

KOEN PLEVOETS, DIRK SPEELMAN & DIRK GEERAERTS

DE VERSPREIDING VAN DE *-E(N)*-UITGANG IN ATTRIBUTIEVE POSITIE

Abstract⁽¹⁾

This article focuses on a well-known aspect of masculine nouns in the Southern Dutch dialects that has not received much attention in the literature, however: the fact that attributive elements are optionally inflected with an *-e(n)*-suffix. The appearance of the *-n* is moreover regulated by a number of phonological conditions. By means of the Corpus of Spoken Dutch and fourteen determiners that were operationalised as linguistic variables, we investigate three questions: the validity of the set of phonological restrictions just mentioned, the problem of the demarcation of the three genders in the light of the masculinisation tendency, and the interplay between internal and external factors in the explanation of the presence of *-e(n)*. It is revealed that, for the first question, the restrictions listed in the literature are to a large extent valid, but that they also exhibit quite a range of variation. For the second question, the difference between the three genders is shown to be statistically significant, rendering the evidence for the masculinisation process in fact problematic. This is substantiated in the analysis of the third question where the *-e(n)*-suffix is found to be a marker of register rather than gender, and that it is furthermore characteristic of speakers of the new elite in Flanders that emerged under the economical growth after World War Two and the accompanying rise in affluence.

1. Inleiding

In de Zuid-Nederlandse dialecten kunnen attributieve elementen bij masculiene nomina in het enkelvoud soms met een *-e(n)*-uitgang vervoegd worden, zoals in *ne groten hond* bijvoorbeeld. Het volledige flectiesysteem staat in Tabel 1:

⁽¹⁾ Koen Plevoets, Dirk Speelman & Dirk Geeraerts. OE Kwantitatieve Lexicologie en Variationale Linguïstiek, Subfaculteit Taalkunde, Faculteit Letteren, KU Leuven. Contactadres: Blijde Inkomststraat 21, bus 3308, 3000 Leuven. <http://www.ling.arts.kuleuven.be/qlvl/>

	m.	v.	o.
enkelv.	-e(n) 'ne grote man', 'ne groten hond'	-Ø 'een grote vrouw'	-Ø 'ee(n) groot kind'
meerv.	-Ø 'de grote mannen', 'de grote honden'	-Ø 'de grote vrouwen'	-Ø 'de grote kinderen'

Tabel 1: Het flectiesysteem in de Zuid-Nederlandse dialecten

De $-e(n)$ -uitgang is historisch gezien een accusatiefuitgang bij het masculiene genus (zie Van Loey 1964: 120). Mettertijd drong hij evenwel ook door in subjectpositie en werd daarmee geherinterpreteerd als een genusuitgang; dit is een ontwikkeling die kaderde in het proces van deflexie (Hermkens 1985, Nijen Twilhaar 1992). Het optreden van de $-n$ wordt daarbij gereguleerd door de fonetische context (Taeldeman 1980, 1983): de $-n$ komt enkel voor als de eerstvolgende klank een vocaal is, een h, d, t, in mindere mate een b, en voor sommige dialecten ook een r (zie De Schutter 1966, 1967 voor een bespreking van de regionale variatie).

Externe variatie vormt een opvallende kentering van de distributie van de $-e(n)$ -uitgang. In een corpusgebaseerde studie waarin de frequentie van bepaalde standaard en niet-standaard features voor enkele extralinguïstische factoren met elkaar werd vergeleken, heeft Plevoets (2008) naast regio ook bij register, leeftijd, opleidingsgraad, beroepsgroep en sekse vastgesteld dat het masculiene genus grotere verschillen vertoont dan de niet-masculiene genera. Deze bevinding werd daarbij gekoppeld aan de tendens tot masculinisering (Geerts 1966, 1988): als aan de ene kant niet-masculiene nomina namelijk de neiging hebben om het masculiene genus over te nemen, en er aan de andere kant een algehele informalisering van het taalgebruik optreedt, die erin bestaat dat dialectelementen in de bovenregionale omgangstaal worden opgenomen (wat aanleiding heeft gegeven tot het ontstaan van de Vlaamse 'tussentaal'), dan zullen het de dialectfeatures van het masculiene genus zijn waarin de determinerende invloed van de relevante factoren het sterkst tot uiting zal komen.

In dit artikel zullen we daarom de verspreiding van de $-e(n)$ -uitgang aan een onderzoek onderwerpen. Het beeld van het attributieve flectiemorfem dat uit dit

alles oprijst is er namelijk één van een complex taalverschijnsel waarin interne factoren zoals masculinisering en fonetische condities interageren met externe factoren zoals regionale spreiding en standaardisering. In concreto zullen we ons toelekken op ten eerste, de fonetische condities die het voorkomen van de *-n* beregelen; ten tweede, het probleem van de genusafbakening en verwant daarmee de masculinisering; en ten slotte, het relatieve aandeel van alle factoren – zowel interne als externe – in het verklaren van het al dan niet optreden van de *-e(n)*-uitgang. Omdat het onderzoek in de eerste plaats exploratief zal zijn, zullen we er ons van onthouden om expliciete probleemstellingen te formuleren. In paragraaf 2 zullen we het corpus bespreken, om in paragraaf 3 de methodologie voor te stellen waarmee we de relevante observaties uit het corpus gehaald hebben. Paragraaf 4 zal dan uit de analyses bestaan, en we sluiten af met enkele concluderende beschouwingen in paragraaf 5.

2. Corpus

De verspreiding van de *-e(n)*-uitgang in attributieve positie zullen we trachten te bestuderen aan de hand van het Corpus Gesproken Nederlands (Oostdijk 2000). Het Corpus Gesproken Nederlands (afgekort: CGN) werd samengesteld in de periode van 1998 tot 2003 en telt ongeveer tien miljoen woorden. Voor onze doeleinden gebruiken we enkel het Vlaamse deel dat daarvan één derde uitmaakt (iets meer dan drie miljoen woorden).

Het CGN is een gestratificeerd corpus, in de zin dat het uit veertien verschillende spreeksituaties bestaat, die “componenten” worden genoemd. De veertien componenten, die in het CGN zelf elk met een letter van het alfabet worden aangeduid, zijn:

- a: spontane conversaties (‘face-to-face’ dialogen)
- b: interviews met leraren Nederlands
- c: telefoondialogen (opgenomen m.b.v. platform)
- d: telefoondialogen (opgenomen m.b.v. minidisc)
- f: interviews en discussies uitgezonden op radio en televisie
- g: (politieke) discussies, debatten, vergaderingen
- h: lessen opgenomen in het klaslokaal
- i: spontane (sport-)commentaren (uitgezonden op radio en tv)
- j: actualiteitsrubrieken en reportages (uitgezonden op radio en tv)
- k: nieuwsbulletins (uitgezonden op radio en tv)

- l: beschouwingen en commentaren (uitgezonden op radio en tv)
- m: missen, lezingen, plechtige toespraken
- n: colleges, voordrachten, lezingen
- o: voorgelezen teksten

Daarbij zijn de eerste zeven componenten meer *dialogische* spreesituaties, terwijl de laatste zeven componenten eerder *monologische* spreesituaties.

Ten slotte is het CGN ook geannoteerd voor allerhande informatie op zowel het niveau van de linguïstische eigenschappen van elk token – een ‘part-of-speech’-tag – als op het metaniveau van de kenmerken van de spreker. Die laatste verschaffen ons de nodige extralinguïstische indicatoren voor ons onderzoek. In concreto zullen we ons toespitsen op vijf ervan, te weten: regio, leeftijd, opleidingsgraad, beroepsgroep en sekse.

Voor regio hanteert het CGN de volgende vierdelige classificatie (de afkortingen zijn van eigen makelij):

- bra: Antwerpen & Vlaams-Brabant
- ovl: Oost-Vlaanderen
- wvl: West-Vlaanderen
- lim: Limburg

Wat de variabele leeftijd betreft, registreert het CGN het geboortjaar van elke spreker. Die hebben we gegroepeerd in een aantal generatiecohorten. Daartoe baseren we ons op de klassieke indeling uit Becker (1992). Dat geeft de volgende vijf leeftijdsgroepen:

- pre: Vooroorlogse generatie (1910-1929)
- sil: Stille generatie (1930-1939)
- pro: Protestgeneratie (1940-1954)
- los: Verloren generatie (1955-1970)
- pra: Pragmatische generatie (1971-...)

