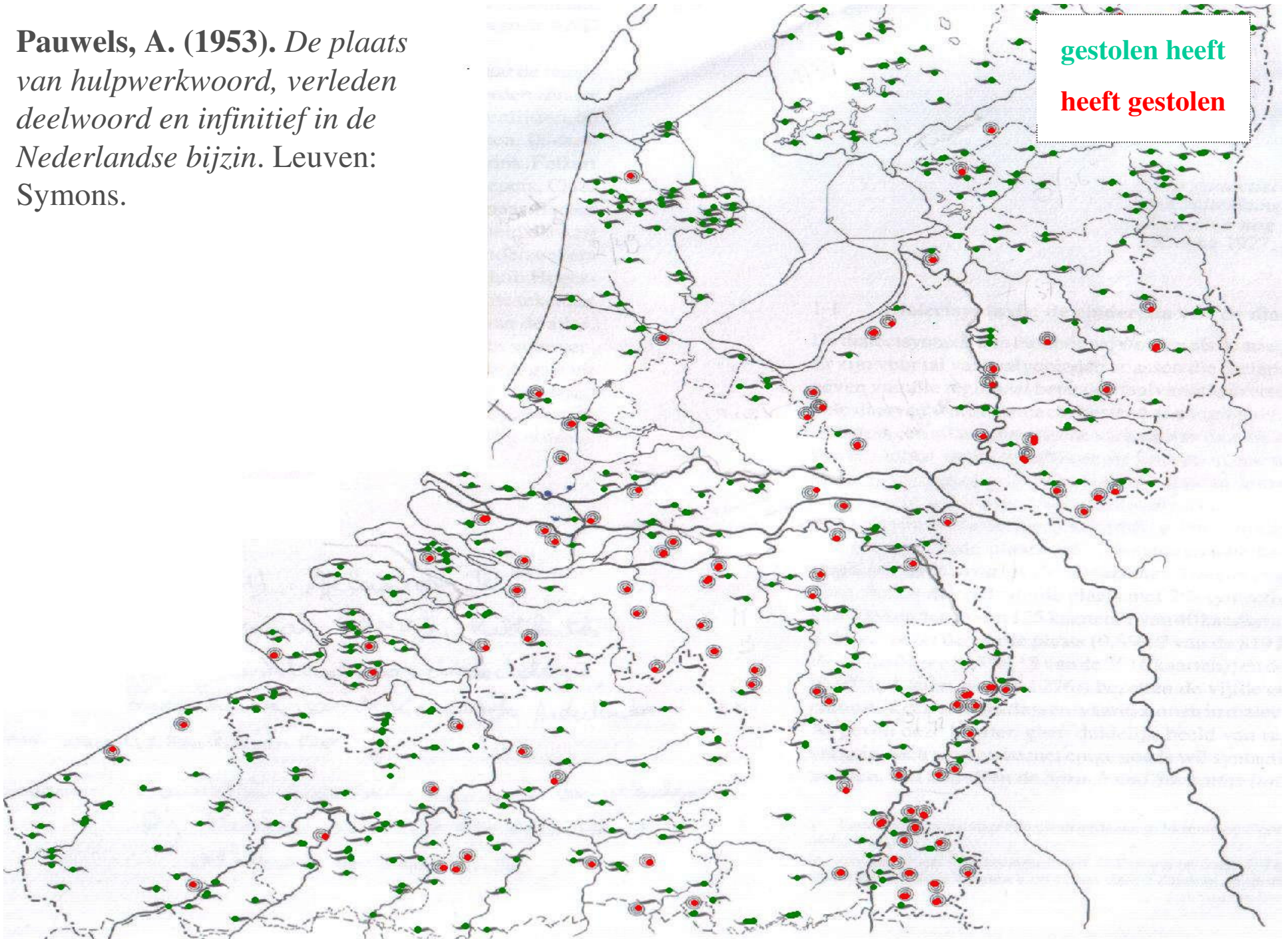
# Vanwaar de termen groen en rood?



**Pauwels, A. (1953).** *De plaats van hulpwerkwoord, verleden deelwoord en infinitief in de Nederlandse bijzin.* Leuven: Symons.



