FACULTY OF ARTS AND PHILOSOPHY

DEPARTMENT OF LINGUISTICS
RESEARCH GROUP: GLIMS

LANGUAGE PRODUCTIVITY @ WORK

HET INTERESSEERT ME GEEN ENE JE-WEET-WEL

THE PRODUCTIVITY OF MINIMIZING CONSTRUCTIONS IN PRESENT-DAY DUTCH

Margot Van den Heede & Peter Lauwers

GHENT UNIVERSITY

OVERVIEW

1. The minimizing construction
2. Productivity
3. Methodology
4. Productivity at the macro-level
5. Productivity of the micro-cx
6. Relations between the two slots
7. Conclusion

1. THE **MINIMIZING** CONSTRUCTION

Ik wil geen seconde missen van de Olympische Spelen.

‘I don’t want to miss a second of the Olympics’

Nina Derwael en Sunisa Lee geven elkaar geen duimbreed toe.

‘Nina Derwael and Sunisa Lee don’t budge an inch’

Het kan hem geen bal schelen.

‘He doesn’t give a damn’

GHENT UNIVERSITY

1. THE **MINIMIZING** CONSTRUCTION

A construction with different slots:

[geen] [MINIMIZING NOUN] [PREDICATE]

↓

the element of negation (quantifier)

GHENT UNIVERSITY

1. THE **MINIMIZING** CONSTRUCTION

A construction with different slots:

[geen] **[MINIMIZING NOUN]** [PREDICATE]

↓

a noun that denotes a small quantity

and that is used to reinforce sentential negation

GHENT UNIVERSITY

1. THE **MINIMIZING** CONSTRUCTION

A construction with different slots:

[geen] [MINIMIZING NOUN] **[PREDICATE]**

GHENT UNIVERSITY

1. THE **MINIMIZING** CONSTRUCTION

A construction with different slots:

[geen] [MINIMIZING NOUN] [PREDICATE]

NOT included (if the minimizer has a lower-level scope)

*U betaalt geen cent **extra*** 'You don't pay a penny extra'

*Hij heeft geen greintje **geduld*** 'He doesn't have a grain (of) patience'

GHENT UNIVERSITY

2. PRODUCTIVITY IN GENERAL

Productivity: a multi-faceted concept

- **Lexical scope**: “Productivity refers to the range of lexical items that may fill the slots of constructions” (Perek 2016: 66) TYPE FREQUENCY
- **Extensibility**: “the extensibility of a pattern to new types” (Barðdal 2008: 29) HAPAX FREQUENCY

GHENT UNIVERSITY

2. PRODUCTIVITY IN GENERAL

Productivity: a multi-faceted concept

- **Lexical scope**: “Productivity refers to the range of lexical items that may fill the slots of constructions” (Perek 2016: 66) TYPE FREQUENCY
- **Extensibility**: “the extensibility of a pattern to new types” (Barðdal 2008: 29) HAPAX FREQUENCY
- “Syntactic productivity is multidimensional – different aspects of productivity are not necessarily correlated (though they often are)” (Zeldes 2012: 135)

2. **PRODUCTIVITY** OF THE MINIMIZING CONSTRUCTION

At different levels of schematicity

The productivity of the **macro-construction**: the range of items that fill the minimizer-slot, the range of predicates and the interaction between the two slots

geen [MINIMIZER] [PRED]

bal	aan vinden
moer	doen
snars	geven om
zier	weten
…	…

The productivity of the **micro-constructions** (= the slot of the minimizer is filled): the range of items that fill the predicate-slot

geen *snars* [PRED]

aan hebben
begrijpen
geloven
snappen
…

GHENT UNIVERSITY

3. METHODOLOGY

Synchronic corpus research: Dutch Web 2014 (nlTenTen14), available on Sketch Engine

↓

Netherlandic Dutch subcorpus: 1.9 billion tokens

[lemma="geen|gene|genen|geene|geenen"] [] {0,2}

[word="bal"] within <s/>

GHENT UNIVERSITY

3. METHODOLOGY

Synchronic corpus research: Dutch Web 2014 (nlTenTen14), available on Sketch Engine

Selection based on literature, exploratory searches in the corpus (“*geen* N”) and searches based on frequent predicates

→ List of 244 minimizers

→ Annotation: 100 tokens per minimizer

(possible for 46 minimizers)

GHENT UNIVERSITY

OVERVIEW

1. The minimizing construction
2. Productivity
3. Methodology
4. Productivity at the macro-level
5. Productivity of the micro-cx
6. Relations between the two slots
7. Conclusion
8. Next steps

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

Dataset: 6224 tokens

Focus: slot of the minimizer

	[MINIMIZER]
Type frequency	244
Hapax frequency	83
Dis legomena	23
Tris legomena	12

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

Dataset: 6224 tokens

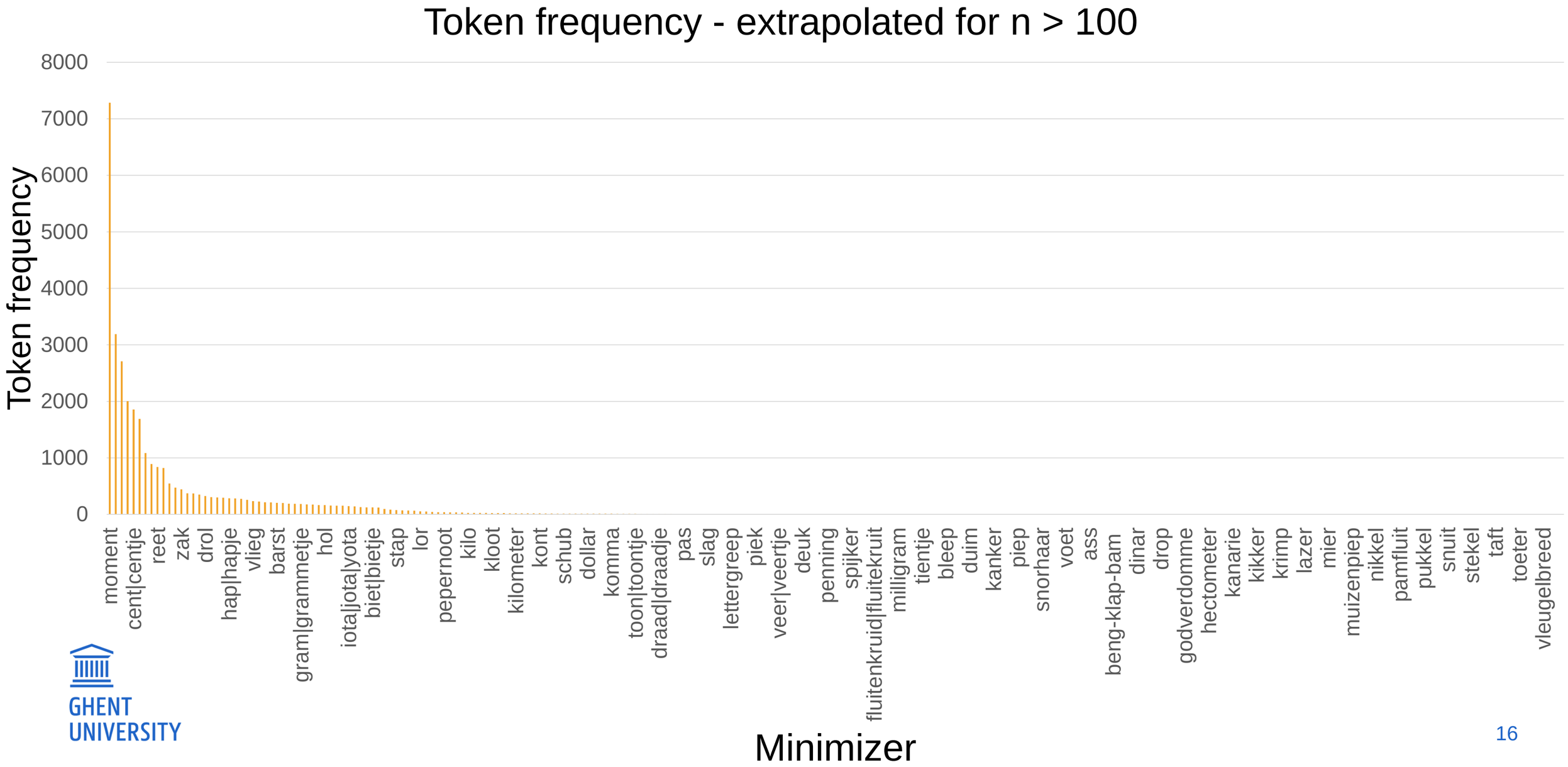
Focus: slot of the minimizer

[MINIMIZER]	
Type frequency	244
Hapax frequency	83
Dis legomena	23
Tris legomena	12