Opleidingsgraad is in het CGN de vertrouwde drieledige onderverdeling:

- low: diploma lager onderwijs
- mid: diploma middelbaar onderwijs
- hie: diploma hoger onderwijs

Voor beroepsgroep hanteert het CGN tien categorieën. Deze zijn:

- occA: werkzaam in het hoger management of bestuur
- occB: beroep dat een hogere opleiding veronderstelt (arts, advocaat, ingenieur, ...)
- occC: werkzaam als onderzoeker of docent aan universiteit of instelling voor hoger onderwijs
- occD: werkzaam in administratieve functie of functie in de zorgsector
- occE: ongeschoold beroep
- occF: werkzaam als zelfstandige
- occG: politicus
- occH: werkzaam in de media (journalist, reporter) of als artiest
- occI: student, stagiair
- occJ: niet beroepsmatig actief

Doordat deze classificatie niet waterdicht is, komen er echter ook een twintigtal combinatiecategorieën voor, zoals bijvoorbeeld: een advocaat én politicus = ‘occB+G’. Elk van deze categorieën hebben we gereduceerd tot wat we intuïtief aanvoelden als de meest representatieve van de tien hoofdcategorieën. In ons voorbeeld betekent dit dat als de persoon in kwestie vooral in politieke praatprogramma’s aan het woord kwam, we hem als een politicus (‘occG’) classificeerden, en anders als een advocaat (‘occB’) – zie Plevoets (2008: 24) voor een exhaustief overzicht van alle reducties.

De laatste extralinguïstische variabele is sekse. Voor de vrouwelijke sprekers zullen daarbij de afkorting ‘F’ gebruiken, en voor de mannelijke sprekers ‘M’.

3. Methodologie

Voor onze studie van de verspreiding van de *-e(n)*-uitgang in attributieve positie hebben we veertien taalvariabelen opgesteld waarbij elke taalvariabele een terminator is die optioneel met de *-e(n)*-uitgang gelecteerd wordt. De veertien taalvariabelen staan in Tabel 2 (het symbool ‘-Ø’ beduidt de niet-gelecteerde variant, ‘-e(n)’ de gelecteerde variant):

	$-\emptyset$	$-e(n)$
1.	de	den
2.	een	ne(n)
3.	geen	gene(n)
4.	iedere	iederēn
5.	elke	elken
6.	deze	dezen
7.	die	dieje(n), diene(n)
8.	mijn	mijne(n)
9.	uw	uwe(n)
10.	zijn	zijne(n)
11.	haar	hare(n)
12.	onze	onzen
13.	hun	hunne(n)
14.	ADJ-e	ADJ-en

Tabel 2: de veertien taalvariabelen met optionele $-e(n)$ -uitgang

Het zal daarbij duidelijk zijn dat variabele 2, 3, 7, 8, 9, 10, 11, en 13 vormelijk anders geled zijn dan variabele 1, 4, 5, 6, 12 en 14: in de eerste set variabelen is de geflecteerde variant (kolom $-e(n)$) een *combinatie* van een vorm op sjwa én een vorm op sjwa plus n , terwijl in de tweede set de ongeflecteerde variant zelf al op een sjwa eindigt en flectie neerkomt op de toevoeging van een n . Beide flectievormen in de eerste set (d.i. sjwa en sjwa plus n) beschouwen we evenwel als één enkele variant (waarvan de precieze realisatie afhangt van fonetische condities – zie 1), en bijgevolg zullen we de observaties ervoor dan ook niet verder uitsplitsen⁽²⁾ (voor de tweede set betekent dat dan weer dat er bij de analyses wel extra rekening gehouden zal moeten worden met de restrictie van de fonetische condities – zie 4.1).

Zodra we de observaties voor elke determinator verzameld hebben, gaan we voor iedere treffer op zoek naar het bijbehorende nomen. Daarbij hebben we voor een praktische oplossing gekozen, en naar het *eerstvolgende* nomen gezocht om vervolgens te checken of het beantwoordde aan de specificaties waarin $-e(n)$ -variatie op kan treden. In het CGN betekent dit dat het nomen voldeed aan de volgende vijf criteria: het moest een soortnaam zijn (geen eigennaam), enkelvoudig (niet meervoudig), niet gediminueerd, niet onzijdig, en geen genitief- of datiefvorm – voor meer details omtrent de part-of-speech-codering, zie Van Eynde (2004).

⁽²⁾ Bij *die* zijn er zelfs vijf flectievormen die we als één variant behandelen, te weten: *dieje*, *diejen*, *dien*, *diene* en *dienen*.

Het algoritme is als volgt:

- Stap 1: Zoek binnen de zin (in het CGN: de ‘annotation unit’) het eerst volgende token met pos-tag N(); als er binnen de zin zo geen token gevonden wordt, ga naar de volgende observatie en herhaal Stap 1.
- Stap 2: Check of de pos-tag van de vorm N(soort,ev,basis,(zijd|genus),stan) is.
- Stap 3: Indien ja, neem het lemma dat na de pos-tag volgt op; indien neen, ga naar de volgende observatie en herhaal Stap 1.

Wat de betrouwbaarheid van het algoritme betreft, leerde een vluchtige, handmatige verkenning van de resultaten dat er louter enkele sporadische missers waren.

Tot slot wordt aan elk nomen een genus toegekend. Daarbij maken we gebruik van de genusclassificatie uit de veertiende editie van Van Dale Groot woordenboek van de Nederlandse taal (Den Boon & Geeraerts 2005) – wat dan ook de reden is waarom we in Stap 3 van het vorige algoritme de lemmavorm van elk token opgenomen hebben. Tabel 3 geeft daarvan enkele cijfers. De Grote Van Dale hanteert met andere woorden veertien genuscategorieën. Op basis van de frequenties zal het duidelijk zijn dat er daarvan drie hoofdcategorieën onderscheiden kunnen worden, te weten: *de (m.)*, *de (v.)* en *de*, oftewel de masculiene, feminiene en algemeen ‘zijdige’ substantieven. Dat de vierde hoofdcategorie, het onzijdige *het*, zo weinig voorkomt heeft er natuurlijk mee te maken dat we deze substantieven er in de tweede stap van het bovenstaande algoritme expliciet uitgefilterd hebben. Verder zijn er nog enkele rest- en mengcategorieën. Bij die eerst gaat het meer bepaald om *zn.* en *mv.*; bij die laatste om substantieven die doorgaans in één categorie thuishoren (en dat is dan het eerste opgegeven label), maar sporadisch ook in een andere categorie kunnen verschijnen – een voorbeeld uit de categorie *de (m.)* & *het* is het woord *salon*. Ten slotte zijn er nog substantieven waarvan het genus van de betekenis van het woord afhangt en die dus ambigu zijn, en er zijn woorden in het corpus waarvan we het lemma niet in de Grote Van Dale teruggevonden hebben. Dat zijn echter meestal samenstellingen of afleidingen.