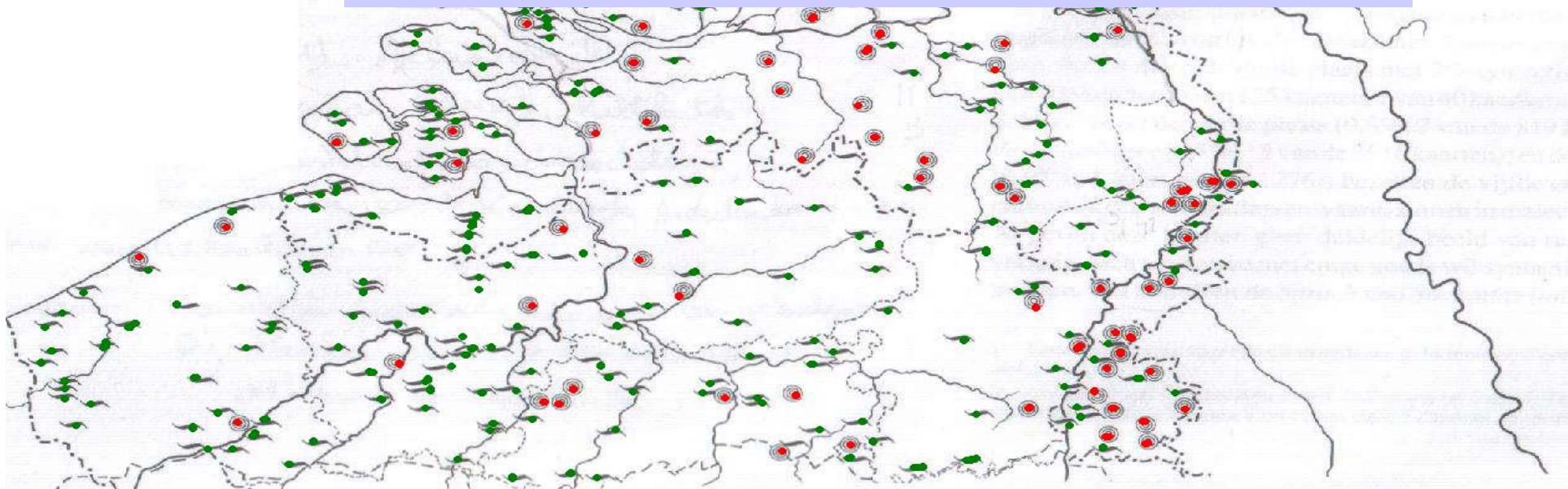
**Pauwels, A. (1953).** *De plaats van hulpwerkwoord, verleden deelwoord en infinitief in de Nederlandse bijzin.* Leuven: Symons.



**Pauwels, A. (1953).** *De plaats van hulpwerkwoord, verleden deelwoord en infinitief in de Nederlandse bijzin.* Leuven: Symons.



Intensief onderzoek sinds jaren 1950, maar nog steeds veel onbeantwoorde vragen...



gestolen heeft

heeft gestolen

# Onderzoeksvragen

- **Welke factoren** zorgen ervoor dat in het concrete taalgebruik nu eens de groene en dan weer de rode volgorde gekozen wordt?
- Waarom bestaan de groene en de rode volgorde naast elkaar? Is er een **cognitief-functioneel onderscheid** aan te wijzen?

gestolen heeft

heeft gestolen

# Onderzoeksvragen

- Welke factoren zorgen ervoor dat in het concrete taalgebruik nu eens de groene en dan weer de rode volgorde gekozen wordt?
- **Doctoraatsonderzoek**

*Rood, groen, corpus! Een taalgebruiksgebaseerde analyse van woordvolgordevariatie in tweeledige werkwoordelijke eindgroepen (2005)*

gestolen heeft

heeft gestolen

# Rood, groen, corpus!

- Empirisch en kwantitatief corpusonderzoek naar rol van taalinterne factoren:
  - Invloed en sterkte van elke factor
  - Relatie tussen de factoren (redundantie?)
  - Simultaan effect van de factoren (verklaring en voorspelling)

gestolen heeft

heeft gestolen

# Rood, groen, corpus!

- Empirisch en kwantitatief corpusonderzoek naar rol van taalinterne factoren:
  - Invloed en sterkte van elke factor
  - Relatie tussen de factoren (redundantie?)
  - Simultaan effect van de factoren (verklaring en voorspelling)
    - Formeel Belgisch Nederlands (Condiv)
    - Tweeledige eindgroepen met een deelwoord en hulpwerkwoord *zijn, hebben* of *worden* in een complementzin ingeleid door *dat*
    - Aantal relevante eindgroepen: 2390
    - Alle resultaten zijn statistisch gevalideerd

gestolen heeft

heeft gestolen

## Rood, groen, corpus!

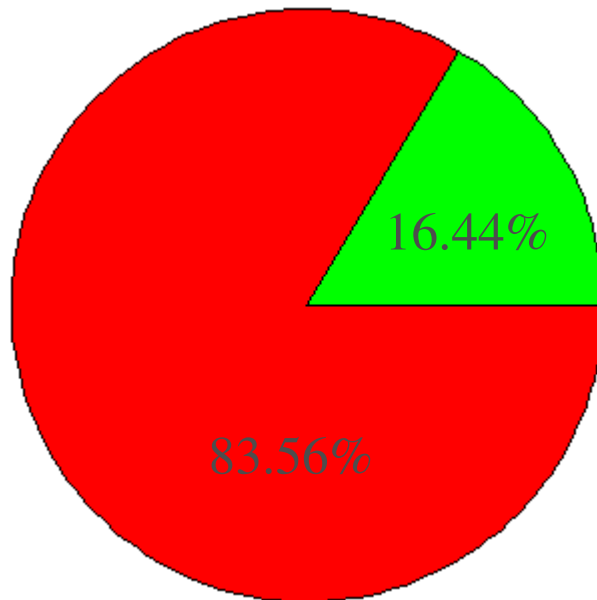
- De rood-groenvariatie wordt simultaan bepaald door 8 verschillende factoren (logistische regressie):
  1. Morfologie van het deelwoord
  2. Frequentie van het deelwoord
  3. Semantisch-categoriale status van het deelwoord
  4. Informatiewaarde van het voorafgaande woord
  5. Inherentie van de voorafgaande constituent
  6. Lengte van het middenstuk
  7. Syntactische functie van de laatste zinsplaats
  8. Syntactische persistentie