Combined with
641
different
predicates

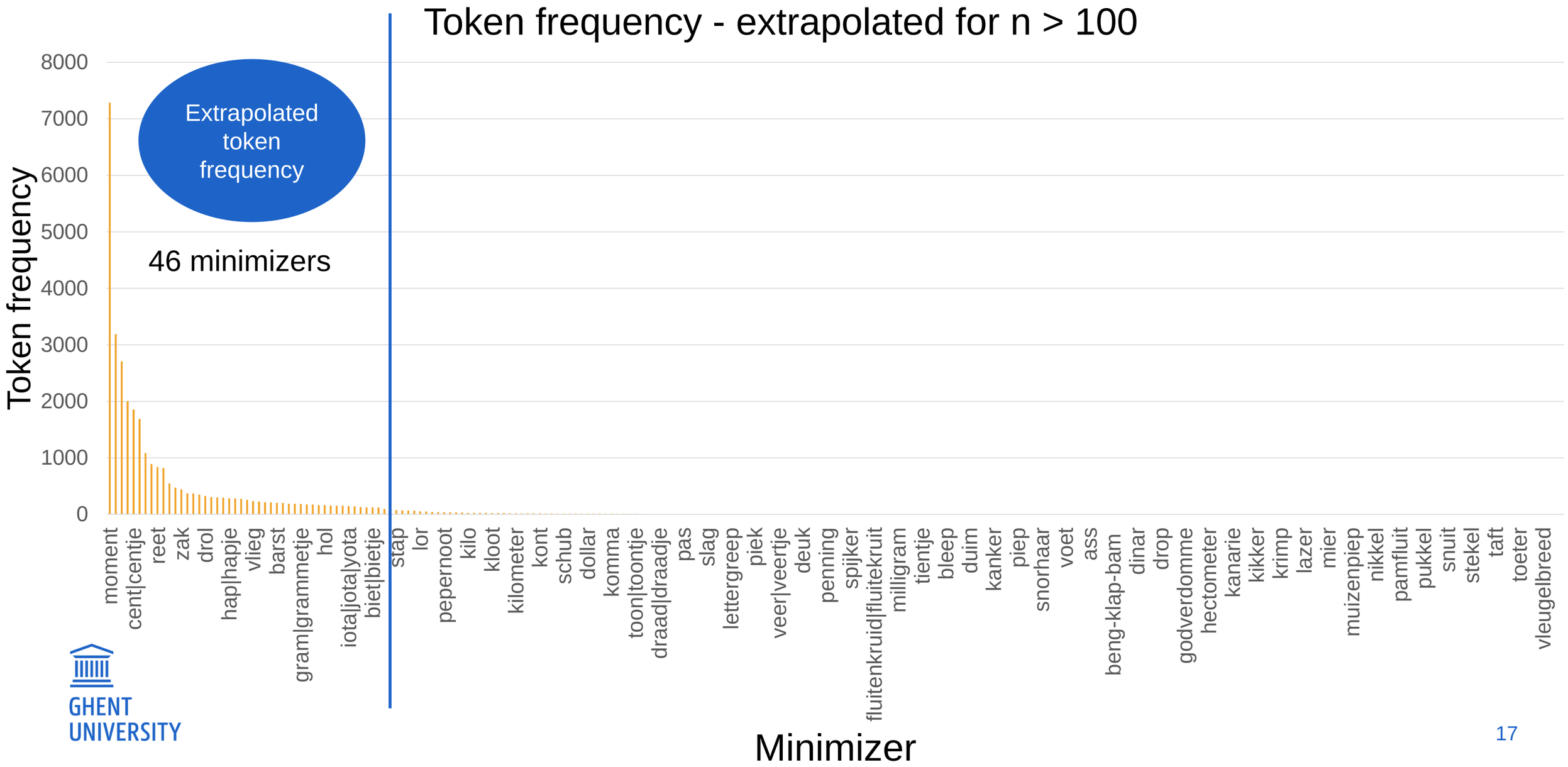
GHENT
UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL



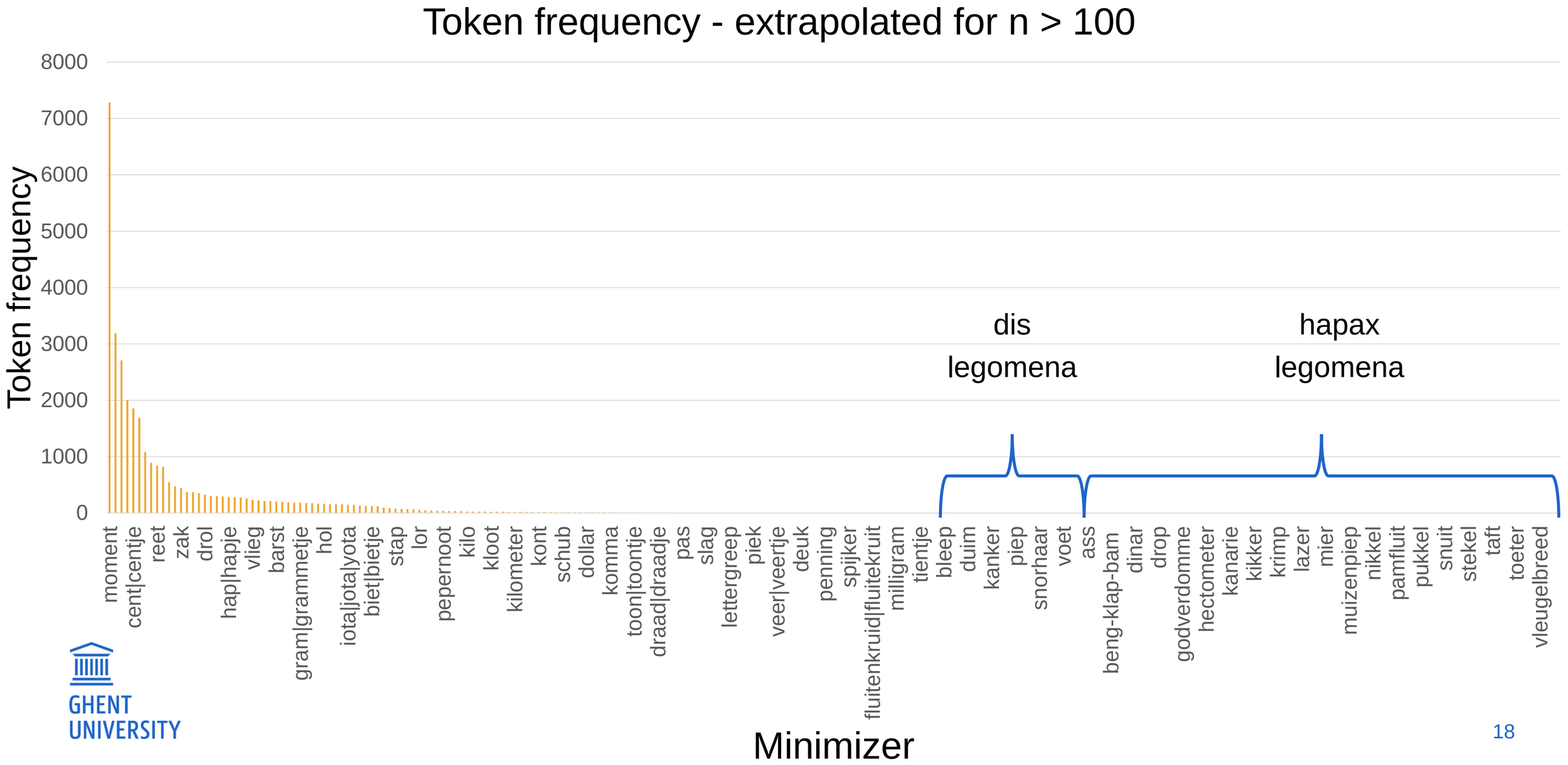
GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL



GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL



GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

Productive?

→ high type frequency of the minimizing slot

→ extensibility of the slot: large number of hapax legomena

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

Productive?

