



# Duurzaamheid in Tijden van Disruptie

Dirk Geeraerts



KU Leuven

RU Quantitative Lexicology and Variational Linguistics

## Uitgangspunt

*duurzaamheid* = een duurzame taalbeheersingspraktijk is evidence-based, d.w.z. gebaseerd op gedegen onderzoek en deugdelijke documentatie; 'degelijk duurt het langst'

*disruptie* = ingrijpende wijzingen in een bepaald (maatschappelijk) terrein als gevolg van technologische ontwikkelingen

i.c. verschuiving (zij het niet abrupt) van theoretische taalkunde naar een taalgebruiksgebaseerd paradigma

## Vraagstelling

(hoe) kan kwantitatief corpuslinguïstisch onderzoek van het type dat wij in onze onderzoeksgroep ontwikkelen bijdragen tot duurzaam taalbeheersingsonderzoek?

# Stappenplan

- 1 achtergrond: de evolutie van het taalkundige onderzoek
- 2 gevalstudies: twee onderwerpen met potentiële consequenties voor taalbeheersing
- 3 discussie: mogelijkheden en beperkingen

# Achtergrond

honderd jaar taalkunde in een timelapseversie:

# Achtergrond

	structurele taalkunde	formele taalkunde	usage-based taalkunde
stelsel vs gedrag			
sociaal vs psychisch			
homogeen vs variabel			

# Achtergrond

	structurele taalkunde	formele taalkunde	usage-based taalkunde
systeem vs gedrag	systeem		
sociaal vs psychisch	sociaal		
homogeen vs variabel	homogeen		

# Achtergrond

	structurele taalkunde	formele taalkunde	usage-based taalkunde
systeem vs gedrag	systeem	systeem	
sociaal vs psychisch	sociaal	psychisch	
homogeen vs variabel	homogeen	homogeen	



# Achtergrond

	structurele taalkunde	formele taalkunde	usage-based taalkunde
systeem vs gedrag	systeem	systeem	systeem & gedrag
sociaal vs psychisch	sociaal	psychisch	sociaal & psychisch
homogeen vs variabel	homogeen	homogeen	intern variabel

# Achtergrond

herkomst van 'usage-based turn'

1° theoretische push:

descriptieve armoede van formalisme

m.n. t.a.v. semantiek en pragmatiek

-> cognitief-functionele benaderingen

2° methodologische pull:

technologische ontwikkelingen (disruptie!)

m.n. digitaal tekstmateriaal en corpustools

-> corpuslinguïstiek

# Achtergrond

consequenties van 'usage-based turn'

de evolutie naar een taalgebruiksgebaseerde taalkunde  
(‘usage based grammar’) brengt fundamenteel en toegepast  
taalkundig onderzoek dichter bij elkaar

## Achtergrond, lokaal

langetermijnonderzoeksprogramma m.b.t. taal als sociale categorisatie

- Geeraerts, Grondelaers & Bakema 1994, *The Structure of Lexical Variation* (Mouton De Gruyter)
- Geeraerts 1997, *Diachronic Prototype Semantics* (OUP)
- Geeraerts, Grondelaers & Speelman 1999, *Convergentie en divergentie in de Nederlandse woordenschat* (Meertens)
- Geeraerts 2018, *Ten Lectures in Cognitive Sociolinguistics* (Brill)