DE VERSPREIDING VAN DE $-E(N)$ -UITGANG IN ATTRIBUTIEVE POSITIE

	Ø		-e(n)		% -e(n)	
	types	tokens	types	tokens	types	tokens
zn.	3	4	0	0	0,00	0,00
de (m.)	3637	39967	1123	7623	23,59	16,02
de (m.&v.)	70	737	19	99	21,35	11,84
de (v.)	3019	38048	119	235	3,79	0,61
het	48	212	21	69	30,43	24,56
het & de (v.)	7	46	0	0	0,00	0,00
het & de	21	121	2	2	8,70	1,63
de	2003	30616	197	688	8,95	2,20
de & het	27	210	2	4	6,90	1,87
de (m.) & het	42	142	16	35	27,59	19,77
de (v.) & het	13	63	3	3	18,75	4,55
het & de (m.)	24	163	6	11	20,00	6,32
de (m./v.)	61	237	11	22	15,28	8,49
mv.	2	3	2	2	50,00	40,00
AMBIGU	638	21232	228	2965	26,33	12,25
NT in GVD-14	4655	9147	582	966	11,11	9,55

Tabel 3: De genusfrequenties voor types en tokens

Omdat we om praktische redenen een genusindeling wilden met een beperkter aantal categorieën, maar anderzijds de proportie nomina met een gereflecteerde determinatorvorm voor sommige mengcategorieën aanzienlijk is (vergelijk de twee kolommen onder % -e(n)), hebben we besloten om elk van deze mengcategorieën te reduceren tot één van de drie hoofdcategorieën – de *het*-woorden hebben we daarbij wegens hun lage frequentie geheel laten vallen. Dat hebben we voor elk lemma in die categorieën afzonderlijk gedaan. Enkel *de (m.&v.)* en *de (m./v.)* hebben we integraal bij de *de*-categorie ondergebracht: voor *de (m.&v.)* is er namelijk geen wezenlijk verschil met *de* – een typisch voorbeeld uit deze categorie is *fiets*; bij *de (m./v.)* betreft het dan weer de zogenaamde ‘gemeenkunnige’ nomina waarvan het genus gelijk is aan het biologische geslacht van de referent, zoals het woord *arts* bijvoorbeeld. Voor de ambigue nomina hebben we het genus gekozen van wat we intuïtief beschouwden als de meest prototypische betekenis, en bij de samenstellingen en afleidingen hebben we het genus laten samenvallen met dat van het rechterlid, respectievelijk het suffix. Dat alles leidt dus tot cijfers over het aandeel van de determinatorreflectie bij de masculiene, feminiene en wat we hier commune substantieven zullen noemen⁽³⁾ De interpretatie van die cijfers stellen we uit tot 4.2.

⁽³⁾ Strikt genomen is de laatste term echter niet helemaal adequaat: ‘commuun’ is namelijk oorspronkelijk een synoniem van ‘gemeenkunnig’, en slaat dus zoals vermeld

4. Analyses

De analyse splitsen we zoals gezegd (zie 1) op in drie delen. Omdat het voorkomen van de *-e(n)*-uitgang afhangt van de fonetische restricties op de *-n*, behandelen we die in 4.1 als eerste. Daarna bespreken we in 4.2 de cijfers met betrekking tot de flectie bij de drie genera. In 4.3 vergelijken we ten slotte het aandeel van elke factor – zowel intern als extern – in het al dan niet optreden van de *-e(n)*-uitgang.

4.1 De *-n*

Om de geopperde bewering over de fonetische restricties op de *-n* (enkel vóór een vocaal, h, d en t, in mindere mate een b, en slechts sporadisch een r) empirisch te toetsen, hebben we ook hier een meer pragmatische benadering gekozen en voor elk van de veertien taalvariabelen het eerstvolgende karakterteken in het corpusmateriaal geteld. Op de keper beschouwd is dat natuurlijk niet helemaal correct aangezien het hier gaat om lettertekens in plaats van klanken. Toch menen we dat deze aanpak verdedigbaar is omdat het CGN bij haar orthografische transcriptie expliciet fonetische principes huldigt – zie Goedertier & Goddijn (2000): afwijkende uitspraken worden afwijkend gespeld. De tellingen hebben we vervolgens uitgesplitst naargelang het feit of de taalvariabele op alleen een sjwa eindigt dan wel op een sjwa plus n – wat betekent dat we voor variabele 1, 4, 5, 6, 12 en 14 de ongefleeteerde en gefleeteerde variant met elkaar vergeleken hebben, maar voor variabele 2, 3, 7, 8, 9, 10, 11, en 13 enkel *binnen* de gefleeteerde variant gekeken hebben (de observaties met de *ongefleeteerde* vorm bij deze acht taalvariabelen hebben we met andere woorden buiten beschouwing gelaten). De resultaten daarvan staan in Tabel 4; ten behoeve van de interpretatie visualiseert Figuur 1 daarbij nog de percentages van de *-n* voor elke letter. We zien dan dat x blijkbaar de hoogste frequentie heeft (40%). Die score is evenwel ruizig omdat het CGN onverstanebare uitingen transcribeert als x en/of xxx, en de x dus op om het even wat kan slaan – zoals manuele controle achteraf heeft bevestigd. Voorts kunnen we duidelijk stellen dat de hypothese klopt: de *-n* treedt vaker op vóór de vijf vocalen (a, e, i, o en u), een h, d, t en b. De lage score voor r weerspiegelt dan weer eveneens het beperkt aantal dialecten waar die klank met een *-n* voorkomt (zie 1). Enkel bij o (7,72%) lijkt het verschil in percentage met n (5,46%) en j (5,29%) te klein om van een duidelijke grens te kunnen spreken. Toch kunnen we de sterkte van associatie van de rijen met de kolommen uit Tabel 3 visualiseren in een plot zoals Figuur 2 (dat de correspondentieanalyse

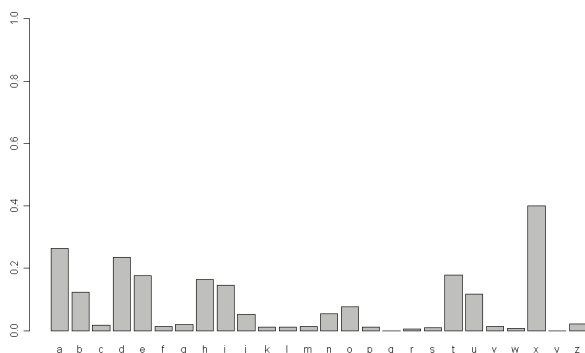
⁽³⁾ vervolg eigenlijk alleen maar op categorie *de (m./v.)*. De lezer vergeve ons deze ambiguïteit.

is uit Plevoets 2008: 38), waar het rechte kruis het nulpunt van de grafiek is en als zodanig de natuurlijke scheidslijn vormt tussen het met $-en$ geassocieerde linkerdeel en het met $-e$ geassocieerde rechterdeel. Het blijkt dan dat o nog net met $-en$ geassocieerd is, terwijl n of j al niet meer. We besluiten dan ook dat er concrete empirische evidentie is om de hele in de literatuur opgesomde reeks van fonetische condities – nogmaals: vocaal, h, d, t en b – te beschouwen als een aparte cluster. Aan de andere kant, en dat is dan de tweede belangrijke conclusie in verband met de $-n$, is het onmiskenbaar dat er binnen die cluster nog heel wat interne variatie is: de verschillen tussen – in aflopende volgorde – a, d, t, e, h, i, b, u en o zijn aanzienlijk (binnen het bestek van dit artikel ontbreekt de mogelijkheid ertoe, maar dit zou nog verder onderzocht kunnen worden).

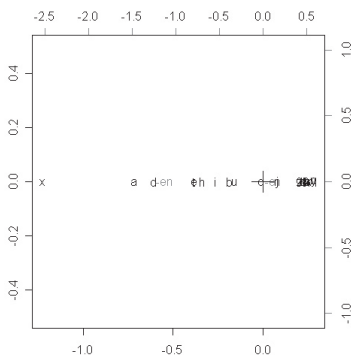
Het gevolg van deze resultaten is wel dat we voor de taalvariabelen waar de $-n$ intrinsiek deel uitmaakt van de flectie – ter herinnering: variabele 1 (*de(n)*), 4 (*iedere(n)*), 5 (*elke(n)*), 6 (*deze(n)*), 12 (*onze(n)*) en 14 (ADJ-*e(n)*) – rekening zullen moeten houden met de fonetische restricties. In wat volgt zullen we voor deze zes variabelen dan ook enkel de observaties behouden vóór een vocaal, h, d, t en b.

