

Wat dragen we vandaag: een hemd met blazer of een shirt met jasje?

Convergentie en divergentie binnen Nederlandse kledingtermen

Jocelyne Daems, Kris Heylen & Dirk Geeraerts

MEM 67 (2): 307–342

DOI: 10.5117/TET2015.2.DAEM

Abstract

What to wear today: *een hemd met blazer* ('a dress shirt with suit jacket') or *een shirt met jasje* ('a dress shirt with suit jacket')? Convergence and divergence in Dutch clothing terminology

This paper reports on a corpus-based investigation into naming preferences in Belgian Dutch and Netherlandic Dutch for fourteen clothing terms. The study is a follow-up of Geeraerts, Grondelaers and Speelman (1999), in which soccer and clothing terminology from 1950, 1970 and 1990 was analysed as an indicator of standardisation in Dutch. This study extends the clothing corpus with new, comparable data from 2012 collected from magazines and shop windows. A profile-based measure of linguistic uniformity quantifies the differences in naming preferences across the 14 concepts between different varieties of Dutch. The results shed new light on the current linguistic situation in the Low Countries. The diachronic convergence between Belgian Dutch and Netherlandic Dutch found in Geeraerts, Grondelaers and Speelman (1999) seems to have come to a halt in present-day Dutch. On the other hand, the recent data confirm that the distance between the language in the lower register shop windows and the standard language in magazines remains largest in Belgian Dutch.

Keywords: profile-based uniformity measure, lexical variation, diachronic variation, register variation, onomasiology, pluricentricity

1 Inleiding¹

Eind vorige eeuw presenteerden Geeraerts, Grondelaers en Speelman (1999) een veelzijdige studie naar de verhouding tussen het Nederlands in België en dat in Nederland². Lexicale variatie bij vijftien kledingconcepten en evenveel voetbalbegrippen werd kwantitatief bestudeerd om na te gaan of er diachroon sprake was van convergentie of divergentie tussen de twee taalvariëteiten. Hiertoe werd krant- en magazinemateriaal uit 1950, 1970 en 1990 verzameld. Tegelijk werd de contemporaine interne taalstructuur van het Belgisch Nederlands en het Nederlands Nederlands vergeleken door het bovengewestelijke schriftelijke standaardtaalgebruik, gerepresenteerd door de kledingtijdschriften, te plaatsen naast dat van het regionaal gekleurde substandaardtaalniveau, zoals te vinden op prijskaartjes en labels in winkelatalages. Een derde pijler van de studie vertrok van attitudele gegevens om het effect van normatieve taalzorg na te gaan en onderzocht daarbij eveneens de waarde van geëliciteerde data naast feitelijk taalgebruik. Methodologisch bestond de uitdaging erin om een nieuwe maat in te voeren om de onderzoeksvragen op een kwantitatief eenduidige wijze te beantwoorden. De hiertoe ontwikkelde profielgebaseerde uniformiteitsmaat laat toe om van geobserveerd taalgebruik te vertrekken om de gestandaardiseerdheid van de taalsituatie in kwestie te onderzoeken. Het achterliggende idee is immers dat uit een gemeenschappelijke linguïstische norm, hier de Nederlandse standaardtaal, gemeenschappelijk taalgebruik volgt. Enerzijds steunt de profielgebaseerde uniformiteitsmaat op het idee dat de overlapping in woordkeuze bij het uitdrukken van een concept tussen twee taalvariëteiten aanduidt hoe uniform die laatste zijn. Anderzijds laat de maat ook toe om te aggregeren over verschillende concepten zodat algemenere uitspraken over lexicale variatie tussen taalvariëteiten mogelijk worden. De twee belangrijkste empirische vaststellingen die volgden uit Geeraerts, Grondelaers en Speelman (1999) zijn de diachrone lexicale convergentie tussen het Nederlands Nederlands en het Belgisch Nederlands, grotendeels op het conto van die laatste, en de synchroon grotere afstand tussen de onderzochte registers in Nederlandstalig België (zie de laatste alinea van sectie 2 voor een meer uitgebreid overzicht van de resultaten).

Sindsdien hebben verschillende studies voortgebouwd op deze basis. Grondelaers et al. (2001b) hernemen enerzijds het kledingtermenonderzoek en voegen een nieuw corpus van internetmateriaal toe als bijkomende vertegenwoordiging van het meer informele niveau. Anderzijds stappen de onderzoekers over van lexicale variatie bij inhoudswoorden

naar die bij functiewoorden. In Grondelaers et al. (2001a) wordt dieper ingegaan op het gebruik van internetmateriaal bij onderzoek naar lexicale standaardisering. Zenner, Speelman en Geeraerts (2012) gebruiken de uniformiteitsberekening om het succes van Engelse leenwoorden na te gaan bij persoonsaanduidende nomina. Daarnaast zijn er verscheidene studies waarbij de vergelijking van verschillende methodes centraal staat, en vorderingen in het onderzoek naar automatische synoniem- en betekenisbepaling laten toe dat de schaal van het onderzoek aanzienlijk vergroot wordt, mede door de toenemende beschikbaarheid van grote corpora. Speelman, Grondelaers en Geeraerts (2003; 2006) vergelijken aan de hand van vijftien profielen (een aangepaste versie van) de profielgebaseerde uniformiteitsmaat uit Geeraerts, Grondelaers en Speelman (1999)³ met niet-profielgebaseerde methodes met als doel om taalvariëteiten zo goed mogelijk te classificeren. Met hetzelfde doel, maar om andere methodes te vergelijken, vertrekken Ruetten, Speelman en Geeraerts (2011) van 1455 profielen met telkens een typisch Belgisch-Nederlands gemarkeerde variant en Heylen en Ruetten (2013) van 218 semiautomatisch geselecteerde concepten.

Deze vervolgstudie keert terug naar de oorsprong en neemt twintig jaar later de draad van de kledingtermen weer op. Voor veertien kledingconcepten zijn opnieuw verscheidene Belgische en Nederlandse (mode)tijdschriften doorzocht en winkelatalages bezocht⁴. De diachrone dimensie wordt zo uitgebreid met een nieuwe sprong van een twintigtal jaar. De convergerende trend tussen Belgisch Nederlands en Nederlands Nederlands kan nu verder onderzocht worden met data uit het tweede decennium van de eenentwintigste eeuw. Daarbovenop, terwijl er in Geeraerts, Grondelaers en Speelman (1999) enkel synchrone data voorhanden waren om variatie binnen respectievelijk het Belgisch en het Nederlands Nederlands te onderzoeken, beschikken we met de toevoeging van het nieuwe etalagemateriaal ook over een diachroon stratificationeel perspectief tussen 1990 en 2012.

Zoals bij Geeraerts, Grondelaers en Speelman (1999) moet dit lexicale variatieonderzoek gezien worden in het licht van de taalpolitieke geschiedenis van het Nederlands in België en de positie die het daarbij inneemt ten opzichte van het Nederlands Nederlands. Voor een uitgebreide beschrijving van de historische en taalpolitieke achtergrond bij het standaardiseringsproces van het Belgisch Nederlands verwijzen we naar Geeraerts, Grondelaers en Speelman (1999, pp. 12-29) (maar zie ook Van Hoof en Jaspers 2012 voor een kritisch overzicht van de standaardiseringsacties in Vlaanderen in de periode 1950-1980). Het volstaat in deze context om te herhalen dat het Belgisch Nederlands een standaardiseringsachterstand

heeft opgelopen onder meer door de invloed van het Frans en dat de taalpolitiek vanaf de jaren zestig van de vorige eeuw bewust aansluiting met de Nederlands-Nederlandse norm nastreefde. Die integrationistische visie staat in contrast met het particularisme, waar er ruimte is voor een eigen Belgisch-Nederlandse standaardtaal. In deze context kan niet onvermeld blijven dat het idee van een standaardiseringsachterstand van het Belgisch Nederlands betwist wordt (Van Hoof & Jaspers 2012, pp. 97-98), of alleszins als problematisch beschouwd wordt (Grondelaers & van Hout 2011a, p. 200). Die laatsten wijzen er immers op dat de beschrijving van de standaardisering van het Belgisch Nederlands in het gedrang komt wanneer de Nederlands-Nederlandse situatie als ijkpunt genomen wordt, doordat beide nationale variëteiten historisch zeer verschillende standaardiseringsprocessen kennen (2011a, p. 200).

Het onderzoek van Geeraerts, Grondelaers en Speelman (1999) neemt de tweedeling, integrationisme tegenover particularisme, als vertrekpunt, maar houdt een descriptieve benadering aan om na te gaan bij welke visie de feitelijke data het best aansluiten. Hetzelfde geldt overigens voor dit artikel. Met de empirische bevindingen in deze studie beogen we niet om uitspraken te doen over de historische standaardiseringsprocessen van het Belgisch en Nederlands Nederlands. Wel gaan we van naderbij bekijken welke verschillen er tussen deze twee nationale variëteiten bestaan in de kledingwoordenschat. Zo zal onder meer blijken dat in de laatste twintig jaar er in het Belgisch Nederlands andere verschuivingen plaatsvinden dan in het Nederlands Nederlands.

In sectie 2 van dit artikel wordt het onderzoek van Geeraerts, Grondelaers en Speelman (1999) samengevat, in het bijzonder het deel met betrekking tot het kledingmateriaal. We lichten ook uitvoerig de profielgebaseerde uniformiteitsmaat toe. Daarna wordt in sectie 3 het huidige vervolgonderzoek voorgesteld en enkele methodologische aanpassingen toegeleucht. Eveneens wordt er uitgebreid ingegaan op de huidige resultaten en wordt de vergelijking gemaakt met de originele studie. Sectie 4 is gewijd aan de conclusies en we bespreken verdere uitbreidingen in het studiedomein.

2 Corpusgebaseerd onderzoek naar convergentie en divergentie in de Nederlandse kledingwoordenschat

Het onderzoek in Geeraerts, Grondelaers en Speelman (1999) vertrekt ruwweg van twee basishypotheses. De diachrone convergentiehypothese

steunt op de verwachting dat in de loop van de standaardiseringsbeweging van het Belgisch Nederlands de twee nationale variëteiten convergeren als gevolg van de expliciete normatieve oriëntatie op het Nederlandse Nederlands, kortom de integrationistische visie. Daarnaast veronderstelt de synchrone stratificatiehypothese een grotere afstand tussen het regionale en supraregionale niveau in België door de veronderstelde standaardisering-sachterstand. Immers, de bovengewestelijke standaardtaal in het Belgisch Nederlands is gebaseerd op een exogene norm en is niet het gevolg van een natuurlijk proces dat vanuit een (uitverkoren) dialect begonnen is.

