



# Doe wat je niet laten kan: A usage-based analysis of Dutch causative constructions

Natalia Levshina

Supervisors: Dirk Geeraerts  
Dirk Speelman



University of Leuven  
RU Quantitative Lexicology and Variational Linguistics

# Outline

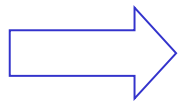
1. Theoretical and methodological background
2. Dutch causative constructions
3. Data and method
4. Semasiological variation (polysemy)
  - semantic and contextual variation of *doen*
  - semantic and contextual variation of *laten*
5. Onomasiological variation (near-synonymy)
  - variation of *doen* and *laten*
  - contextual factors
6. Conclusions and future research

# Empirical Cognitive Semantics

- Meaning is a concept
- Meaning reveals itself in usage
- Meaning is subject to social variation

# Empirical Cognitive Semantics

- Meaning is a concept

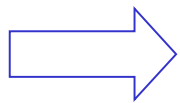


categorization and comprehension  
categories and exemplars

- Meaning reveals itself in usage
  
- Meaning is subject to social variation

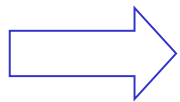
# Empirical Cognitive Semantics

- Meaning is a concept



categorization and comprehension  
categories and exemplars

- Meaning reveals itself in usage

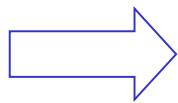


distributional bottom-up approach

- Meaning is subject to social variation

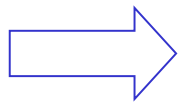
# Empirical Cognitive Semantics

- Meaning is a concept



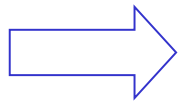
categorization and comprehension  
categories and exemplars

- Meaning reveals itself in usage



distributional bottom-up approach

- Meaning is subject to social variation



lectally diverse data

# State of the Art

- bias towards ‘alternations’ in studies of constructions (cf. Goldberg 2002: each construction should [also] be studied on its own right)
- main focus on generalizations and prototypes, although exemplar effects in human categorization are proved to be powerful
- insufficient attention paid to contextual (social and genre-related) variation; some voices that it is too problematic to be studied as a factor (Gries & Divjak 2010)
- lack of theoretically interpretable yet robust statistical techniques that can integrate different sources of variation

# Outline

1. Theoretical and methodological background
2. Dutch causative constructions
3. Data and method
4. Semasiological variation (polysemy)
  - semantic and contextual variation of *doen*
  - semantic and contextual variation of *laten*
5. Onomasiological variation (near-synonymy)
  - variation of *doen* and *laten*
  - contextual factors
6. Conclusions and future research



# Dutch Causative Constructions

*De tovenaar deed zijn dienaars een kasteel bouwen.  
liet*

The magician	made made/liet	his servants	a castle	build
<b>CAUSER</b>	<b>AUX</b>	<b>CAUSEE</b>	<b>AFFECTEE*</b>	<b>EFFECTED PREDICATE</b>

slot names: Kemmer & Verhagen (1994)

# *doen vs. laten*

<i>doen</i>	<i>laten</i>
Direct causation	Indirect causation
«The initiator produces the effected event directly; there is no intervening energy source 'downstream'»	«Some other force besides the initiator is the most immediate source of energy in the effected event»

Verhagen & Kemmer (1997)

# Other dimensions of variation

Speelman & Geeraerts 2009:

- *doen* is used more frequently in formal registers
- *doen* is used more frequently in Belgian Dutch
- *doen* is more prone to collocational fixation

“*doen* is an obsolescent form with a tendency towards semantic and lexical specialization”

# Outline

1. Theoretical and methodological background
2. Dutch causative constructions
3. **Data and method**
4. Semasiological variation (polysemy)
  - semantic and contextual variation of *doen*
  - semantic and contextual variation of *laten*
5. Onomasiological variation (near-synonymy)
  - variation of *doen* and *laten*
  - contextual factors
6. Conclusions and future research

# Data

- 5768 observations with causative *doen* and *laten*
- corpus
  - formal written genre: articles from Dutch and Belgian quality newspapers in TwNC and LeNC
  - informal written genre: posts from Usenet.nl and Usenet.be in Dutch, spam-checked
  - informal spoken genre: Dutch and Flemish spontaneous face-to-face conversations from CGN
- retrieval: XML parser (newspapers), regular expressions, manual cleaning-up (e.g. *laten we zeggen*)

# 37 Variables, 731 features

- **General context:** sentence type, clause type, clause mood, clause tense, Cx syntactic function
- **Causer:** semantic class, syntactic expression, POS, person, number, definiteness
- **Causee:** see Causer; intentionality, thematic proto-role
- **Affectee:** see Causer
- **Effected Predicate:** semantic class, semantic source and target domain, prefix, 'root', extended transitivity, prepositional complements
- **Other:** coreferentiality, possession, negation, modals, adverbials

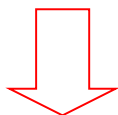
# Some corpus methods in CogSem

	Polysemy	Synonymy	Exemplars	Features
Bi-and multinomial Regression	predef. <sup>+</sup> senses	+	?	+
Collostructional Analysis	+	+	?	?
Behavioral Profiles	predef. <sup>+</sup> senses	+	?	+
<b>Exemplar Spaces</b>	+	+	+	+

# Exemplar Spaces Method

## Example of Data

	Var 1	Var 2	Var 3
Exemplar 1	A	yes	X
Exemplar 2	A	yes	Y
Exemplar 3	B	no	X

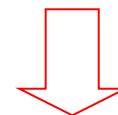


## Distances (Gower)

	Ex. 1	Ex. 2	Ex.3
Exemplar 1	0	0.33	0.66
Exemplar 2	0.33	0	1
Exemplar 3	0.66	1	0

## Partitioning into k Clusters

	Cluster
Exemplar 1	1
Exemplar 2	1
Exemplar 3	2



## Interpretation: Cluster profiles

Cluster	A	B	yes	no	X	Y
Cluster 1	1	0	1	0	0.5	0.5
Cluster 2	0	1	0	1	1	0

distinctive features

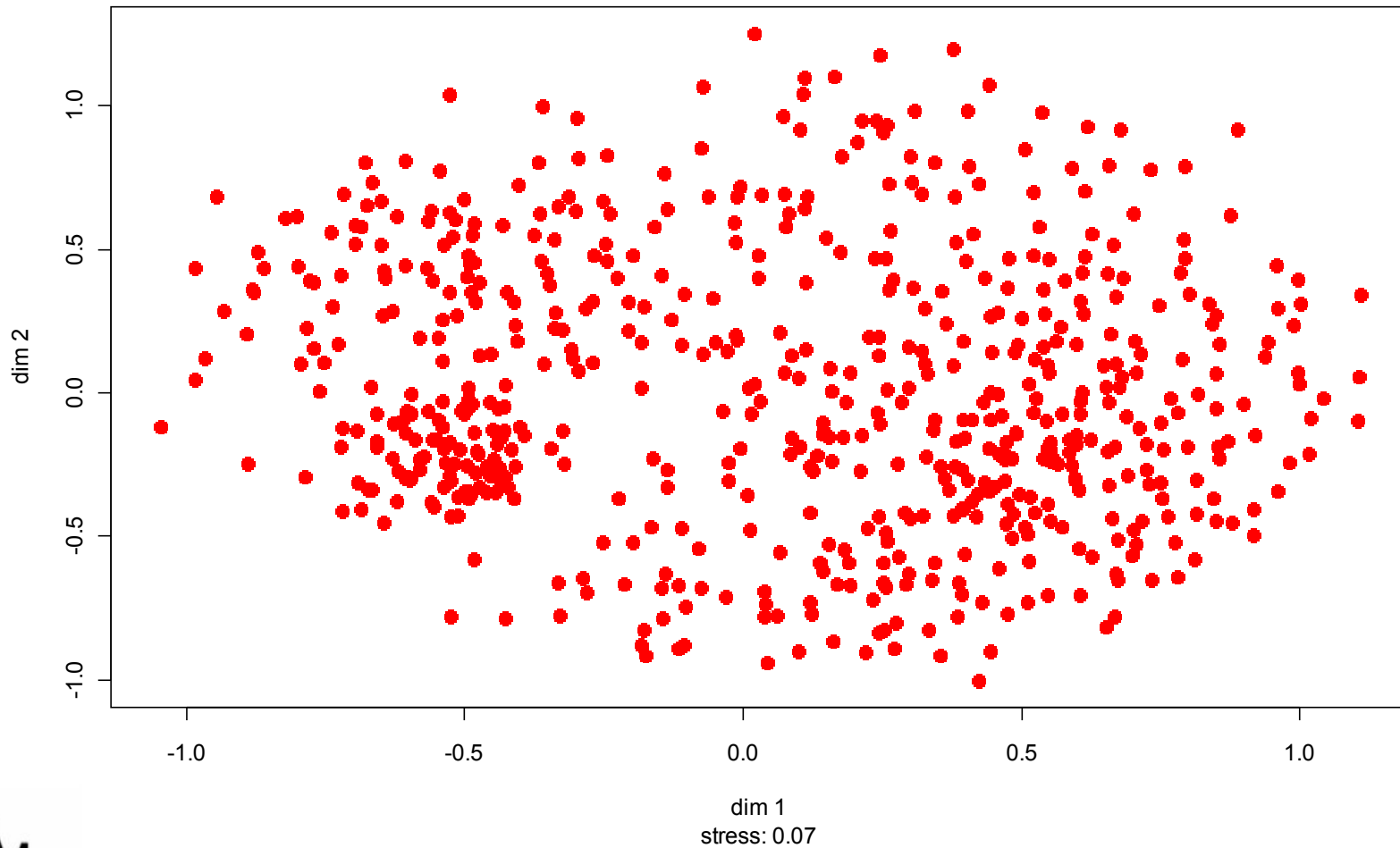


# Outline

1. Theoretical and methodological background
2. Dutch causative constructions
3. Data and method
4. Semasiological variation (polysemy)
  - semantic and contextual variation of *doen*
  - semantic and contextual variation of *laten*
5. Onomasiological variation (near-synonymy)
  - variation of *doen* and *laten*
  - contextual factors
6. Conclusions and future research

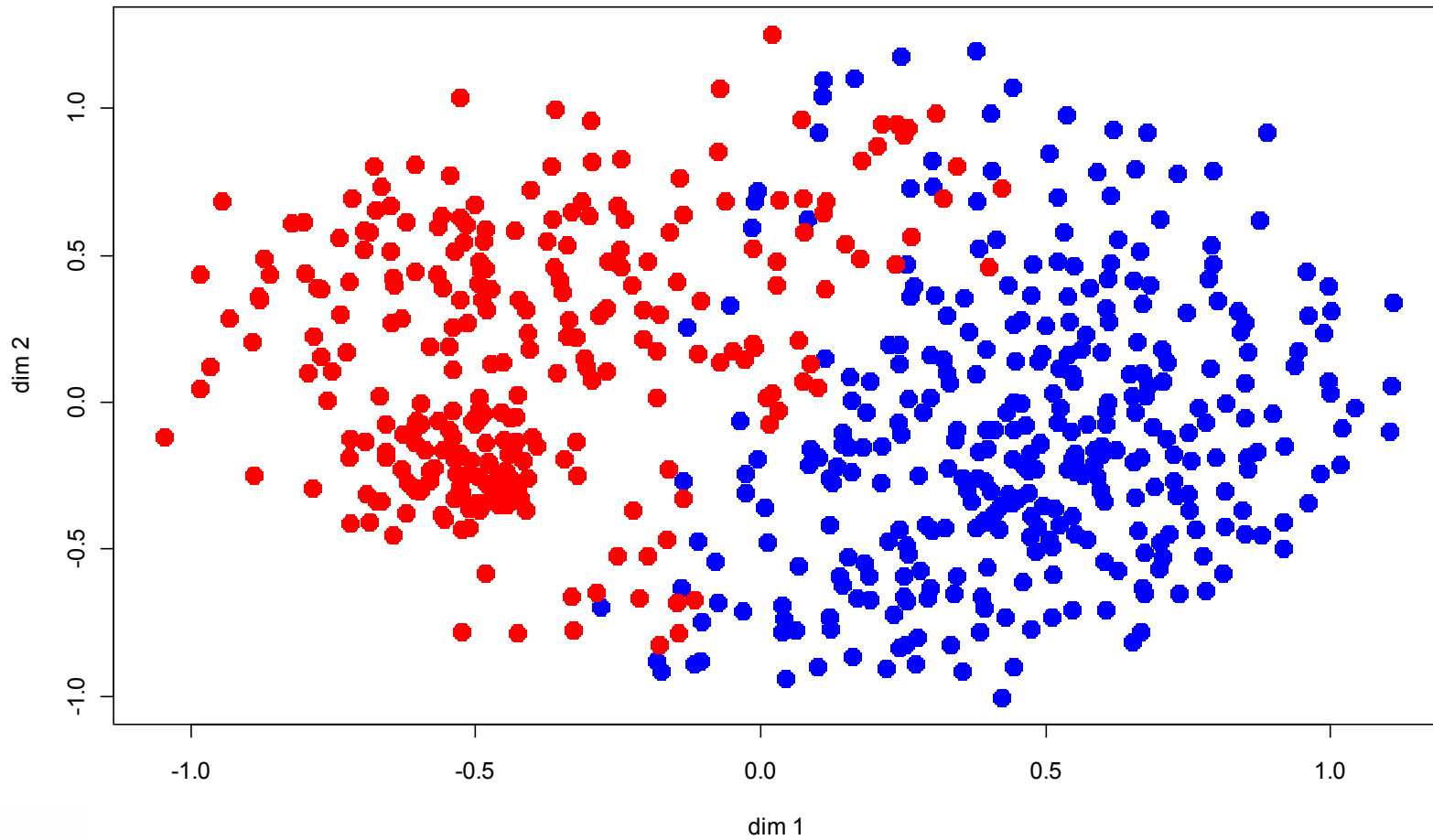
# Exemplar Space of *doen*

SMACOF MDS, *doen*



# 2 Clusters of *doen*

2 clusters of doen



# Clusters of *doen*: Top distinctive features

<b>Red Cluster</b>	<b>Blue Cluster</b>
Psych. Caused Event (fig.)	3 <sup>rd</sup> Person Causee
Psych. Caused Event (lit)	Social Caused Event (fig.)
Human Causee	Nominal Causee
Pronominal Causee	Social Caused Event (lit.)
1 <sup>st</sup> Person Causee	Causee undergoes change
No change occurs	No prepositional complements
Verb class Mental Processes	Explicit Causee
Root <i>denk</i> -	Explicit NP Affectee
Preposition <i>aan</i>	Verb class Motion
Implicit Causee	Indefinite NP Affectee

# Causative *doen*: A dictionary entry

1

Cause a change in the world, especially a social change conceptualized as an abstract objective event.