→ high type frequency of the minimizing slot

→ extensibility of the slot: large number of hapax legomena

Key notions: analogy, context

*Sommige dierenartsen snappen er **geen hondendrol** van*
'Some veterinarians don't understand a dog turd about it'

*Hij lijkt **geen graspol** te geven om het amateurvoetbal*
'He doesn't seem to give a clump of grass about amateur soccer'

*De kamerleden interesseert het **geen ene je-weet-wel***
'It didn't interest the members of the Parliament a single you-know-what'

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

Cluster analysis based on the co-occurring predicates in the dataset

Method: pam (Partitioning Around Medoids), for N = 100 (46 minimizers)

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

Cluster analysis based on the cooccurring predicates in the dataset

Method: pam (Partitioning Around Medoids), for N = 100 (46 minimizers)

	Number of minimizers	**Average silhouette width**	**Minimizers (from high to low silhouette width)**
1	5	0,69	*moment, seconde, minuut, dag, ogenblik*
2	3	0,65	*centimeter, millimeter, duimbreed*
3	17	0,51	*fuck, moer, fluit, zak, bal, hol, reet, biet, donder, ruk, barst, klap, flikker, snars, zier, jota, hout*
4	5	0,49	*sterveling, mens, ziel, kip, hond*
5	5	0,34	*stuiver, euro, cent, brood, drol*
6	11	0,10	*druppel, vlieg, gram, hap, woord, spaan, letter, spat, strobreed, steek, meter*

GHENT UNIVERSITY

R package: “fpc”

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

"The silhouette value measures the degree of confidence in the clustering assignment of a particular observation" (Divjak & Fieller 2014: 432)

	Number of minimizers	Average silhouette width	Minimizers (from high to low silhouette width)
1	5	0,69	*moment, seconde, minuut, dag, ogenblik*
2	3	0,65	*centimeter, millimeter, duimbreed*
3	17	0,51	*fuck, moer, fluit, zak, bal, hol, reet, biet, donder, ruk, barst, klap, flikker, snars, zier, jota, hout*
4	5	0,49	*sterveling, mens, ziel, kip, hond*
5	5	0,34	*stuiver, euro, cent, brood, drol*
6	11	0,10	*druppel, vlieg, gram, hap, woord, spaan, letter, spat, strobreed, steek, meter*

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

	Number of minimizers	Average silhouette width	Minimizers (from high to low silhouette width)	
1	5	0,69	*moment, seconde, minuut, dag, ogenblik*	TIME
2	3	0,65	*centimeter, millimeter, duimbreed*	DISTANCE
3	17	0,51	*fuck, moer, fluit, zak, bal, hol, reet, biet, donder, ruk, barst, klap, flikker, snars, zier, jota, hout*	TABOO…
4	5	0,49	*sterveling, mens, ziel, kip, hond*	PEOPLE
5	5	0,34	*stuiver, euro, cent, brood, drol*	MONEY
6	11	0,10	*druppel, vlieg, gram, hap, woord, spaan, letter, spat, strobreed, steek, meter*	LANGUAGE, WEIGHT…

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

	Number of minimizers	Average silhouette width	Minimizers (from high to low silhouette width)
1	5	0,69	*moment, seconde, minuut, dag, ogenblik*
2	3	0,65	*centimeter, millimeter, duimbreed*
3	17	0,51	*fuck, moer, fluit, zak, bal, hol, reet, biet, donder, ruk, barst, klap, flikker, snars, zier, jota, hout*
4	5	0,49	*sterveling, mens, ziel, kip, hond*
5	5	0,34	*stuiver, euro, cent, brood, drol*
6	11	0,10	*druppel, vlieg, gram, hap, woord, spaan, letter, spat, strobreed, steek, meter*

MONEY

*geen (droog) brood **verdienen***
'to earn not a (dry) bread'

*geen drol **kosten***
'to cost not a turd'

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

	Number of minimizers	Average silhouette width	Minimizers (from high to low silhouette width)
1	5	0,69	*moment, seconde, minuut, dag, ogenblik*
2	3	0,65	*centimeter, millimeter, duimbreed*
3	17	0,51	*fuck, moer, fluit, zak, bal, hol, reet, biet, donder, ruk, barst, klap, flikker, snars, zier, jota, hout*
4	5	0,49	*sterveling, mens, ziel, kip, hond*
5	5	0,34	*stuiver, euro, cent, brood, drol*
6	11	0,10	*druppel, vlieg, gram, hap, woord, spaan, letter, spat, strobreed, steek, meter*

PEOPLE

geen vlieg kwaad doen

'to hurt not a fly'

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

	Number of minimizers	Average silhouette width	Minimizers (from high to low silhouette width)
1	5	0,69	*moment, seconde, minuut, dag, ogenblik*
2	3	0,65	*centimeter, millimeter, duimbreed*
3	17	0,51	*fuck, moer, fluit, zak, bal, hol, reet, biet, donder, ruk, barst, klap, flikker, snars, zier, jota, hout*
4	5	0,49	*sterveling, mens, ziel, kip, hond*
5	5	0,34	*stuiver, euro, cent, brood, drol*
6	11	0,10	*druppel, vlieg, gram, hap, woord, spaan, letter, spat, strobreed, steek, meter*

DISTANCE

geen strobreed in de weg leggen

'to put not a straw in the way'

? type frequency = 37

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

geen centimeter / geen millimeter

De man wijkt geen centimeter 'The man doesn't budge a centimetre'

VS.

geen meter

Het was de eerste taxichauffeur in Azië die een TomTom had. En dat werkte dus voor geen meter. Hij wist gewoon niet waar ons hotel lag.

'It was the first taxi driver in Asia to have a TomTom. And that didn't work for a metre. He just didn't know where our hotel was.'

GHENT UNIVERSITY

4. PRODUCTIVITY AT THE **MACRO**-LEVEL

geen centimeter	7/100 tokens	[voor] + [geen] + [centimeter]
geen millimeter	13/100 tokens	[in/met/op/voor] + [geen] + [millimeter]

VS.

geen meter	91/100 tokens	[voor/van] + [geen] + [meter]

GHENT UNIVERSITY

OVERVIEW

1. The minimizing construction
2. Productivity
3. Methodology
4. Productivity at the macro-level
5. Productivity of the micro-cx
6. Relations between the two slots
7. Conclusion

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

SNARS	
begrijpen	28
snappen	20
geloven	17
interesseren	5
uitmaken	5
weten	5
geven om	2
kloppen	2
te maken hebben	2
terechtkomen	2
veranderen	2
bakken van	1
deugen	1
hebben aan	1
helpen	1
opschieten	1
terugzien	1
uitvoeren	1
vertrouwen	1
vinden aan	1
zien	1

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Token frequency: **100**
Extrapolated token frequency: **371**

SNARS	
begrijpen	28
snappen	20
geloven	17
interesseren	5
uitmaken	5
weten	5
geven om	2
kloppen	2
te maken hebben	2
terechtkomen	2
veranderen	2
bakken van	1
deugen	1
hebben aan	1
helpen	1
opschieten	1
terugzien	1
uitvoeren	1
vertrouwen	1
vinden aan	1
zien	1

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Token frequency: **100**
Extrapolated token frequency: **371**
Frequency of top 1 predicate: **28**
Frequency of top 3 predicates: (28+20+17)/3 = **21,6**

SNARS	
begrijpen	28
snappen	20
geloven	17
interesseren	5
uitmaken	5
weten	5
geven om	2
kloppen	2
te maken hebben	2
terechtkomen	2
veranderen	2
bakken van	1
deugen	1
hebben aan	1
helpen	1
opschieten	1
terugzien	1
uitvoeren	1
vertrouwen	1
vinden aan	1
zien	1

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Token frequency: **100**
Extrapolated token frequency: **371**
Frequency of top 1 predicate: **28**
Frequency of top 3 predicates: (28+20+17)/3 = **21,6**

Type frequency (and type token ratio): **21**
Hapax frequency (and hapax token ratio): **10**
Hapax type ratio: **0,48**

SNARS	
begrijpen	28
snappen	20
geloven	17
interesseren	5
uitmaken	5
weten	5
geven om	2
kloppen	2
te maken hebben	2
terechtkomen	2
veranderen	2
bakken van	1
deugen	1
hebben aan	1
helpen	1
opschieten	1
terugzien	1
uitvoeren	1
vertrouwen	1
vinden aan	1
zien	1

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Token frequency: **100**
Extrapolated token frequency: **371**
Frequency of top 1 predicate: **28**
Frequency of top 3 predicates: (28+20+17)/3 = **21,6**

Type frequency (and type token ratio): **21**
Hapax frequency (and hapax token ratio): **10**
Hapax type ratio: **0,48**

SNARS	
begrijpen	28
snappen	20
geloven	17
interesseren	5
uitmaken	5
weten	5
geven om	2
kloppen	2
te maken hebben	2
terechtkomen	2
veranderen	2
bakken van	1
deugen	1
hebben aan	1
helpen	1
opschieten	1
terugzien	1
uitvoeren	1
vertrouwen	1
vinden aan	1
zien	1

GHENT UNIVERSITY

"Syntactic productivity is multidimensional – different aspects of productivity are not necessarily correlated (though they often are)"
(Zeldes 2012: 135)

5. PRODUCTIVITY OF THE **MICRO**-CX

Correlations between the productivity measures (Pearson correlation)

(based on the minimizers with 100 tokens)

