

Grondelaers, Stefan, Hilde Van Aken, Dirk Speelman & Dirk Geeraerts. 2001. "Inhoudswoorden en preposities als standaardiseringsindicatoren. De diachrone en synchrone status van het Belgische Nederlands". *Nederlandse Taalkunde* 6: 179-202.

Inhoudswoorden en preposities als standaardiseringsindicatoren.

De diachrone en synchrone status van het Belgische Nederlands

Stefan Grondelaers
Hilde Van Aken
Dirk Speelman
Dirk Geeraerts*

Abstract

This paper reports on a series of corpus-based investigations into the diachronic and synchronic status of Belgian Dutch, a variant whose standardisation was delayed in the 16th and 17th C and is as yet incomplete. Building on a newly developed measure of linguistic overlap and an extensive corpus of synonymous designations of 30 clothing and soccer concepts, a first study (Geeraerts, Grondelaers & Speelman 1999) confirmed the alleged diachronic convergence between Belgian and Netherlandic Dutch from 1950 to 1990, as well as the larger distance traditionally assumed between standard and substandard language in Belgian Dutch (a major indication of the latter's delayed standardisation). Whereas the first investigation relied on shopwindow materials to represent substandard language, the two follow-up studies reported on in this paper make use of a new corpus containing Internet materials as representatives of substandard register – viz. Internet Relay Chat and Usenet. In the first follow-up study, the diachronic and synchronic findings are replicated on the basis of clothing terms. In the second, the basis of comparison between different variants of Dutch is extended from content words to function words, viz. prepositions. Although the lexical as well as the prepositional indicators fully confirm the diachronic convergence between Belgian and Netherlandic Dutch and their different synchronic stratification, the present methodology does not allow us to compare their respective impact on the latter.

* Stefan Grondelaers, Dirk Speelman en Dirk Geeraerts zijn verbonden aan het Departement Linguïstiek van de KU Leuven. Hilde Van Aken is OIO bij de Vakgroep Algemene Taalwetenschap van de VU Amsterdam. Stefan Grondelaers en Hilde Van Aken zijn de Leuvense en Amsterdamse uitvoerder van de Leuvens-Amsterdamse pool van het VNC-project "Lexicale variatie in het Standaardnederlands. Convergentie/divergentie en standaardisering/substandaardisering in het Belgische en Nederlandse Nederlands". Correspondentie i.v.m. dit artikel wordt bij voorkeur aan hen gericht: stefan.grondelaers@arts.kuleuven.ac.be en hmm.van.aken@let.vu.nl. Het ruimere VNC-project is een samenwerkingsverband tussen de Katholieke Universiteit Leuven, de Vrije Universiteit Amsterdam, de Universiteit Gent, en het I.N.L. Leiden. Over het onderzoek van de Gents-Leidse pool van het project wordt elders in dit nummer gerapporteerd (zie het artikel van Van Den Heede, Deygers, De Caluwe & Taeldeman).

1 Inleiding

Dit artikel rapporteert over het onderzoek van de Leuvens-Amsterdamse pool van het VNC-project "Convergentie en divergentie in de Nederlandse woordenschat". Dit project wil een empirisch ondersteund antwoord vinden op de vraag of het Belgische Nederlands convergeert met het Nederlandse Nederlands, of een autonome variant van het Nederlands vertegenwoordigt. Deze descriptieve vraag is gekoppeld aan een controversieel taalpolitiek debat: moet het Belgische Nederlands het Nederlandse Nederlands als norm overnemen, of is er plaats voor een officiële Belgische variant van de Nederlandse standaard? Wij beperken ons in dit artikel uitsluitend tot de descriptieve kwestie.

Wie niet op de hoogte is van de geschiedenis van het Belgische Nederlands, herinnere zich dat in de Renaissance in de meeste Europese landen een linguïstische standaardiseringsbeweging op gang kwam. In Vlaanderen echter werd de ontluikende standaardisering geblokkeerd als gevolg van het feit dat Vlaanderen in de Tachtigjarige Oorlog politiek geïsoleerd raakte van Nederland. In plaats van een eigen standaard te ontwikkelen, nam Vlaanderen in toenemende mate het Frans over voor supraregionale communicatiedoelinden. Hoewel het Nederlands bij de hereniging van de Nederlanden (1814-1830) opnieuw tot officiële Vlaamse taal gepromoveerd werd, en – vooral vanaf 1930 - zijn standaardisering hervatte, zijn de meeste linguïsten het erover eens dat de standaardisering van het Belgische Nederlands nog niet voltooid is.

Met betrekking tot de status en de ontwikkeling van het Belgische Nederlands zijn in de linguïstische literatuur twee hypothesen te vinden. Om te beginnen wordt een *diachrone convergentie* tussen het Belgische en Nederlandse Nederlands verondersteld, die in hoofdzaak het gevolg is van veranderingen aan Belgische zijde: de standaardisering van het Belgische Nederlands, zo vermoedt men, wordt gekenmerkt door een expliciete normatieve oriëntatie op het Nederlandse Nederlands. Daarnaast manifesteert de standaardiseringsachterstand van het Belgische Nederlands zich in een grotere *synchrone afstand* tussen het Belgische regionale en supraregionale register dan tussen het Nederlandse regionale en supraregionale register; zelfs voor de meest ongetrainde waarnemer is het duidelijk dat de afstand tussen het discours van de modale Belgische wielrenner en het taalgebruik van de Belgische nieuwslezer veel groter is dan de afstand tussen het register van Nederlandse wielrenners en nieuwslezers. In wat volgt verwijzen we naar deze twee hypothesen als resp. de *diachrone convergentiehypothese* en de *synchrone stratificatiehypothese*.

Omdat de synchrone stratificatiehypothese en de diachrone convergentiehypothese haast uitsluitend gebaseerd zijn op anekdotische observatie en m.b.v. enquêtes geëliciteerde data (die niet noodzakelijk overeen hoeven te stemmen met spontaan taalgebruik, zie onder meer Labov 1972, Geeraerts e.a. 1999:24 en Grondelaers 2000:59-63 en 256), lanceerde de onderzoeksgroep rond Dirk Geeraerts in 1994 een grootschalig corpusgebaseerd onderzoek naar de lexicale standaardisering van het Belgische Nederlands (een volledig verslag is te vinden in Geeraerts, Grondelaers & Speelman 1999, verder Geeraerts e.a. 1999). Dat gebeurde op basis van een 40.000 observaties tellende database met synoniemenreeksen voor 15 kleding- en 15 voetbalbegrippen.¹ In de materiaalverzameling werd niet alleen een *geografische* en een *diachrone* dimensie ingebouwd, waardoor Belgisch en

¹ Appendix 1 en 2 bevatten een overzicht van de gebruikte synoniemenreeksen op het gebied van kleding en voetbal.

Nederlands materiaal uit 1950, 1970 en 1990 vergeleken kon worden; we implementeerden ook een *stilistisch-stratificationele* dimensie, door naast supraregionaal materiaal uit modetijdschriften regionale gegevens uit de etalages van kledingboutiques in twee Belgische en twee Nederlandse steden op te nemen.

In paragraaf 2 van dit artikel worden kort de basisgegevens uit Geeraerts e.a. (1999) samengevat. De rest van het artikel is gewijd aan twee nieuwe studies waarin we m.b.v. een speciaal daarvoor samengesteld tekstcorpus de basisgegevens trachten te repliceren. Paragraaf 3 rapporteert over een lexicale replicatie met de kleding- en voetbaltermen uit Geeraerts e.a. (1999); paragraaf 4 is gewijd aan een replicatiestudie op basis van voorzetselverbindingen. In paragraaf 5, tenslotte, vatten we samen en verkennen we enkele voor de hand liggende parcours voor verder onderzoek.

2 Corpusgebaseerd onderzoek naar de standaardisering van het Belgische Nederlands

2.1 Materiaal

De empirische basis van Geeraerts e.a. (1999) bestond uit 40.000 observaties van spontaan taalgebruik. We verzamelden de verschillende benamingen (en hun respectievelijke frequenties) die gebruikt worden om 15 kledingbegrippen en 15 voetbalbegrippen aan te duiden. Op basis van de resulterende database kunnen we voor een concept BUITENSPEL nagaan hoe in Belgische en Nederlandse publicaties de kwantitatieve verhouding is tussen de term *buitenspel* en het leenwoord *offside*; in het geval van het begrip JURK kunnen we vaststellen of er een voorkeur bestaat voor benoeming met *jurk*, *kleet* of *japon*.

Merk op dat de 30 onderzochte kleding- en voetbalbegrippen niet – zoals in het meeste variatieonderzoek - gepreselecteerd werden op basis van hun variatiegevoeligheid. Door ons onderzoek op een representatieve selectie van begrippen uit twee lexicale velden te baseren, trachten we het risico op beeldvervorming te voorkomen: wanneer de belangstelling vooral gericht is op termen waarvan de variatiegevoeligheid vooraf vermoed wordt, dan heeft men er nog geen idee van wat die termen vertegenwoordigen in het geheel van het taalgebruik. Een uitgesproken Belgische voorkeur voor de purismen *regenscherm*, *duimspijker* en *geldbeugel*, bijvoorbeeld, hoeft nog geen divergentie met het Nederlandse Nederlands te impliceren als die puristische tendens kwantitatief gezien verwaarloosbaar is in een voor de rest identieke woordenschat.