	-e	-en	% -en
a	3610	1299	26,46
b	6256	886	12,41
c	2462	45	1,79
d	4498	1393	23,65
e	2958	634	17,65
f	2367	33	1,38
g	4246	88	2,03
h	3570	703	16,45
i	1630	277	14,53
j	860	48	5,29
k	8128	87	1,06
l	3423	41	1,18
m	6919	93	1,33
n	2163	125	5,46
o	3477	291	7,72
p	4474	51	1,13
q	20	0	0,00
r	5072	28	0,55
s	6506	59	0,90
t	4965	1073	17,77
u	1131	151	11,78
v	6955	99	1,40
w	4568	37	0,80
x	6	4	40,00
y	13	0	0,00
z	2583	56	2,12

Tabel 4: De frequenties van de eerstvolgende letter na -e(n)



Figuur 1: De proporties van -n vóór de eerstvolgende letter



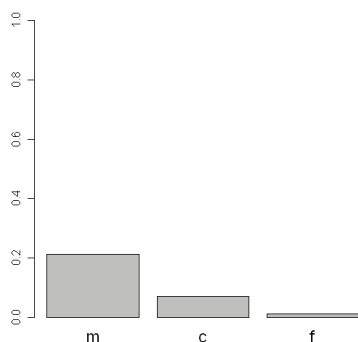
Figuur 2: De associaties van de eerstvolgende letter met $-e(n)$

4.2 De genera

Als zowel de herleiding tot de drie hoofdgenera doorgevoerd als de filtering voor de fonetische omgeving aangebracht is verkrijgen we in Tabel 5 de frequenties van de flectie-uitgang $-e(n)$ bij de drie genera – m staat daarbij voor de masculiene nomina, c voor de commune, en f voor de feminiene (zie 3). Opnieuw worden gemakshalve de flectiepercentages per genus gevisualiseerd in Figuur 3.

	$-\emptyset$	$-e(n)$	% $-e(n)$
m	32631	8828	21,29
c	22004	1654	6,99
f	26279	281	1,06

Tabel 5: De flectiefrequenties per genus



Figuur 3: De flectieproporties per genus

Zoals al bleek uit Tabel 3 (zie 3), komt de *-e(n)*-uitgang naast masculiene nomina (21,29%) ook voor bij commune (6,99%) en feminiene (1,06%). Op het eerste gezicht kunnen we dit opvatten als evidentie voor de werking van de masculiniseringsstendens: dat die tot uiting komt in de anaforische pronomina is al meermaals opgemerkt (zie Geeraerts 1992 en Audring 2006 voor twee tegengestelde standpunten hieromtrent, en De Vogelaer [te versch.] voor een poging tot synthese); nieuw is nu dat dit tevens voor de flectie bij determinatoren geldt. Bij nader inzien rijzen er echter allerhande problemen. Zo weerspiegelen onze cijfers bijvoorbeeld de genusclassificatie van de standaardtaal, aangezien we ons baseren op de lijst van Van Dale. Op het niveau van de dialecten zijn er evenwel tal van nomina met ‘afwijkend’ genus – zie Pauwels (1938) voor een overzicht, in casu niet-masculiene substantieven in de standaardtaal die masculien zijn in het dialect: voor sprekers van zulke dialecten is het gemakkelijker om de determinatoren bij deze woorden met een *-e(n)*-uitgang te realiseren. De oplossing bestaat er dan ook in om voor elk substantief *apart* te testen of het significant vaker met *-e(n)* voorkomt dan op basis van zuiver toeval verwacht mag worden. Dit heet een collocatieanalyse⁽⁴⁾ en leverde als uitkomst 355 masculiene, 119 commune en 83 feminiene nomina op; de volledige opsomming vindt men in de appendix van dit artikel. Een korte blik op die lijst leert dat het meestal blijkt te gaan om nomina die, als men ze onder een gemeenschappelijke noemer zou moeten brengen, verwijzen naar typisch conversationele, dagdagelijkse onderwerpen. In het verlengde daarvan zou het probleem dan ook wel eens kunnen zijn – en dit is het tweede punt van kritiek op de interpretatie in termen van masculinisering – dat de *-e(n)*-uitgang eenvoudigweg niet primair meer fungeert om het masculiene genus te markeren. Het voorkomen van *-e(n)* bij commune en feminiene substantieven die betrekking hebben op onderwerpen uit de dagdagelijkse communicatie zou erop kunnen wijzen dat de uitgang eerder dient om het gesprek een vertrouwelijke en informele sfeer te geven. De *-e(n)*-uitgang is met andere woorden veeleer een *stijlindicator* (geworden). Dat zou weliswaar enkel kunnen bij jongere sprekers, voor wie de dialectintuïties niet volledig ontwikkeld meer zijn. Aan die voorwaarde was evenwel voldaan in de jaren tachtig van de twintigste eeuw met de opkomst van de Vlaamse *tussentaal*, waar de met *-e(n)* geflecteerde determinatorvormen nu eenmaal kenmerkend voor zijn (zie Goossens 2000). Hoe dan ook, het is echter nog te vroeg om nu al over deze aspecten uitsluitel te geven: met deze bespiegelingen

⁽⁴⁾ De techniek waarvan we meer bepaald gebruik gemaakt hebben om de significanties te meten is de *Fisher-exacttest*.

⁽⁵⁾ Dit is getest door middel van een paarsgewijze toets voor gelijkheid van proporties met de Bonferroni-correctie.

is namelijk finaal de overstap gemaakt naar externe factoren, die zo meteen in 4.3 aan bod zullen komen.

Van een totaal andere orde – maar eveneens vraagttekens plaatsend bij de masculiniseringstheorie – is het feit dat het verschil in flectiepercentage tussen de drie genera statistisch significant is⁽⁵⁾: de mate waarmee de $-e(n)$ -uitgang aan (nomina van) elk genus hecht is met andere woorden reëel verschillend voor de drie genera. Ons lijkt dit te betekenen dat, ondanks de verschuivingen die in het Vlaamse genussysteem onmiskenbaar plaatsvinden en de eventuele expressie ervan in de determinatorische flectie, de distinctie in Vlaanderen tussen masculien, commuun en feminien nog stevig in het taalsysteem verankerd zit: het statistisch significante verschil in flectiepercentage op de $-e(n)$ -uitgang vatten we met andere woorden op als een bevestiging van de bekende stelling dat Vlaanderen, in tegenstelling tot Nederland, nog steeds een drie-generasysteem heeft. Dit beweren we evenwel met enig voorbehoud omdat ook hier de bovenstaande tegenwerpingen van kracht zijn die ten opzichte van de masculinisering gemaakt werden: het voorgaande onderzoek over genussystemen in Vlaanderen richt zich vrijwel uitsluitend op de anaforische pronomina terwijl tendenzen wat de determinatorische flectie betreft niet noodzakelijk daarmee gecorreleerd hoeven te zijn; de genusindeling van de Grote Van Dale weerspiegelt die van de standaardtaal en kan aanzienlijk verschillen van die van de dialecten; het is onduidelijk in hoeverre het gebruik van de $-e(n)$ -uitgang door interne dan wel door externe factoren gemotiveerd is. In antwoord hierop zullen we in de nu volgende paragraaf genus samennemen met de externe factoren in het CGN in een multivariaat model om zo het relatieve aandeel van elke factor in de verklaring van het optreden van de $-e(n)$ -uitgang tegen elkaar af te wegen.