	Extrapol. token freq	**Freq top 1**	**Mean freq top 3**	**Type freq**	**Hapax freq**	**Hapax type ratio**
Extrapol. token freq	1	-0,1	-0,1	0,11	0,12	0,11
Freq top 1	-0,1	1	0,94	-0,77	-0,53	-0,02
Mean freq top 3	-0,1	0,94	1	-0,83	-0,6	-0,06
Type freq	0,11	-0,77	-0,83	1	0,93	0,46
Hapax freq	0,12	-0,53	-0,6	0,93	1	0,73
Hapax type ratio	0,11	-0,02	-0,06	0,46	0,73	1

GHENT UNIVERSITY

R package: “corrplot”

5. PRODUCTIVITY OF THE **MICRO**-CX

Correlations between the productivity measures (Pearson correlation)

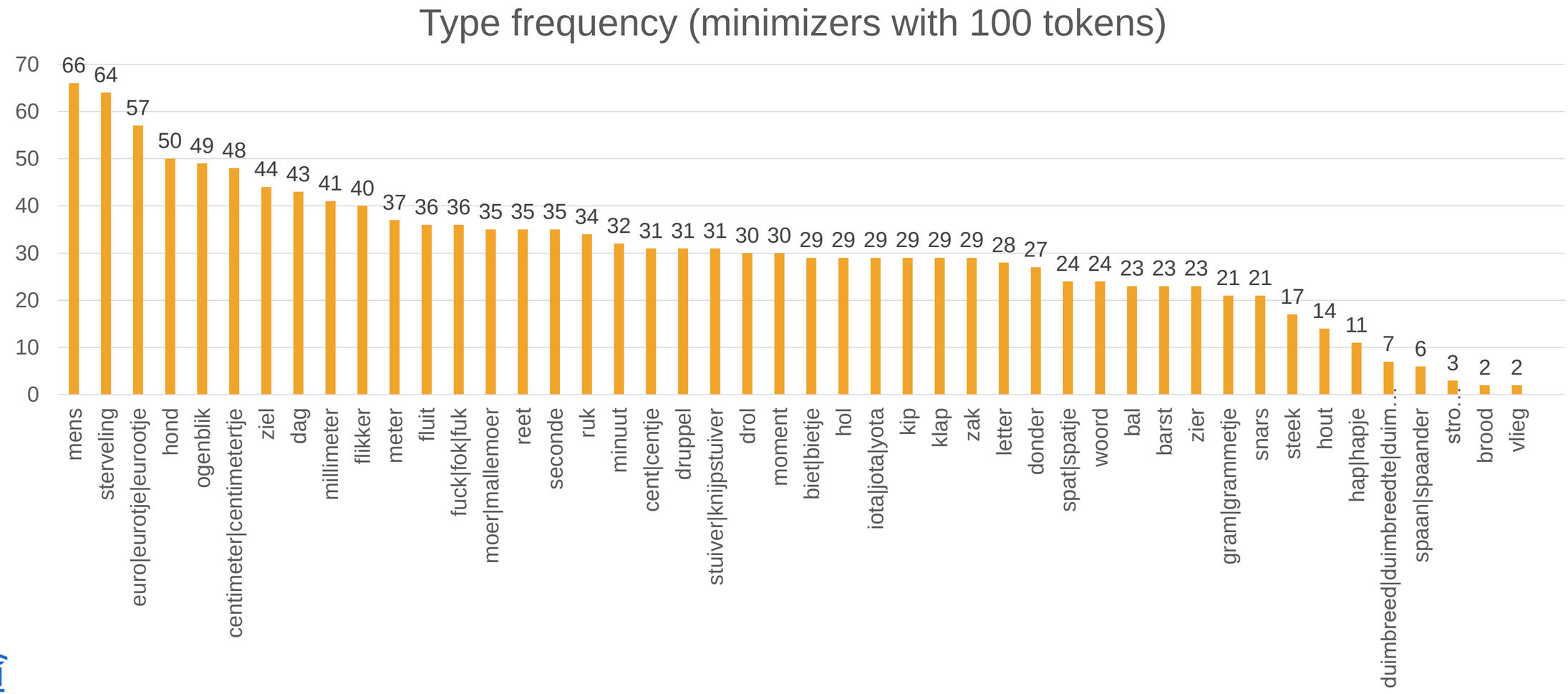
(based on the minimizers with 100 tokens)

	Extrapol. token freq	**Freq top 1**	**Mean freq top 3**	**Type freq**	**Hapax freq**	**Hapax type ratio**
Extrapol. token freq	1	-0,1	-0,1	0,11	0,12	0,11
Freq top 1	-0,1	1	0,94	-0,77	-0,53	-0,02
Mean freq top 3	-0,1	0,94	1	-0,83	-0,6	-0,06
Type freq	0,11	-0,77	-0,83	1	0,93	0,46
Hapax freq	0,12	-0,53	-0,6	0,93	1	0,73
Hapax type ratio	0,11	-0,02	-0,06	0,46	0,73	1

GHENT UNIVERSITY

R package: “corrplot”

5. PRODUCTIVITY OF THE **MICRO**-CX



GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Top 10 - Type frequency

Type frequency	
mens ‘human being’	66
sterveling ‘mortal’	64
euro ‘euro’	57
hond ‘dog’	50
ogenblik ‘instant’	49
centimeter ‘centimetre’	48
ziel ‘soul’	44
dag ‘day’	43
millimeter ‘millimetre’	41
flikker ‘fag’	40

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Top 10 - Type and hapax frequency

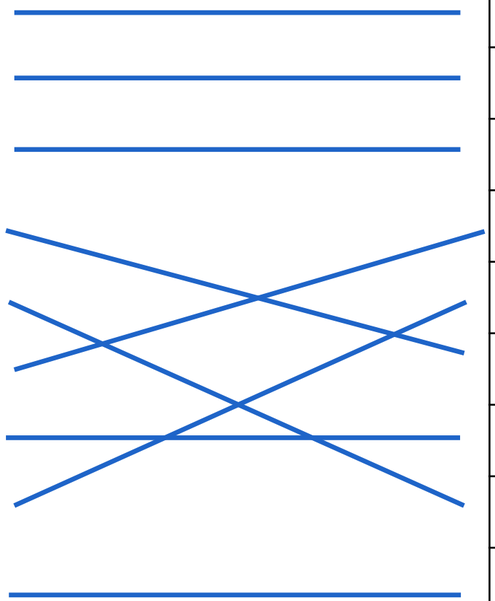
Type frequency	
mens ‘human being’	66
sterveling ‘mortal’	64
euro ‘euro’	57
hond ‘dog’	50
ogenblik ‘instant’	49
centimeter ‘centimetre’	48
ziel ‘soul’	44
dag ‘day’	43
millimeter ‘millimetre’	41
flikker ‘fag’	40

Hapax frequency	
mens ‘human being’	52
sterveling ‘mortal’	50
euro ‘euro’	40
centimeter ‘centimetre’	34
dag ‘day’	33
hond ‘dog’	33
ziel ‘soul’	33
ogenblik ‘instant’	30
millimeter ‘millimetre’	28
druppel ‘drop’	22

5. PRODUCTIVITY OF THE **MICRO**-CX

Top 10 - Type and hapax frequency

Type frequency	
mens ‘human being’	66
sterveling ‘mortal’	64
euro ‘euro’	57
hond ‘dog’	50
ogenblik ‘instant’	49
centimeter ‘centimetre’	48
ziel ‘soul’	44
dag ‘day’	43
millimeter ‘millimetre’	41
flikker ‘fag’	40



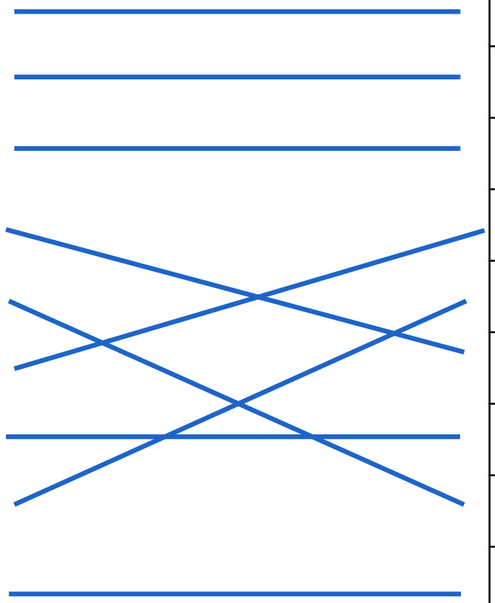
Hapax frequency	
mens ‘human being’	52
sterveling ‘mortal’	50
euro ‘euro’	40
centimeter ‘centimetre’	34
dag ‘day’	33
hond ‘dog’	33
ziel ‘soul’	33
ogenblik ‘instant’	30
millimeter ‘millimetre’	28
druppel ‘drop’	22

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Top 10 - Type and hapax frequency