*Betere booklets zijn er geregeld (maar dat doet de prijs stijgen).*

2

Trigger a psychological reaction.

*Dacht ook dat ik het wist, tot je vraag me deed twijfelen, en terecht...*

*doen denken aan* “remind of”

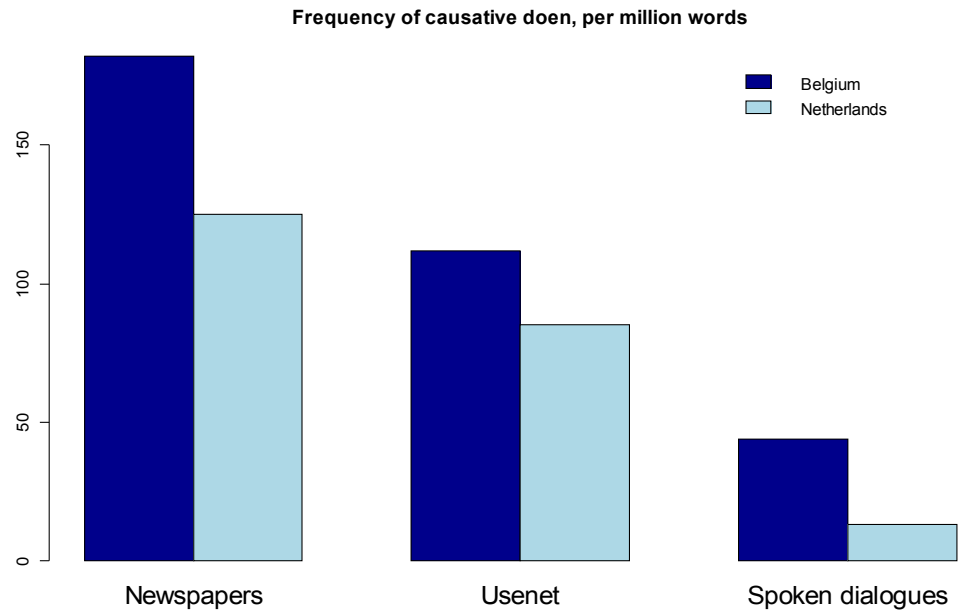
*"Groovin' on E" doet met zijn fris jengelende gitaren en neuzelende fluisterzang denken aan The Kinks.*

# Contextual variation of *doen*

*doen* is more frequent

- in Belgium
- in formal genres

(Chi-test,  $\alpha = 0.05$ )



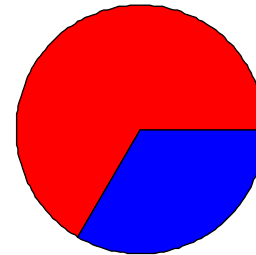
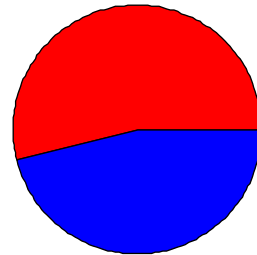
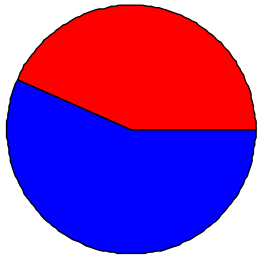
# *doen*: Interaction Context\*Semantics

newspapers

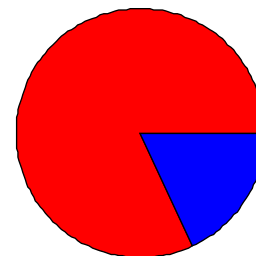
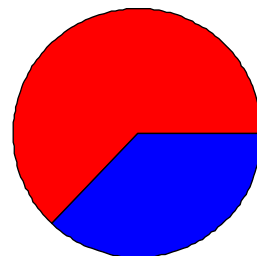
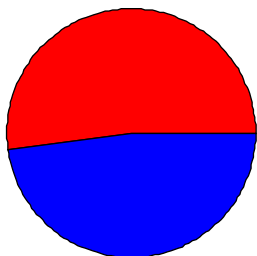
Usenet