4.3 De $-e(n)$ -uitgang

Om een totaalbeeld te verkrijgen van de distributie van de $-e(n)$ -uitgang in attributieve positie, nemen we het genus samen met de zes externe factoren uit het CGN (zie 2) op in een globaal model dat het al dan niet optreden van de flectie-uitgang poogt te verklaren. De techniek die we daartoe gebruiken is *logistische regressie*, die toch al enige staat van dienst heeft als het op de studie van variatieverschijnselen uit de neerlandistiek aankomt – vergelijk

⁽⁶⁾ Een technisch detail is hierbij dat niet zozeer de kans P op een succes gemodelleerd wordt als wel de ‘logit’, oftewel $\log(P/(1-P))$; dit doet men om tegemoet te komen aan allerhande assumpties waarvan regressieanalyse uitgaat – zoals bijvoorbeeld die van normaalverdeeltheid.

bijvoorbeeld: Grondelaers (2000), De Sutter (2005) en Tummers (2005); die laatste focust daarbij trouwens ook op een flectieverschijnsel. Kenmerkend voor logistische regressie is dat de variabele waar het om draait, de *respons*variabele, een *binair* variabele is: een variabele met twee mogelijke waarden: ja/nee, 1/0, succes/mislukking, enzovoorts – zoals het wel of niet voorkomen van de $-e(n)$ -uitgang. Logistische regressie bestaat er dan in dat de kans op een succes⁽⁶⁾ – in ons geval: de aanwezigheid van $-e(n)$ – gemodelleerd wordt in termen van één of meer ‘verklarende’ variabelen, die daarom de *predictor*variabelen genoemd worden. Mathematisch kan dit in de volgende formule uitgedrukt worden – Y stelt daarbij de responsvariabele voor; de X -en zijn de predictoren:

$$Y = \beta_0 + \beta_1 * X_1 + \dots + \beta_k * X_k + \varepsilon$$

De bèta’s zijn de zogenaamde *coëfficiënten* waarvoor de techniek een schatting berekent (op basis van de data), en die op die manier aangeven wat het effect is van elke predictorvariabele op de responsvariabele – de β_0 wordt het *intercept* genoemd, en drukt de verdeling van de responsvariabele uit los van de effecten van de predictoren; de ε is de ‘ruisterm’ die de nog onverbljvende, onverklaarde, *residuele*, variatie kwantificeert. De effecten van de geschatte coëfficiënten hebben een *kwalitatief* en een *kwantitatief* aspect. Het kwalitatieve aspect bestaat erin dat de coëfficiënt positief of negatief kan zijn, wat er respectievelijk op neerkomt dat de predictor in kwestie de kans op een succes bij de responsvariabele doet toenemen dan wel afnemen. Wat het kwantitatieve aspect betreft, geeft de getalswaarde van de coëfficiënt dan weer een indicatie van de grootte van het effect van de predictor. Het is op basis van deze kwantitatieve waarden dat we het relatieve aandeel van de verschillende predictoren in de verklaring van de responsvariabele ten opzichte van elkaar kunnen afwegen.

Naast de coëfficiënt berekent logistische regressie voor elk effect ook een *p*-waarde, die aangeeft in hoeverre het effect statistisch significant is. Dat er een coëfficiënt geschat kan worden voor het verband tussen predictor- en responsvariabele, betekent namelijk nog niet dat die schatting niet gewoon te wijten zou kunnen zijn aan zuiver toevallige variatie, eerder dan aan een reëel effect. Om dat te weten te komen, moeten we de kans bepalen dat de coëfficiënt in kwestie een louter willekeurige waarde is – en dat is dan de *p*-waarde. Als die laag is – en in de praktijk hanteert men daarbij de vuistregel van minder dan vijf procent – dan mogen we de coëfficiënt met voldoende betrouwbaarheid extrapoleren van de steekproef naar de populatie in haar geheel. Om die reden zullen we in

de bespreking die nu volgt dan ook voor elke predictor zowel de coëfficiënt als de p-waarde geven.

Ten slotte moeten we nog opmerken dat er bij regressieanalyse voor categorische variabelen – zoals alle variabelen in onze dataset; *genus* is bijvoorbeeld een categorische variabele met drie mogelijke waarden: masculien, commuun en feminien – altijd met een zogenaamde referentiewaarde gewerkt wordt. Dat betekent dat er van de categorische variabele automatisch één waarde als default-niveau gekozen wordt, en dat alle andere waarden daartegen afgezet worden. Het effect van de defaultwaarde zit dan in het intercept verrekend, terwijl de coëfficiënten van de overige waarden – en voor een categorische variabele met k categorieën zijn dat er dus $(k-1)$ – het *verschil* van deze waarde met die van de referentieniveau uitdrukken (en niet het effect van de waarde in kwestie zelf). Referentiewaarden maken de resultaten misschien minder inzichtelijk, maar vormen een onlosmakelijk onderdeel van logistische regressie. Daarbij zullen we voor de variabele *genus* het masculiene niveau als referentiewaarde gebruiken, voor register component 0 (voorgelezen teksten), voor regio Brabant, voor generatie de Vooroorlogse generatie, voor opleidingsgraad de hoger opgeleiden, voor beroepsgroep de managers (*occA*), en voor sekse de mannen.

Coëfficiënt	Schatting	P-waarde
Intercept	-4.43782	< 2e-16
zijc	-1.49563	< 2e-16
zijf	-3.18711	< 2e-16
coma	3.42583	< 2e-16
comb	0.97568	< 2e-16
come	3.51444	< 2e-16
comd	3.71248	< 2e-16
comf	0.91232	6.94e-15
comg	-0.06776	0.725712
comh	1.46056	< 2e-16
comi	-0.86949	0.000565
comj	0.03186	0.877056
comk	-0.86828	0.003327
coml	-0.23882	0.263731
comm	-1.61149	0.107283
comn	1.90788	< 2e-16
regovl	-0.18118	1.04e-08
regwvl	-1.03148	< 2e-16
reglim	-0.50421	< 2e-16
gensil	0.99779	< 2e-16
genpro	0.95656	< 2e-16
genlos	1.16154	< 2e-16
genpra	0.85451	1.79e-13
edumid	0.14797	0.000319
edulow	0.23052	0.010212
occoccB	0.38349	1.26e-05
occoccC	-0.13534	0.039686
occoccD	0.30764	5.88e-06
occoccH	-0.25026	0.006539
occoccF	0.17178	0.166561
occoccE	-0.05600	0.717062
occoccG	-0.06501	0.709208
occoccI	0.18929	0.022922
occoccJ	0.32317	0.004088
sexF	-0.06934	0.008510

Tabel 6: Coëfficiënten, schattingen en p-waarden

De resultaten van de logistische regressie staan dan in Tabel 6. Voor genus zien we daarbij dezelfde effecten als in 4.2: de geschatte coëfficiënt voor het commune (*zijc*) en het feminiene genus (*zijf*) is telkens negatief, wat betekent dat de proportie waarmee de $-e(n)$ -uitgang in deze genuscategorieën voorkomt

lager is dan bij het masculiene genus; verder zijn de verschillen tussen de drie genera weer statistisch significant. De interpretatie verschilt dan ook geenszins van die uit 4.2.

Bij register vallen vooral de grote waarden voor component a (3,43), c (3,51) en d (3,71) op. Dat vormt dan een confirmatie van de in 4.2 geopperde hypothese dat de $-e(n)$ -flectie niet zozeer meer het genus markeert als wel de stijl, in de zin van de mate van (in)formaliteit van een situatie: zoals aangetoond in Plevoets (2008: 76-77) vormen componenten a (face-to-face conversaties), c (telefoon-dialogen op platform) en d (telefoon-dialogen op minidisc) namelijk de top drie van de meest informele componenten uit het CGN. Dat de coëfficiënten hiervan zo hoog zijn bewijst dat de $-e(n)$ -uitgang vaker optreedt naarmate de gespreks-situatie informeler wordt.