# Resultaten

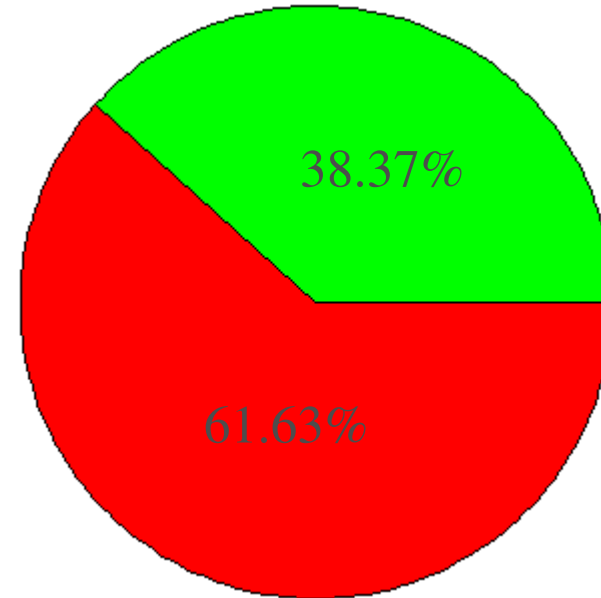
gestolen heeft

heeft gestolen

## 1. Morfologie van het deelwoord



Scheidbaar samengestelde  
deelwoorden  
opgemerkt



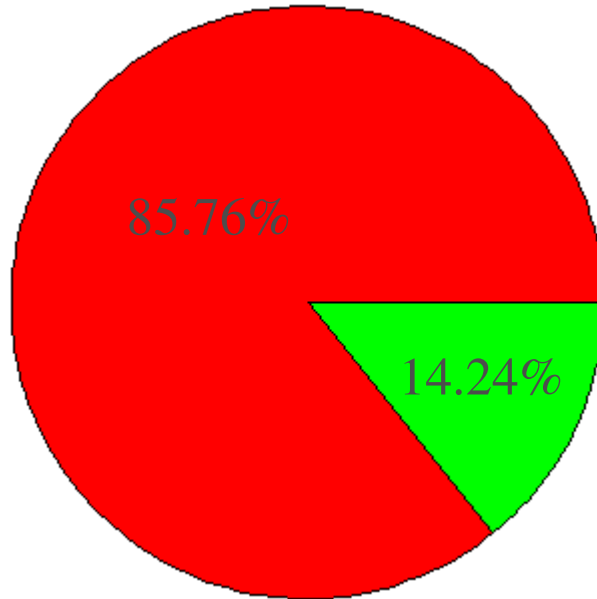
Niet-scheidbaar samengestelde  
deelwoorden  
gemerkt

# Resultaten

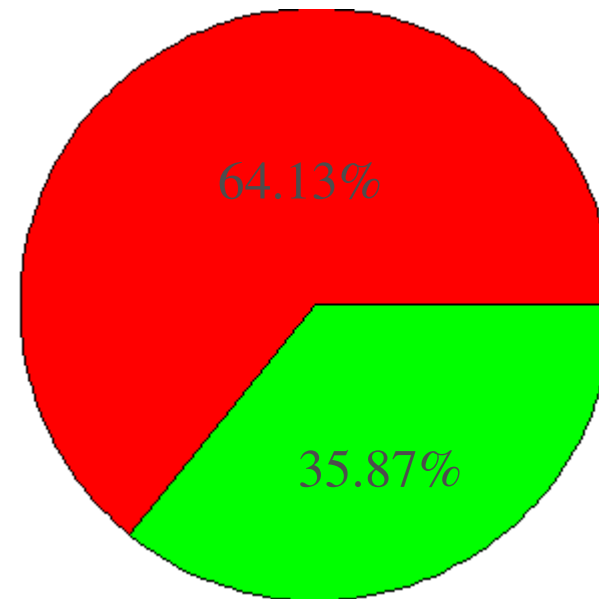
gestolen heeft

heeft gestolen

## 2. Frequentie van het deelwoord



Hoog-frequente deelwoorden (> 36894)



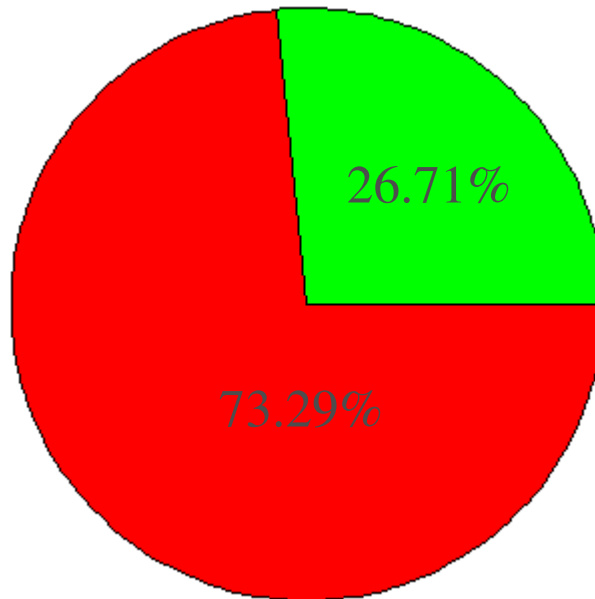
Laag-frequente deelwoorden (< 36894)