dialogues

Belgium

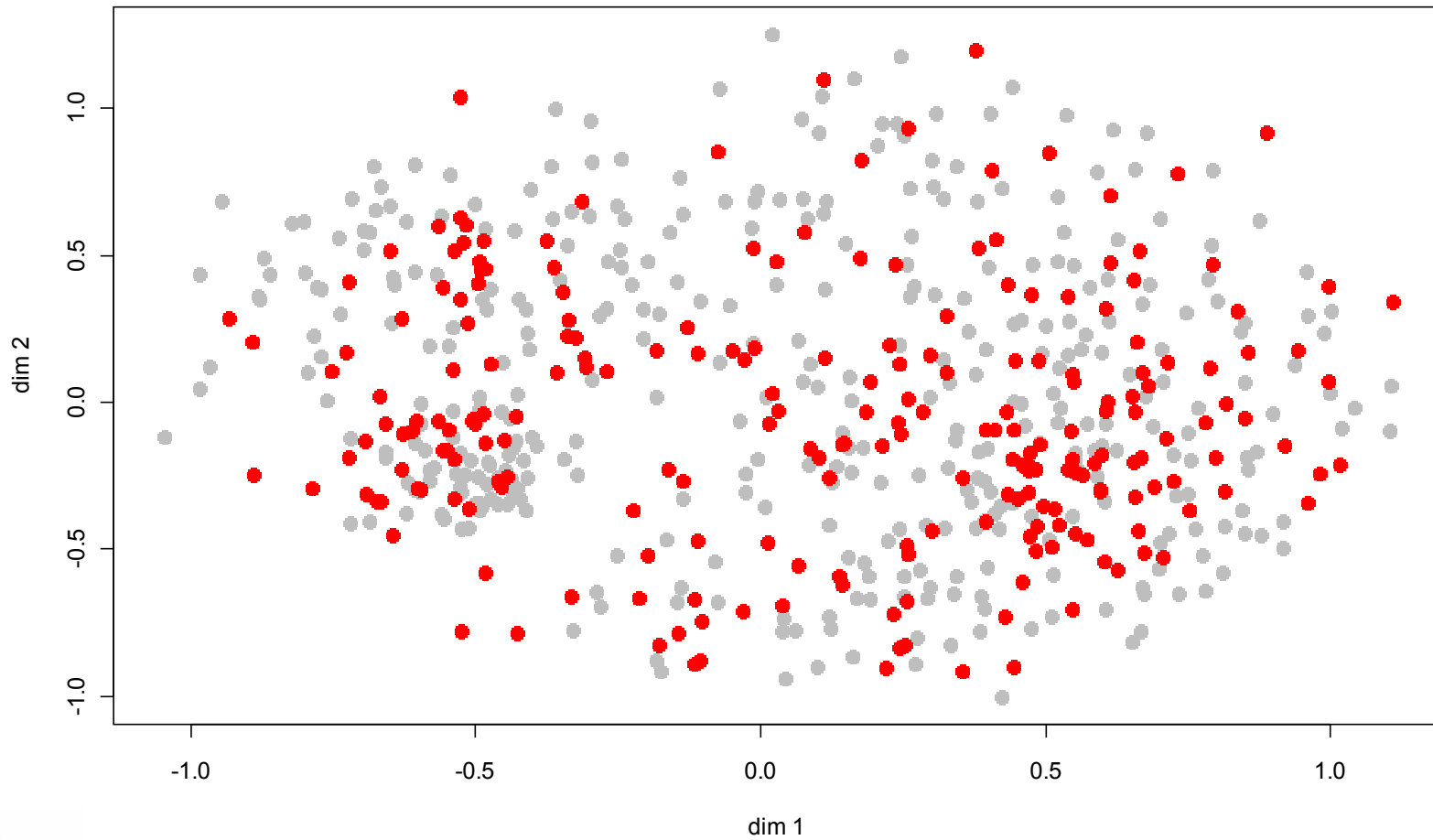


Netherlands



# Belgium, newspapers

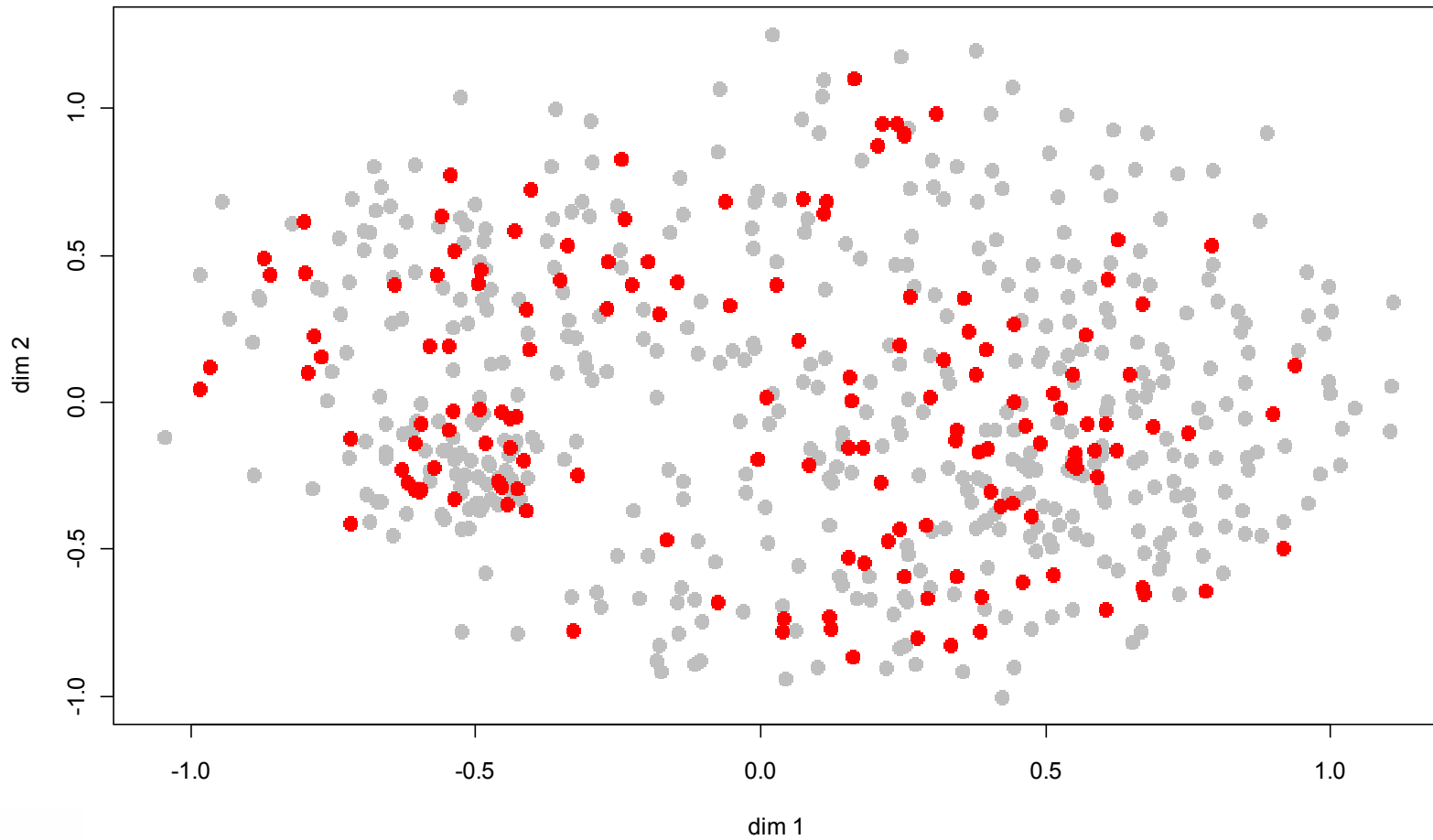
doen, Flemish newspapers





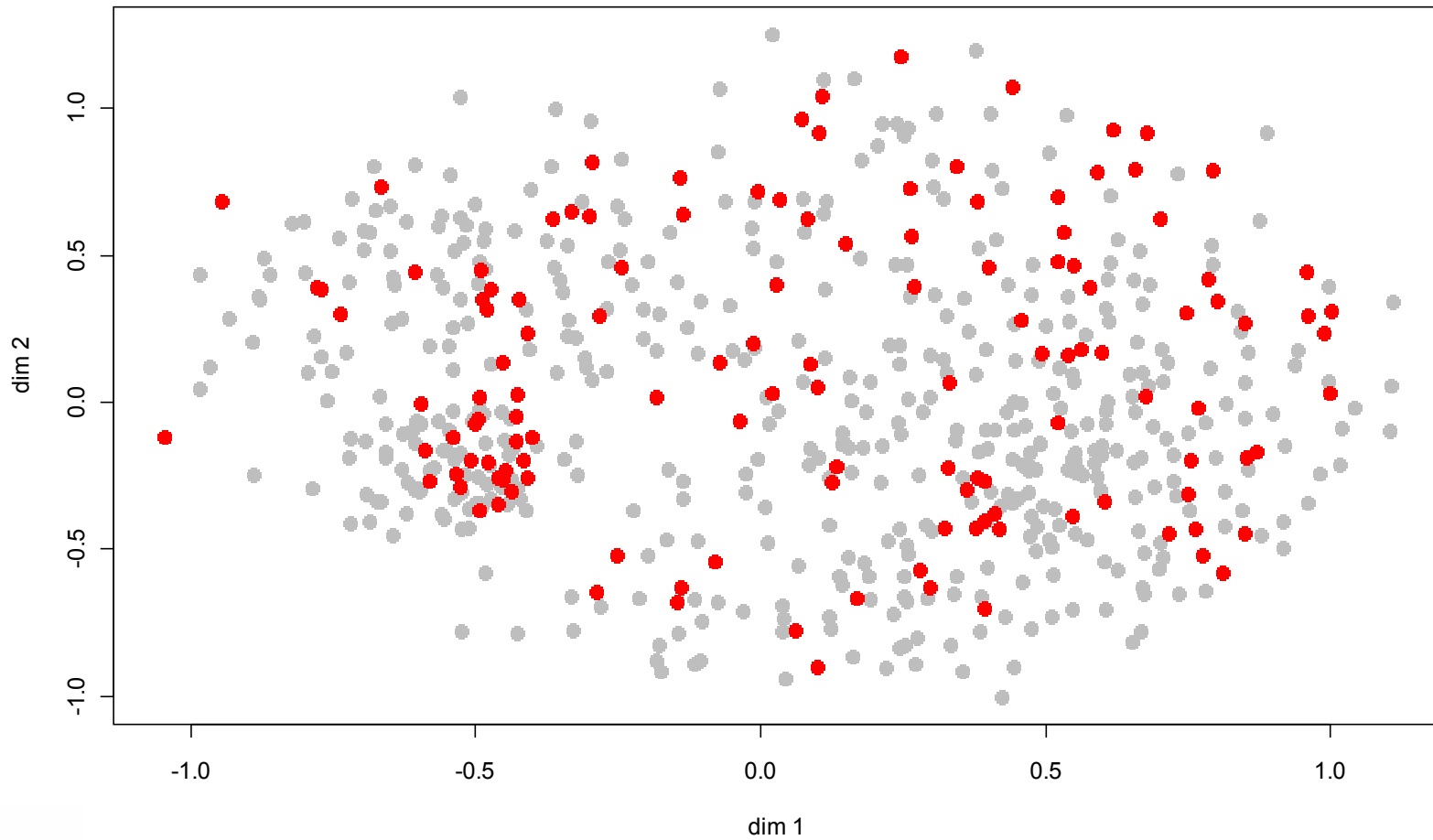
# The Netherlands, newspapers

doen, Dutch newspapers



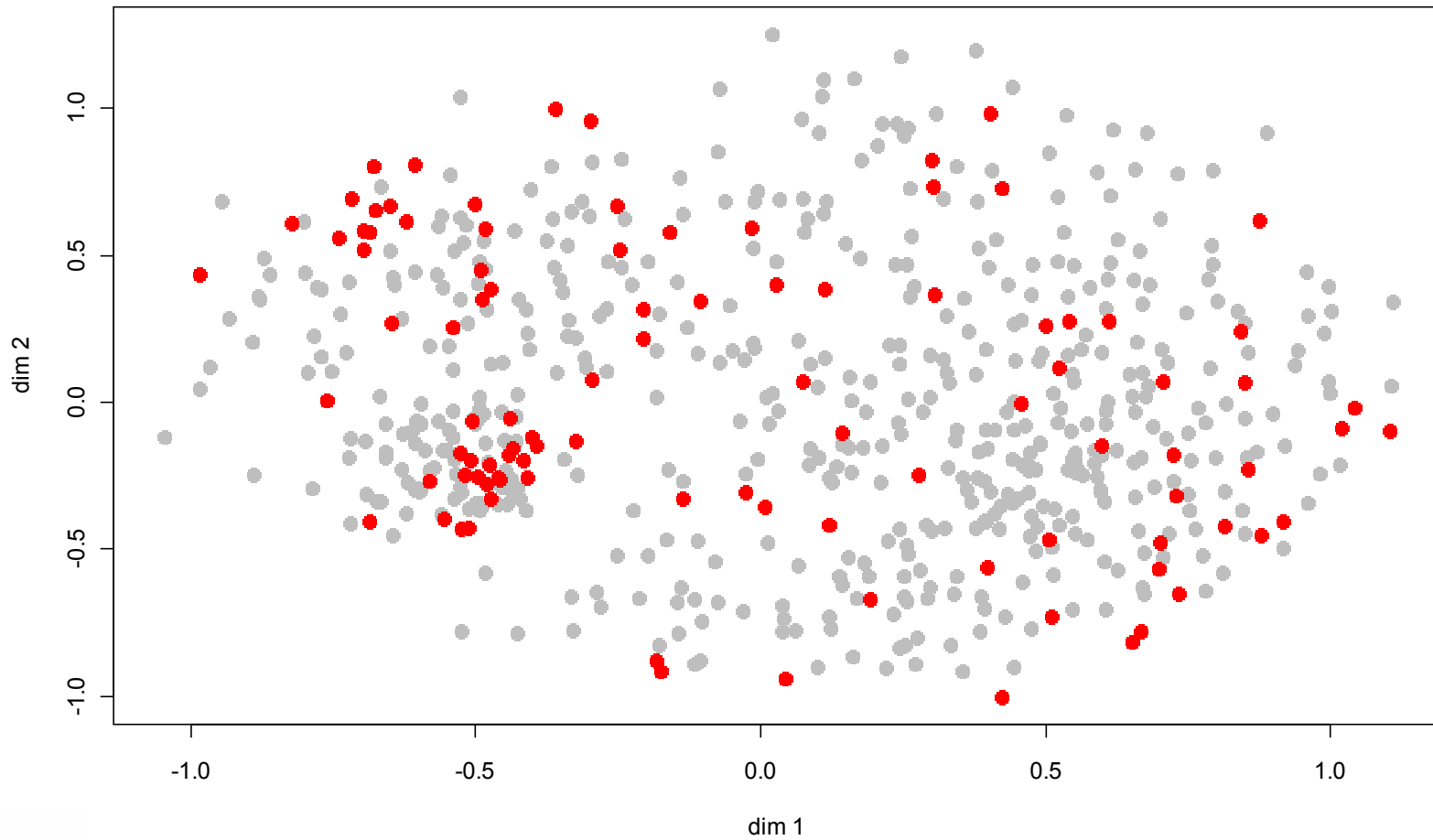
# Belgium, Usenet

doen, Flemish Usenet



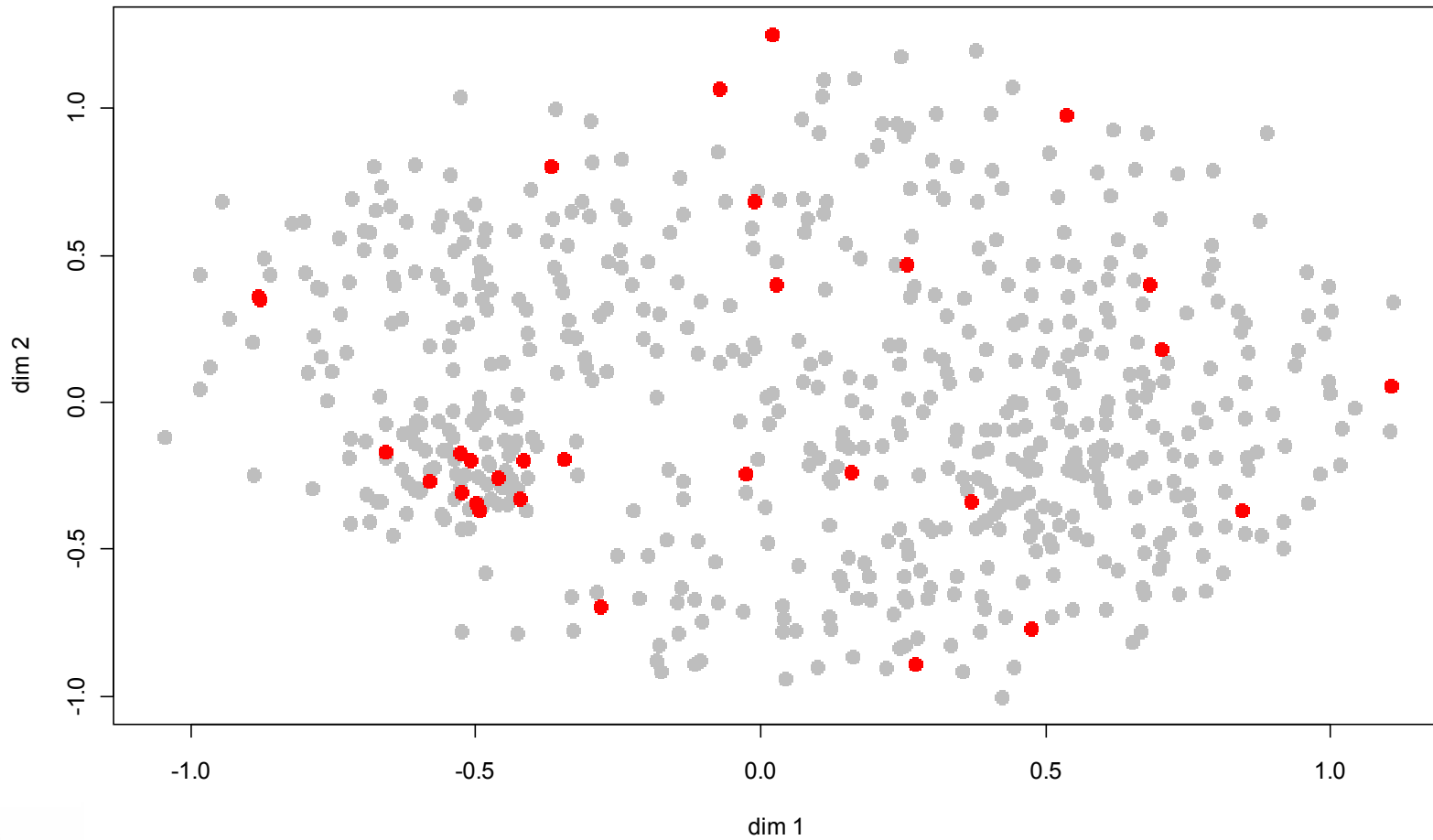
# The Netherlands, Usenet

doen, Dutch Usenet



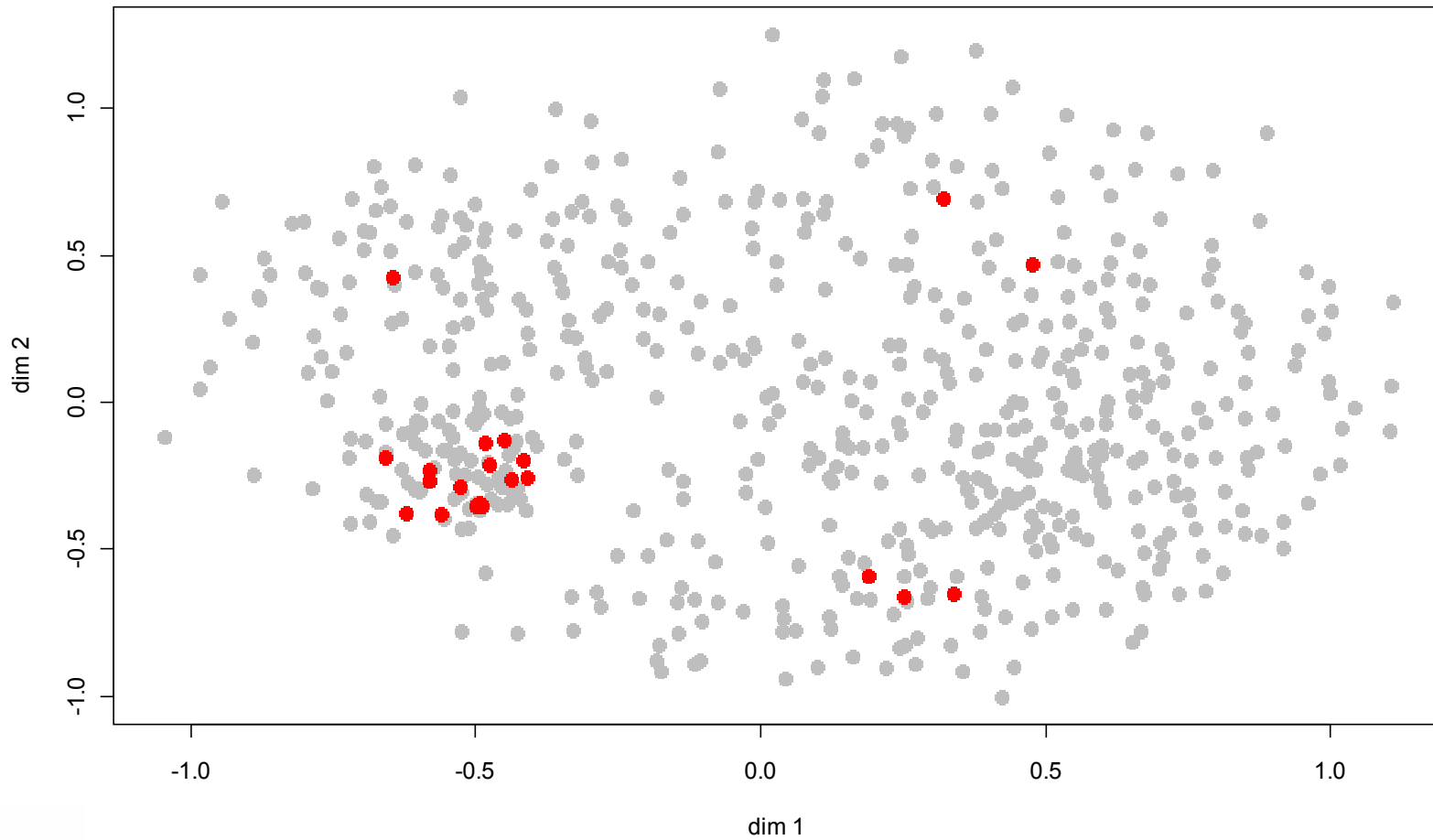
# Belgium, conversations

doen, Flemish conversations



# The Netherlands, conversations

doen, Dutch conversations



# doen: Lectally enriched entry

1

*esp. Belg. and formal* Cause a change in the world, especially a social change conceptualized as an abstract objective event.

*Betere booklets zijn er geregeld (maar dat doet de prijs stijgen).*

2

Trigger a psychological reaction.

*Dacht ook dat ik het wist, tot je vraag me deed twifelen, en terecht...*

doen denken aan “remind of”

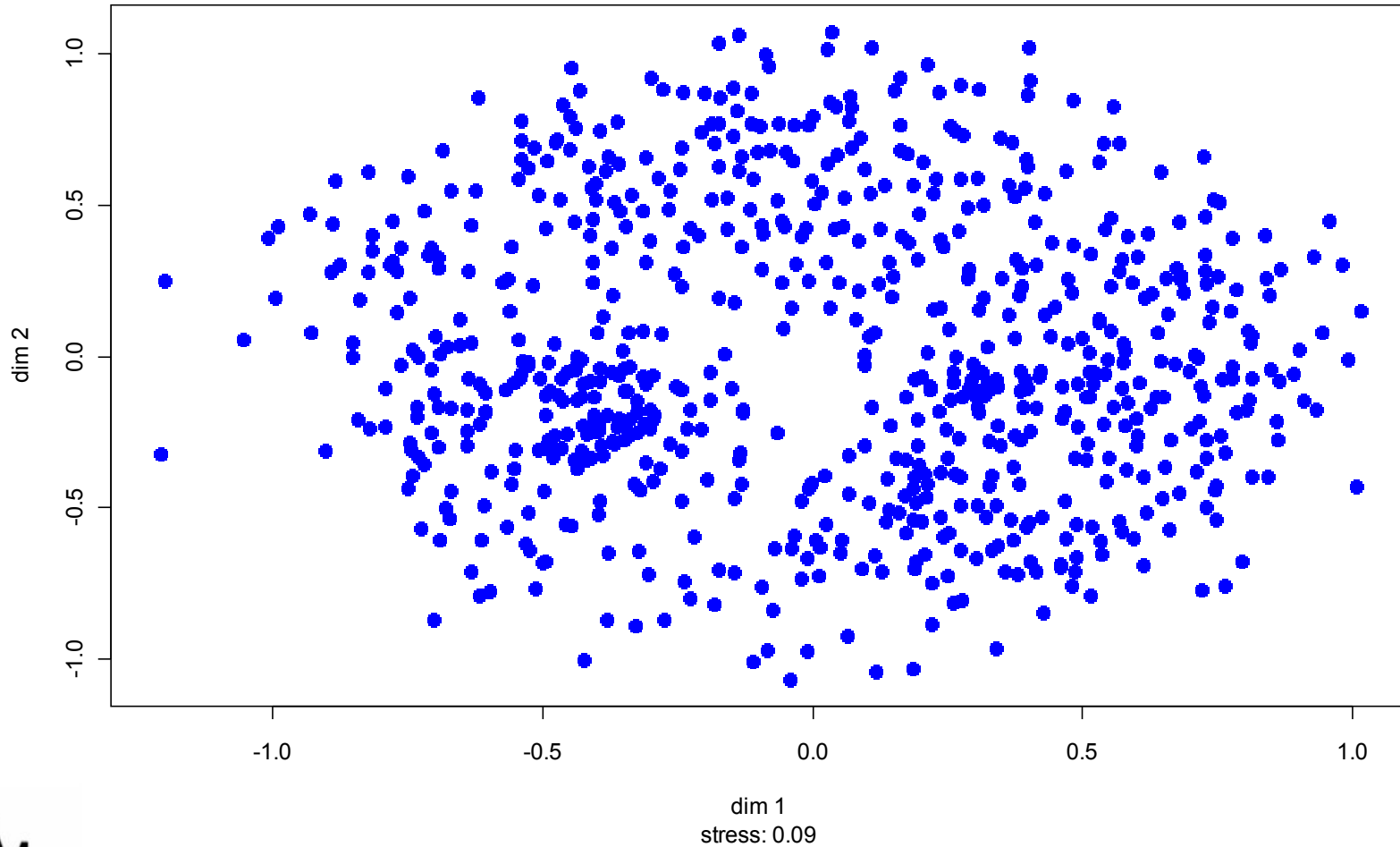
*" Groovin' on E " doet met zijn fris jengelende gitaren en neuzelende fluisterzang denken aan The Kinks.*

# Outline

1. Theoretical and methodological background
2. Dutch causative constructions
3. Data and method
4. Semasiological variation (polysemy)
  - semantic and contextual variation of *doen*
  - semantic and contextual variation of *laten*