De kern van de geobserveerde gegevens bestond uit kleding- en voetbalmateriaal uit 1990. Die kern werd in twee richtingen uitgebreid. Om te beginnen werd vergelijkbaar materiaal verzameld voor 1950 en 1970, hetgeen ons in staat stelde een "real-time"-onderzoek naar lexicale convergentie- of divergentieprocessen uit te voeren. Daarnaast werd rekening gehouden met de stilistische stratificatie van het taalgebruik. Tussen standaard en dialect bevinden zich namelijk verschillende tussenniveaus waarop registers verschillen (zoals een toenemende graad van informaliteit) samengaan met een toenemende geografische specialisatie. Voor een onderzoek naar de linguïstische verhouding tussen het Belgische en het Nederlandse Nederlands zijn die tussenstrata tussen standaard en dialect buitenmate interessant: we kunnen namelijk veronderstellen dat de verschillen tussen Belgisch en Nederlands Nederlands zullen toenemen op dit regiolectische niveau.

In Geeraerts e.a. (1999) werd het tussenliggende taalstratum vertegenwoordigd door termen die we hebben aangetroffen op prijskaartjes en labels in etalages van kledingwinkels in de Belgische steden Leuven en Kortrijk en de Nederlandse steden Leiden en Maastricht. Het beoogde publiek van deze vorm van communicatie is beperkter dan het nationale of binationale publiek van de kledingtijdschriften waaruit de kern van de materiaalverzameling geselecteerd werd. Het feit dat we te maken hebben met geschreven semi-formeel taalgebruik belet anderzijds dat we terecht komen bij het zuiver dialectische eind van het stratificationele continuüm.

2.2 Methode

Hoe kunnen we lexicale convergentie of divergentie tussen twee taalvariëteiten meten? In Geeraerts e.a. (1999) werd een overlappingsmaat gedefinieerd die gebaseerd is op de noties *onomasiologisch profiel* en *uniformiteit*.

Het *onomasiologisch profiel* van een bepaald concept in een bepaalde bron of bronnengroep definiëren we als het geheel van de alternatieve benamingen voor dat concept met een specifieke frequentieverhouding in die bron of bronnengroep. Tabel 1 bevat de onomasiologische profielen voor OVERHEMD in het Belgische en Nederlandse materiaal uit 1990 (resp. B90 en N90 genoemd):

	B90	N90
hemd	31%	17%
overhemd	69%	46%
shirt	0%	37%

Tabel 1: Onomasiologische profielen voor OVERHEMD in het Belgische en Nederlandse materiaal uit 1990

Uniformiteit is een kwantitatieve maat voor de overeenkomst tussen twee onomasiologische profielen. Onze berekening van uniformiteit vertrekt vanuit het idee dat een gemeenschappelijke linguïstische norm gemeenschappelijk taalgedrag genereert. Maximale uniformiteit bij de benoeming van een begrip komt slechts voor wanneer twee taalvariëteiten dezelfde naam voor dat begrip gebruiken, of verschillende namen die identiek gedistribueerd zijn, i.e. dezelfde relatieve frequentie hebben.

Veel frequenter echter dan zulke voorbeelden van maximale uniformiteit zijn partiële overlappingsen zoals in tabel 1. Laat ons, bij wijze van illustratie, veronderstellen dat de relatieve frequenties in tabel 1 100 reële gebruikgevallen in elk van beide profielen vertegenwoordigen. De partiële overlapping tussen de profielen in tabel 1 kwantificeren we dan door in beide profielen de gebruikgevallen te tellen die een tegenhanger hebben in het andere profiel. De ideale scenario's uit de vorige paragraaf – twee taalvariëteiten hebben één term voor hetzelfde concept of meerdere termen maar dan met dezelfde frequentieverhouding - leveren in dit verband een maximale uniformiteit op. In tabel 1, echter, vinden we 14 gevallen van *hemd* in B90 die geen tegenhanger in N90 hebben; 23 Belgische *overhemden* hebben geen Nederlandse tegenhanger, en er zijn geen Belgische tegenhangers voor de 37 Nederlandse *shirts*. Op het totaal van 200 benoemingsinstanties in beide profielen, hebben m.a.w. slechts $200 - (14 + 23 + 37) = 126$ gevallen een tegenhanger in het andere profiel, hetgeen een uniformiteit van $126/2 = 63\%$ oplevert. Ter wille van het

kwantitatieve gemak moet opgemerkt worden dat dat percentage altijd gelijk is aan de som van de kleinste relatieve frequentie voor iedere alternatieve term (die in tabel 1 telkens schuingedrukt is), i.e. $17+46+0=63\%$.

Als we de uniformiteit tussen twee taalvariëteiten op meer dan één begrip willen baseren, definiëren we uniformiteitsindex U als het gemiddelde van de uniformiteitsindexen per individueel begrip. In die berekeningswijze wordt aan een hoogfrequent concept zoals ROK in principe hetzelfde gewicht toegekend als een laagfrequent concept zoals, bijvoorbeeld, OVERSCHOEN. Vanuit een pragmatisch perspectief op standaardisering, echter, zou men kunnen stellen dat een gebrek aan uniformiteit minder vaak tot communicatieve misverstanden zal leiden bij laag- dan bij hoogfrequente concepten. Om bij uniformiteitsberekeningen rekening te houden met de frequentie van de onderzochte begrippen, definiëren we U' als een *gewogen* gemiddelde waarin de relatieve frequentie van elk begrip verrekend wordt.

We illustreren beide berekeningswijzen met een concreet voorbeeld (bijkomende illustraties zijn te vinden in Geeraerts e.a. 1999:41 e.v.). De onderstaande tabel 2 bevat in resp. de tweede en de vierde kolom de grijsgearceerde onomasiologische profielen in het Nederlandse en het Belgische 1970-corpus van de begrippen JACK, TRUI en JURK. De vijfde kolom bevat de minima per onomasiologisch alternatief, op basis waarvan in de zesde kolom de individuele uniformiteitswaarden 37.3, 55 en 76.8 berekend worden. De globale index U onderaan in tabel 2 is het gemiddelde van de drie individuele uniformiteitswaarden.

	N70	rel	B70	rel	<i>min</i>	U	G	U'
jack								
bloes	0	0,0	0	0,0	0,0	37,3	0,03	1,0
blouson(netje)	0	0,0	31	59,6	0,0			
jack/jek	36	97,3	18	34,6	34,6			
jacket	0	0,0	0	0,0	0,0			
jak(je)	0	0,0	0	0,0	0,0			
windbloes(je)	0	0,0	0	0,0	0,0			
windjack	1	2,7	3	5,8	2,7			
windjacket	0	0,0	0	0,0	0,0			
<i>totaal</i>	37		52					
trui								
pull(etje)	3	3,6	105	48,6	3,6	55,0	0,09	4,9
pullover	24	28,6	29	13,4	13,4			
trui(tje)	57	67,9	82	38,0	38,0			
<i>totaal</i>	84		216					
jurk								
japon(netje)	351	25,7	40	2,5	2,5	76,8	0,88	67,9
jurk(je)	1015	74,3	1493	94,7	74,3			
kleed(je)	0	0,0	43	2,7	0,0			
<i>totaal</i>	1366		1576		$U =$	56,4	$U' =$	73,8

Tabel 2: Gewogen en ongewogen uniformiteit tussen Belgische en Nederlandse benamingen uit 1970 voor de begrippen JACK, TRUI en JURK

In deze berekeningswijze van U wordt aan elk van de drie concepten hetzelfde gewicht toegekend, hetgeen gezien de lage individuele uniformiteitswaarden voor JACK en TRUI tot een lage globale uniformiteit $U=56.37$ leidt. Als we er (fictief) van uitgaan dat de gehele Belgische en Nederlandse materiaalverzameling voor 1970 in dit voorbeeld uit benamingen voor JACK, JURK en TRUI bestaat, dan vertegenwoordigen de twee begrippen die voor een lage uniformiteit zorgen nauwelijks 12% van de materiaalverzameling, terwijl het veel uniformer benoemde JURK ($U=76.84$) maar liefst 88% van het materiaal representeert. Als we die asymmetrie in onze berekening tot uiting willen brengen – omdat we geloven dat gebrek aan uniformiteit vooral bij hoogfrequente begrippen tot communicatiestoornissen zal leiden –, dan vermenigvuldigen we de individuele uniformiteitswaarden met de wegingsfactor G uit de voorlaatste kolom, die de relatieve frequentie van de individuele begrippen in de materiaalverzameling reflecteert. Aangezien het begrip JACK in de Belgische en Nederlandse materiaalverzamelingen samen $37+52=89$ keer voorkomt, en de Belgische en Nederlandse materiaalverzamelingen samen 3331 observaties bevatten, is G voor JACK gelijk aan $89/3331=0.03$. Vermenigvuldigd met de ongewogen uniformiteit voor JACK, levert dat een gewogen uniformiteit $U'=0.9971$ op. Dezelfde berekeningswijze levert voor JURK een wegingsfactor G van 0.8832, en een gewogen uniformiteit van maar liefst 67.8688 op! De gewogen uniformiteit U' voor de hele materiaalverzameling krijgen we vervolgens door de gewogen waarden voor de drie begrippen bij elkaar op tellen: $0.9971+4.9499+67.8688=73.82$; dat laatste getal reflecteert veel adequater dan de ongewogen uniformiteit het feit dat het meest frequente concept in ons fictieve voorbeeld het meest uniform benoemd wordt. In wat volgt doen we dan ook uitsluitend uitspraken op basis van U' .