# Resultaten

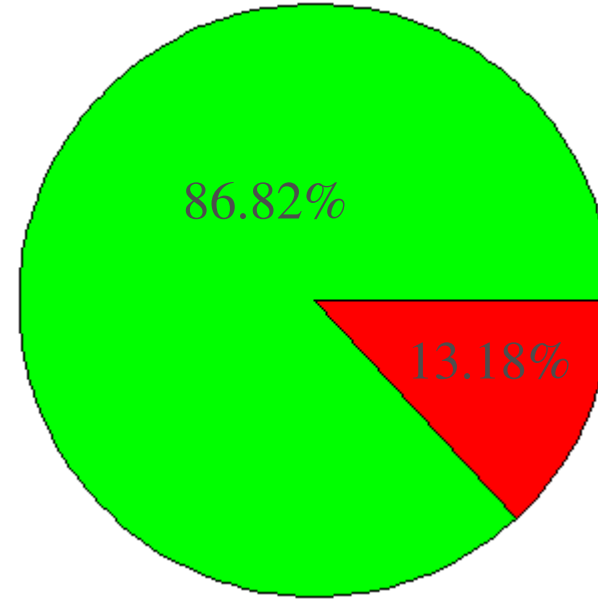
gestolen heeft

heeft gestolen

## 3. Semantisch-categoriale status van het deelwoord



Verbale deelwoorden  
[...] dat de winkel daarnet door  
de eigenaar gesloten is



Adjectivische deelwoorden  
[...] dat de winkel al een hele tijd  
gesloten is

gestolen heeft

heeft gestolen

# Resultaten

## 4. Informatiewaarde van het voorafgaande woord

[...] dat hij met de Vlos  
een vergiftigd geschenk  
heeft doorgegeven aan  
zijn opvolger minister  
Weckx