Om na te gaan in welke mate er uniformiteit is tussen en binnen het Belgisch Nederlands en het Nederlands Nederlands werd een gebalanceerd corpus samengesteld met een geografische, diachrone en stratificationele dimensie. Voor de synoniemen van vijftien kleding- en vijftien voetbalconcepten werden observaties uit Belgische en Nederlandse kranten en tijdschriften uit 1950, 1970 en 1990 verzameld. Daarbovenop werden voor het kledingmateriaal benamingen op labels en etiketten uit etalages in twee Belgische en twee Nederlandse steden opgenomen. Leuven en Leiden gold- en hierbij als vertegenwoordiging van respectievelijk de linguïstische centra Brabant en de Randstad, terwijl Kortrijk en Maastricht dan weer gelegen zijn in economische centra met een secundaire academische functie, en zich daarenboven bevinden in de periferie van het taalgebied. In tegenstelling tot de benamingen in tijdschriften, werden deze labels niet geredigeerd door taalprofessionals en is de relatie met de potentiële koper veel directer. Het lezerspubliek van het etalagemateriaal is bovendien ook beperkter dan dat van het tijdschriftenmateriaal, maar de ruime regionale commerciële functie van de geselecteerde steden belet dat we te maken hebben met zuiver dialectisch taalgebruik. Dit semiformele schriftelijk brontype vormt zo een aanvulling op het nationaal gerichte tijdschriftenmateriaal binnen het stratificationele continuüm van het Nederlands. De twee gekozen lexicale velden bestaan bovendien telkens uit een representatieve selectie van basisbegrippen binnen het veld. Een gedetailleerde omschrijving van de materiaalverzameling is te vinden in Geeraerts, Grondelaers en Speelman (1999, pp. 30-35; pp. 162-164), voor een summier overzicht verwijzen we naar onderstaande tabel:

Tabel 1 Kwantitatief overzicht per lexicaal veld, land en brontype in aantal woorden

			België	Nederland	totaal
Voetbal	1950	magazine/krant	3283	3220	
	1970		4529	1604	
	1990		4850	4215	21.801
Kleding	1950	magazine/krant	2210	3555	
	1970		3858	2972	
	1990		2450	3427	18.472
centrum (Leuven/Leiden)		etalage	3289	1162	
periferie (Kortrijk/Maastricht)			3299	1694	9444
Totaal			27.768	21.849	49.617

Om de lexicale convergentie tussen het Belgisch Nederlands en het Nederlands Nederlands te meten, introduceerden Geeraerts, Grondelaers en Spielman (1999) de profielgebaseerde uniformiteitsmaat. Hierbij is het onomasiologisch profiel van een bepaald concept in een gegeven bron het geheel van alternatieve benamingen voor dat concept uitgedrukt in de frequentieverhouding in die bron. Onderstaande tabel⁵ geeft de onomasiologische profielen weer voor het concept OVERHEMD (MAN) in het Belgische en Nederlandse standaardtaalmateriaal uit 1990 (respectievelijk B90 en N90 genoemd):

Tabel 2 Onomasiologische profielen voor OVERHEMD (MAN) in het Belgische en Nederlandse materiaal uit 1990

overhemd (man)	B90	N90
<i>hemd</i>	31%	17%
<i>overhemd</i>	69%	46%
<i>shirt</i>	0%	37%

De uniformiteitsmaat drukt uit in hoeverre twee onomasiologische profielen overeenkomen. We spreken van maximale uniformiteit wanneer twee taalvariëteiten, door een gedeelde norm, voor eenzelfde begrip dezelfde benaming gebruiken. Het is ook mogelijk dat verschillende benamingen gebruikt worden voor een gegeven concept, maar dan moet de distributie hiervan identiek zijn in de beide variëteiten. Wanneer dit niet het geval is, wanneer de alternatieve lexicalisering binnen een profiel slechts gedeeltelijk hetzelfde gedistribueerd zijn, is er sprake van partiële uniformiteit.

Om de uniformiteit voor een bepaald concept te berekenen, meten we

de overlapping in alternatieve benamingen tussen twee profielen. Anders gezegd, voor iedere alternatieve lexicalisering nemen we de kleinste relatieve frequentie van de twee profielen (gemarkeerd in grijs in Tabel 2) en hiervan berekenen we de som om tot de uniformiteitsmaat te komen. Voor het concept OVERHEMD (MAN) houdt dat in dat de keuze voor de variant *hemd* in 17% van de gevallen overlapt tussen België en Nederland. Voor *overhemd* delen Belgisch Nederlands en Nederlands Nederlands 46% voorkeur en doordat in België *shirt* niet gebruikt wordt om OVERHEMD (MAN) uit te drukken, is er daar geen overlapping. Kortom, voor het begrip OVERHEMD (MAN) in Belgische en Nederlandse standaardtalige bronnen uit 1990 vinden we (17% + 46% + 0% =) 63% uniformiteit. Omgezet in formulevorm ziet deze intuïtieve berekening er zo uit:

$$U_Z(Y_1, Y_2) = \sum_{i=1}^n \min(F_{Z, Y_1}(x_i), F_{Z, Y_2}(x_i))$$

We kijken dus naar de mate van onomasiologische uniformiteit U tussen de bronnen, of deelverzamelingen, Y_1 en Y_2 voor concept Z dat uitgedrukt wordt door de alternatieve benamingen x_1 tot x_n . Deze formule drukt uit dat de graad van uniformiteit berekend wordt door het minimum te nemen van de relatieve frequentie van variant x_i voor concept Z in de materiaalverzameling Y_1 , weergegeven als $F_{Z, Y_1}(x_i)$, en van dezelfde variant voor hetzelfde concept in de deelverzameling Y_2 . Dezelfde berekening wordt herhaald voor alle varianten x_i die concept Z aanduiden en vervolgens worden alle minima opgeteld. Merk op dat voor mogelijke verschillen in conceptfrequentie tussen de deelverzamelingen gecontroleerd wordt door met de relatieve in plaats van de absolute frequenties te werken.

Om te aggregeren over verschillende concepten en te komen tot de gemiddelde uniformiteit voor een lexicaal veld wordt het gemiddelde genomen van de uniformiteitsgraad van alle concepten Z_i in het veld, zoals onderstaande formule illustreert:

$$U(Y_1, Y_2) = \frac{1}{n} \sum_{i=1}^n U_{Z_i}(Y_1, Y_2)$$

Tot slot is er nog een aanvulling op deze grondformule die het relatieve gewicht van het concept in het veld mee in rekening brengt. Immers, in communicatief opzicht is het logisch dat hoogfrequente concepten sterker doorwegen bij het meten van de uniformiteit van een lexicaal veld. De gewogen uniformiteit U' neemt de som van de uniformiteitswaarde van

elk concept vermenigvuldigd met de wegingsfactor G voor dat concept. Die laatste is de relatieve frequentie van concept Z_i in de gehele deelverzamelingen Y_1 en Y_2 . De formule voor de gewogen uniformiteit wordt dan:

$$U'(Y_1, Y_2) = \sum_{i=1}^n U_{Z_i}(Y_1, Y_2) \cdot G_{Z_i}(Y_1 \cup Y_2)$$

Naast de onomasiologische uniformiteit introduceerden Geeraerts, Grondelaers en Spielman (1999) nog twee maten die toelaten een beter inzicht te krijgen in het materiaal. De interne uniformiteit drukt uit in hoeverre in een deelverzameling voor een bepaald concept eenzelfde benaming wordt gekozen. Twee factoren beïnvloeden de interne uniformiteit, het aantal alternatieve termen en de dominantie van een bepaalde term. De interne uniformiteit is de som van de relatieve frequentie in het kwadraat⁶ van ieder van de alternatieve benamingen voor een concept:

$$I_Z(Y) = \sum_{i=1}^n F_{Z,Y}(x_i)^2$$

Voor het voorbeeld van OVERHEMD (MAN) uit Tabel 2 houdt dit in dat de interne uniformiteit voor het concept in het Belgisch Nederlands ($0.31^2 + 0.69^2 =$) 57,22% is en in het Nederlands Nederlands slechts ($0.17^2 + 0.46^2 + 0.37^2 =$) 37,74%. Deze maat drukt met andere woorden uit in hoeverre er een dominante term is om het lexicale concept uit te drukken. Om dan de interne uniformiteit van een lexicaal veld in een bepaalde bron te berekenen nemen we het gemiddelde van de interne uniformiteitswaarde van alle concepten in het veld:

$$I(Y) = \frac{1}{n} \sum_{i=1}^n I_{Z_i}(Y)$$

Net als bij de (externe) uniformiteitsmaat speelt de conceptfrequentie een rol en wordt de wegingsfactor geïntroduceerd om een accurater beeld te krijgen van de interne uniformiteit binnen het lexicale veld:

$$I'(Y) = \sum_{i=1}^n I_{Z_i}(Y) \cdot G_{Z_i}(Y)$$

Tot slot, om het relatieve aandeel van termen met een bepaald kenmerk (zoals 'normatief afgekeurd' of 'van Franse oorsprong') te berekenen in het

bronmateriaal, is de aandeelmaat A geïntroduceerd. Bij deze maat staat K voor de verzameling van alle lexicaliseringën met dat kenmerk. De wegingsfactor W, een waarde van 0 tot 1, bepaalt voor een individuele variant in welke mate het lidmaatschap tot K moet doorwegen. Zo is *jeans* nog steeds ontegensprekelijk herkenbaar als van Engelse oorsprong, terwijl bij *bloes* de afstand tot het Frans iets groter is. Dat onderscheid uit zich in verschillende wegingen bij de berekening van de aandeelmaat⁷. De onderstaande formule toont vervolgens dat de aandeelmaat berekend wordt door de som van de relatieve frequentie van iedere lexicalisering voor een concept vermenigvuldigd met zijn specifieke wegingsfactor:

$$A_{K,Z}(Y) = \sum_{i=1}^n F_{Z,Y}(x_i) \cdot W_{x_i}(K)$$

Als we dan opnieuw gaan kijken wat dat concreet betekent voor het concept OVERHEMD (MAN) uit Tabel 2, vinden we dat enkel de variant *shirt* een Engelse term is. Voor het Belgisch Nederlands is het aandeel voor het Engels dus 0%, want er zijn geen observaties van *shirt*, en voor het Nederlands-Nederlandse materiaal is het 37%. We berekenen vervolgens het gemiddelde om te komen tot het aandeel van een bepaald taalkundig kenmerk in een lexicaal veld:

$$A_K(Y) = \frac{1}{n} \sum_{i=1}^n A_{K,Z_i}(Y)$$

Ook hier is er een gewogen alternatief, waarbij de frequentie van het concept Z in de gehele deelverzameling Y meegerekend wordt:

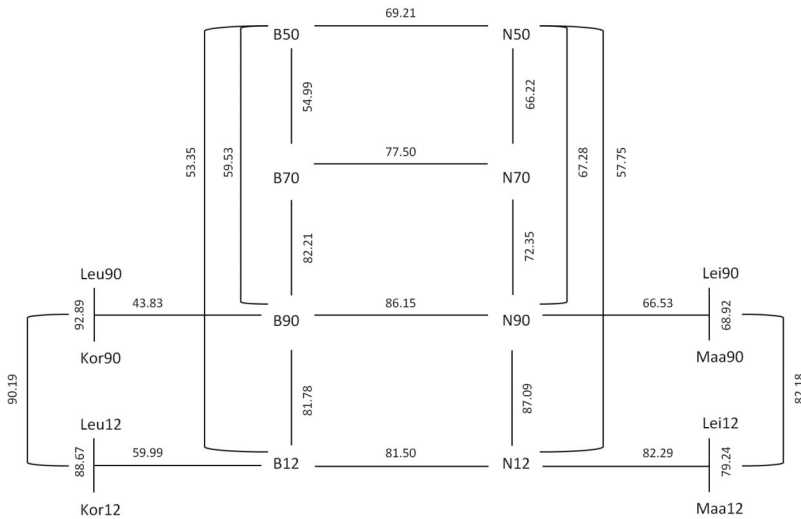
$$A'_K(Y) = \sum_{i=1}^n A_{K,Z_i}(Y) \cdot G_{Z_i}(Y)$$

Vanuit deze systematische methodologische onderbouw is het een kleine stap om de diachrone convergentiehypothese en de synchrone stratificatiehypothese te operationaliseren in termen van uniformiteit. De convergerende beweging tussen het Belgische en Nederlandse Nederlands manifesteert zich in een stijgende uniformiteit over de tijdspanne van 1950 tot 1990. De grotere synchrone afstand tussen standaardtaal en substandaardtaal in België laat zich kwantificeren als een lagere uniformiteit tussen het tijdschrift- en het winkelalagemateriaal in België dan in Nederland. We

geven hierbij de voornaamste resultaten uit Geeraerts, Grondelaers en Speelman (1999) weer, relevant voor dit vervolgonderzoek.

De data in Figuur 1⁸ tonen dat er in de tijdspanne van 1950 over 1970 tot 1990 onmiskenbaar sprake is van lexicale convergentie tussen Belgisch Nederlands en Nederlands Nederlands. De convergerende trend die waargenomen wordt, is bovendien voornamelijk te wijten aan een normatieve oriëntatie van het Belgisch Nederlands op het Nederlands Nederlands: typisch Nederlands-Nederlandse termen worden overgenomen ten koste van typisch Belgisch-Nederlandse uitdrukkingen. Verder geldt algemeen dat de interne uniformiteit van de concepten toeneemt: een dominante term binnen het profiel krijgt duidelijk de voorkeur op de onomasiologische alternatieven. Verder, zowel het Belgisch als het Nederlands Nederlands worden gekenmerkt door een toenemende invloed van Engelse kledingtermen. Voor het aandeel Franse termen wordt de ontwikkeling van het Belgisch Nederlands gekenmerkt door een duidelijk observeerbare afname als historische reactie tegen de Franse invloed, terwijl in het Nederlands-Nederlandse materiaal die afwijzing afwezig is. Opmerkelijk is wel dat het gebruik van Franse termen in Nederland onderhevig lijkt te zijn aan modegrillen: het aandeel “chique” Franse kledingtermen daalt aanzienlijk in 1970, wat toegeschreven kan worden aan de informele tijdsgeest van toen (Geeraerts en Grondelaers, 1999, p. 62). Tot slot, de veronderstelde vertraagde standaardisering van het Belgisch Nederlands wordt weerspiegeld door een beduidend grotere afstand tussen de supraregionale tijdschriftobservaties en het regionale winkelalagemateriaal dan het geval is bij het Nederlands Nederlands.