2.3. Resultaten

Laten we, alvorens de resultaten te bespreken, de diachrone convergentiehypothese en de synchrone stratificatiehypothese operationaliseren in termen van uniformiteit. Diachroon gezien kunnen we convergentie of divergentie tussen variëteiten kwantificeren als resp. stijgende of dalende uniformiteit tussen die variëteiten. De grotere synchrone afstand tussen Belgisch regionaal en supraregionaal taalgebruik die we veronderstellen zal zich – indien onze hypothese correct is – manifesteren in een lagere uniformiteit tussen het etalage- en het tijdschriftenmateriaal in België dan in Nederland.

Figuur 1 bevat de relevante synchrone en diachrone gegevens. De linker- en rechteras in de figuur vertegenwoordigen het Belgische (B) materiaal, terwijl de rechtersas het Nederlandse (N) materiaal representeert ("50", "70", "90", en "shop" staan voor resp. het 1950-, het 1970-, het 1990-, en het etalagemateriaal). De waarde op de lijnen tussen de cirkels vertegenwoordigen de gewogen uniformiteit U' tussen de variëteiten in de cirkels. Bij de vergelijking tussen de meetpunten B50/N50, B70/N70 en B90/N90 geven de cijfers voor de schuine streep de kledinggegevens, en de cijfers na de schuine streep de voetbalgegevens weer.²

² In verband met de materiaalverzameling waarop de replicatie gebaseerd is moeten nog twee opmerkingen gemaakt worden. Om te beginnen moet opgemerkt worden dat de uniformiteitswaarden in figuur 1 berekend werden op basis van 12 kledingconcepten i.p.v. de 15

insereer figuur 1 ongeveer hier

De gegevens in figuur 1 vormen een duidelijke bevestiging van de diachrone convergentiehypothese zowel als de synchrone stratificatiehypothese. De stijgende uniformiteit tussen het Belgische en het Nederlandse Nederlands suggereert om te beginnen een uitgesproken lexicale convergentie tussen beide variëteiten:³

1.	U'(B50,N50)	<	U'(B70,N70)	<	U'(B90,N90)
	KL 69.63	<	75.29	<	86.62
	VO 66.22	<	72.04	<	77.08

Synchroon gezien manifesteert de vertraagde standaardisering van het Belgische Nederlands zich inderdaad in een opvallend lagere uniformiteit tussen het Belgische tijdschriften- en etalagemateriaal dan tussen het Nederlandse tijdschriften- en etalagemateriaal:

2.	U'(B90,shop)	<	U'(N90,shop)
	47.63	<	67.94

3 Naar een lexicale replicatie van de gegevens uit Geeraerts e.a. (1999)

3.1 Materiaal

Hoewel de data in figuur 1 een solide bevestiging van de synchrone stratificatiehypothese en de diachrone convergentiehypothese vormen, is onze zaak voorlopig nog niet waterdicht. Nog afgezien van Van de Velde's (1996) observatie dat er op het vlak van de uitspraak een opvallende diachrone *divergentie* is tussen het Belgische en Nederlandse Nederlands, zijn er in ons eigen onderzoek tekortkomingen die betrekking hebben op de afhankelijke en de onafhankelijke onderzoeksvariabele. Om te beginnen is het empirische fundament van onze bevindingen te smal: twee lexicale velden vormen een te eenzijdige vergelijkingsbasis om veilige uitspraken te kunnen doen over de ontwikkeling en de interne structuur van het Belgische Nederlands. Daarnaast doet de huidige implementering van de stratificatieparameter geen recht aan de synchrone complexiteit van het Belgische Nederlands. Tot nu toe hebben we slechts één tussentaalstratum in het onderzoek betrokken – geïnstantieerd door het etalagemateriaal -, terwijl in de literatuur twee (Geerts 1975), drie (Baetens Beardsmore & Van de Craen 1979), en zelfs vijf (Willemyns 1987)

uit de oorspronkelijke studie (hetgeen de afwijkingen tussen de uniformiteitswaarden in figuur 1 en de cijfers in Figuur 5.2 van Geeraerts e.a. 1999:69 verklaart). In tegenstelling tot Geeraerts e.a. (1999) is de replicatie namelijk gebaseerd op een tekstcorpus zonder afbeeldingen van de kledingstukken waarnaar verwezen wordt, en zonder afbeeldingen is de precieze referentie van de begrippen VEST, COLBERT (M) en BLAZER (V) onvoldoende vaststelbaar. Om de gegevens uit Geeraerts e.a. (1999) veilig met de replicatiegegevens te kunnen vergelijken werden die drie begrippen in dit artikel dan ook uit beide analyses verwijderd. Het ontbreken van synchrone gegevens voor de voetbalbegrippen is daarnaast het logische gevolg van het feit dat voetbalbegrippen zoals BUITENSPEL en HOEKSCHOP (uiteraard) niet in winketalages voorkomen.

³ In Geeraerts e.a. (1999:67) wordt van een interpreteerbaar verschil tussen variëteiten gesproken wanneer de uniformiteitsgegevens voor die variëteiten tenminste 5% verschillen. Verschillen van minder dan 5% vallen binnen een statistische foutenmarge.

stilistisch verschillende variëteiten van het Belgische Nederlands geïdentificeerd werden.

In de rest van dit artikel richten we onze aandacht dan ook op twee bijkomende onderzoeksvragen: (i) kunnen we de bevindingen uit Geeraerts e.a. (1999) met behulp van een ander corpus, en – meer in het bijzonder – een ander type informeel Nederlands bevestigen, en (ii) kunnen we de diachrone convergentiehypothese en synchrone stratificatiehypothese repliceren als we de vergelijkingsbasis uitbreiden van inhoudswoorden zoals kleding- en voetbaltermen naar relationele woorden zoals preposities?

Om die vragen te beantwoorden werd een omvangrijk tekstcorpus van ongeveer 45.000.000 woorden samengesteld. Net zoals het corpus in Geeraerts e.a. (1999), is het CONDIV-corpus gestructureerd op basis van drie types variatie.⁴ Het geografische verschil tussen Belgisch en Nederlands Nederlands accommoderen we door in het corpus een Nederlandse tekstenbank – die in totaal 19.505.008 woorden telt –, en een Belgische tekstenbank – die 25.984.343 woorden telt – te includeren.

Bij de aanleg van de materiaalverzameling werd eveneens rekening gehouden met de stilistische stratificatie van het Nederlands. Omdat het onmogelijk is a priori vast te stellen hoeveel tussenstrata precies onderscheiden moeten worden (cf. supra), gaan we niet in de eerste plaats op zoek naar taalgebruik dat representatief is voor een bepaald stratum, maar vertrekken we van verschillende types *taalsituaties* als onafhankelijke variabele, en beschouwen we het daarin gebruikte *taalgebruik* als afhankelijke variabele. Stratificatievariëteit accommoderen we dan door taalmetaal te vergelijken uit vijf stilistisch verschillende taalsituaties. Om te beginnen bevat het CONDIV-corpus drie soorten krantenmetaal: we onderscheiden *kwaliteitskranten* zoals *De Standaard* en het *NRC Handelsblad* en *massakranten* zoals *Het Laatste Nieuws* en *De Telegraaf*. Er zijn *nationale* massakranten – zoals *Het Laatste Nieuws* en *De Telegraaf* –, maar ook *regionale* massakranten zoals *Het Belang van Limburg*, *De Gazet van Antwerpen* of *De Limburger*.

Daarnaast werden aan het CONDIV-corpus twee types Internetaal toegevoegd als vertegenwoordigers van het informele Nederlands. IRC – Internet Relay Chat – is een Internetmodule waarin *online* en *synchroon* gecommuniceerd kan worden in babbelkanalen zoals #Leuven en #Vlaanderen. Omdat IRC-ers zoveel mogelijk een conversationeel spreek/schrijf-tempo willen benaderen, en aan dat tempo hun spelling-, morfologie-, en grammaticanormen aanpassen, kan IRC beschouwd worden als "geschreven spreektaal"; IRC-ers manifesteert dan ook colloquialismen en dialectismen die in gewone geschreven teksten ontbreken (voor een overzicht van de linguïstische en sociologische kenmerken van IRC, zie onder meer Herring 1996, Ko 1996, Hentschel 1998, Bays 1998, en Paolillo 1999). Voorbeeld (3) bevat een kort extract uit het babbelkanaal #Leuven:

3. <boink> thebuzz, komde gij ier gewoon reclaam maken of wa?
<SarAHken> loooooooooooooooooooooooooooooootje !
<King_John> jaja
<atttila> daar kunde ook al een goeie pc voor kopen zenne
* DeepStar (((shakes))) lotje
<Gonzy> lotjeeeeeeeeeeeeeuuuuuuuuuuuuuuuh

⁴ Het CONDIV-corpus is genoemd naar het VNC-project "Convergentie en divergentie in de Nederlandse woordenschat" waarvoor het werd samengesteld. Een uitgebreide toelichting van het corpus is te vinden in Grondelaers, Deygers, Van Den Heede, Van Aken & Speelman (2000).

<thebuzz> neen, gewoon eens vragen, maar ben nu al weer weg
 <SarAHken> lotje : how are you !? long time no see
 <boink> thebuzz, ge moe daarvoor nie frutten ;)
 <King_John> zijde weg buzz?
 <Grooveke> attila : voor 92000 ballen ebde nog gene goeie pc
 <King_John> allé jong
 <thebuzz> ja, kzijn weg, nog een beke gaan werken

Aangezien Nederlandse en Vlaamse IRC-ers eigen kanalen hebben waarin users uit het andere land niet welkom zijn, en gewoonlijk zelfs automatisch uitgesloten worden (zie onder meer Grondelaers 2000:65 e.v. en Grondelaers, Deygers, Van Den Heede, Van Aken & Speelman 2000:358 e.v.), beschikken we over aparte corpora met zeer informeel Belgisch en zeer informeel Nederlands Nederlands.