Bij regio zijn de schattingen voor alle provincies negatief ten opzichte van het referentieniveau *bra*: in het Brabantse gebied wordt er met andere woorden het vaakst geflecteerd met de $-e(n)$ -uitgang. Op het eerste gezicht zou dit te maken kunnen hebben met het feit dat de $-e(n)$ -flectie oorspronkelijk een Brabants dialectkenmerk is, dat zich van daaruit naar de omliggende gebieden verspreid heeft. Het verschijnsel van de ‘Brabantse expansie’ is al veelvuldig aangehaald om de export van Brabantse dialectvormen naar de naburige niet-Brabantse dialectgebieden te verklaren (zie Goossens 1970, Geerts 1983 en recent nog Vandekerckhove 2009): doordat het Brabantse gebied niet alleen het geografische maar ook het economische, politieke en culturele centrum van Vlaanderen is, oefent Brabant een zekere linguïstische invloed uit op die gebieden, en fungeert het zo als een – officieus – taalcentrum. Toch bewijst de historische oorsprong van de $-e(n)$ -uitgang als Brabants dialectverschijnsel nog niet meteen waarom het uitgerekend de Brabanders zijn die er het meeste gebruik van maken: immers, als de omliggende regio’s de $-e(n)$ -flectie overgenomen hebben, hoe komt het dan dat ze die minder gebruiken dan het Brabantse gebied? Wat dus verklaard moeten worden is eerder waarom het voor Brabanders gemakkelijker is om de dialectische $-e(n)$ -uitgang te realiseren dan voor de andere sprekers. Opnieuw wijst de these van de Brabantse expansie hier evenwel de uitweg. Omdat Brabant het socio-politieke centrum van Vlaanderen is, genieten Brabanders en Brabantse dialectvormen een zeker prestige, waardoor het voor Brabantse sprekers minder problematisch is om een dialectfeature te gebruiken dan voor andere sprekers, die relatief gezien vlugger moeten kiezen voor de standaardtalige variant – en dat is in onze set van veertien variabelen dus de *ongeflecteerde* variant. De hoge

frequentie van de $-e(n)$ -uitgang bij de Brabanders ligt dus niet zozeer aan de endogene oorsprong van het feature, als veeleer aan de sociale status van Brabant in het Vlaamse maatschappijbestel. Om die interpretatie te toetsen hebben we een deelanalyse uitgevoerd waar we de interacties tussen regio en component onderzocht hebben. De resultaten staan in Tabel 7.

DE VERSPREIDING VAN DE $-E(N)$ -UITGANG IN ATTRIBUTIEVE POSITIE

Coëfficiënt	Schatting	P-waarde
Intercept	-4.77680	< 2e-16
reglim	0.88948	6.23e-06
regovl	0.77213	0.000332
regwvl	0.40075	0.057378
coma	3.91477	< 2e-16
comb	0.89652	0.000104
comc	4.21048	< 2e-16
comd	4.35862	< 2e-16
comf	1.32616	1.21e-13
comg	-0.24743	0.364450
comh	2.09640	< 2e-16
comi	-0.26464	0.400005
comj	-0.08837	0.778793
comk	-0.34269	0.371752
coml	0.45340	0.103936
comm	-1.43981	0.154673
comn	1.48541	3.21e-09
reglim:coma	-1.12227	2.96e-08
regovl:coma	-0.98724	6.89e-06
regwvl:coma	-1.40654	8.73e-11
reglim:comb	-0.77281	0.017575
regovl:comb	-0.36458	0.247116
regwvl:comb	-0.61326	0.062193
reglim:comc	-1.64047	2.00e-14
regovl:comc	-0.91727	3.62e-05
regwvl:comc	-1.23150	2.25e-08
reglim:comd	-1.91782	< 2e-16
regovl:comd	-0.72684	0.001177
regwvl:comd	-1.31361	5.74e-09
reglim:comf	-1.27249	0.000373
regovl:comf	-0.98360	0.000365
regwvl:comf	-0.91805	0.003273
reglim:comg	0.42239	0.263919
regovl:comg	-0.74800	0.129234
regwvl:comg	-0.12146	0.824272
reglim:comh	-2.87817	1.35e-08
regovl:comh	-1.79426	1.03e-07
regwvl:comh	-1.35501	7.82e-05
reglim:comi	0.32876	0.589686
regovl:comi	-0.94018	0.234885
regwvl:comi	-10.92538	0.933649
reglim:comj	-0.99413	0.348233
regovl:comj	-0.11662	0.810511
regwvl:comj	0.21184	0.681753
reglim:comk	-1.37580	0.203048
regovl:comk	-1.72568	0.035443
regwvl:comk	-1.24227	0.129526
reglim:coml	-12.13215	0.946830
regovl:coml	-1.81564	0.000982
regwvl:coml	-0.38950	0.479518
reglim:comm	NA	NA
regovl:comm	-10.12159	0.926078
regwvl:comm	NA	NA
reglim:comn	-0.99640	0.004858
regovl:comn	0.17209	0.593405
regwvl:comn	0.34462	0.303687

Tabel 7: Interacties tussen regio en component

We zien dan inderdaad dat het niet algemeen zo is dat de overige regio's de *-e(n)*-uitgang minder gebruiken dan het Brabantse gebied: de coëfficiënten voor de hoofdeffecten (die het *-e(n)*-gebruik van elke regio kwantificeren voor alle componenten samen) van Limburg (*reglim*), Oost-Vlaanderen (*regovl*) en West-Vlaanderen (*regwvl*) zijn alle significant positief – het referentieniveau is hier wel *bra.o*, oftewel de Brabanders uit component o (voorgelezen teksten); ten opzichte van deze sprekers gebruiken Limburgers, Oost-Vlamingen en West-Vlamingen wel degelijk vaker de *-e(n)*-uitgang. Bij de componenten bmerken we voorts weer de hoge coëfficiënten voor component a, c en d, die opnieuw wijzen op het al besproken stijleffect (*-e(n)* komt meer voor naarmate de spreek situatie informeler wordt). Enkel bij de interacties (en dat is dus de reeks van coëfficiënten die begint vanaf *reglim.coma*) worden de effecten voor deze drie componenten voor elke regio significant negatief (er zijn ook significant negatieve effecten bij componenten f en h, maar omdat we primair geïnteresseerd zijn in a, c en d laten we die verder voor wat ze zijn). In het licht van de hoofdeffecten betekent dit dat Limburgers, Oost- en West-Vlamingen weliswaar in het algemeen ook *-e(n)* hanteren (dat is juist het hoofdeffect), maar dat zij dit gebruik in informele registers aanzienlijk terugschroeven. Als de situatie informeler wordt, is een Brabander met andere woorden sneller bereid om de *-e(n)*-uitgang te realiseren dan een Limburger, Oost-Vlaming of West-Vlaming, die dus verhoudingsgewijs eerder voor de ongeffecteerde, standaardtalige realisatie zullen opteren. Dat lijkt ons de hierboven geopperde hypothese te staven.

Als we voor generatie vervolgens teruggaan naar Tabel 6, dan constateren we een op het eerste gezicht oninterpretabel resultaat: ten opzichte van de Vooroorlogse generatie als referentieniveau stijgt het flectiepercentage voor alle generaties (de coëfficiënten zijn positief), alleen is het verloop ervan ronduit grillig: een lichte teruggang na de Stille generatie voor de Protestgeneratie, vervolgens een piek bij de Verloren generatie, en ten slotte een dieptepunt (hoewel nog steeds meer dan de Vooroorlogse generatie) met de Pragmatische generatie. Dit kan evenwel aan een interactie met een andere – vooralsnog onbekende – factor liggen. Als we de regressieanalyse namelijk enkel met generatie als verklarende variabele uitvoeren, dan verkrijgen we, zoals Tabel 8 aantoont, een veel inzichtelijker patroon.

Coëfficiënt	Schatting	P-waarde
Intercept	-1.86519	< 2e-16
gensil	-0.22633	0.0199
genpro	-0.38447	3.59e-05
genlos	-0.46827	5.22e-07
genpra	0.37373	5.16e-05

Tabel 8: Effecten voor enkel generatie

De flectie op $-e(n)$ daalt gestaag van de Vooroorlogse tot en met de Verloren generatie, waarna er bij de Pragmatische generatie een totale omslag volgt en het flectiepercentage fors toeneemt. Dat laatste heeft wellicht te maken met de opkomst van de hedendaagse Vlaamse tussentaal, die zich zoals bekend kenmerkt in een toevloed van dialectische elementen in het bovenregionale spraakgebruik aan het einde van de twintigste eeuw – de periode dat de Pragmatische generatie het openbare leven betreedt (een verklaring voor het grillige verloop in de globale analyse zal een eventuele latere studie moeten uitwijzen).

Voor opleidingsgraad krijgen we dan weer wel een volkomen interpreteerbaar patroon – zie Tabel 6: naarmate de opleidingsgraad daalt (de hoger opgeleiden – *eduhie* – vormen daarbij het referentieniveau), stijgt het flectiepercentage. Hoe hoger opgeleid men met andere woorden is, hoe minder men geneigd is te flecteren.