Hoewel deze bevindingen steunen op een stevige empirische basis, blijft voorzichtigheid geboden bij de interpretatie ervan. Zo steunt het onderzoek op slechts twee lexicale velden en zijn de bevindingen dus niet zonder meer extrapolieerbaar naar andere woordvelden of woordsoorten. Ook de tweeledige stratificatie binnen het Nederlands in het onderzoek is slechts een beperkte representatie van de werkelijke registerdiversiteit (zie sectie 1 voor vervolgstudies die op deze twee punten uitbreidingen hebben voorgesteld). Tot slot, de omvang van het manueel aangelegde corpus laat niet toe om bepaalde statistische testen uit te voeren (met name de lineaire regressieanalyse). Voor de vergelijkingen van uniformiteitsindices wordt dus veiligheidshalve gewerkt met een foutenmarge van 5% alvorens met zekerheid gezegd kan worden of een uniformiteitswaarde groter al dan niet kleiner is dan een andere.



Figuur 1 Gewogen uniformiteitsverhoudingen binnen het kledingmateriaal

Bron: Geeraerts, Grondelaers en Speelman, 1999, p. 69

3 Convergentie en divergentie binnen Nederlandse kledingtermen anno 2012

Om de vraag te beantwoorden of de synchrone stratificatiehypothese en de diachrone convergentiehypothese twintig jaar later nog steeds standhouden, werd het initiële corpus uitgebreid. De originele samenstelling werd hierbij zo goed mogelijk gerespecteerd. Voor het standaardtaalmateriaal werd gezocht naar observaties van kledingtermen in mode- en lifestylemagazines uitgegeven in 2012. Twee aanpassingen waren hierbij nodig. Zo waren sommige van de oorspronkelijk gebruikte tijdschriften niet langer in de handel verkrijgbaar en waren niet alle tijdschriften van de maanden januari en februari 2012 verkrijgbaar (zie Appendix 3 voor een overzicht). Voor het etalagemateriaal werden opnieuw winkelatalages in dezelfde twee Belgische en twee Nederlandse universiteitssteden uitgekamd⁹. Anders dan in de oorspronkelijke studie hebben we het concept MANTELPAKJE (VROUW) achterwege gelaten omwille van dataschaarste. Dit houdt eveneens in dat voor het materiaal uit Geeraerts, Grondelaers en Speelman (1999) alle berekeningen opnieuw uitgevoerd zijn voor veertien in plaats van vijftien concepten. Tabel 3 toont de samenstelling in aantal observaties voor veertien kledingconcepten in alle deelverzamelingen:

Tabel 3 Kwantitatief overzicht voor de kledingtermen per land en brontype in aantal woorden

	België	Nederland	totaal
1950	1918	3114	
1970	3389	2824	
1990	2347	3412	17.004
centrum (Leuven/Leiden)	3240	1141	
periferie (Kortrijk/Maastricht)	3210	1652	9243
2012	2844	4363	7207
centrum (Leuven/Leiden)	504	355	
periferie (Kortrijk/Maastricht)	413	537	1809
totaal	17.865	17.398	35.263

Voor het testen van de hypotheses¹⁰ in deze vervolgstudie vertrekken we van dezelfde methodes als Geeraerts, Grondelaers en Speelman (1999). Aan de basis ligt de uniformiteitsmaat en die wordt aangevuld met de interne uniformiteitsmaat en de aandeelmaat voor linguïstische kenmerken. We zullen onze interpretaties baseren op de gewogen uniformiteit U' omdat die maat de reële verhoudingen in het taalgebruik het best reflecteert.

We brengen in herinnering dat de statistische betrouwbaarheid van de berekeningen uit Geeraerts, Grondelaers en Speelman (1999) beperkt werd door de omvang van het corpus, maar dat die beperking opgevangen werd door een marge van 5% te hanteren bij het vergelijken van de resultaten. In Speelman, Grondelaers en Geeraerts (2003, p. 321; 2006, p. 192) wordt een bijkomende stap geïntroduceerd bij de uniformiteitsberekening, en die nemen wij in deze studie ook. Om na te gaan of een profiel uit de ene variëteit zonder meer vergeleken mag worden met het profiel voor hetzelfde concept in een andere variëteit, passen we eerst de *log likelihood ratio test*¹¹ toe. Deze toets test de aannemelijkheid dat de twee profielen significant van elkaar verschillen. Indien dit niet het geval is, moeten we aannemen dat de uniformiteit maximaal (100%) is. Dat wil zeggen dat we op basis van de beschikbare data moeten concluderen dat de verschillende (absolute) frequenties voor de lexicalisering in de twee profielen te wijten zijn aan het toeval. In het andere geval, als de twee profielen wel significant van elkaar verschillen, kunnen we uitgaan van de betrouwbaarheid van de uniformiteitsmaat. We hanteren het gebruikelijke significantieniveau van 5%¹². Parallel hebben we ook de uniformiteitsberekeningen

gemaakt waarbij we deze voorafgaande aannemelijkheidsquotiënttest niet toepassen om na te gaan wat het effect ervan is. Hoewel de uniformiteit wat verschilt in waarde, zijn de waargenomen tendensen nagenoeg dezelfde. Waar er afwijkingen zijn, vermelden we die ook bij de resultaten. Tabel 6 in Appendix 2 geeft een overzicht van alle berekende uniformiteitswaarden.

Voor we tot de bespreking van de resultaten overgaan, verklaren we eerst de voorstellingswijze van de uniformiteitswaarden. Eerst geven we telkens schematisch de algemene tendens weer, om dan daaronder dezelfde vergelijking te kwantificeren. Helemaal rechts is tussen haken steeds een verwijzing te vinden naar het aantal concepten waarop de berekening gebaseerd is. Immers, in het materiaal van 1950 hebben we geen observaties van drie concepten, namelijk T-SHIRT (MAN/VROUW), LEGGING (VROUW) en JEANS (MAN/VROUW). Voor de diachrone convergentiehypothese vergelijken we daarom slechts uniformiteitswaarden gebaseerd op de gedeelde elf concepten. Wanneer we enkel data bekijken uit 1990 en 2012, zoals voor de stratificatiehypothese, gaan we uit van alle veertien kledingconcepten.

De convergentiebeweging die geconstateerd werd in Geeraerts, Grondelaers en Speelman (1999) zet zich niet voort in het nieuwe materiaal en een divergerende trend lijkt zich te manifesteren¹³:

$$(1) \quad U'(B_{50}, N_{50}) < U'(B_{70}, N_{70}) < U'(B_{90}, N_{90}) > U'(B_{12}, N_{12})$$

69,21 77,50 86,15 81,50 (nC)

De onderlinge toenadering tussen het Belgisch Nederlands en het Nederlands Nederlands in woordkeuze is niet langer af te leiden uit deze dataverzameling. Bijkomend kan dan de vraag gesteld worden aan welke kant de meeste verschuivingen plaatsvinden. Als de uniformiteit tussen de periodes lager is in België kunnen we daaruit afleiden dat de ontwikkelingen daar sterker zijn. In Geeraerts, Grondelaers en Speelman (1999) was de verwachting dat de convergentie kwam door wijzigingen aan de Belgische kant, wat vooral bevestigd werd door de vergelijking op lange termijn in (2). Ondanks de recente divergerende ontwikkeling vinden we in (3) dat de verschuivingen over de periode van 1950-2012 nog altijd enigszins sneller gebeuren aan Belgische kant. Hoewel een interval van deze lengte andere effecten kan maskeren, toont (4) dat ook voor de laatste tijdsafstand de bevindingen hetzelfde zijn. Het is wel zo dat de ongewogen uniformiteitsberekeningen in (5) een licht ander beeld tonen, namelijk dat de uniformiteit in Belgisch Nederlands iets hoger ligt. Wanneer we de data van nader-

bij bekijken, vinden we dat die vertekening grotendeels veroorzaakt wordt door de uniformiteit voor het laagfrequente concept KOSTUUM (MAN). Hier zien we duidelijk het effect van het niet mee in rekening brengen van de frequentie van concepten, wat de keuze om onze interpretaties te baseren op de gewogen uniformiteitsberekeningen bekrachtigt. In grote lijnen kunnen we stellen dat het verschil in ontwikkelingssnelheid met het Nederlands Nederlands geringer wordt en lijkt te stagneren¹⁴.

- | | | |
|-----|---|------------|
| (2) | $U'(B_{50}, B_{90}) < U'(N_{50}, N_{90})$ | |
| | 59,53 | 67,28 (nC) |
| (3) | $U'(B_{50}, B_{12}) < U'(N_{50}, N_{12})$ | |
| | 53,35 | 57,75 (nC) |
| (4) | $U'(B_{90}, B_{12}) < U'(N_{90}, N_{12})$ | |
| | 81,78 | 87,09 (nC) |
| (5) | $U(B_{50}, B_{12}) > U(N_{50}, N_{12})$ | |
| | 49,94 | 45,39 (nC) |

In Geeraerts, Grondelaers en Speelman (1999) werd bovendien met de nodige voorzichtigheid verondersteld dat standaardisering interne variatie reduceert. Immers, als alle taalgebruikers dezelfde benaming kiezen voor een bepaald begrip, is het taalgebruik in die situatie maximaal gestandaardiseerd. Hoewel het niet vaststaat hoeveel interne variatie normaal is voor een gestandaardiseerde taalsituatie, kan de Nederlands-Nederlandse situatie als ijkpunt genomen worden. De interne uniformiteit in (6) fluctueert maar in grote lijnen is er sprake van een toename in het Belgisch Nederlands. De voornaamste ontwikkeling gebeurde hierbij tijdens de eerste periode (1950-1970). Hoewel de interne uniformiteit van het Nederlands Nederlands in (7) over het algemeen hoog blijft, is deze ook onderhevig aan schommelingen. Ook in een gestandaardiseerde taal kunnen immers steeds onomasiologische verschuivingen optreden. Wanneer we kijken naar de interne uniformiteit voor alle veertien concepten bij het recentste materiaal in (8) vinden we zelfs een hogere interne uniformiteitswaarde voor het Belgisch Nederlands. Een mogelijke verklaring is dat het Belgisch Nederlands na de inhaalbeweging stopt met vernieuwen, aangezien er geen externe bron van variatie meer is, namelijk de exonormatieve oriëntatie. Deze stilstand kan dan een indicatie zijn van de mogelijke zoektocht naar een eigen Belgisch-Nederlandse norm, wat aansluit bij de particularistische visie.

- (6) $I'(B_{50}) < I'(B_{70}) > I'(B_{90}) \cong I'(B_{12})$
 61,06 77,68 72,61 77,56 (*nC*)
- (7) $I'(N_{50}) \cong I'(N_{70}) < I'(N_{90}) \cong I'(N_{12})$
 71,05 65,86 74,72 75,34 (*nC*)
- (8) $I'(B_{12}) > I'(N_{12})$
 79,35 73,71 (*14C*)

Algemeen wordt aangenomen dat de ontwikkeling van het Belgisch Nederlands gekenmerkt wordt door een reactie tegen de Franse invloed, een afkeuring die afwezig geacht wordt in het Nederlands Nederlands (Geeraerts en Grondelaers 2000, p. 54). Voor termen van Engelse oorsprong daarentegen wordt een toegenomen aandeel verwacht. Voor zowel het Nederlands in België (9) als in Nederland (10) vinden we dat het aandeel Franse termen daalt, een tendens die overigens nog sterker is bij het ongewogen materiaal waar de laagfrequente concepten harder doorwegen. Zoals we al eerder aanhaalden, kan het grillige patroon voor het Franse aandeel bij het Nederlands-Nederlandse materiaal gerelateerd worden aan de trendgevoeligheid van het Nederlands Nederlands.