5. Onomasiological variation (near-synonymy)
  - variation of *doen* and *laten*
  - contextual factors
6. Conclusions and future research

# Exemplar Space of *laten*

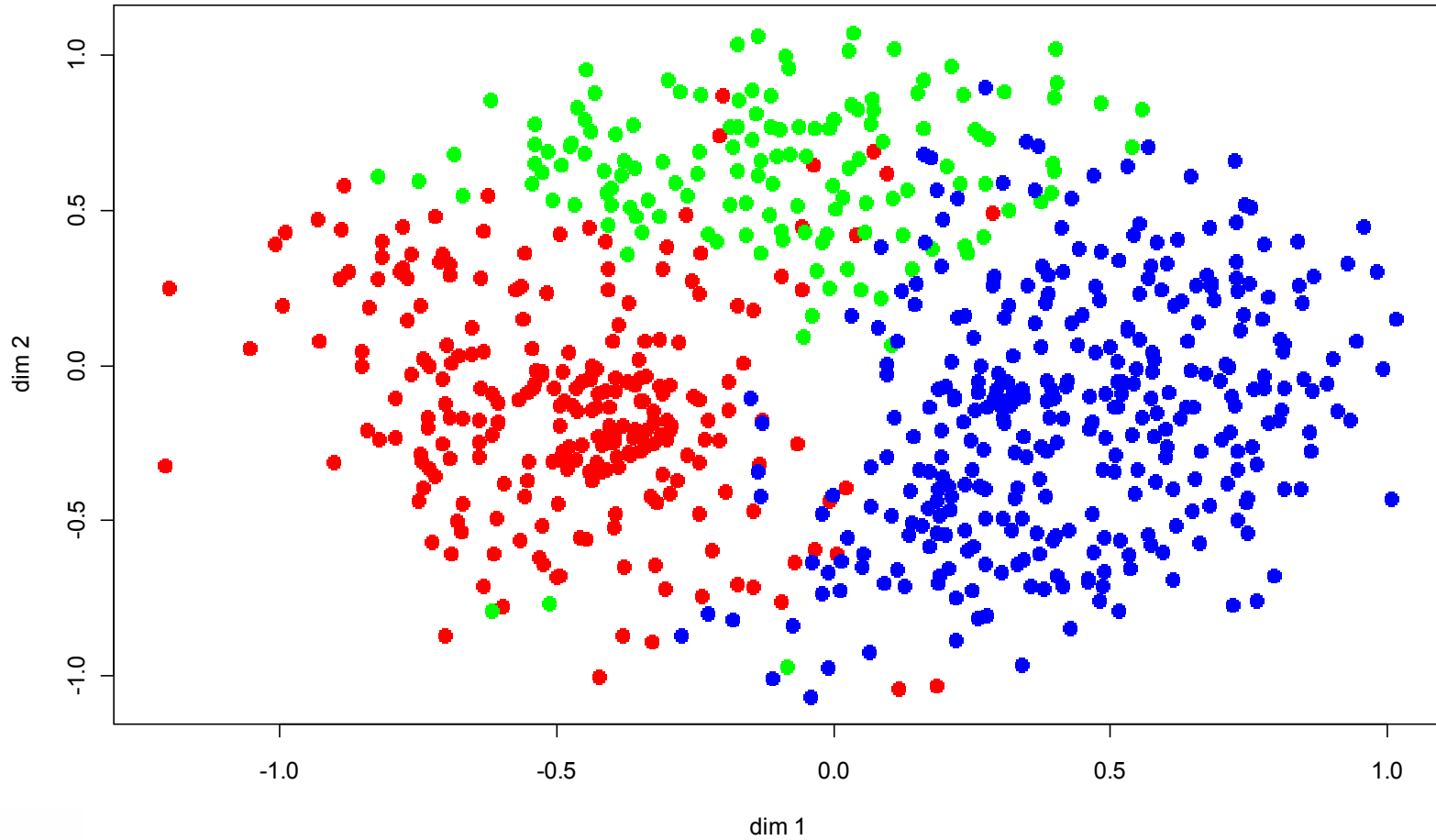
SMACOF MDS, *laten*





# 3 Clusters of *laten*

SMACOF MDS, *laten*



# Clusters of *laten*: Top distinctive features

<b>Red Cluster</b>	<b>Blue Cluster</b>	<b>Green Cluster</b>
Psych. Caused Event (lit)	Explicit NP Causee	Human Undefined Causee
Psych. Caused Event (fig)	Intransitive EP	Intentional Causee
Human Causee	V. Motion	Causee causes change
V. Perception	Abstract Causee	V Creation
Root <i>zie-</i>	Causee is changed	V. Phys Manipulation
Pronominal Causee	Indefinite NP Affectee	V. Social Interaction
V. Factive	V. Existence	V. Change of Possession
Root <i>weet-</i>	Plural Affectee	Material Object Affectee
Affectee Clause	No coreferentiality	Sing. Number Causee
1 <sup>st</sup> Person Causee	Nominal Affectee	Explicit NP Affectee

# Causative *laten*: An entry

1 Stop or fail to carry out impingement, abandon, leave.

*Ook Rombouts laat Van Hecke vallen.*

2 Suffer or benefit from someone's actions, e.g. services.

*En als ik het geschreven heb ga ik het eerst laten registreren bij ISBN .*

*Ref Allaerts laat zich vangen.*

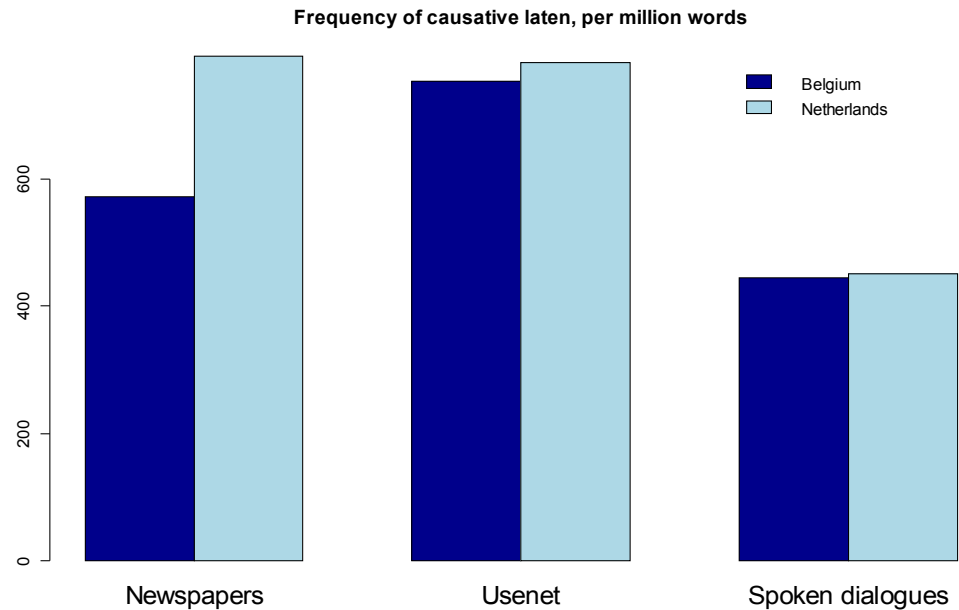
3 Provide information, show.