Hoge  
informatiewaarde

[...] dat Duitsland na  
1945 opnieuw is  
gevormd naar westerse  
waarden

Intermediaire  
informatiewaarde

[...] dat de wapens bij  
hem gedeponereerd zijn  
door een stadsgenoot

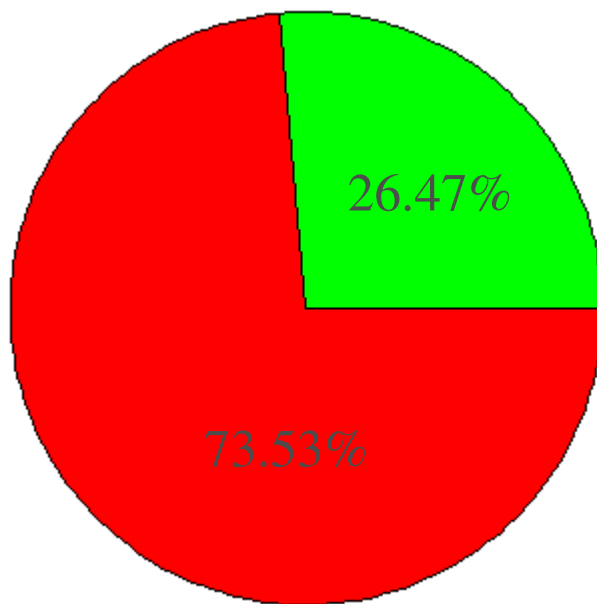
Lage  
informatiewaarde

# Resultaten

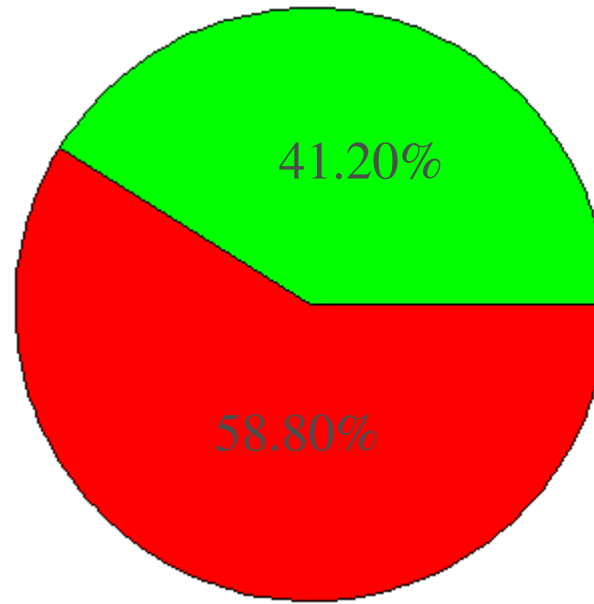
gestolen heeft

heeft gestolen

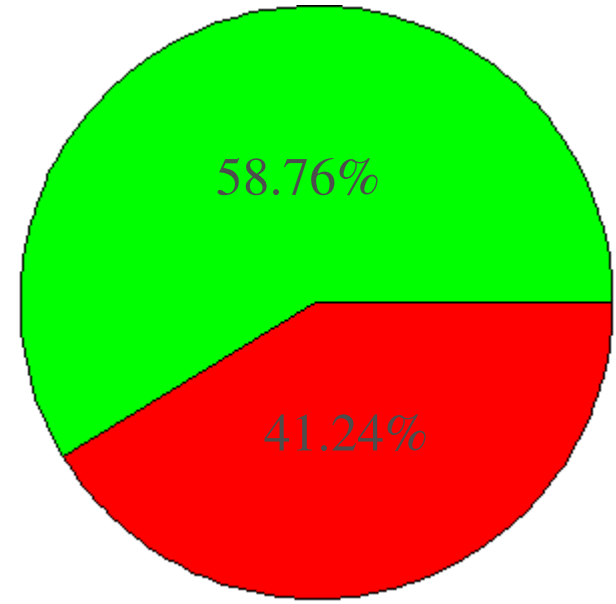
## 4. Informatiewaarde van het voorafgaande woord