Eveneens vanuit taalintern perspectief vinden we dat het aandeel Engelse termen in beide variëteiten in (11-12) stijgt. Bovendien is dit patroon nog sterker wanneer we alle veertien kledingconcepten beschouwen in (13), aangezien JEANS (MAN/VROUW), LEGGING (VROUW) en T-SHIRT (MAN/VROUW) net nieuwe concepten zijn waarvoor aanvankelijk de Engelse term werd overgenomen. In (13) stellen we eveneens vast dat de Engelse invloed minder sterk aanwezig is bij het Nederlands Nederlands dan bij het Belgisch Nederlands. Terwijl de taalzuivering in Nederlandstalig België wegens historische redenen vooral gallicismen in het vizier nam, lijkt de houding tegenover anglicismen permissiever geweest te zijn. Nederland lijkt de Engelse invloed van meet af aan meer te weren.

- (9) $A'_{FR}(B_{50}) > A'_{FR}(B_{70}) \cong A'_{FR}(B_{90}) > A'_{FR}(B_{12})$
 29,00 14,90 17,63 11,83 (*nC*)
- (10) $A'_{FR}(N_{50}) < A'_{FR}(N_{70}) < A'_{FR}(N_{90}) > A'_{FR}(N_{12})$
 39,39 20,78 31,71 19,39 (*nC*)
- (11) $A'_{EN}(B_{50}) \cong A'_{EN}(B_{70}) \cong A'_{EN}(B_{90}) < A'_{EN}(B_{12})$
 2,87 3,29 7,83 15,18 (*nC*)
- (12) $A'_{EN}(N_{50}) \cong A'_{EN}(N_{70}) \cong A'_{EN}(N_{90}) \cong A'_{EN}(N_{12})$
 1,75 2,87 7,80 7,71 (*nC*)
- (13) $A'_{EN}(B_{12}) > A'_{EN}(N_{12})$
 23,92 17,40 (*14C*)

Samenvattend kunnen we zeggen dat de toenadering tussen Belgisch Nederlands en Nederlands niet langer volgt uit het materiaal van 2012. Om de standaardiseringsachterstand in te halen zijn de verschuivingen diachroon altijd sneller geweest aan de Belgische kant, maar recent wordt het verschil met het Nederlands Nederlands minder duidelijk. Ook voor de interne uniformiteit zijn beide variëteiten tot op hetzelfde niveau genaderd. Voor het Belgisch-Nederlandse kan er hier sprake zijn van een normverschuiving weg van het Nederlands Nederlands. Verder lijkt zich naast de algemene afname van de Franse invloed in beide variëteiten, in Nederland eveneens een meer terughoudende reactie te manifesteren ten opzichte van het Engels.

Naast de ontwikkelingen op het standaardtaalniveau, laat de opbouw van de dataverzameling ook toe om de taalsituatie binnen de nationale variëteiten zelf te bestuderen. Het etalagemateriaal vertegenwoordigt het regiolectische niveau en wordt in de bespreking aangeduid met de stadsnaam waarvan de observaties afkomstig zijn. De hoofdhypothese in Geeraerts, Grondelaers en Speelman (1999) luidde dat er door de veronderstelde standaardiseringsachterstand in het Belgisch Nederlands een grotere afstand zou zijn tussen het standaard- en substandaardtaalgebruik. Zoals de berekeningen in (14) en (15) tonen, werd die verwachting destijds bevestigd en blijft dat ook het geval in de huidige dataset. Wel vinden we zowel voor het Belgisch Nederlands (16) als voor het Nederlands Nederlands (17) dat de afstand tussen regionaal en supraregionaal taalgebruik is afgenomen in de periode 1990-2012.

- (14) $U'(B_{90}, LeuKor_{90}) < U'(N_{90}, LeiMaa_{90})$
 46,86 67,72 (14C)
- (15) $U'(B_{12}, LeuKor_{12}) < U'(N_{12}, LeiMaa_{12})$
 63,95 81,65 (14C)
- (16) $U'(B_{90}, LeuKor_{90}) < U'(B_{12}, LeuKor_{12})$
 46,86 63,95 (14C)
- (17) $U'(N_{90}, LeiMaa_{90}) < U'(N_{12}, LeiMaa_{12})$
 67,72 81,65 (14C)

Dat de afstand tussen de twee registers groter is in België dan in Nederland impliceert overigens niet dat de onderlinge variatie tussen de regionale meetpunten ook groter zou zijn in België. Net als in de oorspronkelijke studie in (18), vinden we daarvoor ook nu geen bewijs in (19). Bovendien bevestigen de interne uniformiteitswaarden in (20-23) dat het Belgisch Nederlands niet heterogener is dan het Nederlands Nederlands.

- (18) U'(Leu₉₀,Kor₉₀) > U'(Lei₉₀, Ma₉₀)
 93,56 87,33 (14C)
- (19) U'(Leu₁₂,Kor₁₂) > U'(Lei₁₂, Ma₁₂)
 90,05 82,06 (14C)
- (20) I'(B₉₀) ≅ I'(N₉₀)
 72,10 73,69 (14C)
- (21) I'(LeuKor₉₀) > I'(LeiMa₉₀)
 86,25 73,20 (14C)
- (22) I'(B₁₂) > I'(N₁₂)
 79,35 73,71 (14C)
- (23) I'(LeuKor₁₂) > I'(LeiMa₁₂)
 76,28 69,56 (14C)

Doordat deze vervolgstudie de opbouw van Geeraerts, Grondelaers en Spielman (1999) overneemt, beschikken we voor het standaardtaalgebruik nu ook over een diachrone dimensie. Dit laat toe om na te gaan op welk stratificationeel niveau er de meeste verschuivingen zijn geweest de laatste twintig jaar. We herhalen dat een lagere uniformiteit een indicatie is voor meer ontwikkelingen. In Nederlandstalig België vinden we dat het bovengewestelijke niveau in (24) het meest in beweging is geweest, wat we kunnen linken aan de eerder geobserveerde divergerende tendens op dat niveau. Het Nederlands-Nederlandse materiaal in (25) toont een omgekeerde situatie, waarbij de verschuivingen vooral plaatsvinden op het regionale niveau. Voor de volledigheid signaleren we dat dit niet het geval is bij de ongewogen uniformiteit in (26), maar dat daar de waarden eerder vergelijkbaar zijn.

- (24) U'(B₉₀,B₁₂) < U'(LeuKor₉₀,LeuKor₁₂)
 81,57 87,71 (14C)
- (25) U'(N₉₀,N₁₂) > U'(LeiMa₉₀,LeiMa₁₂)
 88,00 81,93 (14C)
- (26) U(N₉₀,N₁₂) < U(LeiMa₉₀,LeiMa₁₂)
 81,34 83,59 (14C)

Verder was het een van de hypotheses in Geeraerts, Grondelaers en Spielman (1999) dat het standaardtalige Belgisch Nederlands dichter zou staan bij het Nederlands Nederlands dan bij het regionale taalgebruik in België. Deze verwachting werd niet alleen ingelost in Nederlandstalig België in (27), maar bleek ook te gelden voor het Nederlands Nederlands in (30). De *internationale* convergentie tussen de twee taalgebieden was met an-

dere woorden sterker dan de *intranationale* uniformiteit. Voor de recentste data observeren we in (28) een toegenomen uniformiteit tussen het Belgische tijdschriften- en etalagemateriaal, maar het supraregionale niveau staat wel nog steeds dicht bij Nederlands Nederlands dan bij het regionale taalgebruik. De ongewogen uniformiteit in (29) laat echter een ander beeld zien onder invloed van laagfrequente concepten als COLBERT (MAN) en JACK (MAN/VROUW) die binnen het Belgische materiaal zeer uniform zijn, maar duidelijk andere lexicalisatievoorkeuren vertonen in Nederland. Tot slot merken we ook voorzichtig op dat in (31) het Nederlandse intranationale taalgebruik een licht sterkere uniformiteit vertoont.

- (27) U'(B90,N90) > U'(B90,LeuKorgo)
 83,49 46,86 (14C)
- (28) U'(B12,N12) > U'(B12,LeuKor12)
 80,50 63,95 (14C)
- (29) U(B12,N12) < U(B12,LeuKor12)
 70,01 73,72 (14C)
- (30) U'(N90,B90) > U'(N90,LeiMaa90)
 83,49 67,72 (14C)
- (31) U'(N12,B12) < U'(N12,LeiMaa12)
 80,50 81,65 (14C)

Als laatste benaderen we de stratificatiehypothese vanuit taalintern perspectief. Voor de data van 2012 vinden we dat het Franse aandeel het grootst is in het regionale materiaal, wat vooral blijkt uit het ongewogen materiaal in (35) voor het Belgisch Nederlands en uit (37) voor het Nederlands Nederlands. Geeraerts, Grondelaers en Speelman (1999) brachten die observatie in verband met de commerciële functie van de winkelatalages, waar de appreciërende invloed van het Frans meer tot uiting komt. Daarnaast verwachtten Geeraerts, Grondelaers en Speelman (1999, p. 118; p. 126) dat het Franse aandeel het grootst zou zijn in de Leuvense en Kortrijkse etalages door de historische invloed van het Frans. De data wezen echter meer in de omgekeerde richting, wat als een mogelijke aanwijzing beschouwd werd dat de typisch Belgische puristische tendens ook tot het niveau van het minder formele taalgebruik doorgedrongen zou zijn. In navolging dan van de bevindingen in Geeraerts, Grondelaers en Speelman (1999, p. 126) kunnen we opnieuw verwachten dat de invloed van het Frans in (41) sterker is in Nederland dan in Vlaanderen. Daarnaast stellen we ten opzichte van 1990 in (38-39) een daling vast.

- (32) $A'_{FR}(\text{LeuKorgo}) \cong A'_{FR}(\text{B90})$
 19,29 17,62 (14C)
- (33) $A'_{FR}(\text{LeuKor12}) \cong A'_{FR}(\text{B12})$
 12,21 10,60 (14C)
- (34) $A_{FR}(\text{LeuKorgo}) > A_{FR}(\text{B90})$
 33,04 17,81 (14C)
- (35) $A_{FR}(\text{LeuKor12}) > A_{FR}(\text{B12})$
 25,49 12,10 (14C)
- (36) $A'_{FR}(\text{LeiMaag0}) \cong A'_{FR}(\text{N90})$
 33,76 28,94 (14C)
- (37) $A'_{FR}(\text{LeiMaai2}) > A'_{FR}(\text{N12})$
 22,87 17,03 (14C)
- (38) $A'_{FR}(\text{LeuKorgo}) > A'_{FR}(\text{LeuKor12})$
 19,29 12,21 (14C)
- (39) $A'_{FR}(\text{LeiMaag0}) > A'_{FR}(\text{LeiMaai2})$
 33,76 22,87 (14C)
- (40) $A'_{FR}(\text{LeuKorgo}) < A'_{FR}(\text{LeiMaag0})$
 19,29 33,76 (14C)
- (41) $A'_{FR}(\text{LeuKor12}) < A'_{FR}(\text{LeiMaai2})$
 12,21 22,87 (14C)

Voor het Engels bevinden de grootste verschuivingen zich in het Belgisch Nederlands. Het aandeel lexicalisering van (deels) Engelse oorsprong kent een toename in de periode 1990-2012 in (47) en komt zelfs tot op dezelfde hoogte als het aandeel in het Nederlands Nederlands in (50). Het ongewogen Belgisch-Nederlandse materiaal in (44) vertoont een groter aandeel op standaardtaalniveau dan op het niveau van het meer regionaal gekleurde taalgebruik door een toename van Engelse termen als *skirt* en *blazer* bij respectievelijk de minder frequente concepten ROK (VROUW), en COLBERT (MAN) en COLBERT (VROUW).

- (42) $A'_{EN}(\text{LeuKorgo}) > A'_{EN}(\text{B90})$
 21,61 15,03 (14C)
- (43) $A'_{EN}(\text{LeuKor12}) \cong A'_{EN}(\text{B12})$
 27,48 23,92 (14C)
- (44) $A_{EN}(\text{LeuKor12}) < A_{EN}(\text{B12})$
 34,83 41,72 (14C)
- (45) $A'_{EN}(\text{LeiMaag0}) > A'_{EN}(\text{N90})$
 29,61 14,86 (14C)

- (46) $A'_{EN}(LeiMaa12) > A'_{EN}(N12)$
 29,71 17,40 (14C)
- (47) $A'_{EN}(LeuKor90) < A'_{EN}(LeuKor12)$
 21,61 27,48 (14C)
- (48) $A'_{EN}(LeiMaa90) \cong A'_{EN}(LeiMaa12)$
 29,61 29,71 (14C)
- (49) $A'_{EN}(LeuKor90) < A'_{EN}(LeiMaa90)$
 21,61 29,61 (14C)
- (50) $A'_{EN}(LeuKor12) \cong A'_{EN}(LeiMaa12)$
 27,48 29,71 (14C)

Ten slotte kijken we naar de invloed van taalnormering op het hedendaagse Belgische materiaal. We veronderstellen hier uitsluitend een synchroon perspectief omdat de laatste decennia het belang van taalzuiveringswerken is afgenomen (zie ook Van Hoof en Jaspers 2012, p. 112) en taaladviezen minder prescriptief zijn geworden (Hendrickx 2013, p. 26). Bovendien zijn de taalbeheersingsbronnen uit Geeraerts, Grondelaers en Speelman (1999) niet langer actueel. Vertrekkende van de afgekeurde termen in Theissen en Debrabrandere (2004) en Koops et al. (2005), bevestigt de aandeelmaat in (51) dat vooral het regionale, meer informele materiaal afgekeurde kledingbenamingen bevat. Een blik op de data toont dat dit voornamelijk komt door het concept JEANS (MAN/VROUW), waarvoor *spijkerbroek* volgens de normatieve werken de voorkeur heeft op jeans.