Naast IRC werd in het CONDIV-corpus ook Usenetmateriaal geïntegreerd. Usenet is een internetmodule waarin *offline* en *asynchroon* in nieuwsgroepen wordt gediscussieerd d.m.v. e-mailboodschappen.⁵ Aangezien e-mail een veel indirecter communicatiemedium is dan IRC, en academische internetoperatoren uitsluitend "nette" nieuwsgroepen op hun server aanbieden, ligt het voor de hand dat het Nederlands op Usenet formeler zal zijn dan dat op IRC. De linguïstische context waarin IRC ontstaat is in elk geval informeler dan die waarin "Usenetese" tot stand komt: terwijl IRC "immediate" en "online" is (Bays 1998), kunnen Usenetters hun boodschappen dankzij het offline karakter van nieuwsgroepen nakijken alvorens ze te posten (cf. onder meer Harrison 1998 voor een gedetailleerder overzicht van de linguïstische kenmerken van Usenet).

De vijf materiaaltypes die in het corpus opgenomen zijn kunnen op een informaliteitsas uitgezet worden die als meest informele pool het IRC-materiaal heeft, en als meest formele pool het materiaal uit de kwaliteitskranten:

4. IRC > Usenet > regionale massakranten > nationale massakranten > kwaliteitskranten⁶

In de derde plaats bevat ook het CONDIV-corpus een belangrijke diachrone dimensie, waarop materiaal uit 1958, 1978 en 1998 gecontrasteerd wordt. Aangezien IRC en Usenet recente mediafenomenen zijn, kunnen ze niet ingevuld worden voor de oudere periodes, en daardoor bevat het diachrone corpus alleen oudere tegenhangers van de kranten. De afwezigheid van elektronisch beschikbare historische krantenedities noopte ons ertoe krantenmateriaal uit 1958 en 1978 in te scannen. Dat het diachrone corpus qua omvang beperkter is dan het synchrone, is een rechtstreeks gevolg van de hoge kostprijs (en de grote foutenmarge⁷) die met het inscannen van sterk verzuurde, poreuze krantenpagina's gepaard gaat.

⁵ IRC is *online* en *synchroon* omdat users onmiddellijk kunnen reageren op een boodschap die in een IRC-kanaal op hun scherm verschijnt. Usenet daarentegen is *offline* en *asynchroon* omdat er "propagation delay" (Paolillo 1999) is tussen het opstellen, nalezen, verzenden, ontvangen, lezen en beantwoorden van een e-mailboodschap.

⁶ Met deze ordening correleren factoren zoals "geïntendeerd bereik" en "mate van redactionele controle" (zie Grondelaers e.a. 2000:358).

⁷ Het synchrone corpus bevat 22.984.343 woorden Belgisch Nederlands (IRC = 8.207.007; Usenet = 4.980.780; Het Belang van Limburg = 3.012.330; De Gazet van Antwerpen = 3.068.405; Het Laatste Nieuws = 3.486.911; De Standaard = 3.228.910), en 19.505.008 woorden Nederlands

Aangezien het CONDIV-corpus niet, zoals het corpus uit de eerste studie, een verzameling manueel geselecteerde kleding- en voetbalnamen is, maar een digitaal tekstcorpus, gebeurt de extractie van de taalverschijnselen die we willen bestuderen op basis van elektronische bevraging. Daartoe ontwikkelde de derde auteur van dit artikel een browser op linguïstenmaat: *Abundantia Verborum* (Speelman 1997) bevat een module om digitale bestanden tot bruikbare linguïstische corpora te transformeren, een zoekmachine die de targetdata in die corpora identificeert, een module om de resulterende gegevens te labelen, classificeren en manipuleren, en een component waarin statistische analyses op de gegevens uitgevoerd kunnen worden.⁸

3.2 Methode en resultaten

In deze replicatiestudie werd dezelfde kwantitatieve methodologie gehanteerd als in de oorspronkelijke studie. De afhankelijke variabele was uniformiteit in de benaming van 12 kledingconcepten.⁹ De onafhankelijke variabelen waren *regio* – België vs. Nederland – en *stratificatie* – IRC vs. Usenet vs. kranten.

insereer hier ongeveer figuur 2

We brengen nog even in herinnering dat we in dit onderzoek de diachrone convergentie tussen Belgisch en Nederlands Nederlands trachten te repliceren, en de grotere synchrone afstand tussen suprareginaal/formeel en regionaal/informeel Belgisch Nederlands dan tussen suprareginaal/formeel en regionaal/informeel Nederlands Nederlands. Merk op dat de synchrone stratificatiehypothese nog verfijnd kan worden: als de relatief grotere afstand tussen informeel en formeel Belgisch Nederlands in het nieuwe onderzoek inderdaad bevestigd wordt, dan zal het feit dat Usenet minder informeel lijkt dan IRC zich in het Belgische materiaal vertalen in een hogere uniformiteit tussen het Usenet- en het krantenmateriaal, dan tussen het IRC- en het krantenmateriaal.

De relevante gegevens staan in figuur 2, die op dezelfde manier georganiseerd is als figuur 1, met het verschil dat het synchrone deel aan Belgische zowel als Nederlandse kant uit drie vergelijkingspunten is samengesteld. De synchrone gegevens bevestigen trouwens zowel de oorspronkelijke als de verfijnde versie van de synchrone stratificatiehypothese. De grotere Belgische afstand tussen regionale en supraregionale taal manifesteert zich het duidelijkst in IRC:

$$\begin{array}{rcl}
 5. & U'(B98,IRCB98) & < & U'(N98,IRCN98) \\
 & 82.63 & < & 89.82
 \end{array}$$

Nederlands (IRC = 6.965.291 woorden; Usenet = 7.748.436 woorden; De Limburger = 1.680.636; De Telegraaf = 1.590.581; NRC Handelsblad = 1.520.064). Het diachrone corpus telt 3.587.035 woorden Belgisch Nederlands (De Gazet van Antwerpen 1958 = 734.018; De Gazet van Antwerpen 1978 = 699.255; De Standaard 1958 = 1.018.658; De Standaard 1978 = 1.135.104) en 2.079.728 woorden Nederlands Nederlands (Nieuwe Rotterdamsche Courant 1958 = 1.064.001; Nieuwe Rotterdamsche Courant 1978 = 1.015.727).

⁸ Meer informatie over *Abundantia Verborum* is te vinden op <http://www.ling.arts.kuleuven.ac.be/genling/abundant>

⁹ We beperken de replicatie tot de kledingtermen waarvoor in Geeraerts e.a. (1999) en in paragraaf 2 etalagemateriaal voorhanden is.

En de uniformiteit tussen IRC en krantentaal is inderdaad kleiner in het Nederlandse en (vooral) het Belgische Nederlands dan de afstand tussen het Usenet-materiaal en het krantenmateriaal:

$$\begin{array}{rcl}
 6. & U'(B98,IRCB98) & < & U'(B98,UsenetB98) \\
 & 82.63 & < & 90.76 \\
 & & & \\
 & U'(N98,IRCN98) & < & U'(N98,UsenetN98) \\
 & 89.82 & \cong & 92.07
 \end{array}$$

De diachrone convergentiehypothese daarentegen wordt niet rechtstreeks bevestigd door de nieuwe gegevens:

$$\begin{array}{rcl}
 7. & U'(B58,N58) & < & U'(B78,N78) & < & U'(B98,N98) \\
 & 71.2 & > & 64.34 & < & 82.02
 \end{array}$$

In plaats van een progressieve stijging – zoals in de initiële gegevens – tonen de data in figuur 2 een "geknikt" patroon, met een daling in het materiaal uit 1978. Dat die daling geen probleem vormt voor de diachrone convergentiehypothese tonen de data in de derde kolom van tabel 3. Daarin worden de beide bronnen die B58 en B78 constitueren – *De Gazet van Antwerpen* en *De Standaard* – in de tweede en de derde balk apart met de enige Nederlandse bron – *NRC Handelsblad* – vergeleken; in de eerste balk zijn ter vergelijking nog eens de globale cijfers opgenomen. De gegevens in tabel 3 laten zien dat de initieel vastgestelde convergerende beweging wél opduikt in de uitgesplitste *Standaard*- en *Gazet*-gegevens.

	Abs Freq staal A	Abs Freq staal B	U' globaal	U' n=5
U'(B58,N58)	101	54	71,2	71,67
U'(B78,N78)	109	45	64,34	60,9
U'(B98,N98)	890	391	82,02	82,15
U'(GVA58,NRC58)	36	54	45,27	37,36
U'(GVA78,NRC78)	64	45	46,82	43,36
U'(GVA98,N98)	241	391	81,49	81,49
U'(STA58,NRC58)	65	54	71,38	72,2
U'(STA78,NRC78)	45	45	82,18	78,42
U'(STA98,N98)	103	391	81,37	80,02

Tabel 3: Absolute frequenties en gewogen uniformiteit bij kledingnamen in Belgische en Nederlandse kranten uit 1958, 1978, en 1998.