*Nou dat heeft ie laten zien.*

# Contextual variation of *laten*

## Frequency of *laten*

- NL > BE newspapers
- overall written > spoken  
(Chi-test,  $\alpha = 0.05$ )



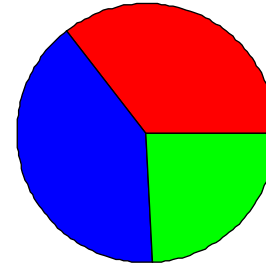
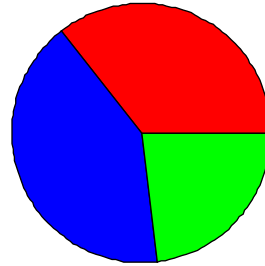
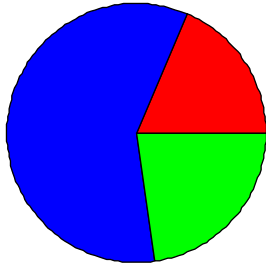
# *laten*: Interaction Context\*Semantics

newspapers

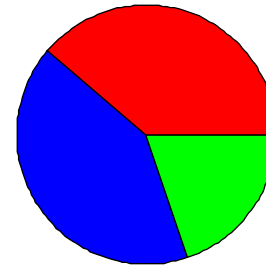
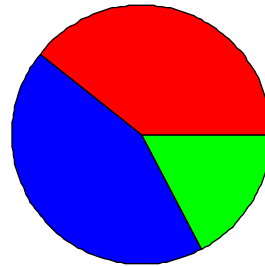
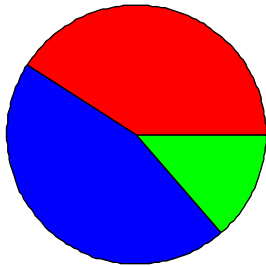
Usenet