- (51) $A'_{TZ}(LeuKor12) > A'_{TZ}(B12)$
 23,83 10,20 (14C)

In het kort kunnen we stellen dat de oorspronkelijke stratificatiehypothese in grote lijnen bevestigd is voor het hedendaagse materiaal. De afstand tussen beide onderzochte registers blijft het grootst in het Belgisch Nederlands, maar dat maakt die variëteit intern niet heterogener dan het Nederlands Nederlands. In Nederlandstalig België vonden in de periode 1990-2012 de meeste verschuivingen plaats op het standaardtaalniveau, terwijl in Nederland dat het geval was op regionaal niveau. Het Nederlandse Standaardnederlands staat nu zelfs dichter bij het lagere register dan dat het bij het Standaard Belgisch Nederland staat. Ook de resultaten voor de aandeelmaten stemmen overeen met de traditionele verwachtingen. Het aandeel Franse termen is afgenomen en is groter in het regionale register, vooral dan in de Nederlands-Nederlandse data. Voor het Engels vinden we een inhaalbeweging bij het Belgisch Nederlands, waardoor het tot op de-

zelfde hoogte als het Nederlands Nederlands komt wat dit aandeel betreft. Ten slotte merken we de meeste afgekeurde termen op in het etalagemateriaal in België, maar kunnen we tegelijk een vraagteken plaatsen bij de huidige relevantie van taalzuiveringswerken.

4 Discussie en conclusies

De vervolgstudie die in dit artikel voorgesteld wordt, bouwt voort op de basis van Geeraerts, Grondelaers en Speelman (1999). Voor veertien kledingconcepten hebben we onderzocht of er voor de periode 1950-2012 sprake is van convergentie of divergentie tussen het Belgisch Nederlands en Nederlands Nederlands en of die eerste variëteit bovendien een grotere afstand vertoont tussen het regionale en supraregionale taalregister als gevolg van de exogene normgerichtheid. Hiertoe gaan we uit van een corpus van reëel taalgebruik dat op drie niveaus gelaagd is. De geografische dimensie weerspiegelt het pluricentrische karakter van het Nederlands, waarbij zowel de Nederlands-Nederlandse als de Belgische-Nederlandse variëteit vertegenwoordigd zijn. Het oorspronkelijke corpus bestond uit observaties van standaardtaalgebruik zoals aangetroffen in tijdschriften in 1950, 1970 en 1990 en daar is voor dit artikel een nieuwe deelverzameling aan toegevoegd met gelijkaardig materiaal uit 2012. Daarbovenop zijn er ook observaties opgenomen van kledingbenamingen uit winketalages in vier universiteitssteden in Vlaanderen en Nederland als weergave van het meer informele, regionaal gekleurde taalgebruik.

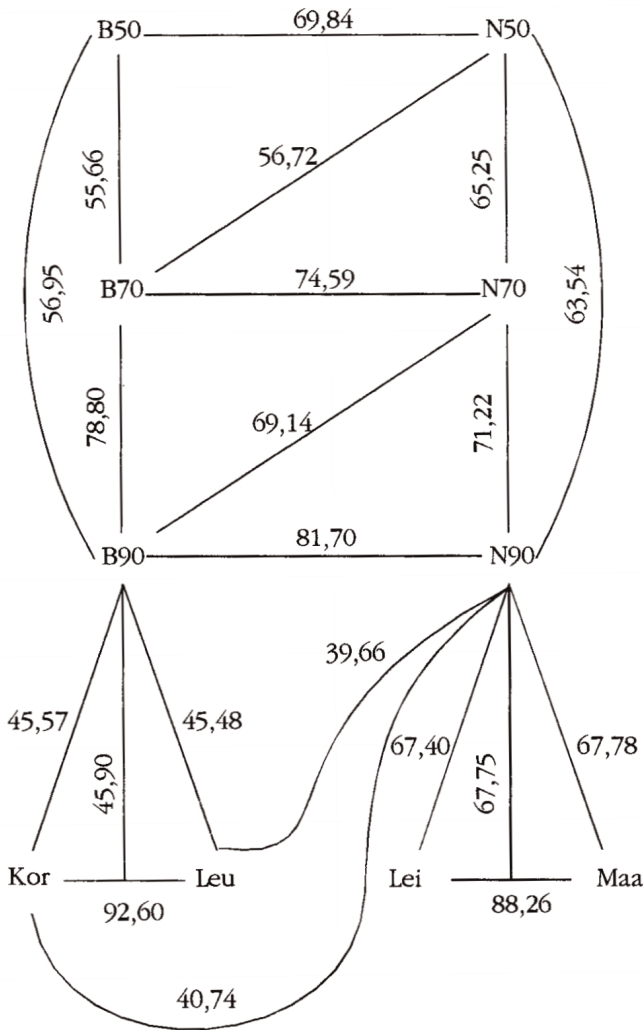
De methodologische vernieuwing uit Geeraerts, Grondelaers en Speelman (1999), met name de introductie van de profielgebaseerde uniformiteitsmaat, reikt een eenduidige berekeningswijze aan om de overlapping tussen de verschillende taalvariëteiten te meten. Op die manier hebben we de meervoudige bevindingen uit de oorspronkelijke studie gerepliceerd, wat verschillende nieuwe inzichten heeft opgeleverd. Figuur 2 geeft een overzicht van de resultaten voor de gewogen uniformiteitswaarden op basis van elf kledingconcepten¹⁵. In wat volgt vatten we eerst de voornaamste conclusies samen in verband met de verhouding *tussen* de nationale variëteiten en vervolgens bespreken we de evoluties *binnen* die nationale variëteiten in het licht van het destandaardiseringsdebat.

4.1 De verhouding *tussen* de nationale variëteiten: het Belgisch en Nederlands Nederlands groeien niet langer naar elkaar toe

Wanneer we over de periode 1950-2012 kijken naar de verhouding tussen het Belgisch Nederlands en het Nederlands Nederlands, kunnen we niet langer spreken van convergentie tussen het Belgisch Nederlands en het Nederlands Nederlands, maar observeren we een divergerende trend voor de periode 1990-2012. De langetermijntrend naar diachrone convergentie, zoals vastgesteld in Geeraerts, Grondelaers en Speelman (1999), lijkt dus omgebogen te zijn. Deze vaststelling bevestigt nu ook aan de hand van kwantitatieve metingen (zij het op een bescheiden steekproef) wat reeds elders in de literatuur geopperd werd. Waar de Belgisch-Nederlandse geschreven pers, de modebladen in deze studie, zich in de tweede helft van de twintigste eeuw qua woordgebruik steeds dichter naar de Nederlandse “norm” toebevoog, zien we nu, in het begin van de eenentwintigste eeuw, voor het eerst een lichte verwijdering optreden. Op de oorzaken van deze convergentiestop gaan we hier niet verder in, maar we verwijzen naar de in de literatuur (zie, o.a., Grondelaers en Van Hout 2011a, p. 228) reeds aangehaalde factoren zoals het grotere culturele en talige zelfbewustzijn in Nederlandstalig België, een verminderde aandacht voor taalzuivering en belicismenvermijding, en het verminderde contact van de Vlaming met de Nederlandse media: allemaal factoren die de gerichtheid van het Belgisch Nederlands op de exogene Nederlands-Nederlandse norm deden afnemen. Interessant is nog dat de toenemende invloed van het Engels, in zekere zin een nieuwe exogene “norm”, deze keer inwerkend op beide nationale variëteiten, niet sterk genoeg is om een tegengewicht te bieden voor divergentie, zelfs niet in een trendgevoelig lexicaal veld als mode.

4.2 Evoluties *binnen* de nationale variëteiten: standaardisering én (in)formalisering?

De tweede hoofdhypothese uit de oorspronkelijke studie betrof de grotere afstand tussen het bovengewestelijke formele en het regionaal gekleurde informele niveau in het Belgisch Nederlands in vergelijking met het Nederlands Nederlands. De uniformiteitsmaten uit 2012 bevestigen dat dit nog steeds het geval is. We vinden zelfs dat de intranationale uniformiteit, d.i. de overlapping tussen het supraregionale en het regionale register, in Nederland groter is dan de supranationale overeenkomst tussen het Nederlands Nederlands en het Belgisch Nederlands. De data duiden bovendien op een veranderende taalinterne situatie in beide nationale variëteiten. In het Nederlands Nederlands observeren we verschuivingen in het lagere register, terwijl in Nederlandstalig België net het omgekeerde het geval is.



Figuur 2 Gewogen uniformiteitsverhoudingen binnen het kledingmateriaal (elf concepten)

Tot slot vermelden we dat we in het Nederlands Nederlands een grotere terughoudendheid observeren ten opzichte van het Engels dan in het Belgisch Nederlands, alsook dat er een afname van de Franse invloed is op beide taalniveaus in de twee variëteiten.

Hoewel het de hoofdpzpet van dit artikel was om de bevindingen van deze replicatiestudie (in opvolging van Geeraerts, Grondelaers en Speelman 1999) te beschrijven en de vergelijking te maken met de oorspronkelijke studie, bespreken we ook kort hoe de vaststellingen in deze studie in

verband gebracht kunnen worden met het lopende onderzoek naar standaardisering en demotisering van het Nederlands en het tussentaaldebat (zie o.a. Auer en Spiekermann 2011 voor een algemene beschouwing; Grondelaers en Van Hout 2011a; Absillis, Jaspers en Van Hoof 2012; Cajot 2012; Van Hoof en Jaspers 2012). Hierbij is het van belang om termen als (de) standaardisering te vertalen naar de verschillende processen die we bestuderen vanuit een profielgebaseerd perspectief. We maken een conceptueel onderscheid tussen drie dimensies en volgen hierin Geeraerts en Spielman (2014). De eerste dimensie gaat kijken of de afstand tussen twee taalregisters toe- of afneemt. We spreken dan van **convergentie of divergentie**¹⁶ en meten dit aan de hand van de uniformiteitsmaat. Als de afstand tussen twee taalregisters verandert, kan dat in twee richtingen. Wanneer het hogere taalregister zich naar het lagere beweegt, spreken we van **informalisering**. Het omgekeerde gebeurt bij **formalisering**. Tot slot, de variabiliteit binnen een bepaald register (voornamelijk, het hogere taalniveau) kan ook af- of toenemen. We spreken dan van **(de)homogenisering**. Een register wordt homogener als de variabiliteit afneemt en dus de interne uniformiteit groter wordt. Standaardisering, zoals opgevat door Auer en Spiekermann (2011), kan dan vanuit een profielgebaseerd perspectief in verband gebracht worden met convergentie, formalisering en homogenisering. Echter, deze drie dimensies zijn onafhankelijk van elkaar en kunnen dus in verschillende combinaties voorkomen.