Bij beroepsgroep moeten we vooral letten – in Tabel 6 – op de cluster van de hoger opgeleiden (*occB*) en de bedienden (*occD*) aan de ene kant, die elk een positief effect vertonen, versus de cluster van de academici (*occC*) en medi-afiguren (*occH*) aan de andere kant, die ieder een negatieve waarde hebben (positieve coëfficiënten vinden we ook nog bij *occI* en *occJ*, maar omdat dit nogal ‘bijzondere’ beroepsgroepen zijn – respectievelijk studenten en werklozen – gaan we er niet verder op in). Het sluit aan bij het onderscheid, dat in Plevoets (2008) gemaakt is, tussen een *economische* elite enerzijds en respectievelijk een *intellectuele* of *culturele* elite anderzijds. De economische elite, waartoe we hier ook het referentieniveau *occA* (managers) rekenen, bedraagt een laag van sprekers die op basis van in de eerste plaats *materiële* middelen aan de top van de Vlaamse maatschappij staat. Het is een klasse die wat Vlaanderen betreft voornamelijk tijdens de tweede helft van de twintigste eeuw ontstaan is uit de grote welvaartsgroei na de Tweede Wereldoorlog en de emancipatie- en verburgerlijgingsprocessen die daarmee gepaard gingen. De culturele of intel-

lectuele elite is daarentegen een klasse van sprekers die, zoals de naam het zelf zegt, vanwege hun culturele en intellectuele bagage een zeker aanzien genieten. Het is een laag die eerder een *symbolische* positie in de Vlaamse samenleving bekleedt. Dat is misschien nog het duidelijkst voor de mediafiguren (occH): doordat de belichaming van het Standaardnederlands in Vlaanderen het VRT-Nederlands is, zijn zij als vanouds de vaandeldragers van de standaardtaal; de academici (occC) zijn op hun beurt dan weer de exponenten van de ‘geleerde cultuur’. De sprekers in deze twee beroepsgroepen zijn ontegensprekelijk meer op de standaardtaal georiënteerd, zoals hun negatieve coëfficiënten laten zien. In contrast daarmee laten de economische elitegroepen zich al sneller verleiden om de *-e(n)*-uitgang te realiseren, getuige hun positieve coëfficiënten. De interpretatie die we daaraan hechten is dat dit uit hun sociale rang volgt: zij die al een verworven positie in de samenleving hebben weten te bereiken, kunnen het zich minder genegen laten liggen om de prestigevarianten te gebruiken – zij permitteren zich enige laksheid, omdat zij de nodige welstand genieten om hun status te handhaven.

Dat de managers (occA) overigens ook tot de economische elite behoren, kunnen we aantonen door een andere referentiewaarde te kiezen. Met wat we zonet gezegd hebben over de mediafiguren als de typische representanten van het Standaardnederlands in Vlaanderen, ligt het voor de hand om hen hiervoor te kiezen. Zo krijgen we voor elke beroepsgroep een indicatie hoezeer het taalgebruik afwijkt van de standaardtaal. Tabel 9 geeft daarvan de resultaten. We zien dan dat, op één uitzondering na, alle beroepsgroepen een significant positieve score hebben en dus de *-e(n)*-uitgang vaker gebruiken – enkel de politici (occG) doen het nog minder dan de mediafiguren. De managers (occA) situeren zich daarbij op het niveau van de hoger opgeleiden (occB), en zelfs nog een tikkeltje meer. Het bewijst in aanvulling op wat we net over de economische elitegroepen gezegd hebben dat de vorming van een Vlaamse bovenklasse na de Tweede Wereldoorlog niet onverwijd geleid heeft tot een opschuiving naar het Standaardnederlands.

Coëfficiënt	Schatting	P-waarde
Intercept	-3.71211	< 2e-16
occoccA	1.94024	< 2e-16
occoccB	1.83484	< 2e-16
occoccC	1.60152	< 2e-16
occoccD	2.54684	< 2e-16
occoccF	2.10382	< 2e-16
occoccE	2.34904	< 2e-16
occoccG	-0.78958	4.43e-09
occoccI	2.25934	< 2e-16
occoccJ	2.83343	< 2e-16

Tabel 9: Beroepsgroep met occH als referentieniveau

Uiteindelijk lijkt ons dat ook de reden voor de omslag die de Pragmatische generatie vertoont. De totstandkoming van een elite geeft namelijk aanleiding tot een zekere opwaardering van dialectische elementen. Dat zagen we bijvoorbeeld al bij de analyse van de Brabantse expansie, waar het voor Brabanders door het prestige van hun regio inderdaad gemakkelijker is om dialectvarianten te realiseren dan voor sprekers uit de andere regio's (die relatief gezien eerder moeten overschakelen op de standaardtaal). Als de naoorlogse welvaartstoename nu met zich meebrengt dat de Vlaamse samenleving zich emancipeert waarbij er een maatschappelijke toplaag ontstaat, dan verwachten we ook een effect in de tijd – meer bepaald: de 'apparent time'-effecten die we gezien hebben. Voor alle oudere generaties was een betere positie nog een te verwerven ideaal, en het gebruik van de standaardtaal was daar het middel bij uitstek toe. Voor de Pragmatische generatie is emancipatie echter een voldongen feit, zodat deze sprekers zich minder geremd voelen om onbekommerd 'af te glijden' in minder gestandaardiseerd taalgebruik – zie Plevoets (te versch.) voor een meer uitgewerkte analyse van deze en soortgelijke facetten.

Bij sekse – Tabel 6 – zien we ten slotte in de negatieve coëfficiënt voor de vrouwen (de mannen vormen het referentieniveau) de bevestiging van het bekende sociolinguïstische principe dat vrouwelijke sprekers sterker neigen naar standaardtalige varianten dan mannelijke sprekers (Coates 1986). Die coëfficiënt is misschien klein, maar de p-waarde is significant. De flectie op $-e(n)$ is met andere woorden eerder een mannelijke aangelegenheid.

5. Besluit

In dit artikel zijn we de verspreiding van de *-e(n)*-uitgang in attributieve positie nagegaan. Na een bondige formulering van onze probleemstelling hebben we ons corpus – het CGN – voorgesteld, waarbij we bijzondere aandacht hebben besteed aan de tot extralinguïstische factoren te operationaliseren kenmerken ervan. Vervolgens hebben we veertien taalvariabelen gedefinieerd – adnominale determinatoren die in Vlaanderen potentieel met *-e(n)* geflecteerd worden – en hebben we de methodologie uit de doeken gedaan waarmee we het bijhorende nomen voor elk van deze determinatoren opgezocht hebben evenals daaraan het nominale genus toegekend (ons daarbij baserend op de genusindeling uit de veertiende editie van Van Dale Groot Woordenboek van de Nederlandse taal). De analyses hebben we dan opgesplitst in drie delen. Voor de verbindings-*n* hebben we daarbij om te beginnen vastgesteld dat de *-n* inderdaad optreedt voor een eerstvolgende vocaal, h, b, d en t (maar niet r), zoals een oude stelling beweert, maar dat er tussen deze fonetische contexten nog veel variatie heerst. Daarna heeft de flectieproportie van de *-e(n)*-uitgang per genuscategorie aangetoond dat er statistisch significante evidentie is voor een onderscheid tussen het masculiene, commune en feminiene genus, hoewel er indicaties zijn dat de *-e(n)*-uitgang vandaag de dag veeleer als *stijl-* dan als genusmarkeerder fungeert. Tot slot hebben we het genus samen met de externe factoren uit het CGN opgenomen in een globaal model om het al dan niet voorkomen van de *-e(n)*-uitgang te verklaren; daaruit is gebleken dat de *-e(n)*-uitgang inderdaad typisch is voor informele en/of spreektaalige registers, met bovendien een interactie met regio zodat het voor Brabantse sprekers vanwege het – historische – prestige van hun regio gemakkelijker is om de *-e(n)*-variant te realiseren dan voor sprekers uit de overige regio's, die eerder (moeten) kiezen voor de standaardtalige, ongeflecteerde variant; voorts is het zo dat de *-e(n)*-uitgang gestaag afneemt van de Vooroorlogse tot en met de Verloren generatie, waarna er een omslag in de tegenovergestelde richting volgt, dat ze afneemt naarmate de opleidingsgraad stijgt, dat ze het sjibbolet is van een economische elite in tegenstelling tot een culturele elite die zich meer op de standaardtalige variant zonder flectie verlaat, en dat ze ten slotte eerder een mannelijk sprekerskenmerk is.