# Theoretische basis: Formele en conceptuele onomasiologie

FORMAL LAYER

*breeches*

CONCEPTUAL LAYER

trousers ending  
just below the knee

DENOTATIONAL LAYER



# Theoretische basis: Formele en conceptuele onomasiologie

FORMAL LAYER

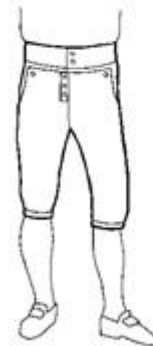
*knickers*

*breeches*

CONCEPTUAL LAYER

trousers ending  
just below the knee

DENOTATIONAL LAYER



# Theoretische basis: Formele en conceptuele onomasiologie

FORMAL LAYER

*trousers*

*knickers*

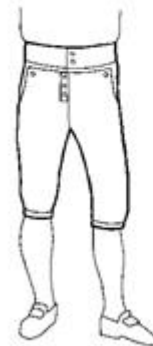
*breeches*

CONCEPTUAL LAYER

garment extending from the waist to the knee or ankle, covering each leg separately

trousers ending just below the knee

DENOTATIONAL LAYER



# Theoretische basis: Formele en conceptuele onomasiologie

FORMAL LAYER

*trousers*      *pants*

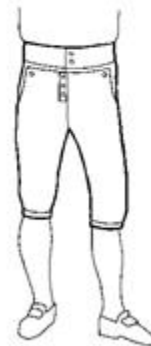
*knickers*      *breeches*

CONCEPTUAL LAYER

garment extending from the waist to the knee or ankle, covering each leg separately

trousers ending just below the knee

DENOTATIONAL LAYER





# Theoretische basis: Formele en conceptuele onomasiologie

FORMAL LAYER

*trousers*

*pants*

*knickers*

*breeches*

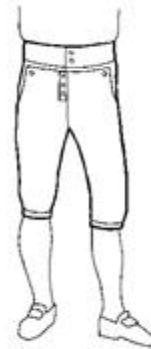
CONCEPTUAL LAYER

garment extending from the waist to the knee or ankle, covering each leg separately

underpants worn by women

trousers ending just below the knee

DENOTATIONAL LAYER



# Theoretische basis: Formele en conceptuele onomasiologie

semasiologie

onomasiologie

woorden

~ concepten

concepten

~ referenten

# Theoretische basis: Formele en conceptuele onomasiologie

	semasiologie	onomasiologie
woorden ~ concepten	polysemie	synonymie
concepten ~ referenten		

# Theoretische basis: Formele en conceptuele onomasiologie

	semasiologie	onomasiologie
woorden ~ concepten	polysemie	synonymie
concepten ~ referenten	prototypiciteit	

# Theoretische basis: Formele en conceptuele onomasiologie

	semasiologie	onomasiologie
woorden ~ concepten	polysemie	synonymie
concepten ~ referenten	prototypiciteit	categorisatie

# Theoretische basis: Formele en conceptuele onomasiologie

	semasiologie	onomasiologie
woorden ~ concepten	polysemie	formele onomasiologie
concepten ~ referenten	prototypiciteit	conceptuele onomasiologie

# Gevalstudie 1: Formele onomasiologische variatie

methodologische basis

1° behandel lexicale variatie als een sociolinguïstische variabele in de zin van Labov, en definieer een *onomasiologisch profiel* als een verzameling synoniemen met hun respectieve frequenties

2° kwantificeer de (veranderende) verhouding tussen lecten aan de hand van de mate van gelijkheid tussen hun onomasiologische profielen

# Gevalstudie 1: Formele onomasiologische variatie

replicatie Geeraerts, Grondelaers & Speelman 1999

*Convergentie en divergentie in de Nederlandse  
woordenschat* -> lexicale variatie in