Met de data waarover we nu beschikken, vinden we dat er zowel binnen het Belgisch Nederlands als binnen het Nederlands Nederlands convergentie (de afstand tussen de registers vernauwt, want de uniformiteit stijgt) en homogenisering (de interne uniformiteit stijgt in het hoogste register) optreedt. Deze tendensen voor twee van de drie dimensies wijzen dus op een klassiek standaardiseringsproces. Bij de derde, (in)formaliseringsdimensie, zien we in de twee nationale variëteiten echter tegenovergestelde processen. In het Belgisch Nederlands beweegt het hogere register zich naar het lagere¹⁷: de traditionele exonormatieve oriëntatie maakt plaats voor een endogene, informelere norm. Deze verschuiving kan als demotisering in de zin van Kristiansen (2009) geïnterpreteerd worden. Meteen wordt in de Belgisch-Nederlandse situatie dan ook duidelijk dat aspecten van standaardisering en demotisering elkaar niet noodzakelijk uitsluiten. Bij het Nederlands in Nederland observeren we echter een verschuiving van het lagere taalniveau richting het hogere. In de Nederlandse situatie wijzen de drie dimensies voor de periode 1990-2012 dus op een klassiek standaardiseringsproces, en niet op aan de gang zijnde de-standaardisering. Een mogelijke verklaring voor dit verrassende resultaat

is dat in Nederland woordenschat historisch minder een *markeerder* is geweest van een lagere status dan in Nederlandstalig België, waar het net in het middelpunt stond in de taalverzorgingsliteratuur. Dit kan ook de tegenstelling met de bevindingen van Grondelaers en Van Hout (2011a) verklaren, doordat uitspraak en grammatica in Nederland misschien meer expliciete taalattitudes teweegbrengen. Hoe dan ook zijn verder onderzoek en een kritische beschouwing van termen als (de)standaardisering en (in)formalisering vanuit een profielgebaseerd perspectief (zoals in Geeraerts en Speelman (2014)) nodig om conclusies te trekken over de standaardiseringsstatus van het Nederlands.

4.3 Verder onderzoek: grotere corpora, meer registers en andere (lexicale) velden

Met de bevindingen in deze studie is nog niet alles gezegd over de taalsituatie in de Lage Landen. Verscheidene uitbreidingen zijn nodig om de vaststellingen uit dit onderzoek, hoe robuust ook, door te trekken met enige mate van zekerheid naar conclusies over de status en de structuur van het Nederlands. In eerste instantie is de uitbreiding van het corpus een cruciale stap. Zelfs voor slechts veertien kledingconcepten is het een tijdrovend proces om manueel data te verzamelen in tijdschriften en winkelatalages. De automatisering van het tellen van observaties in een elektronisch corpus zoals in Ruetten, Speelman en Geeraerts (2011) en Heylen en Ruetten (2013), is hiervoor deels een oplossing en kan een significante schaalvergroting opleveren, ook al leidt deze automatisering tot het verlies van de referentiële informatie uit afbeeldingen die de oorspronkelijke verzamelingswijze oplevert. Toch is het een voor de hand liggende keuze om voor de verzameling van concepten uit andere lexicale velden of van grammaticale constructies die lijn te vervolgen. Een corpusgebaseerde studie naar variatie zoals in dit onderzoek voorgesteld is, kan immers bijdragen aan het standaardtaalonderzoek. Daarnaast is de stratificationele opbouw van dit artikel slechts een beperkte weerspiegeling van de talige realiteit van het Nederlands. De toevoeging van andere registers zou toelaten om de complexiteit van het brede spectrum tussen de standaardtaal en de dialecten, in zowel geschreven als gesproken vorm, beter te beschrijven. Tot dan biedt deze studie enkel een gedeeltelijke inkijk in de lectale structuur van het Nederlands.

Noten

1. Graag bedanken we masterstudente Tine De Cnodder (academiejaar 2012-2013) voor het verzamelen van de data uit zowel de tijdschriften en de winkeletalages. We danken ook de anonieme reviewers voor hun suggesties bij een eerdere versie van dit artikel. Het onderzoek waarover dit artikel rapporteert wordt gesteund door het OT-project 3H110243 van de KU Leuven.
2. In dit artikel geldt, alsook het geval is in Geeraerts, Grondelaers en Speelman (1999, pp. 6-7), dat we zullen spreken van *Belgisch Nederlands* en *Nederlands Nederlands* om te verwijzen naar de twee geografisch gescheiden variëteiten van het Nederlands. In deze context vermijden we dus de term *Vlaams*, die ook verwijst naar een verzameling dialecten, evenals de spelling met koppelteken *Belgisch-Nederlands*, welke vooronderstelt dat het een aparte taalvariëteit betreft. Wanneer deze termen bijvoeglijk gebruikt worden, is een koppelteken onvermijdelijk.
3. De *City-block distance* (CBD) uit Speelman, Grondelaers en Geeraerts (2003; 2006) meet niet de overlapping of uniformiteit (U) tussen twee taalvariëteiten, maar het verschil ertussen. Een simpele rekensom laat de overgang van de ene maat naar de andere toe, namelijk $CBD = 1 - U$, of anders genoteerd: $U = 1 - CBD$.
4. Appendix 1 bevat een overzicht van de veertien kledingconcepten en hun frequentie in aantal woorden in de verschillende deelverzamelingen. Appendix 3 somt de tijdschriftbronnen voor het materiaal van 2012 op.
5. De absolute frequenties van de kledingtermen in deze twee profielen zijn te vinden in Tabel 4 in Appendix 1.
6. Door te kwadrateren wordt meer gewicht gegeven aan de dominantie van een één bepaalde term, wat niet het geval is door louter het gemiddelde te nemen. Een kleine vergelijking illustreert dit goed. In bron A wordt een bepaald concept aangeduid met vier synoniemen, in de verhouding 50%, 25%, 15% en 10%. Datzelfde concept kent in bron B een andere distributie, namelijk 85%, 5%, 5% en 5%. Door de interne uniformiteit te berekenen op basis van de som van de kwadraten, betekent dit dat in bron A, met 34,5%, de uniformiteit veel lager ligt dan in bron B, met 73%. Het gemiddelde in beide gevallen is evenwel 25%, en dus nietszeggend met betrekking tot de homogeniteit van de taalsituatie in bron A en die in bron B.
7. Deze wegingsfactor wordt momenteel manueel bepaald door de onderzoeker, maar een andere mogelijkheid is het experimenteel meten van de herkenbaarheid van gallicismen en anglicismen. Voor de manuele toekenning onderscheiden we op verschillende gronden de waarden 0 (geen aandeel), 1 (volledig aandeel) en enkele tussenliggende gevallen. Kledingtermen waarvan zowel de schrijfwijze als de uitspraak nog dezelfde is als in de brontaal krijgen waarde 1, zoals *caleçon* en *leggings* voor LEGGING_V voor respectievelijk de Franse en Engelse aandeelmaat. Vernederlandste termen qua uitspraak en/of schrijfwijze die niet onmiddellijk als termen van vreemde herkomst herkenbaar zijn, ontvangen de score 0,5. Dit is bijvoorbeeld het geval bij *vest* voor VEST_MV. Aan samenstellingen met een Nederlands grondwoord en een grondwoord van Franse of Engels herkomst wordt waarde 0,5 toegekend, zoals bij *jeansbroek* voor JEANS_MV. Aan *windbloes* voor JACK_MV kennen we slechts 0,25 toe, op grond van het vernederlandste grondwoord *bloes*. Voor het aandeel van de taalzuiveringswerken geldt in eerste instantie een binaire toekenning, 0 (niet opgenomen in het taalzuiveringswerk) en 1 (afgekeurd). Vervolgens wordt het gemiddelde genomen van de waarden gebaseerd op de twee geraadpleegde werken.

8. De uniformiteitswaarden in Figuur 1 zijn gebaseerd op de berekening voor elf kledingconcepten en kunnen dus voor het hedendaagse materiaal verschillen van de waarden die we opgeven bij de vergelijkingen (14-19; 24-31), omdat we daar uitgaan van alle veertien kledingconcepten.
9. Net zoals voor de oorspronkelijke materiaalverzameling, zijn er vooral kleinere winkels (en niet grotere winkelketens) onderzocht (Geeraerts, Grondelaers en Speelman, 1999, p. 34). Het is echter aannemelijk dat een groot aantal van de boetieks in het straatbeeld plaats hebben gemaakt voor net die grotere winkelketens, wat verklaart waarom het aantal observaties voor het etalagemateriaal in 2012 ligt lager dan in 1990.
10. Niet alle subhypotheses uit Geeraerts, Grondelaers en Speelman (1999) worden hier herhaald en getest met de nieuwe data uit 2012, maar de vergelijkingen zijn repliceerbaar met de tabellen uit Appendix 2.
11. Voor het uitvoeren van de *log likelihood ratio test* gebruiken we de implementatie van Pete Hurd in R, genaamd de 'g-test', waarvoor de code te vinden is op: <http://www.psych.ualberta.ca/~phurd/cruft/g.test.r>.
12. De significantiegrens van 5% die we hier vermelden verwijst naar de *log likelihood ratio test* en is dus niet hetzelfde als de foutenmarge gehanteerd in Geeraerts, Grondelaers en Speelman (1999) voor het vergelijken van twee uniformiteitswaarden. Bij het vergelijken van berekeningen die steunen op de interne uniformiteitsmaat en de aandeelmaat gebruiken we wel de statistische foutenmarge van 5%.
13. De uniformiteitsindices voor het materiaal van 1950, 1970 en 1990 verschillen van die in Geeraerts, Grondelaers en Speelman (1999) omdat het concept MANTELPAKJE (VROUW) omwille van dataschaarste niet mee opgenomen is 2012. De originele uniformiteitswaarden werden daarom herrekend voor 14 concepten.
14. Ook wanneer we de tijdsblokken opsplitsen, zien we een geleidelijke toename van de uniformiteit wat de plausibiliteit verhoogt van de interpretatie dat de ontwikkelingen vooral aan Belgische kant voorvallen. In dit opzicht, net als in Geeraerts, Grondelaers en Speelman (1999, p. 74), zijn er ook geen retardatie-effecten te observeren. We verwijzen hier slechts beknopt naar deze berekeningen omdat ze qua interpretatie niet verschillen van de oorspronkelijke studie en geen nieuwe inzichten opleveren. Appendix 2 geeft wel de relevante berekeningen weer om alle analyses uit de oorspronkelijke studie na te gaan.
15. Doordat Figuur 2 een globaal overzicht geeft waarin zowel de basisbevindingen voor de diachrone convergentiehypothese als voor de stratificatiehypothese af te lezen zijn, geven we de uniformiteitswaarden weer voor elf kledingconcepten. Hierdoor verschillen de waarden voor onder andere het regionale materiaal met die van de vergelijkingen in sectie 3 van dit artikel, omdat we daar rekenen met veertien kledingconcepten.
16. De termen convergentie en divergentie worden door de focus van Geeraerts, Grondelaers en Speelman (1999) voornamelijk geassocieerd met respectievelijk een vernauwing en een verwijding van de afstand tussen *geografische* variëteiten (cf. de diachrone convergentie tussen het Belgisch en Nederlands Nederlands), maar dit hoeft niet het geval te zijn. Convergentie betekent meer algemeen de vernauwing van de afstand tussen twee taalvariëteiten, zijnde geografisch, stratificationeel... In een discussie over standaardisering, wordt de term convergentie gebruikt om de afstand tussen twee verschillende registers te bestuderen. Merk hierbij op dat dit verschilt van de terminologie in Geeraerts en Speelman (2014), waar de dimensie 'convergentie-divergentie' aangeduid wordt met '(de)standaardisering'.
17. In een profielgebaseerde aanpak operationaliseren we het concept 'informalisering' aan de hand van de volgende vergelijking: $U(L_i, H_{i+1}) > U(H_i, L_{i+1})$. Hierbij staat 'L' voor het

lagere register, en 'H' voor hogere taalniveau, voor de twee tijdstippen 'i' en 'i+1'. Voor de data uit deze studie wordt dat:

(52)	U'(LeuKor90,B12) 56,22	>	U'(B90,LeuKor12) 52,29	(14C)
(53)	U'(LeiMaa90,N12) 59,23	<	U'(N90,LeiMaa12) 85,54	(14C)