Het geknikte patroon in de globale gegevens blijkt in belangrijke mate het gevolg van frequentieasymmetrieën tussen *De Gazet van Antwerpen* en *De Standaard*: de globale Belgische uniformiteitswaarden in 1958 en 1978 correleren namelijk met de individuele uniformiteitswaarde van de Belgische bron die qua frequentie het zwaarst

doorweegt op die punten.¹⁰ Dat de globale Belgische uniformiteit in 1958 ($U'=71.2$) naar de 1958-uniformiteit tussen *Standaard* en *NRC* ($U'=71.38$) neigt is een gevolg van het feit dat de Belgische materiaalverzameling uit 1958 65 observaties uit *De Standaard* en slechts 36 observaties uit *De Gazet* bevat. Het gecombineerde effect van *Gazet*-overwicht in 1978 (64 observaties t.o.v. 45 observaties uit *De Standaard*) en een lage individuele uniformiteit tussen *Gazet* en *NRC* ($U'=46.82$) brengt de globale Belgische uniformiteit in 1978 flink omlaag ($U'=64.34$). In 1998 zijn de afzonderlijke krantenstalen qua omvang evenmin in evenwicht, maar dat is wel het geval voor de uniformiteitscijfers, die resp. $U'=81.49$ en $U'=81.37\%$ bedragen, en tot een globale uniformiteit van $U'=82.02$ leiden.

In verband met de cijfers in tabel 3 moeten nog drie belangrijke opmerkingen gemaakt worden. Om te beginnen kan de materiaalschaarste op de diachrone meetpunten enigszins gecompenseerd worden door in een alternatieve telling alleen begrippen te betrekken die in beide vergeleken stalen tenminste vijf keer voorkomen ($n=5$). De resultaten van deze uitgezuiverde methode - die in de vierde kolom van tabel 3 staan - bevestigen de niet-uitgezuiverde gegevens op elk vlak, zodat we voorlopig (met enige mate van voorzichtigheid en in afwachting van verdere replicatie) kunnen stellen dat we met interpreteerbare tendensen te maken te hebben.

Daarnaast moeten we nog even stilstaan bij de verschillen in diachrone ontwikkeling van *De Standaard* en *De Gazet van Antwerpen*. Hoewel beide bronnen globaal gezien een stijgende uniformiteit met het *NRC*-materiaal en een nagenoeg identieke 1998-uniformiteit vertonen (U' resp. 81.49 en 81.37), verschilt hun specifieke ontwikkeling op twee belangrijke punten. *De Standaard* manifesteert van meet af aan een veel hogere uniformiteit met het *NRC*-materiaal dan *De Gazet* ($U'=71.38 > 45.27$), en kent zijn sterkste stijging op het tweede meetpunt. *De Gazet* is in 1958 zowel als 1978 opvallend laag uniform met *NRC*, en begint pas in 1998 uitgesproken lexicale overlapping met *NRC* te vertonen. Binnen het korte bestek van dit artikel, waarin we ons noodgedwongen tot het externe onderzoeksperspectief uit Geeraerts, Grondelaers & Spielman (1999) beperken, kunnen we niet nagaan welke variabelen verantwoordelijk zijn voor de vertraagde uniformisering in *De Gazet*.¹¹ Wat de hogere beginuniformiteit van *De Standaard* t.o.v. *De Gazet* wél suggereert, is dat de individuele lexicale evolutie van de Belgische bronnen bepaald lijkt te worden door de stratificationele status van die bronnen: de hogere aanvangsuniformiteit van *De Standaard* t.o.v. *De Gazet* zou een gevolg kunnen zijn van de hogere positie die de *De Standaard* als kwaliteitskrant op het informaliteitscontinuüm in (4) bekleedt. Wegens het ontbreken van historische gegevens over de lexicale evolutie in de informeelste bronnen kan deze hypothese uitsluitend op basis van replicatie geverifieerd worden. We voegen deze hypothese dan ook toe aan het lijstje van verwachtingen die in de volgende case studie getest moeten worden.

Tenslotte moet nog opgemerkt worden dat de synchrone uniformiteit tussen IRC en krantenmateriaal aan Belgische zowel als Nederlandse zijde onintuïtief hoog

¹⁰ Tabel 3 bevat in de eerste en tweede kolom de absolute frequenties van de stalen die in de uniformiteitsberekening vergeleken worden. Staal A is daarbij de bron voor de komma en staal B de bron na de komma.

¹¹ In Geeraerts e.a. (1999:10) wordt een onderscheid gemaakt tussen een *extern* onderzoeksperspectief - waarbij de geanalyseerde taalvariëteiten als één geheel in het onderzoek betrokken worden - en een *interne* benadering, waarbij binnen de onderzochte variëteiten naar specifieke linguïstische variabelen gekeken wordt die met de externe convergentie of divergentie correleren.

ligt (resp. 82.63 en 89.82).¹² Die opvallende mate van lexicale overlapping tussen IRC en krantentaal betekent niet alleen dat IRC-ers in weerwil van hun voorliefde voor idioom en jargon relatief conformistisch zijn in hun benoeming van alledaagse begrippen zoals kledingstukken. De hoge uniformiteit tussen IRC en de kranten suggereert tevens dat de globale informaliteit van IRC niet in de eerste plaats door lexicale factoren bepaald wordt.¹³ Een correcte inschatting van de linguïstische verhouding tussen formeel en informeel Nederlands is m.a.w. bij uitstek gebaat bij een uitbreiding van de vergelijkingsbasis naar niet-lexicale parameters. In de volgende paragrafen richten we onze aandacht daarom op de rol van voorzetsels als standaardiseringsindicatoren.

4 Preposities als standaardiseringsindicatoren

4.1 Selectie van de data

De rest van dit artikel wijden we aan prepositionele variatie in voorzetselverbindingen zoals *zich interesseren aan/voor* of *langs/via het internet*. Hoewel preposities geen typische functiewoorden zijn, verschillen ze van inhoudswoorden doordat ze geen "zelfstandigheid" (ANS 1997:139) aanduiden (zoals de kleding- en voetbaltermen), maar de aard van de relatie tussen twee zelfstandigheden specificeren (idem:507).

Om de digitale extractie van bruikbare prepositionele verbindingen te vergemakkelijken, hebben we de selectie beperkt tot verbindingen die syntagmatisch sterk afgebakend zijn, en daardoor relatief makkelijk opspoorbaar zijn in een elektronisch corpus. Dat zijn in de eerste plaats "vaste verbindingen" (ANS 1997:1169), waarbij een voorzetsel een vaste combinatie vormt met een specifiek werkwoord, adjectief, of substantief. De verbindingen die ons in het licht van dit onderzoek interesseren zijn natuurlijk degene die met meer dan één voorzetsel kunnen voorkomen, zoals *zich interesseren voor/aan* (ibidem). Een tweede groep van voor onze doeleinden interessante voorzetselcombinaties zijn verbindingen waarin de prepositie een beperkte en daardoor makkelijk opspoorbare set van complementen selecteert, zoals "Hij betrad het huis *langs/via het raam/het venster/de deur/de poort*, etc."

Met het gebruik van preposities als standaardiseringsindicatoren zijn twee belangrijke methodologische problemen verbonden (zie ook Van Aken & Van Den Heede: ter perse). Om te beginnen ligt de "veldaanpak" die we in 2.1 ter voorkoming van statistische beeldvervorming propageerden bij preposities moeilijker. Nog afgezien van het feit dat het minder zinvol is in verband met de onderlinge relaties tussen preposities van een "lexicaal veld" te spreken,¹⁴ zijn we voor het opstellen van

¹² Het is verleidelijk om deze hoge mate van lexicale overlapping te interpreteren aan de hand van de opvallend lage uniformiteit tussen het etalagemateriaal en het tijdschriftenmateriaal in figuur 1 (B 47.63 en N 67.94). Echter, IRC en etalagemateriaal verschillen linguïstisch gezien zo sterk van elkaar dat ze onmogelijk op een veilige manier met elkaar vergeleken kunnen worden: terwijl etalageverwijzingen naar kledingstukken exclusief lexicaal zijn, manifesteert IRC namelijk een globale informaliteit die door lexicale zowel als niet-lexicale parameters bepaald wordt.

¹³ De beperkte impact van lexicale factoren op de globale informaliteit van een bron werd onder meer geobserveerd in Goossens (2000:9 e.v.), waarin beweerd wordt dat de linguïstische eigenheid van het Belgische Nederlands in belangrijker mate door morfo-grammaticale parameters (in het bijzonder de vorm en flexie van enkele determineerders) dan door lexicale factoren bepaald wordt.

¹⁴ Volgens Van Aken & Van Den Heede: ter perse is het in principe mogelijk om spatiële voorzetsels die doel, herkomst, pad e.d. uitdrukken als een lexicaal veld te zien. Dergelijke aanpak is bij

een exhaustieve lijst van variabele preposities in hoofdzaak aangewezen op overzichten uit de taalbeheersingsliteratuur (bijvoorbeeld Penninckx & Buyse 1997 en Paardekooper 1990), die niet zelden opgesteld zijn om de gebruiker te waarschuwen voor te mijden regionale variatie zoals *zoeken achter* (de "foute" variant die beperkt is tot het Belgische Nederlands) vs. *zoeken naar* (de "juiste" variant). Omdat we in dit stadium van het onderzoek niet kunnen nagaan hoe representatief dergelijke gevallen zijn voor het geheel van het taalgebruik (maar cf. infra), lopen we met het gedwongen gebruik van regionale-variatiegevoelige indicatoren het risico een onrealistisch divergent beeld van de relatie tussen het Belgische en het Nederlandse Nederlands te schetsen. De prepositionele gegevens die verderop besproken worden moeten m.a.w. met de grootst mogelijke omzichtigheid benaderd worden.