Type frequency	
mens ‘human being’	66
sterveling ‘mortal’	64
euro ‘euro’	57
hond ‘dog’	50
ogenblik ‘instant’	49
centimeter ‘centimetre’	48
ziel ‘soul’	44
dag ‘day’	43
millimeter ‘millimetre’	41
flikker ‘fag’	40



Hapax frequency	
mens ‘human being’	52
sterveling ‘mortal’	50
euro ‘euro’	40
centimeter ‘centimetre’	34
dag ‘day’	33
hond ‘dog’	33
ziel ‘soul’	33
ogenblik ‘instant’	30
millimeter ‘millimetre’	28
druppel ‘drop’	22

Next: semantic structuring of the predicates

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Top 10 - Type and hapax frequency + cluster analysis

Type frequency		
mens 'human being'	4_PEOPLE	66
sterveling 'mortal'	4_PEOPLE	64
euro 'euro'	5_MONEY	57
hond 'dog'	4_PEOPLE	50
ogenblik 'instant'	1_TIME	49
centimeter 'centimetre'	2_DISTANCE	48
ziel 'soul'	4_PEOPLE	44
dag 'day'	1_TIME	43
millimeter 'millimetre'	2_DISTANCE	41
flikker 'fag'	3_TABOO	40

Hapax frequency		
mens 'human being'	4_PEOPLE	52
sterveling 'mortal'	4_PEOPLE	50
euro 'euro'	5_MONEY	40
centimeter 'centimetre'	2_DISTANCE	34
dag 'day'	1_TIME	33
hond 'dog'	4_PEOPLE	33
ziel 'soul'	4_PEOPLE	33
ogenblik 'instant'	1_TIME	30
millimeter 'millimetre'	2_DISTANCE	28
druppel 'drop'	6_OTHER	22

4_PEOPLE: These minimizers have different syntactic functions (object, subject)

→ one of the next steps in my research

GHENT UNIVERSITY

5. PRODUCTIVITY OF THE **MICRO**-CX

Correlations between the productivity measures

(based on the minimizers with 100 tokens)

	Extrapol. token freq	**Freq top 1**	**Mean freq top 3**	**Type freq**	**Hapax freq**	**Hapax type ratio**
Extrapol. token freq	1	-0,1	-0,1	0,11	0,12	0,11
Freq top 1	-0,1	1	0,94	-0,77	-0,53	-0,02
Mean freq top 3	-0,1	0,94	1	-0,83	-0,6	-0,06
Type freq	0,11	-0,77	-0,83	1	0,93	0,46
Hapax freq	0,12	-0,53	-0,6	0,93	1	0,73
Hapax type ratio	0,11	-0,02	-0,06	0,46	0,73	1

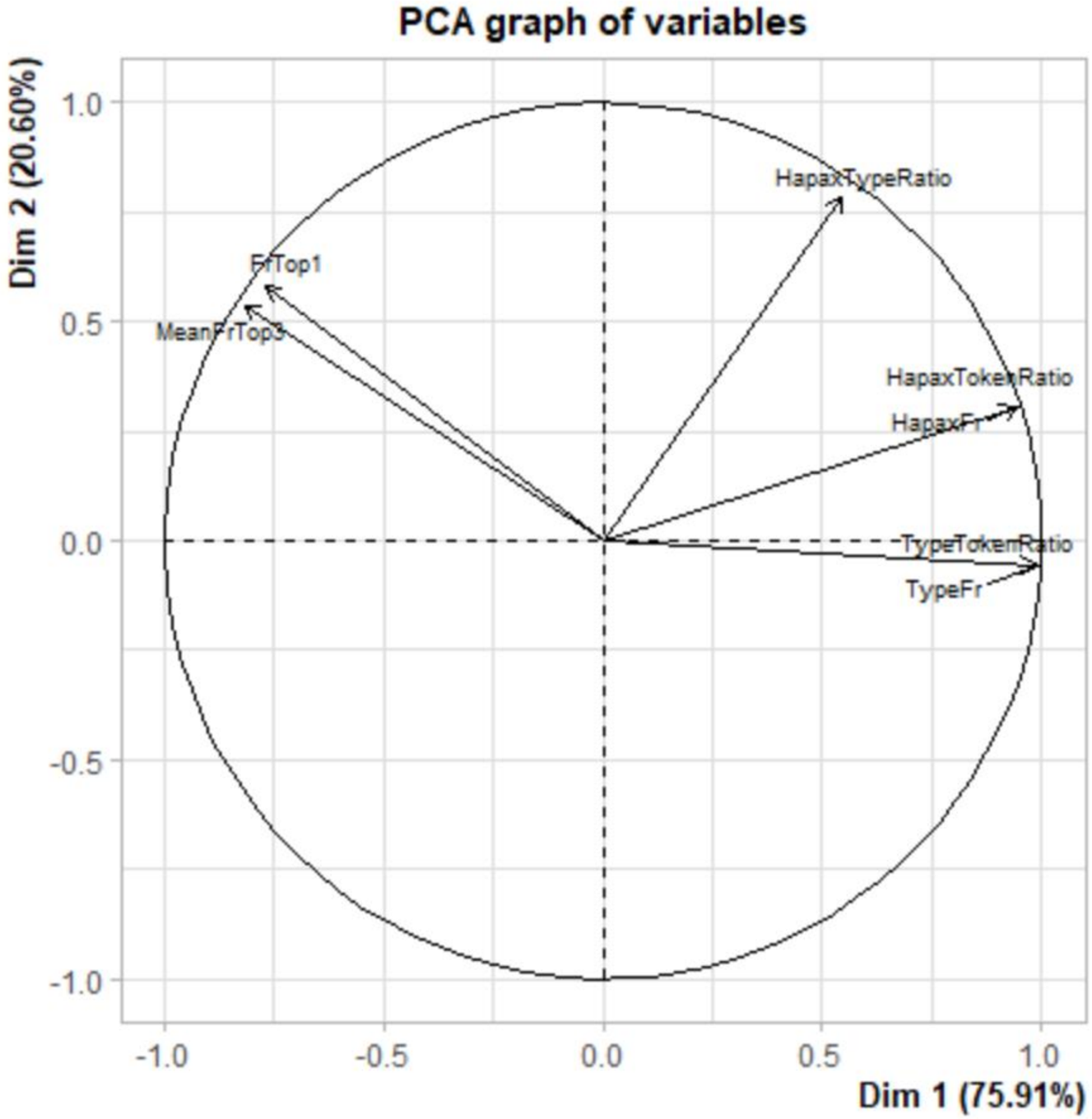
5. PRODUCTIVITY OF THE **MICRO**-CX

Correlations between the productivity measures

(based on the minimizers with 100 tokens)

	Extrapol. token freq	**Freq top 1**	**Mean freq top 3**	**Type freq**	**Hapax freq**	**Hapax type ratio**
Extrapol. token freq	1	-0,1	-0,1	0,11	0,12	0,11
Freq top 1	-0,1	1	0,94	-0,77	-0,53	-0,02
Mean freq top 3	-0,1	0,94	1	-0,83	-0,6	-0,06
Type freq	0,11	-0,77	-0,83	1	0,93	0,46
Hapax freq	0,12	-0,53	-0,6	0,93	1	0,73
Hapax type ratio	0,11	-0,02	-0,06	0,46	0,73	1