Hoge  
informatiewaarde



Intermediaire  
informatiewaarde



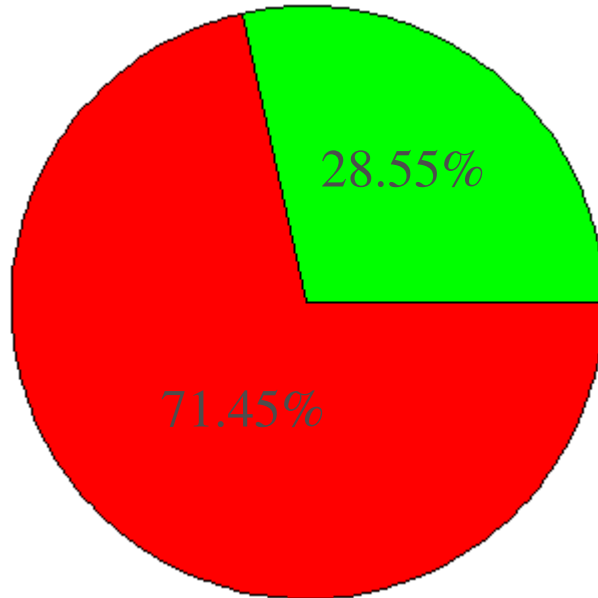
Lage  
informatiewaarde

# Resultaten

gestolen heeft

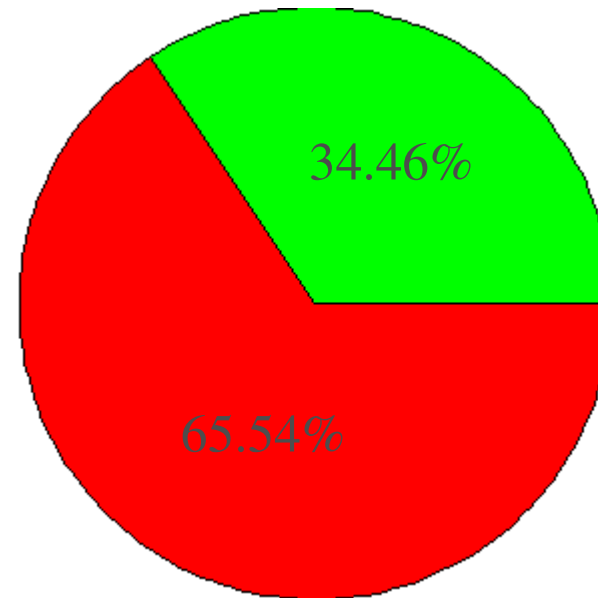
heeft gestolen

## 5. Inherentie van de voorafgaande constituent



Vaste verbinding

[...] dat geen **aandacht** werd **besteed**  
aan de financiering van het plan



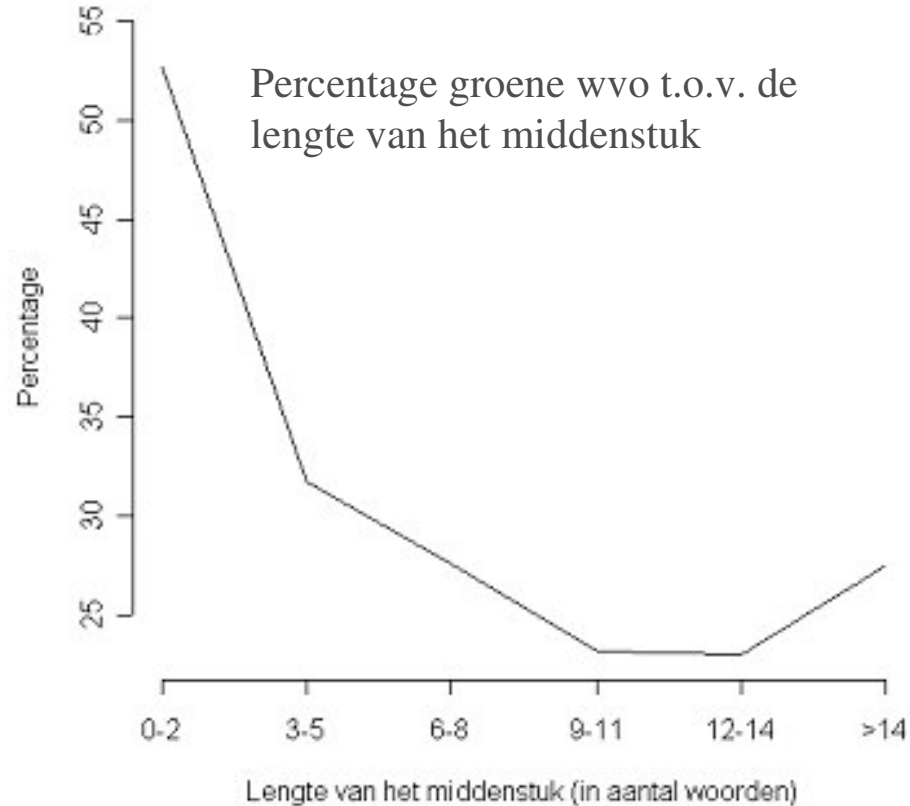
Geen vaste verbinding

[...] dat de burger met alle middelen  
**fiscaal** wordt **uitgeperst**

gestolen heeft  
heeft gestolen

# Resultaten

## 6. Lengte van het middenstuk (*dat ... ww. eindgroep*)



gestolen heeft

heeft gestolen

# Resultaten

## 7. Syntactische functie van de laatste zinsplaats

[...] dat de naam van de club  
door het slijk gehaald wordt net  
op het ogenblik dat hij nieuwe  
sponsors aantrekt

Adjunct

[...] dat men heeft  
aangenomen dat het om een  
gelijkwaardig universitair  
diploma gaat

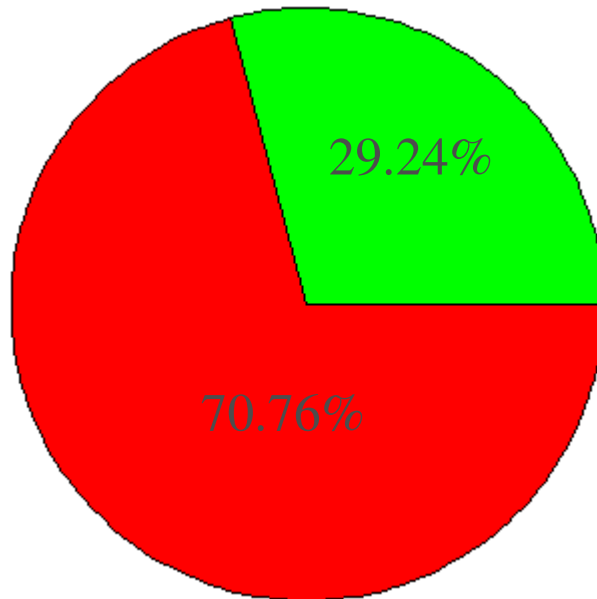
Complement

# Resultaten

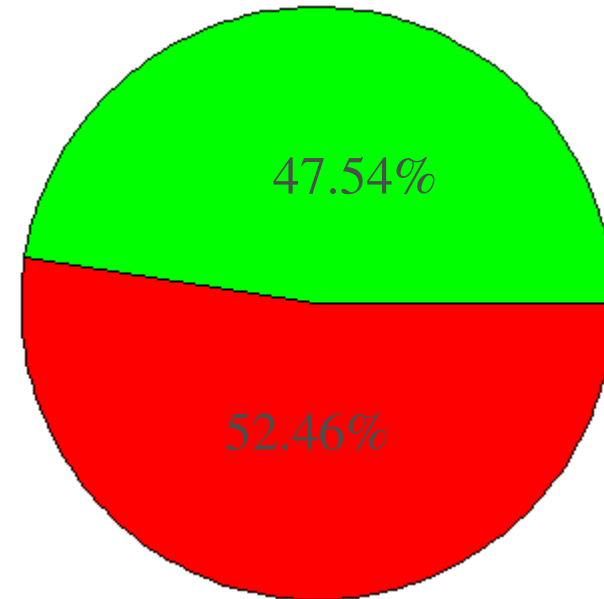
gestolen heeft

heeft gestolen

## 7. Syntactische functie van de laatste zinsplaats



Adjunct



Complement

gestolen heeft

heeft gestolen

# Resultaten

## 8. Syntactische persistentie

[...] dat men te goeder trouw **heeft aangenomen** dat het om een gelijkwaardig universitair diploma gaat.

Wellicht moet de Raad van State de zaak nu verder onderzoeken. Gewezen minister van Cultuur, Patrick Dewael, reageerde op ons bericht dat hij met de Vlos een vergiftigd geschenk **heeft doorgegeven** [...].