Minstens even problematisch is de vaststelling dat het ondanks de rigide syntagmatische afbakening van de gekozen verbindingen moeilijk blijft om volledige synonymie tussen de prepositionele varianten in een onomasiologisch profiel te garanderen. Is denotationele synonymie bij inhoudswoorden relatief gemakkelijk vast te stellen,¹⁵ bij preposities kan betekenisgelijkheid in veel mindere mate gegarandeerd worden. Cuyckens (1991:182 e.v.) geeft een aantal voorbeelden van bijna-synoniemen bij preposities en laat zien dat de keuze voor het ene dan wel het andere voorzetsel heel vaak terug te voeren is op een *verschillende conceptualisering* van de uitgedrukte situatie. De keuze tussen de voorzetsels *in* en *op* in de zin (8), bijvoorbeeld, hangt af van de manier waarop de respectievelijke landmarks *veld* en *kasteel* worden waargenomen.¹⁶ Als het landmark gepercipieerd wordt als een onbegrensde tweedimensionele oppervlakte, dan is *op* het aangewezen voorzetsel. Wordt het landmark daarentegen geconcipieerd als een *begrensde* tweedimensionele oppervlakte, dan is *in* geschikter (1991:183):

8. Het elftal staat *in/op* het veld

Dergelijke conceptuele verschillen zijn veel moeilijker op het spoor te komen wanneer de relatie tussen de verschillende entiteiten figuurlijk is in plaats van ruimtelijk:

9. Hij is leraar *aan/op* de Militaire School

10. *Aan/op* onze universiteit gebeurt veel wetenschappelijk onderzoek

In (9) en (10) (idem:198) is de keuze voor *aan* of *op* wellicht niet semantisch bepaald, en dus lijken beide voorzetsels verwisselbaar (en bruikbaar voor onze onderzoeksdoeleinden). Voorzichtigheid blijft hoe dan ook geboden, want specifieke

voorzetsels die louter op basis van variabiliteit en opzoekbaarheid geselecteerd werden natuurlijk veel minder evident.

¹⁵ Omdat Geeraerts, Grondelaers & Bakema (1994) gebaseerd is op een semasiologisch verrijkt corpus (waarin de referenten van de bestudeerde kledingnamen afzonderlijk toegankelijk zijn), konden we het onderzoek in Geeraerts, Grondelaers & Speelman (1999) beperken tot begrippen waarvoor namen bestaan die tenminste denotationeel synoniem zijn.

¹⁶ Het begrippenpaar "trajector-landmark" is uit de Cognitieve Linguïstiek afkomstig (zie onder meer Langacker 1987:217). In de prepositionele relatie die uitgedrukt wordt door een zin als *Het glas staat op tafel*, is *het glas* de "trajector", i.e. het gelokaliseerde object, terwijl *de tafel* het "landmark" is dat fungeert als referentiepunt voor de lokalisering van de trajector.

semantische of syntactische contexten kunnen die verwisselbaarheid blokkeren. Dat illustreert Taylor (1988:312-313) met de Engelse preposities *over* en *above*:

11. He lives *over* me
12. He lives *above* me

Hoewel informanten nauwelijks verschil merken tussen (11) en (12), bewijst Taylor dat er wel degelijk een semantisch onderscheid is, dat zichtbaar wordt wanneer preciezer informatie over de afstand tussen trajector en landmark aan de zin wordt toegevoegd:

13. He lives ten floors *above* me
14. *He lives ten floors *over* me

Volgens Taylor is *over* onmogelijk in (14) omdat de afstand tussen trajector en landmark te groot is voor invloed van de trajector op het landmark. Dat laatste blijkt nochtans een voorwaarde voor het gebruik van *over*: waar het voorzetsel *above* uitsluitend signaleert dat de trajector zich op een hoger niveau bevindt dan het landmark, bezit *over* de extra betekenisnuance dat de trajector het landmark op de een of andere wijze moet kunnen beïnvloeden.

In de analyse werden alleen verbindingen betrokken waarvan de prepositionele varianten daadwerkelijk onderling verwisselbaar bleken. Na uitzuivering van ongewenste semasiologische variatie werden 16 bruikbare verbindingen geselecteerd die met 2 of 3 prepositionele varianten per profiel in totaal 7937 keer in het totale materiaal voorkwamen.¹⁷

4.2. Resultaten

Op basis van die gegevens werden drie reeksen diachrone uniformiteitscijfers berekend, die in tabel 4 gepresenteerd worden.

	globaal	gva-nrc	sta-nrc
U' 1958	58.66	49.46	66.02
U' 1978	61.97	61.02	63.36
U' 1998	67.14	73.62	73.26

Tabel 4: Gewogen uniformiteit bij voorzetsels uit Belgisch en Nederlands krantenmateriaal uit 1958, 1978 en 1998

In de linkerkolom bevinden zich de globale Belgische gegevens die gebaseerd zijn op alle beschikbare bronnen per meetpunt (i.e. *De Standaard* en *De Gazet van Antwerpen*). Aangezien de globale ontwikkeling in de vorige case studie sterk contrasteerde met de individuele ontwikkeling in de constituerende bronnen, worden de gegevens in de tweede en derde kolom uitgesplitst voor, respectievelijk, *De Gazet van Antwerpen* en *De Standaard*. Alvorens de resultaten te bespreken brengen we nog even in herinnering dat we op basis van de gegevens in tabel 4 twee diachrone convergentiehypothesen over de status van het Belgische Nederlands trachten te

¹⁷ De 16 prepositionele profielen worden in Appendix 3 met een voorbeeld toegelicht.

bevestigen: we zoeken niet alleen bijkomende confirmering voor de verwachting dat het Belgische en Nederlandse Nederlands progressief convergeren, maar we trachten ook de correlatie tussen de aard van de convergerende beweging in een bepaalde bron en de stratificationele status van die bron te repliceren.

Aangezien het prepositiemateriaal op de verschillende meetpunten geen vertekenende absolute-frequentieverschillen laat zien¹⁸ - zoals bij de inhoudswoorden – zijn de globale prepositiedata in de linkerkolom interpreteerbaar. Getuige die gegevens bevestigen de prepositionele data i.t.t. de lexicale replicatiedata ook globaal de convergerende ontwikkeling tussen het Belgische en het Nederlandse Nederlands die in Geeraerts e.a. (1999) gevonden werd. Afgezien van de lichte terugval in *De Standaard* 1978 (die weliswaar binnen de vijf percent-marge valt), bevestigen de individuele ontwikkelingen per bron eveneens de convergerende tendens.

De interessantste observatie die in verband met tabel 4 gemaakt kan worden is de verschillende aard van de convergerende ontwikkeling in *De Gazet* en *De Standaard*, die in belangrijke mate het patroon bij de inhoudswoorden bevestigt. Net zoals bij de kledingtermen heeft *De Gazet* in 1958 een markante uniformiseringsachterstand op *De Standaard*, maar eindigen beide bronnen in 1998 op een nagenoeg identieke uniformiteitshoogte van $U' =$ resp. 73.62 en 73.26! Hoewel de interpretatie van absolute cijfers in deze case studie wegens de huidige grofkorreligheid van het meetinstrument gevaarlijk is, en hoewel de gegevens van slechts twee kranten in de berekeningen betrokken werden, suggereren de dubbel gerepliceerde individuele ontwikkelingskenmerken van *De Gazet* en *De Standaard* dat diachrone ontwikkeling en stratificationeel-stilistische status wel degelijk correleren in de Belgische bronnen: de bulk van de convergerende ontwikkeling tussen het Belgische en het Nederlandse Nederlands komt niet op rekening van de Belgische kwaliteitskranten (die van meet af aan al vrij uniform zijn met de Nederlandse kranten en slechts beperkt verder convergeren), maar van de Belgische massakranten, die aanvankelijk opvallend afwijken van de Nederlandse kranten, maar in 40 jaar tijd hun uniformiteitsachterstand op de kwaliteitskranten inhalen.

Tabel 5 bevat de synchrone uniformiteitsgegevens ("Use", "IRC", "Inf" en "Kr" staan voor het Usenet-materiaal, het IRC-materiaal, het totale informele materiaal (Usenet + IRC) en het krantenmateriaal):

	B	NL
U'(Use,Kr)	81.57	96.58
U'(IRC,Kr)	75.21	93.85
U'(Inf,Kr)	80.02	96.74

Tabel 5: Gewogen uniformiteit bij voorzetsels uit Belgisch en Nederlands kranten- en Internetmateriaal uit 1998

Met deze gegevens trachten we de twee synchrone bevindingen uit de vorige casusstudie te bevestigen. Dat de afstand tussen formeel en informeel Belgisch

¹⁸ Behalve dan op het meetpunt '78, waar het *Standaard*-materiaal 263 observaties bevat, terwijl het *Gazet*-materiaal slechts 164 observaties telt. Omdat de uniformiteit tussen *Standaard* en *NRC* en *Gazet* en *NRC* in 1978 resp. 61.02 en 63.36 bedraagt, levert dat frequentieverschil geen afwijkende globale uniformiteit tussen het beschikbare Belgische en Nederlandse materiaal op.