GHENT UNIVERSITY



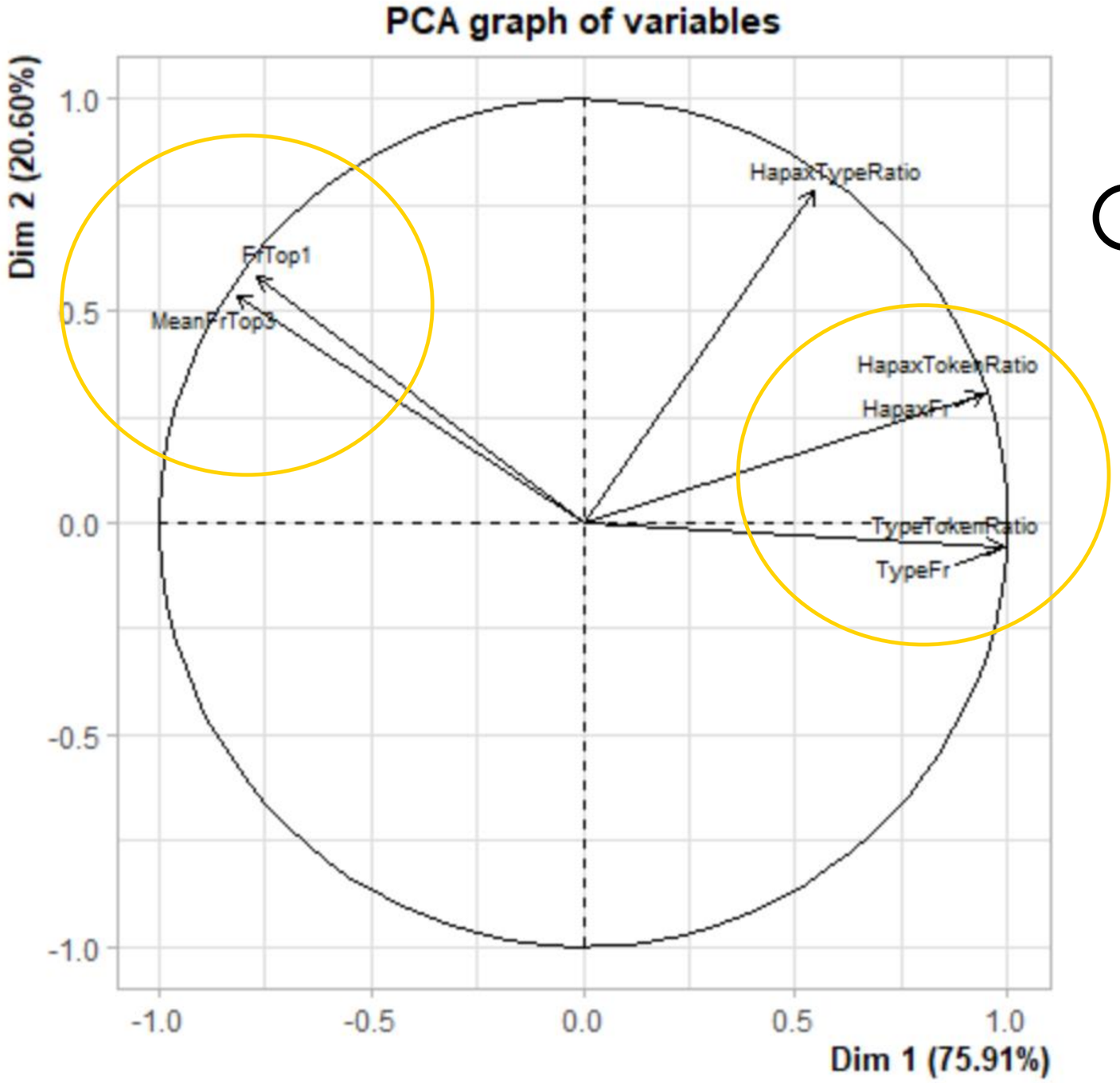
Principal Component Analysis

Allows to reorient the data so that the first few dimensions or principal components account for as much of the available information as possible

Based on the minimizers with 100 tokens

(Zeldes 2012; Van Wettere 2018)

R Package: “FactoMineR”



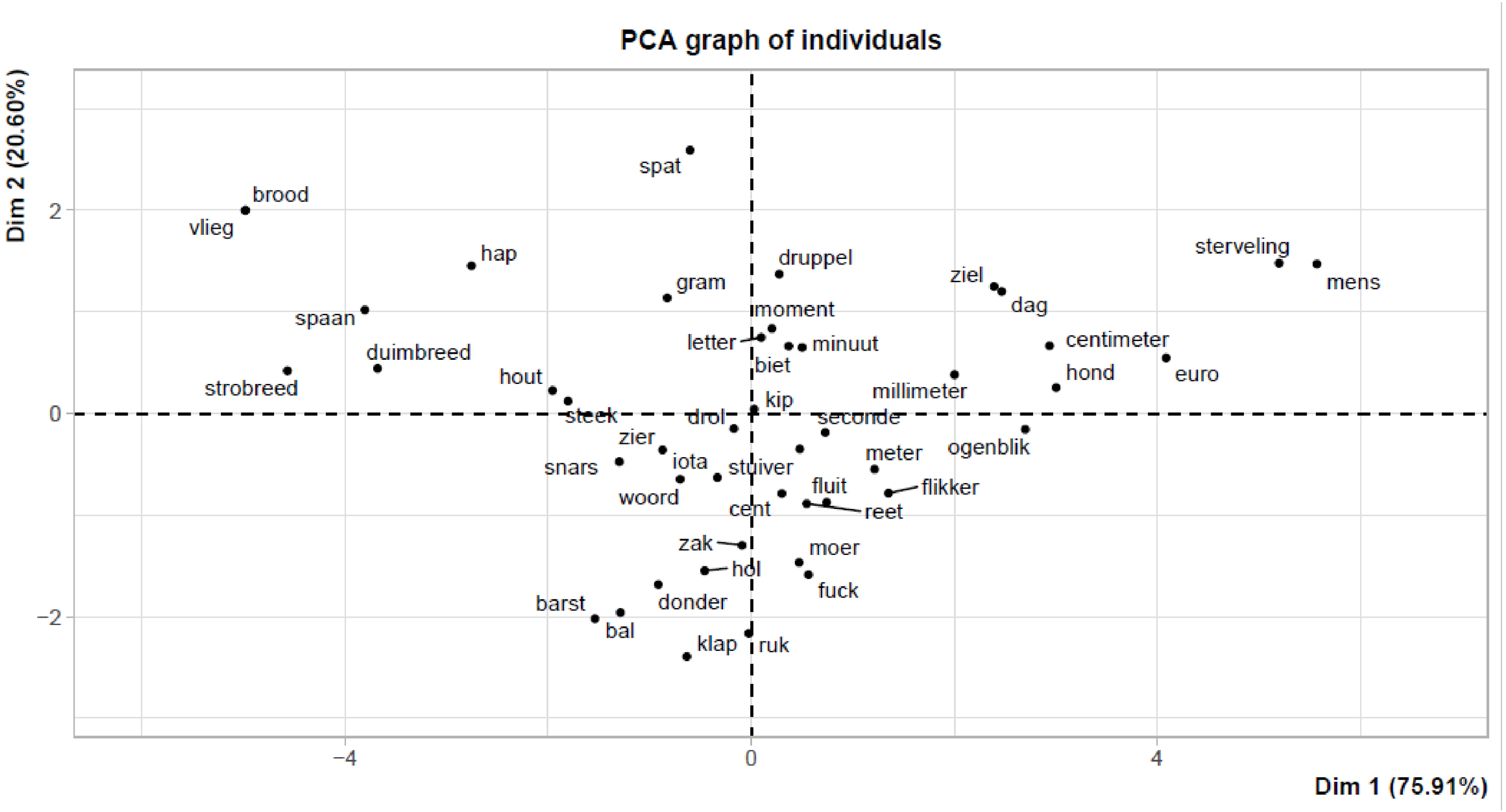
Principal Component Analysis

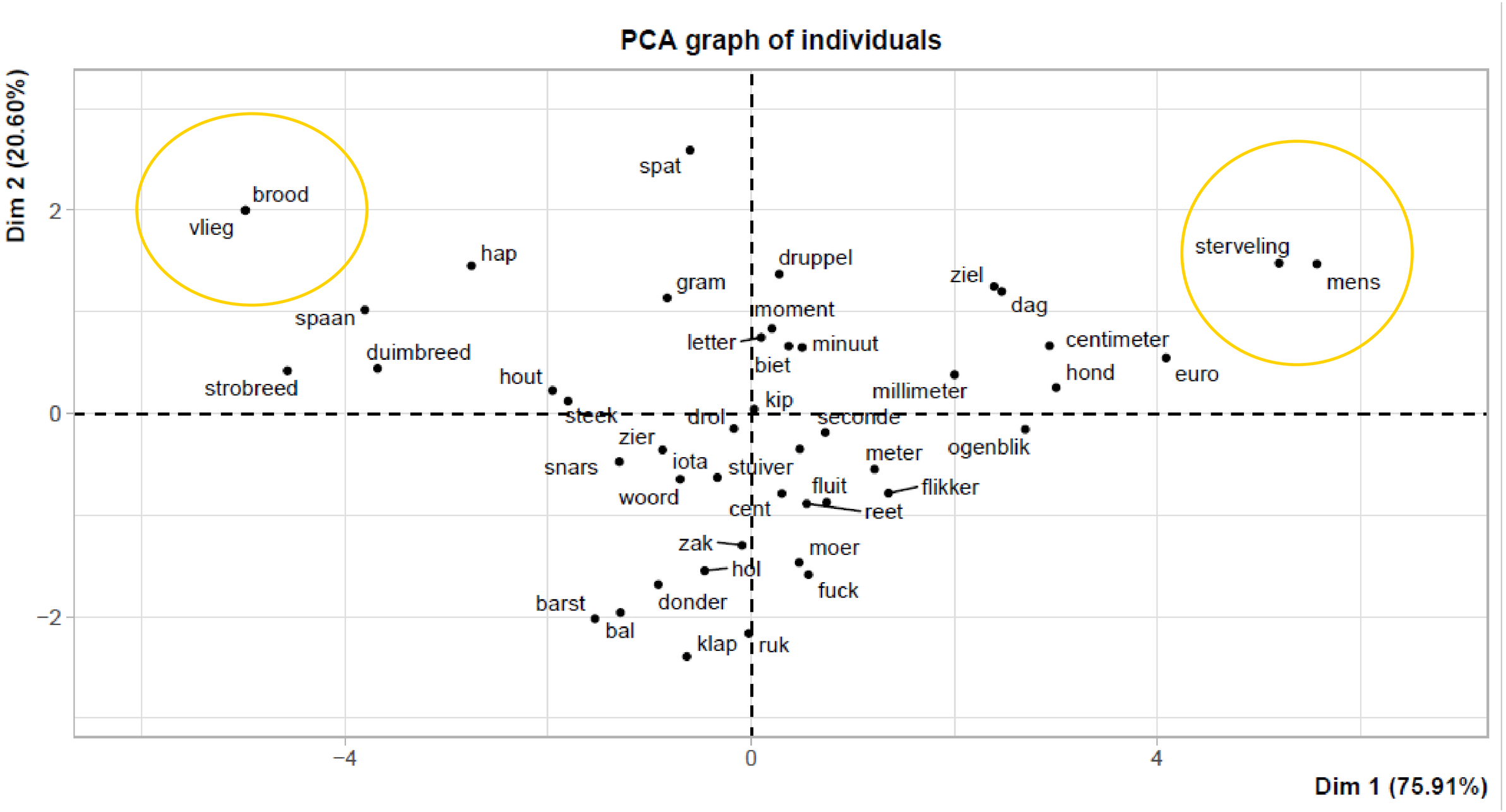
Allows to reorient the data so that the first few dimensions or principal components account for as much of the available information as possible