Bibliografie

- Absillis, Kevin, Jürgen Jaspers en Sarah Van Hoof (red.), *De manke usurpator. Over Verkavelingsvlaams*, Gent: Academia Press, 2012.
- Auer, Peter en Helmut Spiekermann, 'Demotisation of the standard variety or de-standardisation? The changing status of German in late modernity (with special reference to southwestern Germany)', In: Tore Kristiansen en Nikolas Coupland (red.), *Standard Languages and Language Standards in a Changing Europe*, Oslo: Novus Press, 2011, pp. 161-176.
- Cajot, José, 'Waarom het Verkavelingsvlaams onvermijdelijk was. De ontwikkeling van een informele omgangstaal in Vlaanderen', In: Kevin Absillis, Jürgen Jaspers en Sarah Van Hoof (red.), *De manke usurpator. Over Verkavelingsvlaams*, Gent: Academia Press, 2012, pp. 39-66.
- De Cnodder, Tine, 'Convergentie en divergentie in de Nederlandse woordenschat anno 2012. Een onderzoek naar kledingtermen', Meesterproef onder begeleiding van Dirk Geeraerts, Leuven: KU Leuven, 2012-2013.
- Geeraerts, Dirk en Dirk Speelman, 'A lectometric definition of demotisation and destandardisation', Presentatie op *Taal & Tongval 2014* (28 november 2014), 2014.
- Geeraerts, Dirk en Stefan Grondelaers, 'Purism and fashion. French influence on Belgian and Netherlandic Dutch', *Belgian Journal of Linguistics*, 13 (1999), pp. 53-68. doi:10.1075/bjl.13.04gee.
- Geeraerts, Dirk, Stefan Grondelaers en Dirk Speelman, *Convergentie en divergentie in de Nederlandse woordenschat: een onderzoek naar kleding- en voetbaltermen*. Amsterdam: P.J. Meertens-Instituut, 1999.
- Goossens, Jan, 'De toekomst van het Nederlands in Vlaanderen', *Ons Erfdeel* 43 (2000), pp. 3-13.
- Grondelaers, Stefan, Dirk Geeraerts, Dirk Speelman en José Tummers, 'Lexical standardisation in internet conversations. Comparing Belgium and The Netherlands', In: Josep M. Fontana, Louise McNally, M. Teresa Turell en Enric Vallduví (red.), *Proceedings of the First International Conference on Language Variation in Europe*, Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada, Unitat de Investigació de Variació Lingüística, 2001a, pp. 90-100.
- Grondelaers, Stefan, Hilde Van Aken, Dirk Speelman en Dirk Geeraerts, 'Inhoudswoorden en preposities als standaardiseringsindicatoren. De diachrone en synchrone status van het Belgische Nederlands', *Nederlandse Taalkunde*, 6-3 (2001b), pp. 179-202.
- Grondelaers, Stefan en Roeland van Hout, 'The Standard Language Situation in the Low Countries: Top-Down and Bottom-Up Variations on a Diaglossic Theme', *Journal of Germanic Linguistics* 23-3 (2011a), pp. 199-243.
- Grondelaers, Stefan en Roeland van Hout, 'The standard language situation in The Netherlands',

- In: Tore Kristiansen en Nikolas Coupland (red.), *Standard Languages and Language Standards in a Changing Europe*, Oslo: Novus press, 2011b, pp. 113-118.
- Hendrickx, Els, 'Het effect van lexicale taaladvisering op het Belgisch-Nederlandse taalgebruik. Een diachroon corpusonderzoek naar factoren van invloed', Dissertatie, Leuven: KU Leuven, 2013.
- Heylen, Kris en Tom Ruetten, 'Degrees of semantic control in measuring aggregated lexical distances'. In: Lars Borin en Anju Saxena (red.), *Approaches to Measuring Linguistic Differences*, (Trends in Linguistics. Studies and Monographs 265), Berlin: Mouton de Gruyter, 2013, pp. 353-374. doi:10.1515/9783110305258.353.
- Kristiansen, Tore, 'The nature and role of language standardisation and standard languages in late modernity', Projectvoorstel, 2009. <http://dgcss.hum.ku.dk/exploratoryworkshops/proposal/> (18 mei 2015).
- Koops, Bert-Jaap, Pim Slop, Paul Uljé, Kees Vermeij en Dick van Zijderveld, '2400 x liever Nederlands. Woordenlijst onnodig Engels', Stichting Nederlands, 2005. <http://snn.vvb.org/woordenlijst/a.htm> (22 augustus 2013).
- Ruette, Tom, Dirk Speelman en Dirk Geeraerts, 'Measuring the lexical distance between registers in national varieties of Dutch', In: Augusto Soares da Silva, Amadeu Torres en Miguel Gonçalves (red.), *Línguas Pluricêntricas. Variação Linguística e Dimensões Sociocognitivas*, Braga: Publicações da Faculdade de Filosofia, Universidade Católica Portuguesa, 2011, pp. 541-554.
- Speelman, Dirk, Stefan Grondelaers en Dirk Geeraerts, 'A profile-based calculation of region and register variation: the synchronic and diachronic status of the two main national varieties of Dutch'. In: Andrew Wilson, Dawn Archer en Paul Rayson (red.), *Corpus Linguistics around the World*, Amsterdam: Rodopi, 2006, pp. 181-194.
- Speelman, Dirk, Stefan Grondelaers en Dirk Geeraerts, 'Profile-based linguistic uniformity as a generic method for comparing language varieties', *Computers and the Humanities*, 37-3 (2003), pp. 317-337. doi:10.1023/A:1025019216574.
- Theissen, Siegfried en Peter Debrabandere. *Woordenboek voor correct taalgebruik*. Deurne: Wolters Plantyn, 2004.
- Van Hoof, Sarah en Jürgen Jaspers. 'Hyperstandaardisering', *Tijdschrift voor Nederlandse Taal- en Letterkunde*, 128-2 (2012), pp. 97-125.
- Zenner, Eline, Dirk Speelman en Dirk Geeraerts, 'Cognitive Sociolinguistics meets loanword research: Measuring variation in the success of anglicisms in Dutch', *Cognitive Linguistics* 23-4 (2012), pp. 749-792. doi:10.1515/cog-2012-0023.

Over de auteurs

Jocelyne Daems is doctoraatsbursaal aan de KU Leuven in de onderzoeksgroep *Quantitative Lexicology and Variational Linguistics (QLVL)* en is de corresponderende auteur: Jocelyne.

E-mail: Daems@kuleuven.be

Kris Heylen is postdoctoraal onderzoeker aan de KU Leuven in de onderzoeksgroep *Quantitative Lexicology and Variational Linguistics (QLVL)*.

Dirk Geeraerts is gewoon hoogleraar Algemene Taalkunde aan de KU Leuven en is aldaar verbonden aan de onderzoeksgroep *Quantitative Lexicology and Variational Linguistics (QLVL)*.

Appendices

Appendix 1 Overzicht observaties in alle subcorpora in aantal woorden

Tabel 4 Kwantitatief overzicht voor de veertien kledingconcepten in het standaardtaal-materiaal in aantal woorden

concept	variant	B12	N12	B90	N90	B70	N70	B50	N50
overhemd_M	<i>hemd</i>	27	29	11	9	33	21	40	5
overhemd_M	<i>overhemd</i>	0	19	24	25	38	62	23	23
overhemd_M	<i>shirt</i>	2	107	0	20	1	11	2	1
bloes_V	<i>bloes(je)</i>	188	146	46	51	51	174	149	56
bloes_V	<i>bloeze/blouse/blouze</i>	112	324	176	473	172	67	291	322
T-shirt_MV	<i>T-shirt</i>	139	136	57	27	1	4	0	0
T-shirt_MV	<i>shirt</i>	3	80	1	18	0	0	0	0
trui_MV	<i>pull(etje)</i>	2	0	0	5	105	3	0	0
trui_MV	<i>pullover</i>	2	7	2	9	29	24	49	44
trui_MV	<i>trui(tje)</i>	202	251	377	341	82	57	13	9
trui_MV	<i>sweater</i>	28	32	0	0	0	0	0	0
vest_MV	<i>cardigan</i>	130	3	90	2	13	0	0	0
vest_MV	<i>gil(l)et</i>	7	1	41	24	1	0	3	0
vest_MV	<i>jasje</i>	100	140	142	176	11	6	14	8
vest_MV	<i>vest(je)</i>	72	328	166	246	62	66	59	22
vest_MV	<i>golf</i>	0	0	0	0	0	0	0	0
broek_MV	<i>broek</i>	304	602	243	388	390	232	59	291
broek_MV	<i>pantalon</i>	35	74	12	39	154	162	11	124
broek_MV	<i>pant</i>	0	0	0	0	0	0	0	0
broek_MV	<i>pants</i>	4	1	0	0	0	0	0	0
jeans_MV	<i>jeans</i>	121	185	64	81	8	2	0	0
jeans_MV	<i>spijkerbroek</i>	0	67	2	34	1	3	0	0
jeans_MV	<i>jeansbroek</i>	4	2	0	0	0	0	0	0
legging_V	<i>caleçon</i>	0	0	40	0	0	0	0	0
legging_V	<i>legging</i>	28	62	26	113	0	0	0	0
legging_V	<i>leggings</i>	0	0	39	25	0	0	0	0
legging_V	<i>leggingbroek</i>	1	0	0	0	0	0	0	0
rok_V	<i>rok(je)</i>	441	423	328	392	429	405	373	642
rok_V	<i>skirt</i>	8	1	0	0	0	0	0	0
jurk_V	<i>japon(netje)</i>	0	0	0	2	40	351	131	792
jurk_V	<i>jurk(je)</i>	621	891	167	184	1493	1015	104	305
jurk_V	<i>kleed(je)</i>	0	0	0	0	43	0	376	1
jurk_V	<i>dress</i>	3	16	0	0	0	0	0	0

WAT DRAGEN WE VANDAAG: EEN HEMD MET BLAZER OF EEN SHIRT MET JASJE?

concept	variant	B12	N12	B90	N90	B70	N70	B50	N50
colbert_M	<i>blazer</i>	6	12	4	9	1	7	0	0
colbert_M	<i>k/colbert(je)</i>	0	6	3	109	1	21	30	275
colbert_M	<i>colbertjasje</i>	0	0	1	0	0	0	0	3
colbert_M	<i>jas</i>	0	0	1	0	12	2	4	42
colbert_M	<i>jasje</i>	0	49	19	10	29	18	2	7
colbert_M	<i>vest(je)</i>	1	0	2	0	5	1	10	0
colbert_M	<i>veston</i>	0	0	0	0	0	0	1	0
colbert_V	<i>blazer</i>	185	81	62	129	5	19	0	0
colbert_V	<i>colbert(je)</i>	0	91	7	43	0	0	0	0
colbert_V	<i>colbertjas(je)</i>	0	4	0	0	0	0	1	0
colbert_V	<i>jas</i>	0	0	2	1	1	0	0	0
colbert_V	<i>jasje</i>	19	98	117	77	49	11	20	23
colbert_V	<i>mantel(tje)</i>	0	0	0	1	0	0	10	10
colbert_V	<i>tailleurmantel(tje)</i>	0	0	0	0	0	0	0	8
colbert_V	<i>vest(je)</i>	4	4	14	0	5	0	39	0
colbert_V	<i>kostuumvest</i>	2	0	0	0	0	0	0	0
jack_MV	<i>bloes</i>	0	0	0	0	0	0	4	0
jack_MV	<i>blouson(netje)</i>	13	5	13	31	31	0	0	0
jack_MV	<i>jack(je)/jek</i>	11	70	15	141	18	36	1	0
jack_MV	<i>jacket</i>	12	0	0	1	0	0	0	3
jack_MV	<i>jak(je)</i>	0	0	0	0	0	0	0	3
jack_MV	<i>windbloes(je)</i>	0	0	0	0	0	0	7	1
jack_MV	<i>windjack</i>	1	1	1	0	3	1	8	26
jack_MV	<i>windjacket</i>	0	0	0	0	0	0	2	0
kostuum_M	<i>colbertcostuum</i>	0	0	0	0	0	0	8	19
kostuum_M	<i>k/costuum</i>	2	1	7	119	63	39	14	49
kostuum_M	<i>pak</i>	4	14	25	57	9	4	60	0
	TOTAAL	2844	4363	2347	3412	3389	2824	1918	3114