dialogues

Belgium

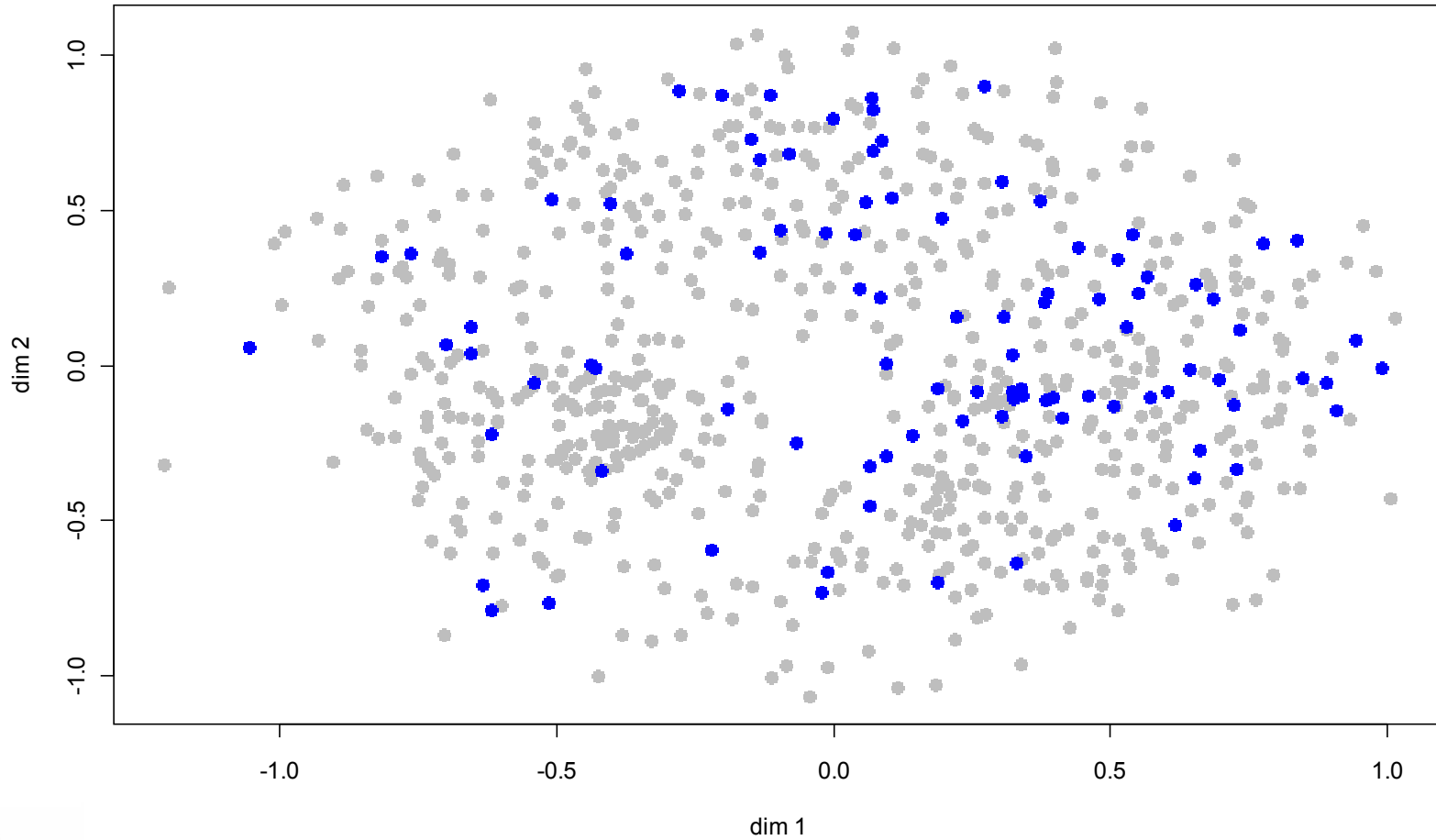


Netherlands



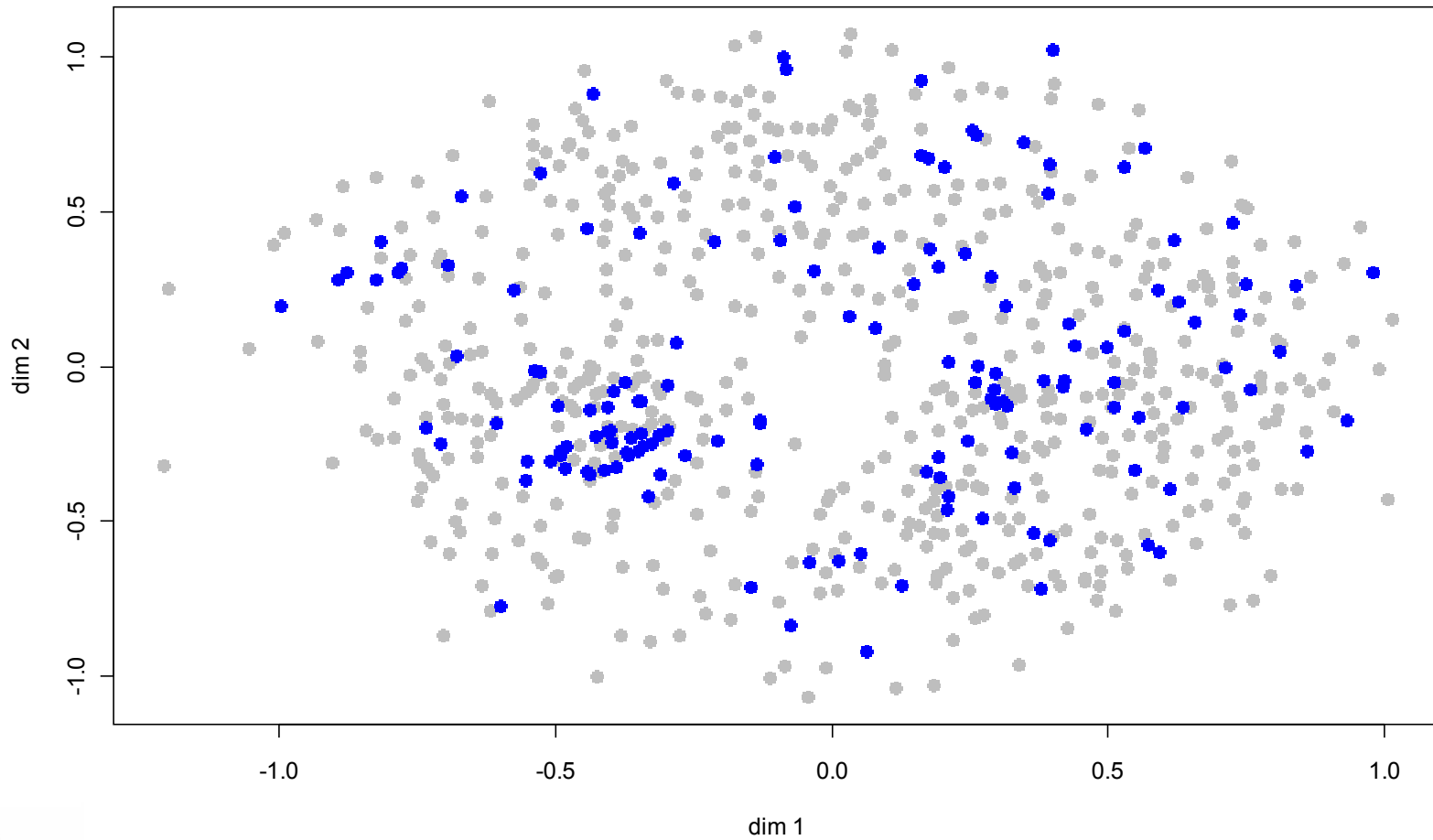
# Belgium, newspapers

laten, Flemish newspapers



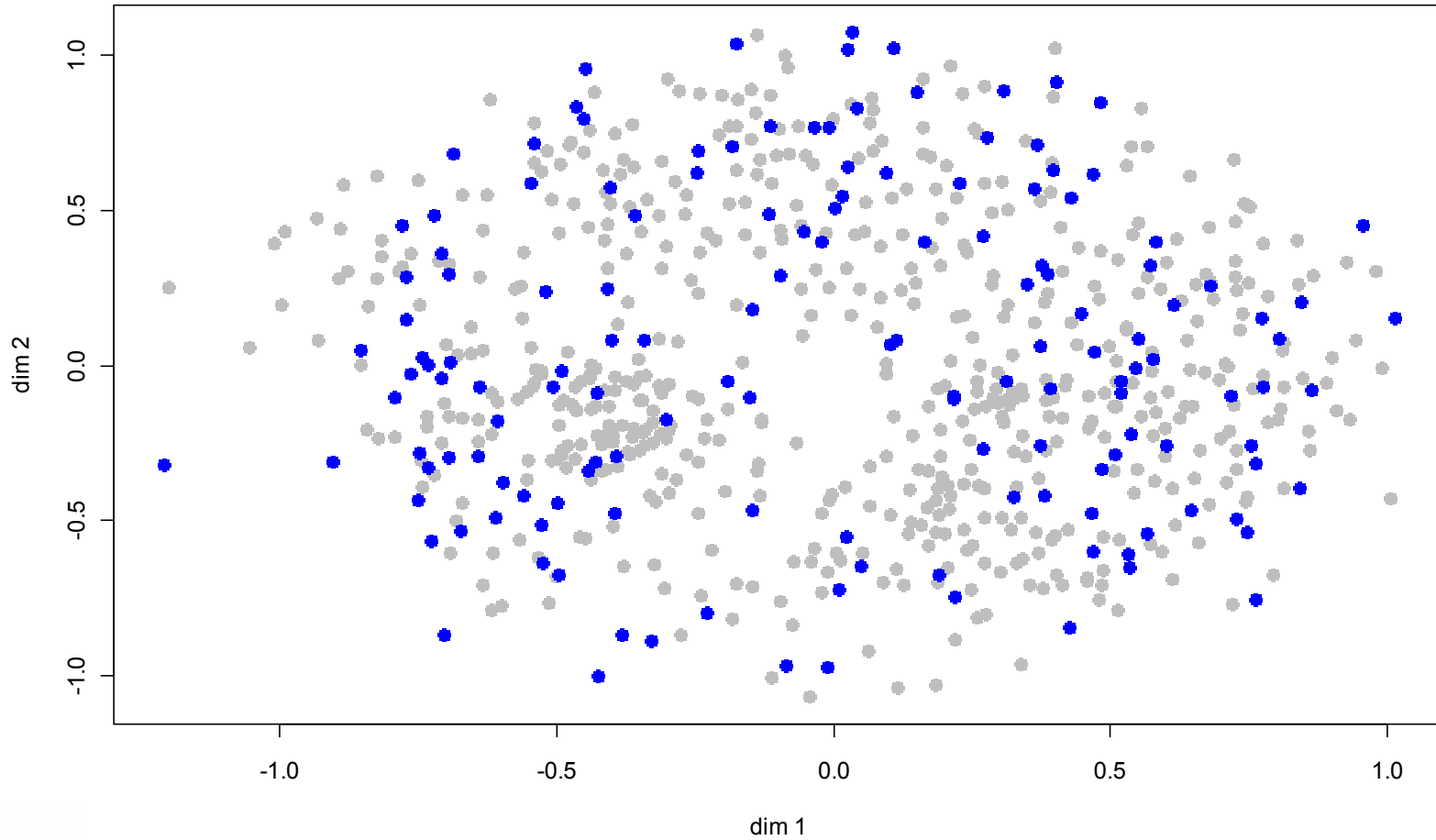
# The Netherlands, newspapers

laten, Dutch newspapers



# Belgium, Usenet

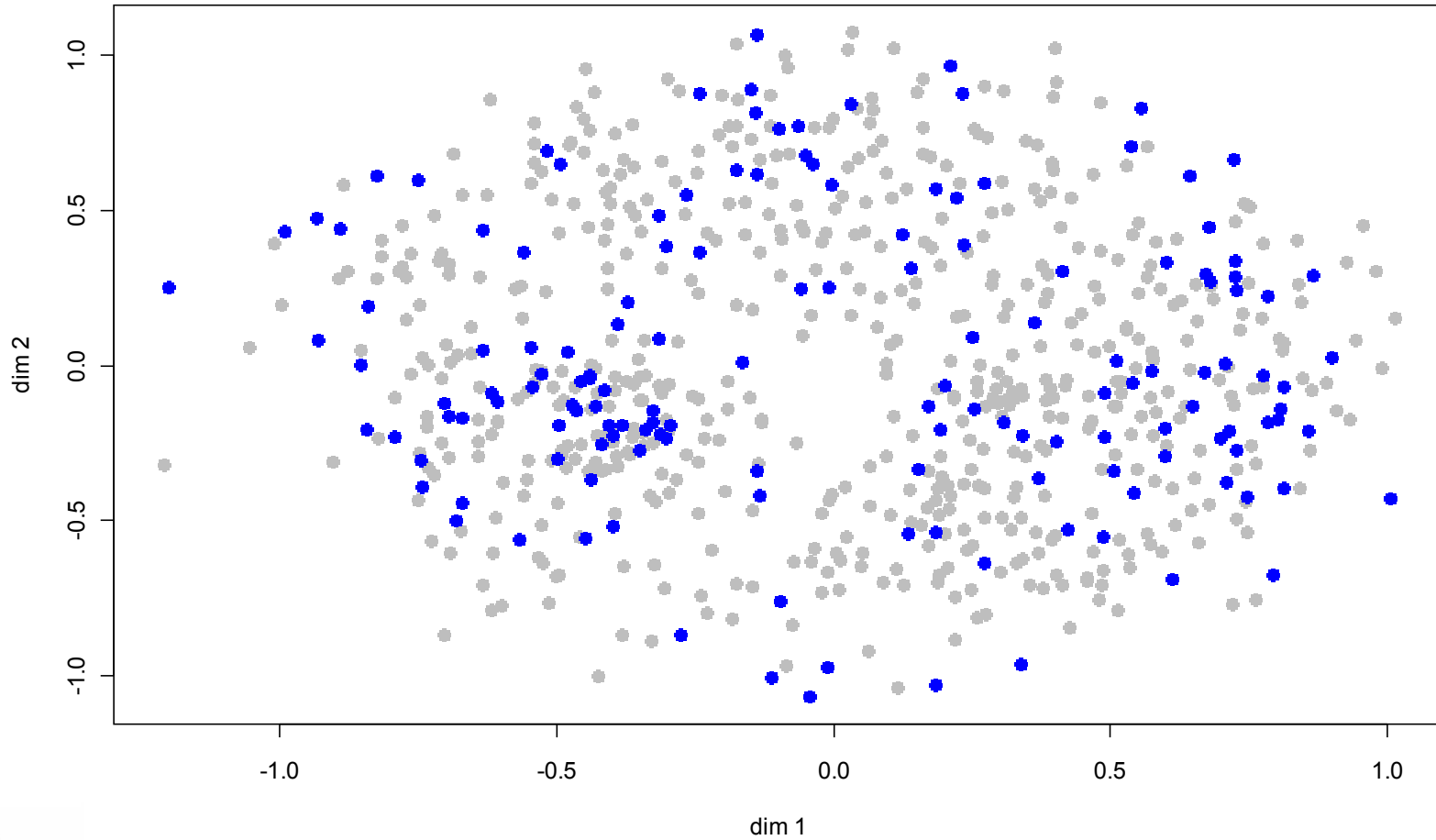
laten, Flemish Usenet





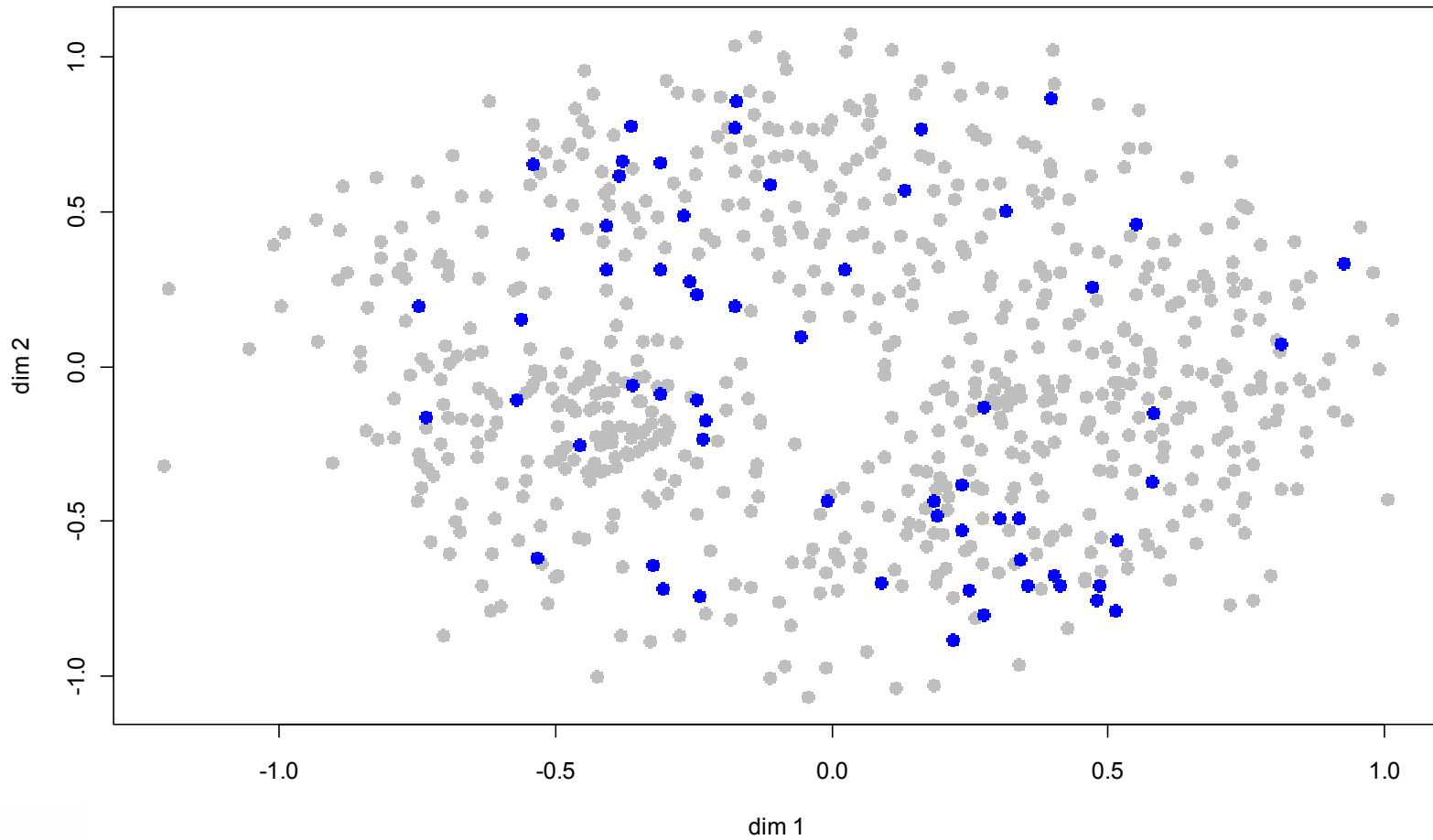
# The Netherlands, Usenet

laten, Dutch Usenet



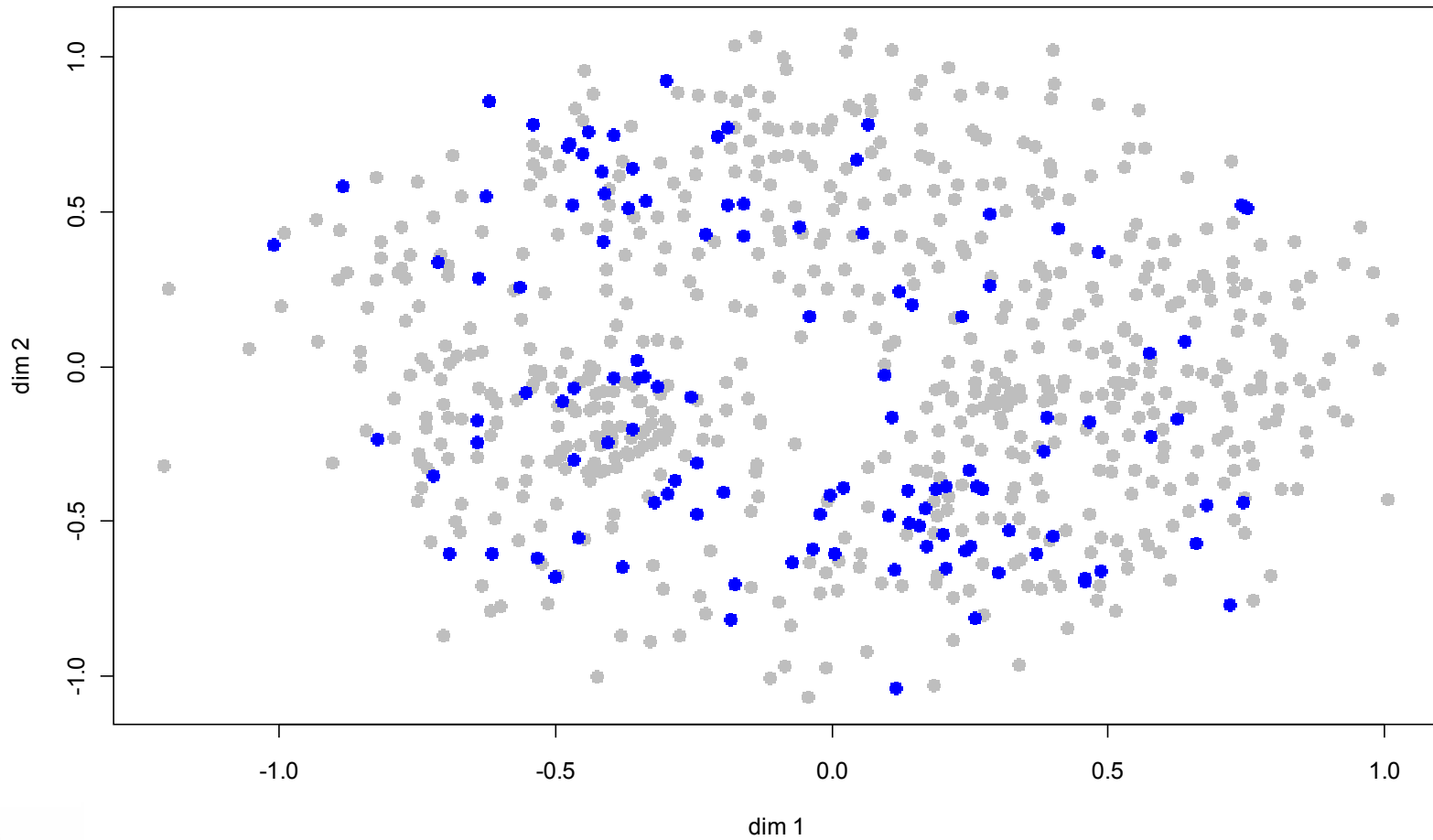
# Belgium, conversations

laten, Flemish conversations



# The Netherlands, conversations

laten, Dutch conversations



## *laten*: Lectally enriched entry

1 Stop or fail to carry out impingement, abandon, leave.

*Ook Rombouts laat Van Hecke vallen.*

2 Suffer or benefit from someone's actions, e.g. services.

*En als ik het geschreven heb ga ik het eerst laten registreren bij ISBN .*

*Ref Allaerts laat zich vangen.*

3 *esp. Neth. Dutch* Provide information, show.