Voorafgaande rode cluster

[...] omdat ze de takken van zijn vogelkersen die boven hun afsluiting uit reikten, **gesnoeid hadden**. De vrederechter bracht oneindig veel geduld op om de reclamant, een Antwerpenaar die elk groen blad adoreerde, diets te maken dat de Canadese vogelkers in de Kempen de bospest **genoemd wordt**

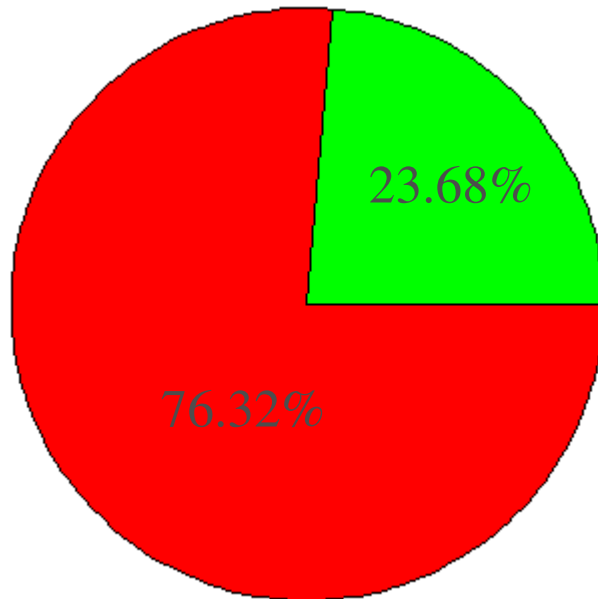
Voorafgaande groene cluster

# Resultaten

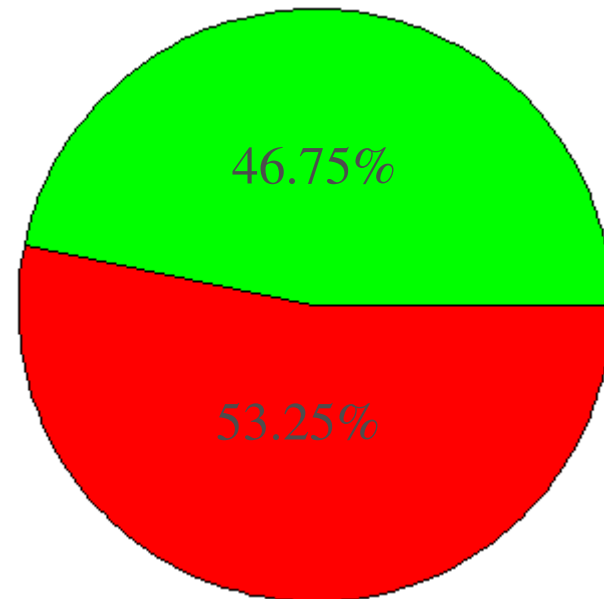
gestolen heeft

heeft gestolen

## 8. Syntactische persistentie



Voorafgaande rode cluster



Voorafgaande groene cluster

gestolen heeft

heeft gestolen

# Tussentijdse conclusie

- Een systematisch empirisch en kwantitatief-corpusgebaseerd onderzoek geeft een betrouwbaar inzicht in de complexiteit van het reële taalgebruik, m.n. de keuze tussen grammaticale alternatieven:
  - Verschillende factoren bepalen de woordvolgorde *simultaan*

gestolen heeft

heeft gestolen

## Tussentijdse conclusie

- Een systematisch empirisch en kwantitatief-corpusgebaseerd onderzoek geeft een betrouwbaar inzicht in de complexiteit van het reële taalgebruik, m.n. de keuze tussen grammaticale alternatieven:
  - Verschillende factoren bepalen de woordvolgorde *simultaan*
  - Sommige factoren hebben een grotere invloed

gestolen heeft

heeft gestolen

# Tussentijdse conclusie

## Hiërarchie van beïnvloedende factoren:

1. Semantisch-categoriale status van het deelwoord
2. Morfologie van het deelwoord
3. Syntactische persistentie
4. Lengte van het middenstuk
5. Inherentie van de voorafgaande constituent
6. Syntactische functie van de laatste zinsplaats
7. Informatiewaarde van het voorafgaande woord
8. Frequentie van het deelwoord

gestolen heeft

heeft gestolen

## Tussentijdse conclusie

- Een systematisch empirisch en kwantitatief-corpusgebaseerd onderzoek geeft een betrouwbaar inzicht in de complexiteit van het reële taalgebruik, m.n. de keuze tussen grammaticale alternatieven:
  - Verschillende factoren bepalen de woordvolgorde *simultaan*
  - Sommige factoren hebben een grotere invloed
  - Op basis van het complex van 8 factoren kan het grootste deel van de variatie adequaat beschreven en verklaard worden
  - Op basis van het complex van 8 factoren kan de woordvolgorde in hoge mate voorspeld worden

gestolen heeft

heeft gestolen

# De coëxistentievraag

- Waarom hebben de effecten het effect dat ze hebben? Met andere woorden, hoe kunnen we de verschillende effecten op een meer linguïstische manier **interpreteren**?
- Kunnen bepaalde significante factoren onder een grotere noemer / **principe** gevat worden (bv. topic-focus)?
- Wat is het **functioneel-cognitieve verschil** tussen de groene en de rode woordvolgorde? Waarom bestaan de beide volgordes naast elkaar?