- twee variëteiten: Belgisch Nederlands / Nederlands  
Nederlands

- vier tijdstippen: 1950 1970 1990 2010

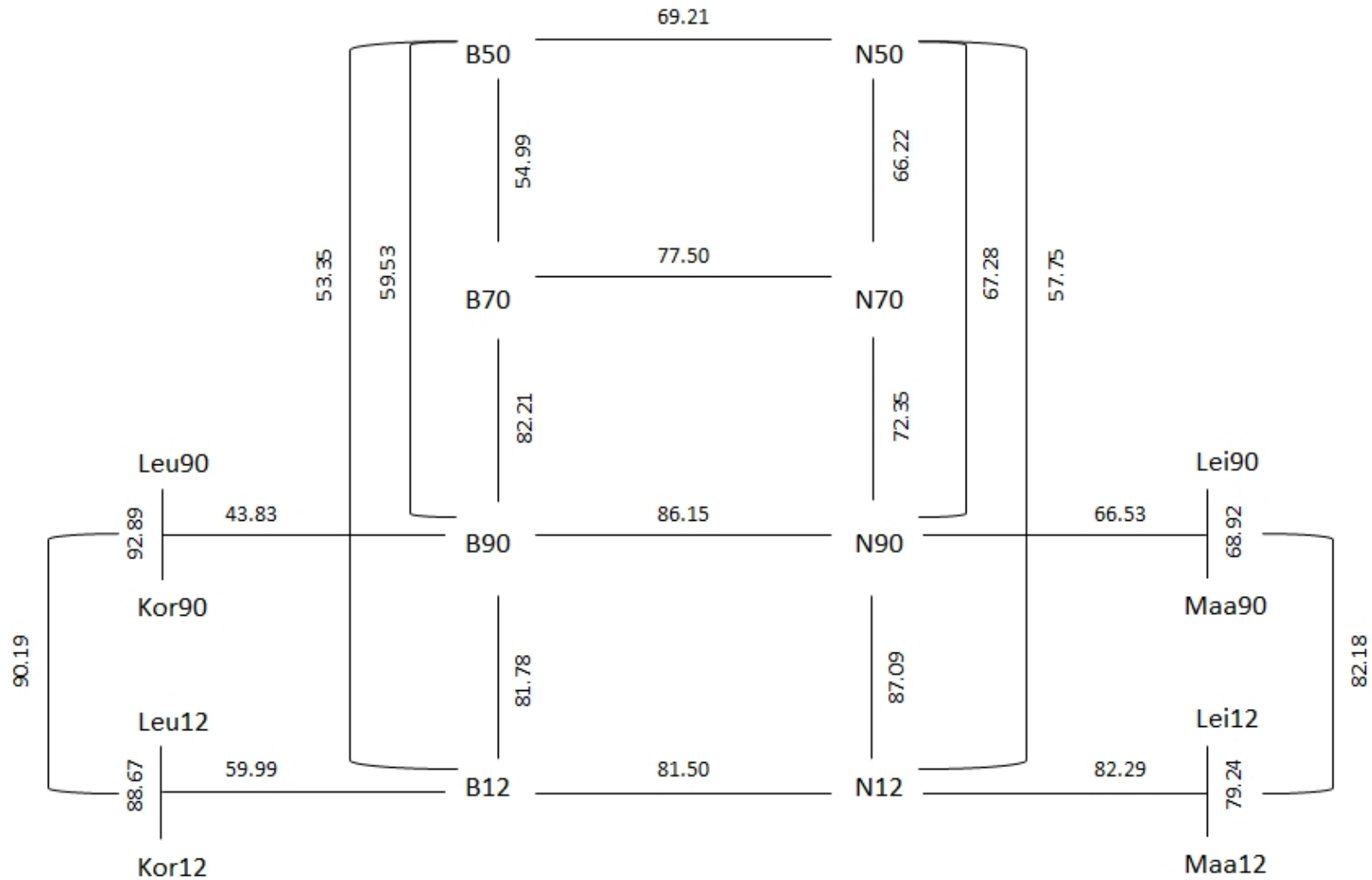
- twee stratificationele niveaus (in 1990 en 2010):  
magazines vs etalages

- veertien kledingconcepten (shirt<sub>M</sub>, shirt<sub>V</sub>, t-shirt<sub>MV</sub>,  
trui<sub>MV</sub>, broek<sub>MV</sub>, jeans<sub>MF</sub>, leggings<sub>V</sub>, rok<sub>V</sub>, jurk<sub>F</sub>, jasje<sub>M</sub>,  
jasje<sub>V</sub>, vest<sub>MV</sub>, pak<sub>MV</sub>)