Tabel 5 Kwantitatief overzicht voor de veertien kledingconcepten in het substandaard-taalmateriaal in aantal woorden

concept	variant	Leu12	Kor12	Lei12	Ma-a12	Leu90	Kor90	Lei90	Ma-a90
overhemd_M	<i>hemd</i>	91	46	0	46	367	364	7	126
overhemd_M	<i>overhemd</i>	0	0	0	0	0	0	20	1
overhemd_M	<i>shirt</i>	0	0	37	14	0	0	112	79
bloes_V	<i>bloes(je)</i>	13	20	0	2	245	285	1	6
bloes_V	<i>bloeze/blouse/blouze</i>	2	1	10	31	83	12	114	155
T-shirt_MV	<i>T-shirt</i>	24	24	6	17	49	33	21	20
T-shirt_MV	<i>shirt</i>	4	1	11	25	0	0	3	0
trui_MV	<i>pull(etje)</i>	61	45	0	4	783	754	19	113
trui_MV	<i>pullover</i>	0	5	10	3	0	0	70	36
trui_MV	<i>trui(tje)</i>	7	19	17	57	3	25	96	174
trui_MV	<i>sweater</i>	3	7	1	0	0	0	0	0
vest_MV	<i>cardigan</i>	37	13	0	0	62	19	0	0
vest_MV	<i>gil(l)et</i>	13	30	0	0	135	132	0	0
vest_MV	<i>jasje</i>	0	1	12	15	3	0	0	0
vest_MV	<i>vest(je)</i>	16	4	44	58	28	31	48	84
vest_MV	<i>golf</i>	1	0	0	0	0	0	0	0
broek_MV	<i>broek</i>	86	74	30	80	296	356	27	34
broek_MV	<i>pantalon</i>	1	6	29	18	31	28	109	159
broek_MV	<i>pant</i>	0	0	3	0	0	0	0	0
broek_MV	<i>pants</i>	0	0	0	1	0	0	0	0
jeans_MV	<i>jeans</i>	31	25	30	23	112	152	74	71
jeans_MV	<i>spijkerbroek</i>	0	0	0	0	0	0	16	2
jeans_MV	<i>jeansbroek</i>	0	0	0	0	0	0	0	0
legging_V	<i>caleçon</i>	0	0	0	0	132	106	0	0
legging_V	<i>legging</i>	1	2	1	8	7	17	40	28
legging_V	<i>leggings</i>	0	0	0	0	0	0	0	0
legging_V	<i>leggingbroek</i>	0	0	0	0	0	0	0	0
rok_V	<i>rok(je)</i>	18	16	13	32	464	437	69	162
rok_V	<i>skirt</i>	1	0	0	1	0	0	0	0
jurk_V	<i>japon(netje)</i>	0	0	1	3	0	0	27	23
jurk_V	<i>jurk(je)</i>	7	9	10	28	5	7	7	22
jurk_V	<i>kleed(je)</i>	48	39	0	0	108	101	0	0
jurk_V	<i>dress</i>	0	0	0	1	0	0	0	0
colbert_M	<i>blazer</i>	9	3	3	5	15	57	5	6
colbert_M	<i>k/colbert(je)</i>	0	0	23	6	0	0	66	95
colbert_M	<i>colbertjasje</i>	0	0	0	0	0	0	1	0
colbert_M	<i>jas</i>	0	0	0	0	29	2	0	0

WAT DRAGEN WE VANDAAG: EEN HEMD MET BLAZER OF EEN SHIRT MET JASJE?

concept	variant	Leu12	Kor12	Lei12	Ma-a12	Leu90	Kor90	Lei90	Ma-a90
colbert_M	<i>jasje</i>	0	0	0	0	0	0	0	0
colbert_M	<i>vest(je)</i>	7	4	0	0	56	51	0	0
colbert_M	<i>veston</i>	0	0	0	0	0	0	0	0
colbert_V	<i>blazer</i>	4	5	5	33	25	70	37	78
colbert_V	<i>colbert(je)</i>	0	0	2	3	0	0	2	5
colbert_V	<i>colbertjas(je)</i>	0	0	0	0	0	0	0	0
colbert_V	<i>jas</i>	0	0	0	0	0	0	0	5
colbert_V	<i>jasje</i>	0	0	2	3	0	0	6	7
colbert_V	<i>mantel(tje)</i>	0	0	0	0	0	0	0	0
colbert_V	<i>tailleurmantel(tje)</i>	0	0	0	0	0	0	0	0
colbert_V	<i>vest(je)</i>	10	2	0	0	142	103	0	0
colbert_V	<i>kostuumvest</i>	0	0	0	0	0	0	0	0
jack_MV	<i>bloes</i>	0	0	0	0	0	0	0	0
jack_MV	<i>blouson(netje)</i>	2	3	0	0	19	13	0	0
jack_MV	<i>jack(je)/jek</i>	0	0	32	12	1	6	80	82
jack_MV	<i>jacket</i>	0	0	0	3	0	4	0	0
jack_MV	<i>jak(je)</i>	0	0	0	0	0	0	0	0
jack_MV	<i>windbloes(je)</i>	0	0	0	0	0	0	0	0
jack_MV	<i>windjack</i>	0	0	0	0	0	0	0	0
jack_MV	<i>windjacket</i>	0	0	0	0	0	0	0	0
kostuum_M	<i>colbertcostuum</i>	0	0	0	0	0	0	0	0
kostuum_M	<i>k/costuum</i>	7	9	23	5	38	45	64	79
kostuum_M	<i>pak</i>	0	0	0	0	2	0	0	0
	TOTAAL	504	413	355	537	3240	3210	1141	1652

Appendix 2 Uniformiteitswaarden, interne uniformiteitswaarden en aandeelmaten

Tabel 6 Uniformiteitswaarden voor elf en veertien kledingconcepten met en zonder de *log likelihood ratio* (LLR) significantietest

	11C		11C_LL		14C		14C_LL	
	U'	U	U'	U	U'	U	U'	U
B12_N12	80.79	64.61	81.50	67.40	79.83	67.57	80.50	70.01
B12_LeuKor12	58.72	58.17	59.99	67.21	62.64	66.14	63.95	73.72
B12_Leu12	56.07	56.38	57.97	65.90	59.95	64.38	61.82	72.34
B12_Kor12	58.96	58.93	63.07	72.29	63.07	67.12	67.02	78.23
N12_LeiMaa12	80.63	69.31	82.29	71.08	80.20	72.24	81.65	73.62
N12_Lei12	77.25	67.13	80.47	71.10	77.05	70.26	79.87	73.37
N12_Maa12	81.61	66.59	85.06	70.30	81.16	70.20	84.18	73.12
LeuKor12_LeiMaa12	41.67	41.55	41.70	41.61	45.99	50.39	46.02	50.44
N12_Leu12	41.78	32.81	44.00	41.76	46.13	43.64	48.08	50.67
N12_Kor12	44.40	34.46	45.07	34.81	47.94	44.21	48.53	44.48
Lei12_Maa12	73.73	78.09	79.24	86.85	76.95	82.41	82.06	89.66
Leu12_Kor12	84.47	85.97	88.67	93.46	85.77	88.24	90.05	94.87
B90_B12	81.61	69.42	81.78	70.50	81.39	70.34	81.57	71.22
B50_B12	52.52	47.96	53.35	49.94	-	-	-	-
N90_N12	86.33	76.61	87.09	77.90	87.07	79.89	88.00	81.34
N50_N12	57.43	45.02	57.75	45.39	-	-	-	-
B12_N90	75.18	63.03	76.74	66.45	74.79	64.84	76.20	67.53
LeuKor90_LeuKor12	87.31	87.41	90.19	93.93	85.11	82.94	87.71	88.07
LeiMaa90_LeiMaa12	80.04	82.10	82.18	85.05	80.02	81.27	81.93	83.59
B90_N90	84.33	75.73	86.15	78.99	81.83	72.22	83.49	74.78
B90_LeuKor90	43.39	41.70	43.83	42.03	46.44	50.09	46.86	50.47
B90_Leu90	43.24	40.26	43.79	40.69	46.00	48.66	46.63	49.34
B90_Kor90	42.84	41.12	43.47	44.00	46.24	49.96	46.96	52.57
N90_LeiMaa90	66.53	66.98	66.53	66.98	67.72	69.07	67.72	69.07
N90_Lei90	65.70	65.40	66.09	66.12	67.45	68.71	68.33	70.12
N90_Maa90	67.08	67.48	67.52	68.29	67.70	68.38	68.10	69.02
Leu90_Kor90	92.19	86.16	92.89	90.50	92.57	88.50	93.56	92.54
Lei90_Maa90	84.93	88.66	86.92	92.47	85.34	89.13	87.33	93.01
N90_Leu90	39.98	35.86	40.03	35.90	39.85	37.86	39.89	37.88
N90_Kor90	40.69	38.38	40.94	38.55	41.06	40.46	41.29	40.59
B50_N50	67.58	66.58	69.21	68.98	-	-	-	-
B70_N70	77.44	68.36	77.50	68.65	-	-	-	-
B50_B90	59.25	54.49	59.53	55.37	-	-	-	-
N50_N90	65.86	59.64	67.28	61.31	-	-	-	-
B70_N50	59.44	59.77	59.67	59.92	-	-	-	-

WAT DRAGEN WE VANDAAG: EEN HEMD MET BLAZER OF EEN SHIRT MET JASJE?

	11C		11C_LLRL		14C		14C_LLRL	
	U'	U	U'	U	U'	U	U'	U
B90_N70	69.80	64.33	69.80	64.33	-	-	-	-
B50_B70	54.54	54.80	54.99	56.38	-	-	-	-
N50_N70	65.63	59.93	66.22	62.81	-	-	-	-
B70_B90	81.47	74.93	82.21	78.12	-	-	-	-
N70_N90	71.50	71.71	72.35	73.34	-	-	-	-

Tabel 7 Interne uniformiteitswaarden voor elf en veertien kledingconcepten

	11C		14C	
	I'	I	I'	I
B12	77.56	69.68	79.35	74.96
B90	72.61	67.62	72.10	69.21
B70	77.68	63.22	-	-
B50	61.06	57.07	-	-
N12	75.34	71.04	73.71	71.06
N90	74.72	70.74	73.69	68.48
N70	65.86	66.15	-	-
N50	71.05	66.23	-	-
LeuKor12	74.11	76.15	76.28	80.04
LeuKor90	85.56	77.05	86.25	80.78
Leu12	79.48	78.85	80.56	81.63
Leu90	85.39	79.76	86.33	83.41
Kor12	71.49	76.96	74.61	81.35
Kor90	87.29	78.09	87.59	81.08
LeiMaa12	68.43	74.91	69.56	76.89
LeiMaa90	71.77	78.12	73.20	80.50
Lei12	74.21	78.67	75.51	79.98
Lei90	76.34	81.58	76.77	81.88
Maa12	73.50	75.60	73.33	77.39
Maa90	72.85	78.71	74.61	82.89

Tabel 8 Aandeelmaten voor de kenmerken ‘Frans’, ‘Engels’ en ‘afkeuring in taalzuiveringswerken’ voor veertien kledingconcepten

	Frans				Engels				taalzuivering	
	11C		14C		11C		14C		14C	
	A'	A	A'	A	A'	A	A'	A	A'	A
B12	11.83	15.40	10.60	12.10	15.18	26.13	23.92	41.72	10.20	21.89
N12	19.39	16.98	17.03	13.34	7.71	16.24	17.40	32.28	11.98	19.10
B90	17.63	19.20	17.62	17.81	7.83	8.34	15.03	25.05	-	-
N90	31.71	29.59	28.94	23.25	7.80	12.77	14.86	29.35	-	-
B70	14.90	28.50	14.86	22.39	3.29	7.57	3.55	19.44	-	-
N70	20.78	27.12	20.72	21.31	2.87	15.36	3.07	22.07	-	-
B50	29.00	27.39	29.00	21.52	2.87	9.12	2.87	7.16	-	-
N50	39.49	35.48	39.49	27.88	1.75	10.48	1.75	8.23	-	-
LeuKor12	13.91	32.45	12.21	25.49	17.39	17.06	27.48	34.83	23.83	29.13
LeiMaa12	26.46	32.71	22.87	25.70	18.68	20.99	29.71	37.92	22.76	26.82
LeuKor90	17.23	33.80	19.29	33.04	17.53	13.50	21.61	25.55	-	-
LeiMaa90	37.45	41.84	33.76	32.87	22.64	20.70	29.61	36.91	-	-

Appendix 3 Overzicht van de bronnen voor het materiaal uit 2012

Onderstaande bronnen werden geraadpleegd door Tine De Cnodder in de stedelijke bibliotheken van Herentals, Leuven en Eindhoven, alsook in de Koninklijke Bibliotheken van België (Brussel) en Nederland (Den Haag). De eerste twee maanden werden weggelaten omwille van de onbeschikbaarheid van de magazines Flair en Feeling uit die periode in de Koninklijke Bibliotheek van België.

- 1 Kledingnamenbronnen voor het vergelijkingspunt Nederland 2012 (N12)
 - a. Burda
 - b. KNIP
 - c. Margriet
 - d. Esquire
 - e. Cosmopolitan
 - f. Libelle (NL)
- 2 Kledingnamenbronnen voor het vergelijkingspunt België 2012 (B12)
 - a. Flair
 - b. Feeling
 - c. Libelle (BE)