*Nou dat heeft ie laten zien.*

# Outline

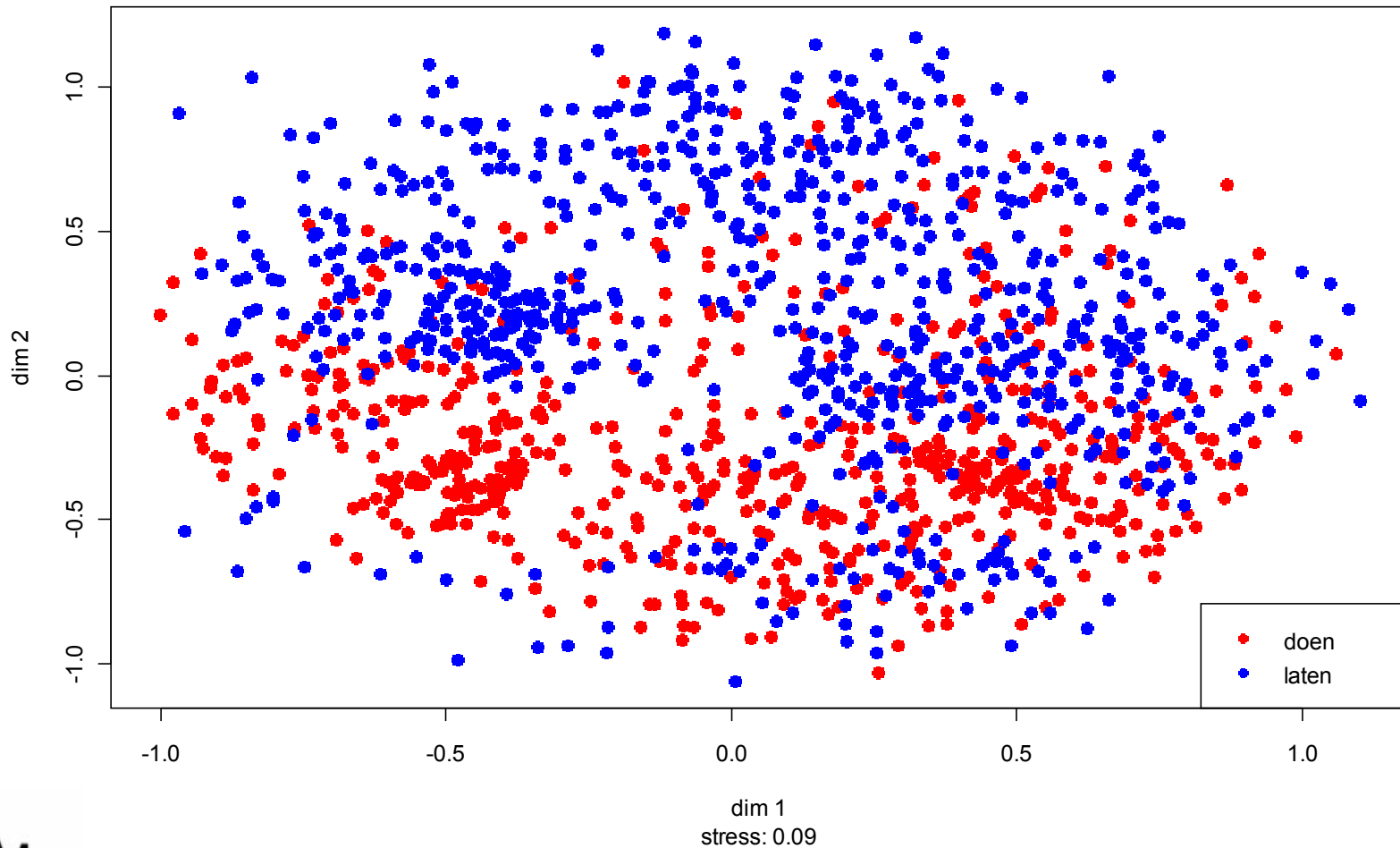
1. Theoretical and methodological background
2. Dutch causative constructions
3. Data and method
4. Semasiological variation (polysemy)
  - semantic and contextual variation of doen
  - semantic and contextual variation of laten
5. Onomasiological variation (near-synonymy)
  - variation of doen and laten
  - contextual factors
6. Conclusions and future research

# A Challenge

“Het is moeilijk een scherp onderscheid te maken tussen het gebruik van *doen* en dat van *laten*, al zijn er wel enkele tendensen aan te wijzen. Bij de keuze voor één van beide werkwoorden kunnen niet alleen betekenisverschillen, maar ook geografische en stilistische verschillen een rol spelen. In ieder geval zijn *doen* en *laten* niet altijd zonder meer uitwisselbaar.” ANS 1997

# doen vs laten: Onomasiological Variation

SMACOF MDS, doen and laten



# doen vs laten: Top distinctive features

<b>doen</b>	<b>laten</b>
abstract Causer	human Causer
intransitive predicates	transitive predicates
3rd person Causer	pronominal Affectee
V. denoting Mental Processes	no prepositional complements
root <i>denk-</i>	intentional Causee
nominal Affectee	3rd person Causee
preposition <i>aan</i>	explicit NP Affectee
explicit NP Causee	V. of Perception
abstract Affectee	implicit Causee
1st person Causee	causing (agentive) Causee



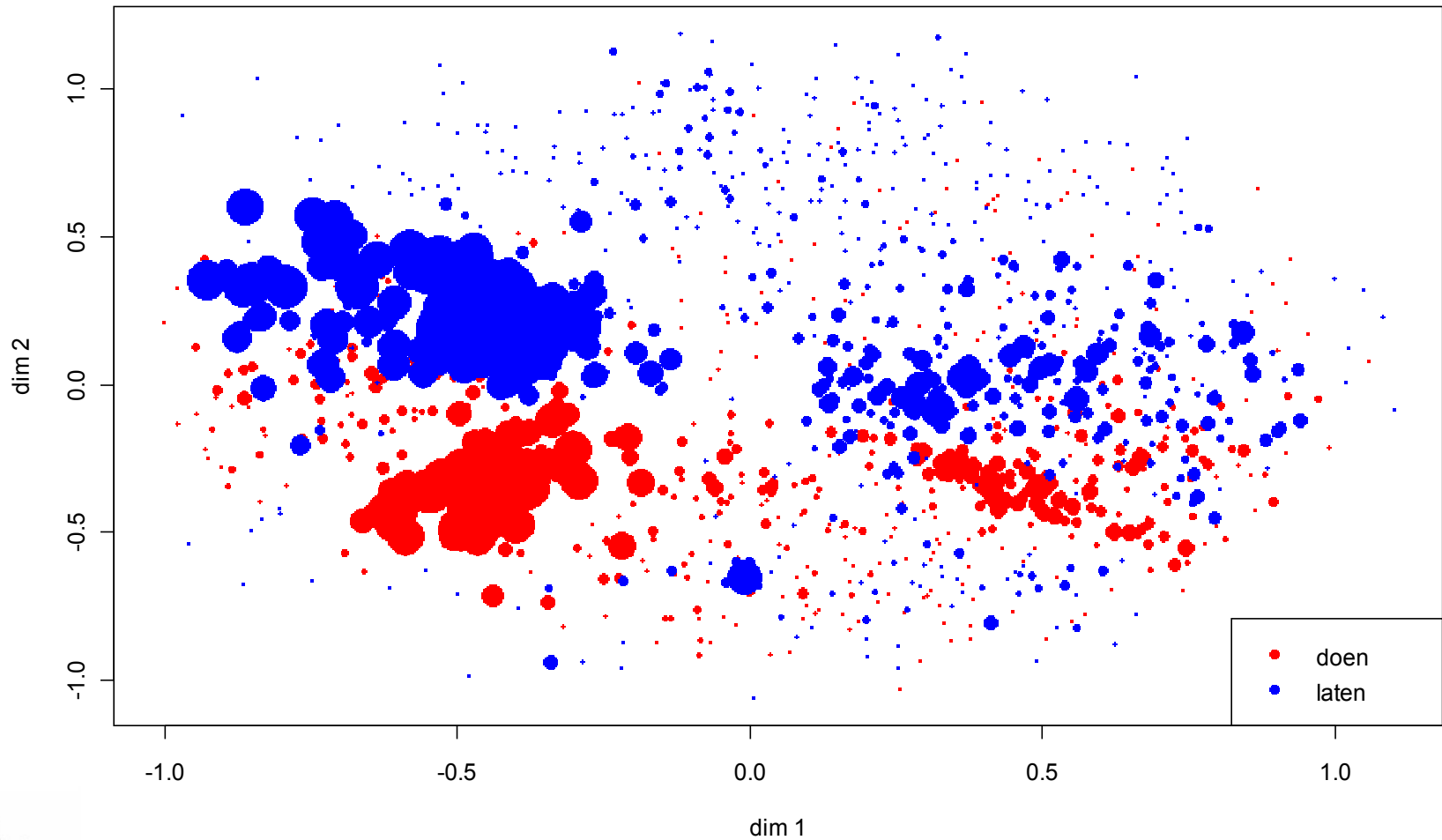
# Confirmatory Regression: Design

- Response: *doen* or *laten*
- General semantic predictors:
  - human/abstract/other Causer
  - transitive/intransitive effected predicate
  - Causee causes or undergoes a change
  - intentional or unintentional Causee
  - implicit or explicit Causee
- Contextual variables: Country and Genre
- Exemplar effects?

# Modelling Exemplar Effects

- Speelman & Geeraerts (2009): lexical fixation operationalized with the help of colostruational analysis (Stefanowitsch & Gries 2003)
- Bybee (2010): rare collexemes can be semantically similar to frequent collexemes and thus typical of Cx; see also Schmid (2010)
- solution (cf. exemplar model by Nosofsky 1986):  
    similarity to neighbours =  $e^{-d^2}$   
    where d is distance between exemplars of the same category (*doen* or *laten*)

# proxDoen and proxLaten



# Main Effects

	Estimate	Std. Error	Pr(> z )
(Intercept)	-4.02	0.31	< 2e-16 ***
<b>CrSem=Abstr</b>	<b>1.86</b>	<b>0.17</b>	<b>&lt; 2e-16 ***</b>
CrSem=Oth	2.1	0.24	< 2e-16 ***
<b>proxDoen</b>	<b>0.06</b>	<b>0.004</b>	<b>&lt; 2e-16 ***</b>
<b>proxLaten</b>	<b>-0.01</b>	<b>0.001</b>	<b>&lt; 2e-16 ***</b>
<b>Genre=spont_dial</b>	<b>-0.91</b>	<b>0.24</b>	<b>&lt;0.001***</b>
<b>Genre=Usenet</b>	<b>-0.16</b>	<b>0.13</b>	<b>0.23</b>
<b>CeEnergy=Change</b>	<b>1.16</b>	<b>0.32</b>	<b>&lt;0.001***</b>
CeEnergy=Oth	1.26	0.29	<0.001***
<b>CeIntent=No</b>	<b>0.62</b>	<b>0.18</b>	<b>&lt;0.001***</b>
CeIntent=Undef	-0.04	0.24	0.85
<b>Country=BE</b>	<b>0.33</b>	<b>0.12</b>	<b>0.008**</b>
<b>CeSynt=Expl</b>	<b>0.37</b>	<b>0.17</b>	<b>0.029*</b>
EPTrans = Intr	0.01	0.17	0.964

full model with interactions: C=0.96 Gamma=0.91 pseudo-R<sup>2</sup>=0.68



# Interpretation of the Model

- *doen* involves direct causation, *laten* indirect?
  - abstract Causers (*doen*) can be only direct 'causes', whereas human Causers (*laten*) can affect the situation indirectly
  - implicit Causees (*laten*) are often intermediaries
  - agent-like (intentionally causing) Causees (*laten*) allow the Causer to influence the world indirectly
- *doen* is more popular in formal genres and Belgium