(17865 +17398 Netherlandic Dutch observaties)



# Gevalstudie 1: Formele onomasiologische variatie



## Gevalstudie 1: Formele onomasiologische variatie

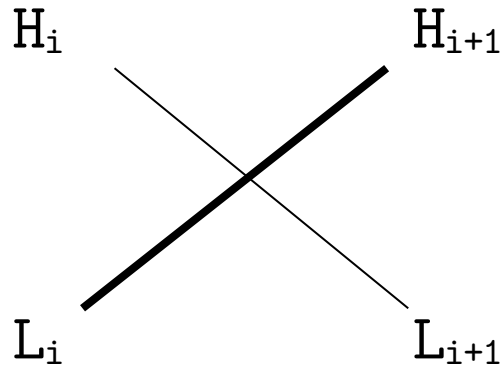
- lexicale toenadering tussen 1950 en 1990,  
met een terugval tussen 1990 en 2010

- groeiende interne homogeniteit van de twee  
hoofdvariëteiten

-> in welke richting verloopt de interne homogenisering?  
groeit hoog naar laag toe of omgekeerd?

# Gevalstudie 1: Formele onomasiologische variatie

b.v. er is een verschuiving van hoog naar laag indien  
 $U(L_i, H_{i+1}) > U(H_i, L_{i+1})$



# Gevalstudie 1: Formele onomasiologische variatie

Nederland

$$U'(\text{NL}_{1990}, \text{NH}_{2012}) = 59.23$$

$$U'(\text{NH}_{1990}, \text{NL}_{2012}) = 85.54$$

Vlaanderen

$$U'(\text{BL}_{1990}, \text{BH}_{2012}) = 56.22$$

$$U'(\text{BH}_{1990}, \text{BL}_{2012}) = 52.29$$

-> historische interpretatie in functie van Vlaamse taalpolitiek en groeiende autonomie

## Gevalstudie 2: Conceptuele onomasiologische variatie

methodologische basis

1° beantwoord de 'Big Data challenge' door betekenis te modelleren aan de hand van de elementen in de context van een woord

2° pas deze distributioneel-semantische aanpak niet alleen toe op lexemen (synonymiedetectie) maar ook op individuele attestaties (polysemiedetectie)

## Gevalstudie 2: Conceptuele onomasiologische variatie

voorbeeld:

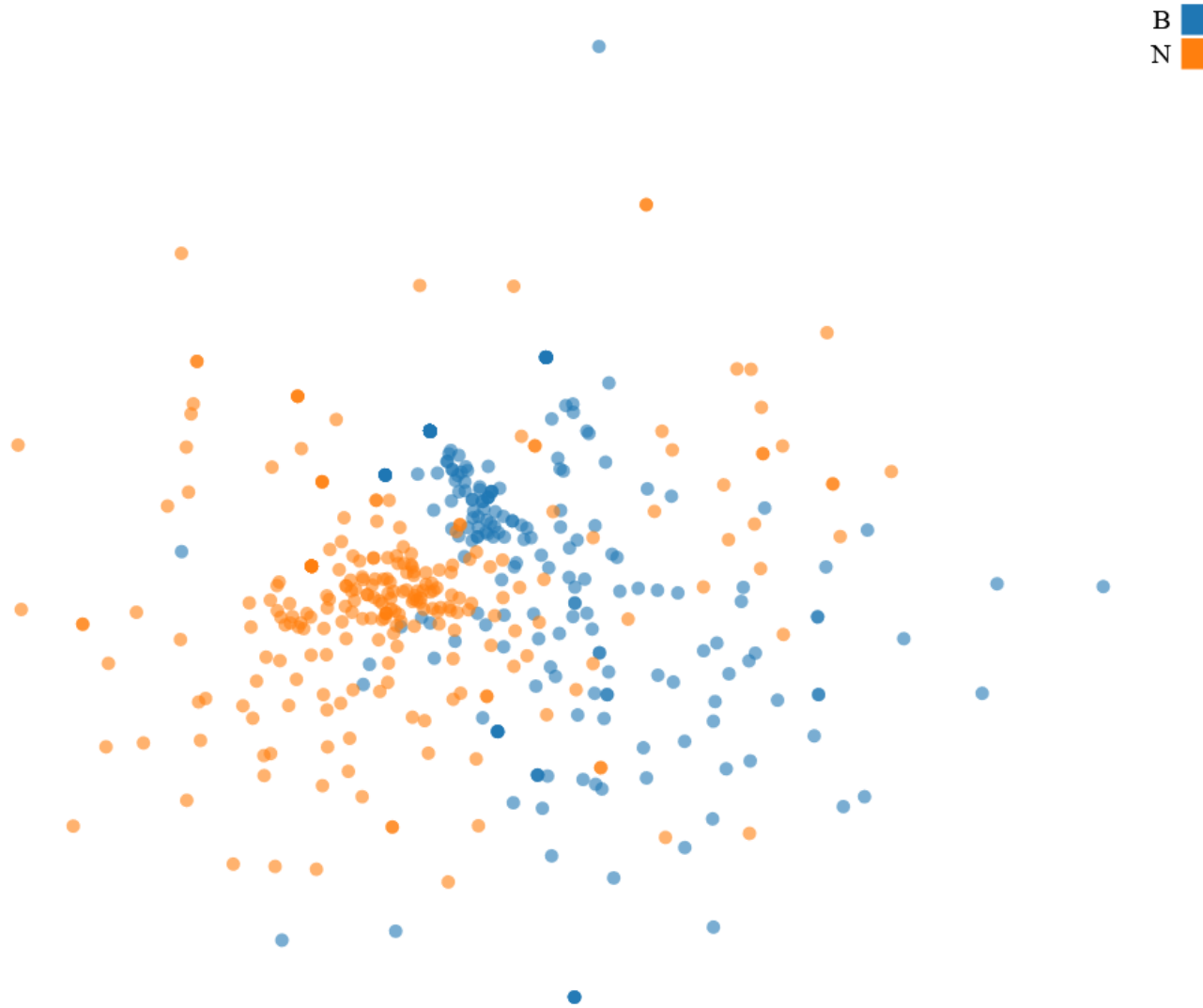
*kleedje* NL 'tapijtje' ~ kleurig, tafel, spreiden

*kleedje* VL 'jurkje' ~ winkel, atelier, kopen  
(en fig. in een nieuw kleedje)

een distributionele methode detecteert de  
betekenisverschillen

# Gevalstudie 2: Conceptuele onomasiologische variatie

---



## Gevalstudie 2: Conceptuele onomasiologische variatie

verandert de verhouding tussen de bijna-synoniemen  
*allochtoon / migrant* ?

studie van Kris Heylen op LeNC (1999–2005)

1° bouw een matrix met collocatiewaarden waaruit de bijna-synonymie blijkt

	<i>jobs</i>	<i>racisme</i>	<i>integratie</i>	<i>misdaad</i>	<i>stemrecht</i>	<i>suiker</i>	<i>zon</i>	<i>hond</i>
allochtoon	5.3	7.9	6.5	4.0	0.8	0.6	0.0	0.0
migrant	4.3	8.1	5.7	3.2	6.2	0.5	0.0	0.1



## Gevalstudie 2: Conceptuele onomasiologische variatie

2° pas de distributionele methode vervolgens toe op de sterkste collocaten, om de onderlinge samenhang daarvan te bepalen

-> een dozijn categorisatiedimensies: o.a. ARBEIDSMARKT, ILLEGALITEIT, STEMRECHT, ONDERWIJS