Nederlands groter is dan de afstand tussen formeel en informeel Nederlands. Nederlands blijkt om te beginnen eveneens uit het prepositiemateriaal, en zelfs in veel hogere mate dan bij de inhoudswoorden:

$$\begin{array}{rcl}
 15. & U'(\text{Inf B,Kr B}) & < & U'(\text{Inf N,Kr N}) \\
 & 80.02 & < & 96.74^{19}
 \end{array}$$

En ook de verfijnde synchrone stratificatiehypothese over de hogere uniformiteit in het Belgische materiaal tussen Usenet en kranten dan tussen IRC en kranten wordt door de prepositiegegevens bevestigd; in het Nederlandse materiaal overlappen Usenet en IRC in vergelijkbare mate met het krantenmateriaal:

$$\begin{array}{rcl}
 16. & U'(\text{Use B,Kr B}) & > & U'(\text{IRC B,Kr B}) \\
 & 81.57 & > & 75.21 \\
 & & & \\
 & U'(\text{Use N,Kr N}) & > & U'(\text{IRC N,Kr N}) \\
 & 96.58 & \equiv & 93.85
 \end{array}$$

Als uit het globale overzicht van lexicale en positionele resultaten één observatie onthouden moet worden, dan is het wel de vaststelling dat er tussen de geobserveerde ontwikkelingspatronen bij inhoudswoorden en preposities uitgesproken parallellieën bestaan. In afwachting van veiligere meetinstrumenten (cf. infra) rechtvaardigen die parallellieën voorlopig de grofkorreligheid van de positionele tellingen.

5 Conclusies en consequenties

We vatten samen. Aan de hand van 12 kledingbegrippen, 15 voetbalbegrippen en 16 voorzetselverbindingen hebben we in dit artikel met behulp van de kwantitatieve methodologie uit Geeraerts, Grondelaers & Speelman (1999) onderzocht of er in de periode van 1958 tot 1998 sprake is (i) van een reële convergentie of divergentie tussen het Belgische en het Nederlandse Nederlands, en (ii) van een grotere afstand tussen het formele en informele Belgische Nederlands dan tussen het formele en informele Nederlandse Nederlands. Anders dan in het meeste traditionele variatieonderzoek, is onze benadering niet gebaseerd op geëliciteerde gegevens, maar op spontaan tot stand gekomen taalgebruik uit een omvangrijk corpus dat niet alleen regionaal en historisch, maar ook stilistisch-stratificationeel gecontroleerd werd: in het corpus zijn naast formele taal uit massa- en kwaliteitskranten namelijk ook twee soorten Internetaal – IRC en Usenet – opgenomen als vertegenwoordigers van het informele Nederlands. Met die nadruk op feitelijk taalgebruik hebben we

¹⁹ Aangezien bij 11 van de 16 prepositionele verbindingen daadwerkelijk variatiegevoeligheid werd vastgesteld (omdat één of meer varianten per profiel wel in België maar niet in Nederland voorkomt), is enige voorzichtigheid geboden bij de interpretatie van de uitzonderlijk hoge synchrone uniformiteitscijfers in de Nederlandse kolom van tabel 5: uniformiteitscijfers die in 11 van de 16 gevallen gebaseerd zijn op profielen met Belgische varianten zullen logischerwijze hoger liggen voor het Nederlandse materiaal (waarin die Belgische varianten ontbreken) dan cijfers die op basis van een kleiner aantal niet-variatiegevoelige verbindingen tot stand kwamen. Dat de gegevens in tabel 5 niettemin betrouwbaar zijn toont een herrekening die beperkt is tot de 6 niet-variatiegevoelige verbindingen in het materiaal (met prepositionele alternatieven die zowel in België als in Nederland voorkomen): noch voor Nederland noch voor België levert die reanalyse namelijk uitgesproken lagere cijfers op: $U'(\text{IRC N,Kr N}) = 92.43$; $U'(\text{Use N,Kr N}) = 93.26$; $U'(\text{Inf N,Kr N}) = 94.51$ en $U'(\text{IRC B,Kr B}) = 73.88$; $U'(\text{Use B,Kr B}) = 80.80$; $U'(\text{Inf B,Kr B}) = 78.33$.

vertekeningen trachten te voorkomen die het gevolg zouden kunnen zijn van de kloof tussen wat taalgebruikers doen met taal en wat ze denken te doen met taal.

Onze analyse levert de volgende empirische bevindingen op:

- Er zijn onweerlegbare kwantitatieve aanwijzingen dat het Belgische en het Nederlandse Nederlands in de periode 1958-1998 in belangrijke mate naar elkaar toegroeien op het vlak van de inhoudswoorden zowel als de preposities.
- Die globale convergentie moet echter enigszins genuanceerd worden. Zowel de analyse op basis van de inhoudswoorden als die op basis van de preposities suggereert namelijk dat de precieze diachrone convergentietoename per Belgische bron afhankelijk is van de stilistisch-stratificationele status van die bron. Individuele convergentiemetingen tussen *De Gazet van Antwerpen* en *NRC* en *De Standaard* en *NRC* tonen m.n. aan dat het taalgebruik van Belgische kwaliteitskranten in 1958 in hoge mate aansluit bij de Nederlandse krantentaal, terwijl de taal van Belgische regionale massakranten aanvankelijk flink afwijkt van Nederlandse krantentaal. Het is in dit licht opvallend dat kwaliteits- zowel als massakranten op lexicaal zowel als prepositioneel vlak naar precies hetzelfde convergentieniveau evolueren.
- I.v.m. de synchrone stratificatie van het Belgische Nederlands bevestigen inhoudswoorden zowel als preposities dat de afstand tussen supraregionaal/formeel en regionaal/informeel taalgebruik in België beduidend groter is dan in Nederland. De aanwezigheid in het corpus van twee soorten Internetmateriaal met een verschillende stratificationele status stelt ons in staat de net genoemde synchrone conclusie te preciseren: de intermediaire positie tussen IRC en krantentaal die "Usenetese" stilistisch gezien lijkt in te nemen blijkt uit een grotere afstand tussen IRC en krantentaal dan tussen Usenet en krantentaal.

Met deze meervoudig gerepliceerde bevindingen is de zaak van het Belgische Nederlands echter nog niet rond. Op korte termijn zijn m.n. twee types vervolgonderzoek nodig als we de uitgangsvragen van dit artikel over de status en de structuur van het Belgische Nederlands met enige mate van vertrouwen willen beantwoorden.

Om te beginnen moeten een aantal praktische en methodologische problemen aangepakt worden die met het beschreven onderzoek verbonden zijn. Zo is het diachrone corpus dat we momenteel hanteren te klein voor onderzoek naar inhoudswoorden; dat de globale diachrone uniformiteitscijfers in tabel 3 oninterpreteerbaar zijn heeft in de eerste plaats met extreme materiaalschaarste te maken, en dus moet gezocht worden naar haalbare alternatieven voor de dure en momenteel nog weinig efficiënte methode van het krantenscannen.

Een ernstiger methodologisch defect is de basis waarop de prepositionele indicatoren geselecteerd werden. Dat de uit taalbeheersingsoverzichten geselecteerde verbindingen haast per definitie variatiegevoelig zijn is vooral problematisch omdat in dit stadium van het onderzoek niet nagegaan kan worden hoe representatief de gekozen indicatoren voor het geheel van het taalgebruik zijn. In een volgende fase van het onderzoek pakken we dit probleem aan door *Abundantia Verborum* op basis van een zo exhaustief mogelijke adjectief- en prepositielijst (uit bijv. CELEX) op zoek te laten gaan naar *alle* adjectief-prepositieverbindingen waarvan het collocatiegehalte – berekend op basis van formele collocatiematen – een

bepaalde drempel overschrijdt.²⁰ Het spreekt vanzelf dat deze methode de kwestie van de variatiegevoeligheid én het probleem van de representativiteit meteen oplost.²¹ Aangezien dergelijke "veilige" data momenteel nog niet voorhanden zijn, moeten de bestaande prepositionele gegevens met de grootst mogelijke voorzichtigheid benaderd worden. Dat het methodologisch gevaarlijk zou kunnen zijn individuele prepositionele uniformiteitsgegevens in absolute zin te interpreteren, neemt echter niet weg dat er tussen de geobserveerde patronen bij inhoudswoorden en preposities een opvallende en geruststellende parallelie bestaat, die in zekere zin als een a posteriori rechtvaardiging van de verzamelde prepositionele gegevens beschouwd kan worden.

De hoge uniformiteit tussen IRC-ees en krantentaal die inhoudswoorden en preposities tonen suggereert tenslotte dat de globale informaliteit van IRC wellicht niet in de eerste plaats door specifieke woord- of prepositiekeuzes bepaald wordt. En dus dient onze vergelijkingsbasis niet alleen van lexicale naar niet-lexicale standaardiseringsindicatoren uitgebreid te worden, maar moeten we ook kwantitatieve methodes vinden om de respectievelijke impact van lexicale en niet-lexicale factoren op de globale informaliteit van IRC en andere bronnen te berekenen. Dat laatste is een noodzakelijke voorwaarde voor inzicht in de complexe stratificationele structuur van het Belgische Nederlands; dat inzicht is op zijn beurt een essentiële voorwaarde voor een empirisch gefundeerd antwoord op de uitgangsvraag waarmee we dit artikel begonnen. Hoe overtuigend ook, de hier gepresenteerde gegevens beantwoorden die vraag (nog) niet.

Bibliografie

- Aken, H. Van & V. Van Den Heede (2001).** Measuring lexical variation: a corpus-based study of prepositions and verbs in Standard Dutch. Verschijnt in *Language variation in Europe. Proceedings of the First International Conference on Language Variation in Europe*.
- Baetens Beardsmore, H. & P. Van de Craen (1979)** . The development of triglossia in Flemish Belgium. In: M. van de Velde & W. Vandeweghe (red.), *Sprachstruktur, Individuum und Gesellschaft. Akten des 13. Linguistischen Kolloquiums 1*. Tübingen: Niemeyer, 191-200.
- Bays, H. (1998).** Framing and face in internet exchanges: a socio-cognitive approach. *Linguistik Online* 98/1. Online publicatie op <http://viadrina.eu.v-frankfurt-o.de/~wjjournal/bays.html>
- Cuyckens, H. (1991).** *The semantics of spatial prepositions in Dutch: A cognitive-linguistic exercise*. Doctorale dissertatie Universiteit Antwerpen.
- Geeraerts, D., S. Grondelaers & P. Bakema (1994).** *The structure of lexical variation. Meaning, naming, and context*. (Cognitive Linguistics Research 5). Berlijn: Mouton de Gruyter.