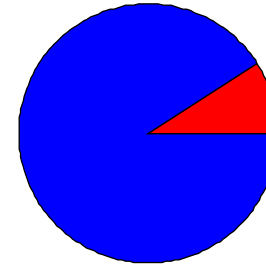
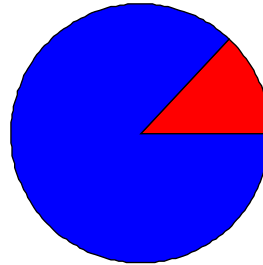
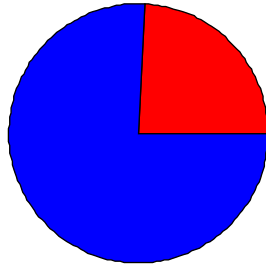
# Interaction Country\*Genre

newspapers

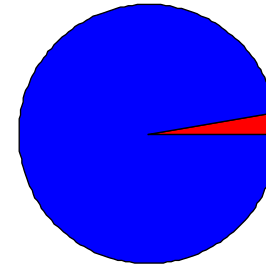
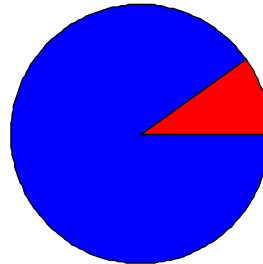
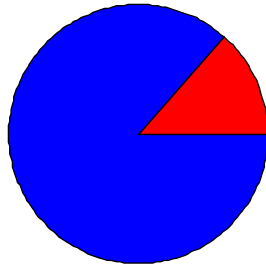
Usenet

dialogues

Belgium



Netherlands

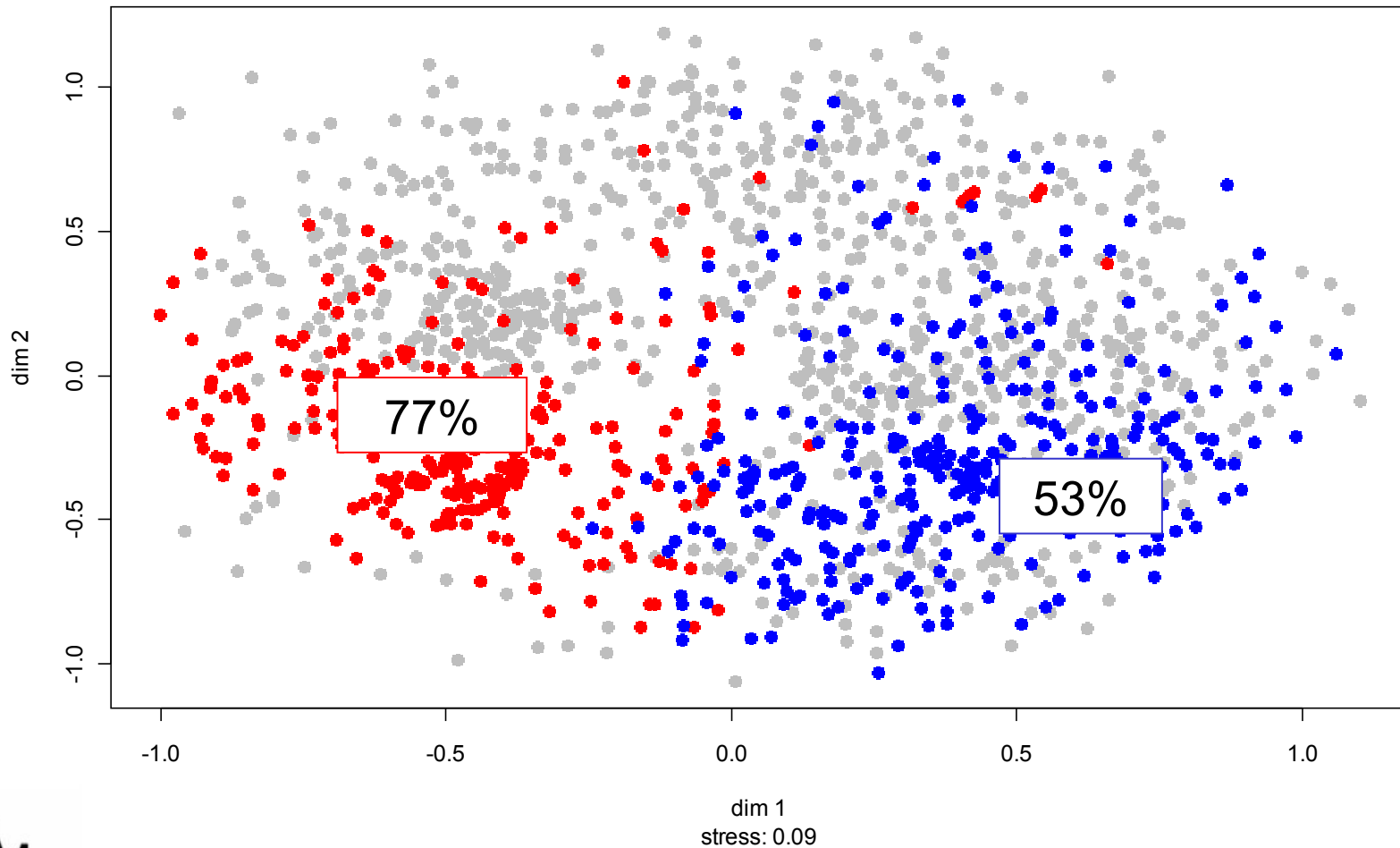


doen  
laten



# Average probability of *doen*

clusters of *doen* in common space



# doen: Onomasiologically enriched entry

1

*esp. Belg. and formal* Cause a change in the world, especially a social change conceptualized as an abstract objective event. **Syn. laten**

*Betere booklets zijn er geregeld (maar dat doet de prijs stijgen).*

2

Trigger a psychological reaction.

*Dacht ook dat ik het wist, tot je vraag me deed twifelen, en terecht...* **Syn. laten (?)**

**doen denken aan** “remind of”

*"Groovin' on E" doet met zijn fris jengelende gitaren en neuzelende fluisterzang denken aan The Kinks.*



## Cf. Van Dale 2005

**doen** als oorzakelijk hulpww.

**synoniem: laten**

*ik doe opmerken dat enz.*

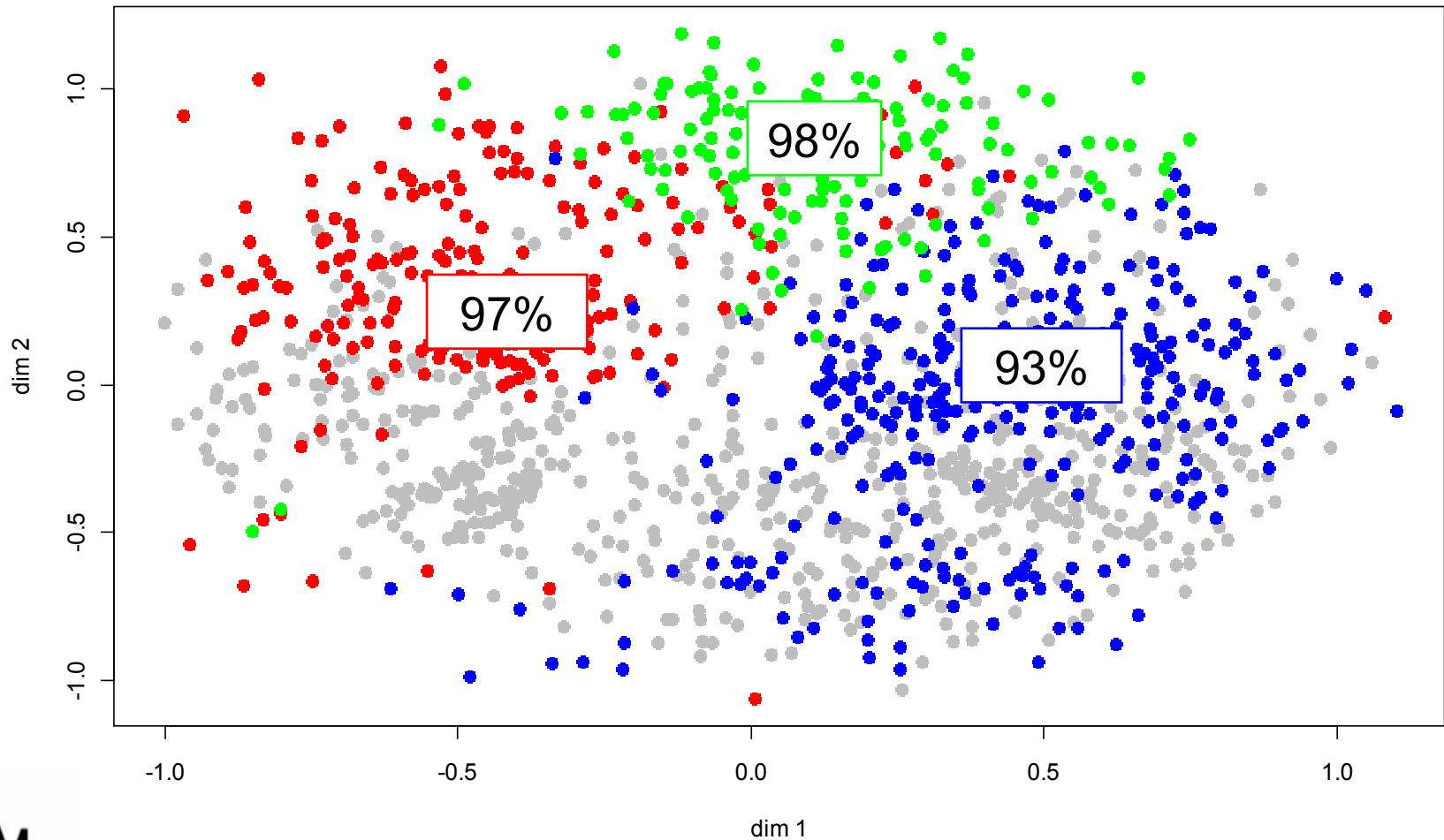
*iets doen of doen doen*

*een steen deed hem struikelen*

*doen te weten*

# Average probability of *laten*

clusters of *laten* in common space



# Cf. Van Dale 2005

## laten (both auxiliary and notional)

...

16. maken dat het object de werking verricht die door de onbep. ww. wordt uitgedrukt. **synoniem: doen**

*zij liet mij vragen of ik meeding,*

*iem. laten struikelen*

*zijn gedachten, zijn oog laten gaan*

*iem. iets laten weten*

?

# Outline

1. Theoretical and methodological background
2. Dutch causative constructions
3. Data and method
4. Semasiological variation (polysemy)
  - semantic and contextual variation of *doen*
  - semantic and contextual variation of *laten*
5. Onomasiological variation (near-synonymy)
  - variation of *doen* and *laten*
  - contextual factors
6. Conclusions and future research

# *doen*: Summary

Global semasiological variation:

- two main senses related to semantic domain of causation (cf. Verhagen & Kemmer 1997, but rather social than ‘physical’ causation)

Contextual semasiological variation:

- *doen* is quantitatively and qualitatively poorer in informal genres and in Netherlandic Dutch (cf. Speelman & Geeraerts 2009), biased towards affective causation, esp. fixed *doen denken aan*

# A historical perspective

- Verhagen 2000: *doen* used to have a more diverse semantic palette

Sense	18th C.	Millenium	
		Belgium	Netherlands
Interpersonal	+	??	??
'Physical'	+	+	?
Affective	+	+	+

- Belgian Dutch is 'lagging behind' (cf. Geeraerts et al. 1999)

# *laten*: Summary

Global semasiological variation:

- three main senses related to different semantic roles of Causer and Causee (cf. Verhagen & Kemmer 1997)

Contextual semasiological variation:

- sense *laten3* (*laten zien/weten/horen*) is unpopular in Belgian newspapers

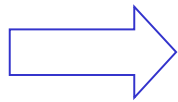
## *doen* vs. *laten*: Summary

- On the general semantic level, the difference between *doen* and *laten* can be formulated in terms of causation (in)directness (cf. Kemmer & Verhagen).
- Cf. polysemy: in all three senses of *laten* the Causer does not actually affect, and the Causee is not affected (indirectness); in the two senses of *doen* the Causer is the (inevitable) cause/stimulus (directness).
- Exemplar effects play an important role, too.



## *doen vs. laten: Summary*

- Non-mental intransitive *doen* and *laten* are more prone to synonymy than the other senses. At the same time, *doen* can be readily replaced by *laten*, but not the other way round.



Laat wat je doen kan,  
maar doe niet wat je laten kan!

# Conclusions

- unsupervised bottom-up exemplar spaces are a way of integrating different interrelated types of variation in one semantic model (complete recontextualization)
- the semasiological and onomasiological perspectives provide specific yet compatible semantic information
- both exemplar effects and more general features are found to be important in categorization
- contextual variation has a large influence on the semasiological structure, but less so on the division of labour between *doen* and *laten*

# Future research

- experimental confirmation of findings
- comparison with unsupervised distributional methods (Vector Space Models)
- a diachronic study of exemplar space evolution

Vragen? Suggesties? Laat je maar gaan!



for further information:

<http://www.ling.arts.kuleuven.be/qlvl>  
[natalia.levshina@arts.kuleuven.be](mailto:natalia.levshina@arts.kuleuven.be)