gestolen heeft

heeft gestolen

# De coëxistentievraag

- Psycholinguïstische verklaringshypothese

gestolen heeft

heeft gestolen

# De coëxistentievraag

- Psycholinguïstische verklaringshypothese
- Echter:
  - Embryonaal, nadruk op hypothese
  - Niet contra andere verklaringen (semantisch-pragmatische hypothese van Pardoën, prosodische hypothese van De Schutter)

gestolen heeft

heeft gestolen

# Verklaringshypothese

## Overzicht van beïnvloedende factoren:

- Semantisch-categoriale status van het deelwoord
- Morfologie van het deelwoord
- Syntactische persistentie
- Lengte van het middenstuk
- Inherentie van de voorafgaande constituent
- Syntactische functie van de laatste zinsplaats
- Informatiewaarde van het voorafgaande woord
- Frequentie van het deelwoord

gestolen heeft

heeft gestolen

# Verklaringshypothese

## Overzicht van beïnvloedende factoren:

- Semantisch-categoriale status van het deelwoord
- Morfologie van het deelwoord
- Syntactische persistentie
- Lengte van het middenstuk
- Inherentie van de voorafgaande constituent
- Syntactische functie van de laatste zinsplaats
- Informatiewaarde van het voorafgaande woord
- Frequentie van het deelwoord

gestolen heeft

heeft gestolen

# Verklaringshypothese

- Het effect van de frequentie van het deelwoord
  - STAP 1: Rode woordvolgorde is dominant wanneer het deelwoord hoog-frequent is
  - STAP 2: Hoog-frequente items zijn cognitief **meer verankerd** en beter toegankelijk en worden daarom sneller verwerkt
  - STAP 3: Aldus, rode volgorde wordt gekozen in omstandigheden waar taalgebruikers **geen of weinig moeilijkheden** hebben in het selecteren en verwerken van het deelwoord
  - STAP 4: Mutatis mutandis, groene volgorde wordt gekozen in omstandigheden waar taalgebruikers **meer moeilijkheden** hebben in het selecteren en verwerken van het deelwoord

gestolen heeft

heeft gestolen

# Verklaringshypothese

- Het effect van de syntactische functie van de LZP:
  - STAP 1: Groene woordvolgorde is dominant wanneer de eindgroep gevolgd wordt door een complement
  - STAP 2: Complementen worden **geselecteerd door het hoofd** en bevatten derhalve cruciale syntactico-semanticke en pragmatische info
  - STAP 3: Complementen die tot na de eindgroep in het geheugen moeten gehouden worden, kosten **meer cognitieve krachten**
  - STAP 4: Aldus, groene volgorde wordt gekozen in omstandigheden waar taalgebruikers **moeilijkheden** ondervinden bij de encoding van de taaluiting
  - STAP 5: Mutatis mutandis, rode volgorde wordt gekozen in omstandigheden waar taalgebruikers **minder moeilijkheden** hebben bij de encoding van de taaluiting



gestolen heeft

heeft gestolen

# Verklaringshypothese

- Assumpties:
  - De geïntendeerde boodschap is niet volledig grammaticaal geëncodeerd vooraleer aangevangen wordt met de articulatie of het schrijven ervan (cf. Bock et al.)
  - De geheugencapaciteiten zijn niet onbeperkt (center-embedding)
  - Het verschil tussen complementen en adjuncten is psychologisch reëel (garden-path)
  - Complementen in het geheugen houden kost cognitief meer (cf. SPLT van Gibson)

gestolen heeft

heeft gestolen

# Conclusie

- Groene woordvolgorde is de woordvolgorde die gekozen wordt in omstandigheden waar taalgebruikers moeilijkheden ondervinden bij de encodering van de taaluiting
  - => groene woordvolgorde is de basiswoordvolgorde (cf. kindertaalverwerving, dialectologie)
- Rode volgorde wordt gekozen in omstandigheden waar taalgebruikers minder moeilijkheden hebben bij de encodering van de taaluiting
  - => rode woordvolgorde is de gemarkeerde woordvolgorde (sociostilistische redenen)

gestolen heeft

heeft gestolen

# Toekomstplannen

- Verder uitwerken van de multifactoriële corpusanalyse
  - Interactie tussen factoren
  - Insereren van nieuwe factoren
  - Statistische en linguïstische factorreductie
- Verificatie van het verklaringsmodel
  - Psycholinguïstische verificatie – hypothese: groene woordvolgorde wordt sneller verwerkt dan de rode
  - Corpusanalytische verificatie – hypothese: groene woordvolgorde wordt verkozen wanneer de eindgroep gevolgd wordt door een focale (vs. een topicale) constituent



Meer weten? Opmerkingen of suggesties?

[gert.desutter@arts.kuleuven.be](mailto:gert.desutter@arts.kuleuven.be)

<http://www.ling.arts.kuleuven.ac.be/qlvl/>