## Gevalstudie 2: Conceptuele onomasiologische variatie

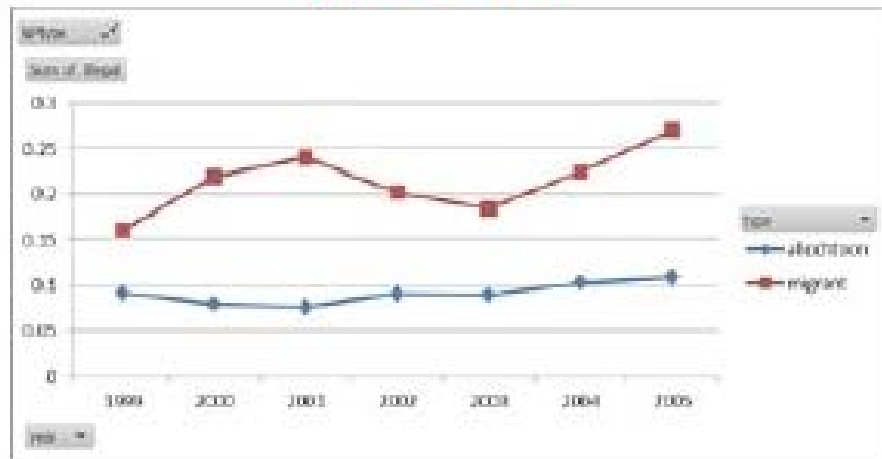
3° bouw een matrix met doelwoorden en dimensies, met als meetpunten combinaties van jaren en tekstsoorten (kwaliteitskranten / populaire kranten), en breng dan de associatiesterkte tussen doelwoorden en dimensies in kaart

# Gevalstudie 2: Conceptuele onomasiologische variatie

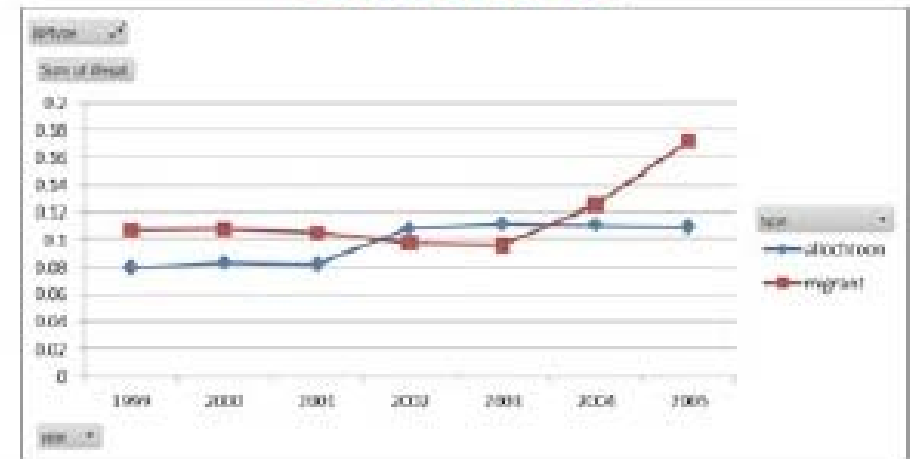
aanwijzingen voor differentiatie:

*migrant* wordt sterker op de dimensie ILLEGALITEIT

## QUALITY



## POPULAR

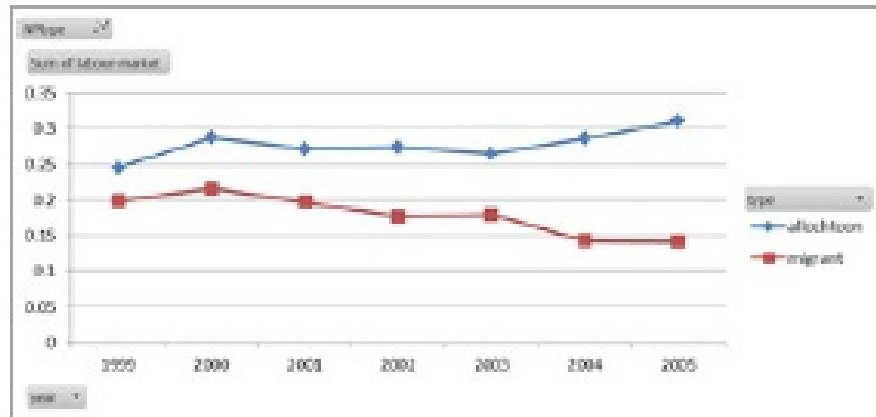


# Gevalstudie 2: Conceptuele onomasiologische variatie

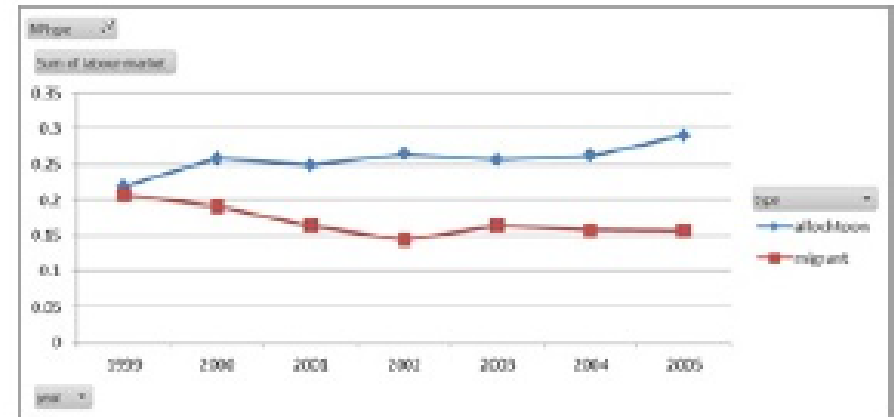
aanwijzingen voor differentiatie:

*allochtoon* wordt sterker op de dimensie ARBEIDSMARKT

## QUALITY NP



## POPULAR NP



## Discussie

onderzoek van dit type biedt antwoord op vragen die relevant zijn vanuit taalbeheersingsperspectief

zowel t.a.v. individuele taalelementen

is *allochtoon* een negatief geladen term?

is *kleedje* 'jurk' ABN of ABN?

als t.o.v. taalvariëteiten en -veranderingen

groeien VL en NL van elkaar weg?

wordt het taalgebruik informeler?

## Discussie

in een taalgebruiksgebaseerd kader is een grotere convergentie van fundamenteel en toegepast onderzoek mogelijk, maar er zijn ook stoorfactoren:

Directieve Driestheid

Documentair Deficit

Disruptieve Dreiging

# Discussie

## Directieve Driestheid

duurzaam taaladvies neigt naar duidelijkheid en vereenvoudiging, maar duurzaam onderzoek laat complexiteiten en nuances zien

-> kan de praktijk voldoende recht doen aan de complexe realiteit?

# Discussie

## Documentair Deficit

een volledig beeld van de ontwikkelingen vraagt meer dan wat we tot nu toe weten

bv. het VL/NL-onderzoek:

slechts één lexicaal veld -> andere velden, en niet alleen woordenschat

uitsluitend productie -> ook perceptie- en attitudeonderzoek



# Discussie

## Disruptieve Dreiging

de technologische veranderingen creëren mogelijkheden voor onderzoek, maar zorgen ook voor communicatieve verschuivingen die het onderzoek kunnen bemoeilijken

- een onoverzichtelijke overvloed aan materiaal met (vermoedelijk) snellere wijzigingen in het taalgebruik
- veranderingen in de aard van CMC die vergelijkbaarheid in de tijd bemoeilijken (IRC, WhatsApp, Twitter ...)

Bij wijze van conclusie

Duurzaam duurt het langst.

dirk.geeraerts@kuleuven.be