Based on the minimizers with 100 tokens

(Zeldes 2012; Van Wettere 2018)

R Package: “FactoMineR”


PCA graph of individuals
Dim 2 (20.60%)
Dim 1 (75.91%)
2
0
−2
−4
0
4
spat
brood
vlieg
hap
druppel
gram
sterveling
mens
ziel
dag
spaan
moment
letter
minuut
centimeter
duimbreed
biet
hond
euro
strobreed
hout
millimeter
kip
steek
drol
seconde
zier
ogenblik
meter
snars
iota
stuiver
fluit
flikker
woord
cent
reet
zak
moer
hol
fuck
donder
barst
bal
klap
ruk


PCA graph of individuals
Dim 2 (20.60%)
Dim 1 (75.91%)
2
0
−2
−4
0
4
brood
vlieg
hap
spat
druppel
gram
ziel
dag
sterveling
mens
spaan
moment
duimbreed
letter
minuut
centimeter
biet
hond
euro
strobreed
hout
millimeter
kip
steek
drol
seconde
zier
ogenblik
meter
snars
iota
stuiver
fluit
flikker
woord
cent
reet
zak
moer
hol
fuck
donder
barst
bal
klap
ruk

5. PRODUCTIVITY OF THE **MICRO**-CX

👉 The higher the type frequency and hapax frequency: the more productive

👉 The token frequency of the predicate can detract from productivity

"The logic behind this assumption is that highly entrenched tokens do not contribute to schematization" (Barðdal 2008: 29; Van Wettere 2018: 610, De Smet 2020)