²⁰ Zie Stubbs (1996) voor een overzicht van collocatiematen.

²¹ Merk op dat de automatische methode die we hier voorstellen bij voorkeur met adjectief-voorzetselcombinaties zoals *gek op/van* uitgevoerd wordt, als we tenminste de verwarring tussen bepalingen en voorzetselvoorwerpen willen vermijden die frequent optreedt bij werkwoord-prepositieverbindingen. De computer maakt namelijk geen onderscheid tussen het voor onze doeleinden interessante Belgische voorzetselvoorwerppatroon *Wij zoeken achter onze portefeuille* en het oninteressante bepalingenpatroon *wij zoeken achter de kast*.

- Geeraerts, D., S. Grondelaers & D. Speelman (1999).** *Convergentie en divergentie in de Nederlandse woordenschat. Een onderzoek naar kleding- en voetbaltermen.* Amsterdam: Meertensinstituut.
- Geerts, G. (1975).** *Voorlopers en varianten van het Nederlands. Een gedocumenteerd dia- en synchroon overzicht.* Leuven: Acco.
- Goossens, J. (2000).** De toekomst van het Nederlands in Vlaanderen. *Ons Erfdeel* 43, 3-13.
- Grondelaers, S. (2000).** *De distributie van niet-anaforisch er buiten de eerste zinsplaats. Sociolexicologische, functionele en psycholinguïstische aspecten van er's status als presentatief signaal.* Doctorale dissertatie K.U. Leuven.
- Grondelaers, S. & D. Geeraerts (1998).** Vagueness as a euphemistic strategy. In: A. Athanasiadou & E. Tabakowska (red.), *Speaking of emotions: conceptualization and expression.* Berlin: Mouton de Gruyter, 357-374.
- Grondelaers, S., K. Deygers, H. Van Aken, V. Van Den Heede & D. Speelman (ter perse).** Het CONDIV-corpus geschreven Nederlands. *Nederlandse Taalkunde* 5/4, 356-363.
- Harrison, S. (1998).** E-mail discussions as conversation: moves and acts in a sample from a listserv discussion. *Linguistik Online* 98/1. Online publicatie op <http://viadrina.euv-frankfurt-o.de/~wjjournal/harrison.htm>
- Hentschel, E. (1998).** Communication on IRC. *Linguistik Online* 98/1. Online publicatie op <http://viadrina.euv-frankfurt-o.de/~wjjournal/irc.html>
- Herring, S. (1996).** *Computer-mediated communication: linguistic, social and cross-cultural perspectives.* Amsterdam: Benjamins.
- Ko, Kwang-Kyu (1996).** Structural characteristics of computer-mediated language: a comparative analysis of Interchange discourse. *Electronic Journal of Communication* 6/3. Online publicatie op <http://ublib.buffalo.edu/libraries/eresources/ejournals/records/ejc.html>
- Labov, W. (1972).** Some principles of linguistic methodology. *Language in Society* 1, 97-120.
- Langacker, Ronald W. (1987).** *Foundations of cognitive grammar. Volume I: Theoretical prerequisites.* Stanford, CA: Stanford University Press.
- Paardekooper, P.C. (1990).** *ABN gids: taalnormboek.* Antwerpen: Standaard.
- Paolillo, J. (1999).** The virtual speech community: social network and language variation on IRC. *Journal of Computer-Mediated Communication* 4/4. Online publicatie op <http://jcmc.huji.ac.il/paolillo.htm>.
- Penninckx, W. & P. Buyse (1997).** *Correct taalgebruik.* Kortrijk-Heule: UGA
- Speelman, D. (1997).** *Abundantia Verborum. A computer tool for carrying out corpus-based linguistic case studies.* Doctorale dissertatie K.U. Leuven.
- Stubbs, M. (1996).** *Text and corpus linguistics.* Oxford: Blackwell.
- Taylor, J.R. (1988).** Contrasting prepositional categories: English and Italian. In: B. Rudzka-Ostyn (red.), *Topics in Cognitive Linguistics.* Amsterdam: Benjamins, xxx-xxx.
- Velde, H. van de (1996).** *Variatie en verandering in het gesproken Standaard-Nederlands (1935-1993).* Doctorale dissertatie K.U. Nijmegen.
- Willemyns, R. (1987).** Norm en variatie in Vlaanderen. In: J. de Rooij (red.), *Variatie en norm in de standaardtaal.* Amsterdam: Meertensinstituut, 142-164.

Appendix 1: onomasiologische profielen voor kledingconcepten

overhemd (m)

hemd
overhemd
shirt

bloes (v)
bloes(je)
bloeze/bl(oe/ou)se

t-shirt (m/v)
T-shirt
shirt

trui (m/v)
pull(etje)
pullover
trui(tje)

broek (m/v)
broek
pantalon

jeans (m/v)
jeans
spijkerbroek

legging (v)
caleçon
legging
leggings

rok (v)
rok(je)

jurk (v)
japon(netje)
jurk(je)
kleed(je)

jack (m/v)
bloes
blouson(netje)
jack/jek
jacket
jak(je)
windbloes(je)
windjack
windjacket

kostuum (m)
colbertcostuum
costuum
pak

mantelpak (v)
complet
costuum(pje)
deux-pièces
ensemble
geheel
mantelcostuum(pje)

mantelpak(je)
pak(je)
tailleur(tje)
tailleurcostuum
tweestuks

Appendix 2: onomasiologische profielen voor voetbalconcepten

aanvaller

aanvaller
forward
puntspeler
speerpunt
spits
spitspeler
voorhoedespeler
voorman
voerspeler
voorwaarts

aftrap

aftrap
kick-off

buitenspel

buitenspel
off-side

doelman

doelkeeper
doelman
doelverdediger
doelwachter
goal-keeper
goalie
keeper

doelpunt

doel
doelpunt
goal
treffer

handspel

handbal
handfout
hands
handspel
handsspel

hoekschop

corner
hoekschop
hoekschot
hoekshot

overtreding

foul
fout

overtreding

scheidsrechter

arbiter
kampleider
match-leader
ref(eree)
scheids
scheidsrechter
spelleider
voetballeider
wedstrijdleider

strafschop

elfmeter
elfmetertrap
penalty
strafschop

tegenaanval

counter
tegenaanval
tegenactie
tegenoffensief
tegenstoot

verdediger

achterhoedespeler
achterman
achterspeler
back
back-speler
verdediger

voorzet

center
centerschot
voorzet

vrije trap

free kick
vrij schot
vrije schop
vrije trap
vrijschop
vrijschot

wedstrijd

confrontatie
duel
kamp
match
ontmoeting
partij
treffen
wedstrijd

Appendix 3: onomasiologische profielen voor voorzetselverbindingen

interesse hebben [in/voor]

'Wolf beweert dat de interesse [in/voor] het satanisme dit jaar flink is gegroeid'

[op/aan] **het einde**

'Pas [op/aan] het einde van het seizoen waren die groeipijnen voorbij'

zeggen [aan/tegen/tot]

'Ze kan dan rechtstreeks [aan/tegen/tot] kardinaal Danneels zeggen dat de paus zich wat moderner moet kleden'

kant [op/uit] **gaan**

'Om maar te zeggen dat 't met Canal+ inderdaad niet de goeie kant [op/uit] gaat'

[aan/in/tegen] **tempo**

'Bij het Toneelhuis volgen de voorstellingen elkaar aan een razend tempo op'

[langs/via] **weg**

'Misschien kan ik je [langs/via] deze weg helpen'

[langs/via] (de) **radio/telefoon/internet/...**

'[langs/via] de telefoon kan zij dan onmiddellijk antwoorden en zien wat zij voor de patiënt kan doen'

[langs/via] **raam/venster/deur/poort**

'Telkens drongen ze de woningen binnen [langs/via] de achterdeur'

[aan/van/langs] **alle/beide/geen/sommige/twee... kanten/zijden**

'Djeezus mja, da piept hier [aan/van/langs] alle kanten'

[aan/langs] **weerszijden/weerskanten**

'In de Zeelaan stonden de wagens [aan / langs] weerszijden gestationeerd'

[aan/tegen/voor] ... **prijs/prijzen**

'Dan zijn er in de meeste studentenkroegen consumpties [aan / tegen / voor] scherpe prijzen te bekomen'

kopen/verkopen...[aan/tegen/voor] + **geldbedrag**

'Eind vorig jaar hadden we in voorverkoop kaarten [aan/tegen/voor] 450 frank gekocht'

[aan/tegen/voor] + **geldbedrag per stuk/eenheid**

'Toen sommige multinationals begonnen, kregen ze grond [aan/tegen/voor] 50fr. per vierkante meter'

allergisch/[over][on]gevoelig [aan/voor]

'Ik ben allergisch [aan/voor] snorren'

zich interesseren [aan/voor]

'Het kongres staat open voor alle dames en heren die zich [aan/voor] deze problemen interesseren'

[langs/aan/van] **de ene kant/de andere kant**

'De kranten zijn onafhankelijker geworden. Maar [langs/aan/van] de andere kant is er de toenemende invloed van de commercie'