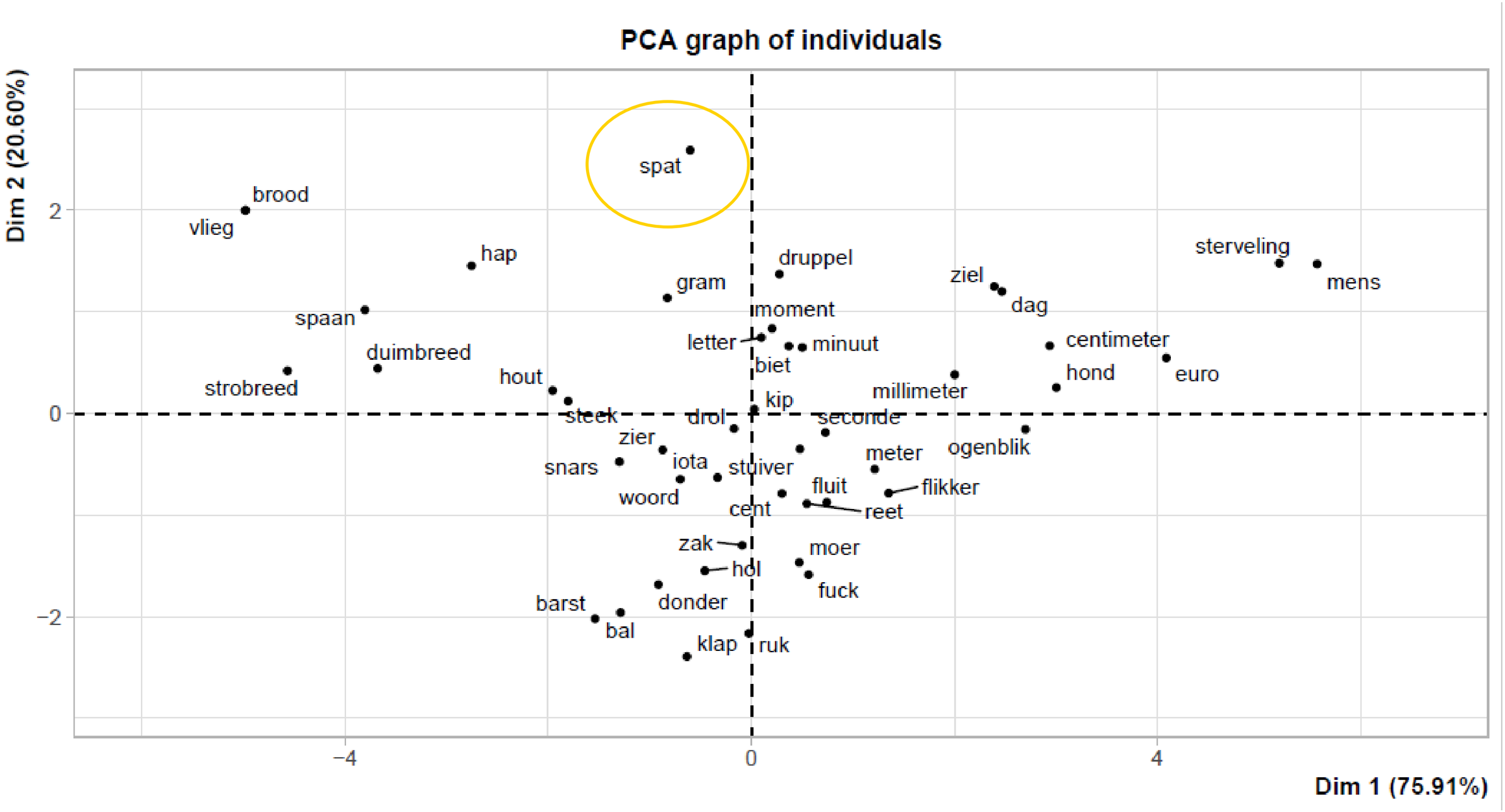
5. PRODUCTIVITY OF THE **MICRO**-CX

👉 The higher the type frequency and hapax frequency: the more productive

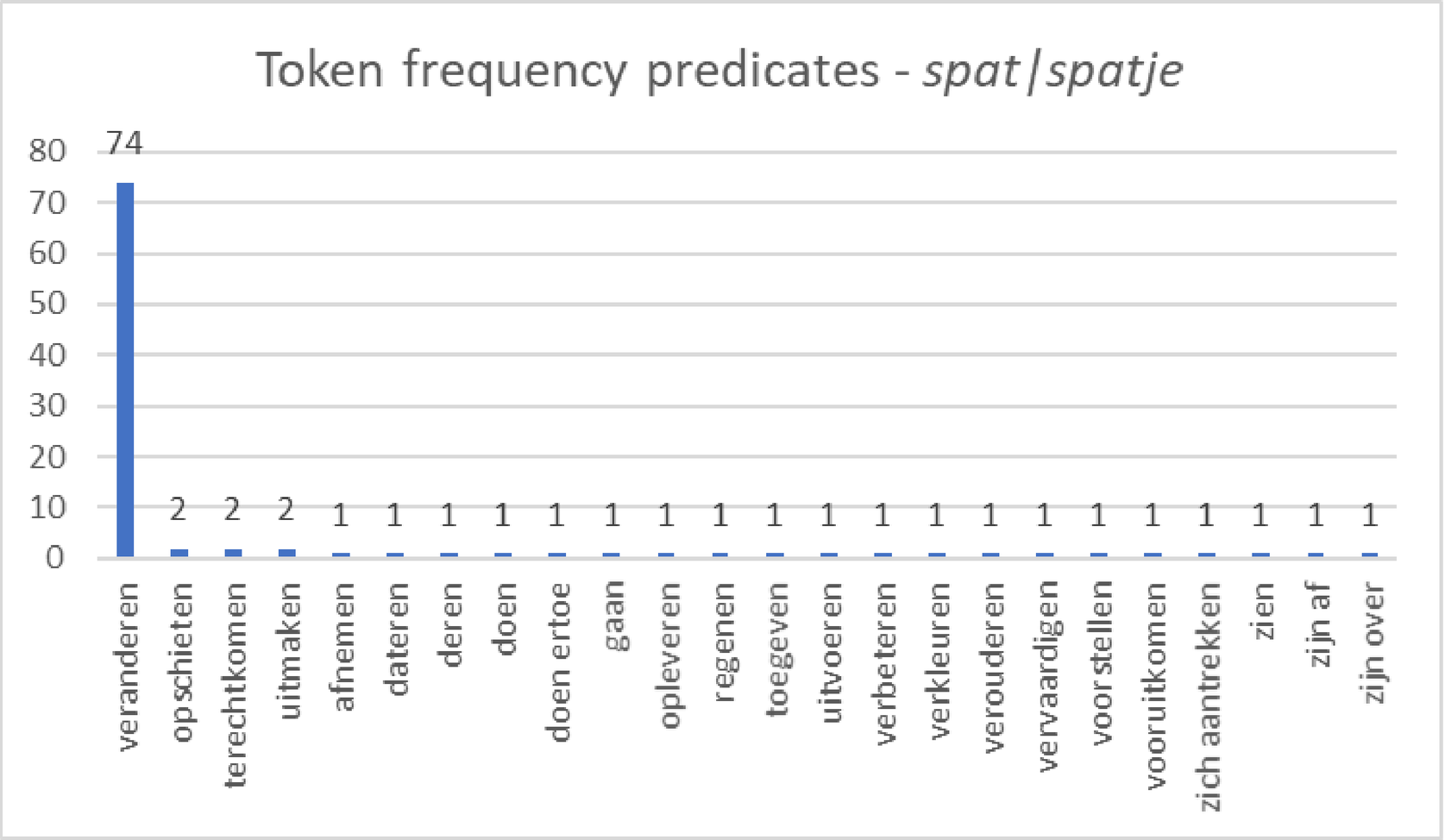
👉 The token frequency of the predicate can detract from productivity

"The logic behind this assumption is that highly entrenched tokens do not contribute to schematization" (Barðdal 2008: 29; Van Wettere 2018: 610; De Smet 2020)

BUT: "I argue that high token frequency also contributes to productivity, although only at the lower end of the productivity cline, which involves analogy, as high token frequency makes individual lexical items more entrenched and thus more suitable as models for analogical extensions." (Barðdal 2008: 34)


PCA graph of individuals
Dim 2 (20.60%)
Dim 1 (75.91%)
2
0
−2
−4
0
4
spat
brood
vlieg
hap
druppel
gram
sterveling
mens
ziel
dag
spaan
moment
letter
minuut
centimeter
duimbreed
biet
hond
euro
strobreed
hout
millimeter
kip
steek
drol
seconde
zier
ogenblik
meter
snars
iota
stuiver
fluit
flikker
woord
cent
reet
zak
moer
hol
fuck
donder
barst
bal
klap
ruk

5. PRODUCTIVITY OF THE **MICRO**-CX



Type frequency = 24

Hapax frequency = 20

GHENT UNIVERSITY

OVERVIEW

1. The minimizing construction
2. Productivity
3. Methodology
4. Productivity at the macro-level
5. Productivity of the micro-cx
6. Relations between the two slots
7. Conclusion

GHENT UNIVERSITY

6. Relations between the two slots

In my dataset of 6224 tokens:

2195 minimizer-predicate combinations

1395 hapaxes (combinations that only occur once in the dataset)

GHENT UNIVERSITY

6. RELATIONS BETWEEN THE TWO SLOTS

Covarying collexeme analysis				
Minimizer – frequency in the cx	**Predicate – frequency in the cx**	**Observed frequency**	**Expected frequency**	**Coll. strength**
(geen) vlieg 'fly' – 100	*kwaad doen* 'do wrong' – 107	99	1,72	205,79
(geen) strobreed 'strawbreadth' – 100	*in de weg leggen* 'to put in the way'– 113	97	1,82	191,46
(geen) spaan 'chip' – 100	*heel laten* 'leave intact' – 95	87	1,53	170,57
(geen) vin 'fin' – 84	*verroeren* 'to move' – 107	83	1,44	164,45
(geen droog) brood 'bread' – 100	*verdienen* 'to earn' – 214	99	3,44	154,41
(geen) hap 'bite' – 100	*door de keel krijgen* 'to get it down the throat/to eat' – 78	74	1,25	143,06
(met geen) vinger 'finger' – 65	*aanraken* 'to touch' – 58	57	0,61	128,08
(geen) duimbreed 'inch' – 100	*toegeven* 'to give in' – 121	76	1,94	120,49
(geen) spat 'splash/drop' – 100	*veranderen* 'to change' – 191	74	3,07	95,27
(geen) druppel 'drop' – 100	*drinken* 'to drink' – 59	49	0,95	83,28

GHENT UNIVERSITY

Gries, Stefan Th. 2014. Coll.analysis 3.5. A script for R to compute perform collostructional analyses.

6. Relations between the two slots

⇆ Attraction in both directions

"bidirectional dependency" (Desagulier 2018: 11)

Covarying collexeme analysis			
Minimizer – frequency in the cx	**Predicate – frequency in the cx**	**Delta P (word 1 to word 2)**	**Delta P (word 2 to word 1)**
(geen) vlieg 'fly' – 100	*kwaad doen* 'do wrong' – 107	0.9887	0.9251

GHENT UNIVERSITY

6. Relations between the two slots

⟵ “One word is a better cue of the other than vice versa”

“directional dependency” (Desagulier 2018: 11)

Covarying collexeme analysis			
Minimizer – frequency in the cx	**Predicate – frequency in the cx**	**Delta P (word 1 to word 2)**	**Delta P (word 2 to word 1)**
(geen) druppel ‘drop’ – 100	*drinken* ‘to drink’ – 59	0.4884	0.8222

GHENT UNIVERSITY

OVERVIEW

1. The minimizing construction
2. Productivity
3. Methodology
4. Productivity at the macro-level
5. Productivity of the micro-cx
6. Relations between the two slots
7. Conclusion

GHENT UNIVERSITY

7. Conclusion

At the macro-level:

- high type and hapax frequency of the minimizing slot (attested and potential productivity)
- the minimizers are recruited from different semantic categories

At the micro-level:

- varying productivity of the micro-constructions
- correlations between the different productivity measures:
    - the more types, the more hapaxes
    - very frequent lexicalisations, highly correlated items reduce productivity → productivity impeded by collostruction strength?

GHENT UNIVERSITY

FACULTY OF ARTS AND PHILOSOPHY

DEPARTMENT OF LINGUISTICS
RESEARCH GROUP: GLIMS

Margot Van den Heede
mcvdnhee.vandenheede@ugent.be

Peter Lauwers
peter.lauwers@ugent.be

Thank you for your attention!

Part of the GOA project (BOF UGent)
Language Productivity at Work
https://www.languageproductivity.ugent.be/

Supervisors: P. Lauwers (PI), J. Barðdal, R. Enghels, T. Colleman, R. Hartsuiker, M. Taverniers, L. De Cuypere, A. Ghyselen

GHENT UNIVERSITY

REFERENCES

Barðdal, J. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: Benjamins.

Desagulier, G. 2016. A lesson from associative learning: asymmetry and productivity in multiple-slot constructions. *Corpus Linguistics and Linguistic Theory 12*(2), 173-219.

De Smet, H. 2020 (aop). What predicts productivity: From theory to individuals. *Cognitive Linguistics*.

Divjak, D., and N. Fieller. 2014. Cluster Analysis. Finding structure in linguistic data. In D. Glynn and J. Robinsion (Ed.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 405–41.

Gries, S. Th. 2014. Coll.analysis 3.5. A script for R to compute perform collostructional analyses.

Perek, F. 2016. Recent change in the productivity and schematicity of the *way*-construction: a distributional semantic analysis. *Corpus Linguistics and Linguistic Theory 14*(1), 65-97.

Van Wettere, N. 2018. *Copularité et productivité: une analyse contrastive des verbes attributifs issus de verbes de mouvement en français et en néerlandais. .*

Zeldes, Amir. 2012. *Productivity in argument selection: From morphology to syntax*. Berlin: De Gruyter Mouton.

Corpus:

Jakubíček, M., A. Kilgarriff, V. Kovář, P. Rychlý, & V. Suchomel. 2013. The TenTen corpus family. In *7th International Corpus Linguistics Conference CL*, Lancaster, 125-127.

GHENT UNIVERSITY