Bibliografie

- AUDRING, J.
(2006). "Pronominal gender in spoken Dutch". *Journal of Germanic Linguistics* 18, 85-116.
- BECKER, H.
(1992). *Generaties en hun kansen*. Amsterdam, Meulenhoff.
- COATES, J.
(1986). *Women, men and language*. Londen, Longman.
- DEN BOON, T. & D. GEERAERTS
(2005). *Van Dale Groot Woordenboek van de Nederlandse taal*. Antwerpen, Van Dale Lexicografie.
- DE SCHUTTER, G.
(1966). "De verbuiging van het onbepaald lidwoord en van een aantal bijvoeglijke woorden in de Zuidnederlandse dialecten". *Taal en Tongval* 18 (3), 71-99.
- DE SCHUTTER, G.
(1967). "De verbuiging van het onbepaald lidwoord en van een aantal bijvoeglijke woorden in de Zuidnederlandse dialecten". Gent: Gentse bijdragen 1967, 48-76.
- DE SUTTER G.
(2005). *Rood, groen, corpus! Een taalgebruiksgebaseerde analyse van woordvolgordevariatie in tweeledige werkwoordelijke eindgroepen*. Leuven, Doctorale dissertatie.
- DE VOGELAER, G.
(te versch.). "Changing pronominal gender in Dutch: transmission or diffusion?"
In: Pavlou, P. et al. (reds.). *Studies in Language Variation*. Amsterdam & Philadelphia, John Benjamins.
- GEERAERTS, D.
(1992). "Pronominale masculiniseringsparameters in Vlaanderen". In: H. Bennis & J. De Vries (reds.) *De binnenbouw van het Nederlands*. Dordrecht, ICG Publications. p. 73-84.
- GEERTS, G.
(1966). *Genus en geslacht in de Gouden Eeuw. Een bijdrage tot de studie van de nominale klassifikatie en daarmee samenhangende adnominale flexievormen en pronominale verschijnselen in Hollands taalgebruik van de zeventiende eeuw*. Brussel, Belgisch Interuniversitair Centrum voor Neerlandistiek.

- GEERTS, G.
(1983). “Brabant als centrum van de standaardtaalontwikkeling in Vlaanderen”. *Forum der Letteren* 24 (1), 55-63.
- GEERTS, G.
(1988). “Genusfouten: hollanditis in Vlaanderen?”. *Verslagen en mededelingen van de Koninklijke Academie voor Nederlandse Taal- en Letterkunde* 1, 68-78.
- GOEDERTIER, W. & S. GODDIJN
(2000). *Protocol voor Orthografische Transcriptie*. http://lands.let.kun.nl/cgn/doc_Dutch/topics/version_1.0/annot/orthography/ort_prot.pdf.
- GOOSSENS, J.
(1970). “‘Belgisch beschaafd Nederland’ en Brabantse expansie”. *De Nieuwe Taalgids* 63, 54-70.
- GOOSSENS, J.
(2000). “De toekomst van het Nederlands in Vlaanderen”. *Ons Erfdeel* 43 (1), 3-13.
- GRONDELAERS, S.
(2000). *De distributie van niet-anaforisch er buiten de eerste zinsplaats. Socio-lexicologische, functionele en psycholinguïstische aspecten van er’s status als presentatief signaal*. Leuven, Doctorale dissertatie.
- HERMKENS, H.
(1985). “Van casus-n naar genus-n”. *De Nieuwe Taalgids* 78 (6), 537-538.
- NIJEN TWILHAAR, J.
(1992). “Deflexie en genusverlies”. *De Nieuwe Taalgids* 85 (2), 147-160.
- PAUWELS, J. L.
(1938). *Bijdrage tot de kennis van het geslacht der substantieven in Zuid-Nederland. Uitslagen van het onderzoek in 1872 ingesteld door de Zuidnederlandsche maatschappij van taalkunde*. Tongeren, Michiels.
- PLEVOETS, K.
(2008). *Tussen spreek- en standaardtaal. Een corpusgebaseerd onderzoek naar de situationele, regionale en sociale verspreiding van enkele morfosyntactische verschijnselen uit het gesproken Belgisch-Nederlands*. Leuven, Doctorale dissertatie.
- PLEVOETS, K.
(te versch.). “Verkavelingsvlaams als de voertaal van de verburgerlijking van Vlaanderen”. *Studies van de BKL* 3.

OOSTDIJK, N.

(2000). "Het Corpus Gesproken Nederlands". *Nederlandse Taalkunde* 5 (3), 280-284.

TAELEDEMAN, J.

(1980). "Inflectional aspects of adjectives in the dialects of Dutch-speaking Belgium". In: W. Zonneveld, F. Van Coetsem & O. W. Robinson (reds.) *Studies in Dutch phonology*. Den Haag, Nijhoff. p. 223-254.

TAELEDEMAN, J.

(1983). "Adjektieffleksie in de Zuidnederlandse dialekten". In: J. Stroop (red.) *Nederlands dialectonderzoek. Artikelen uit de periode 1927-1982*. Amsterdam, Huis aan de drie grachten. p. 211-229.

TUMMERS, J.

(2005). *Het naakt(e) adjectief. Kwantitatief-empirisch onderzoek naar de adjectivische buigingsalternantie bij neutra*. Leuven, Doctorale dissertatie.

VANDEKERCKHOVE, R.

(2009). "Dialect loss and dialect vitality in Flanders". *International Journal of the Sociology of Language* 196-197, 73-97.

VAN EYNDE, F.

(2004). *Part Of Speech tagging en lemmatisering van het Corpus Gesproken Nederlands*. http://lands.let.kun.nl/cgn/doc_Dutch/topics/version_1.0/annot/pos_tagging/tg_prot.pdf.

VAN LOEY, A.

(1964). *Schönfeld's Historische grammatica van het Nederlands: klankleer, vormleer en woordvorming*. Zutphen, Thieme.

Appendix: nomina die significant vaker met *-e(n)* voorkomen, opgesplitst per genus

Masculien (355)

aas	brol	engelsman	kapitein
achterbuur	brommer	etter	kapper
achterkant	buggy	expert	kelder
afdronk	buikstoot	explorer	kinesist
afrikaan	buiten	fauteuil	klant
aftrekker	cakevorm	februari	kloot
afwas	camino	feestdag	klop
aldi	camion	fietser	koek
alma	cd-speler	fnac	koeler
ambras	cd-writer	frak	koffie
amerikaan	cent	fransman	koffiezet
anesthesist	chatbox	gast	koi
apotheker	chief	gentenaar	kop
appeltuin	chinees	gloeidraad	kozijn
april	clown	griek	kruimeldief
arend	cola	gsm	kruiwagen
assam	computer	gsm-lader	laptop
audi	crémant	gynaecoloog	lattenbodem
augustus	cursus	haardroger	leestoets
auto	dag	hans	leffe
autobus	datum	heksenketel	levensgenieter
automobiel	decascoop	helm	lichtbak
avond	december	hoebel	lidl
baas	delhaize	hof	limburger
back-up	dendermondsesteenweg	hollander	linguïst
bak	diesel	honger	linkerpompbak
barbecue	dijk	hoofddepot	liter
basket	dinsdag	hoofdgang	loodgieter
bazaar	diplom	hoop	luster
bebop	directeur	hoopkeszetter	maandag
beenhouwer	directieraad	hr-manager	maart
beep	disney-film	hubo	maïssnijder
belg	dobber	huidspecialist	mcdonald's
bic	dokter	huisbaas	mei
bieldenring	dokteur	huisdokter	mens
bierdrinker	doktoor	ikea	mercedes
bigmac	donderdag	inboedel	microgolf
bloeddruk	donderdagavond	inch	microgolfoven
blok	donker	indiër	migrant
blokkeerder	doop	indruk	missiepost
bm	draadman	ingenieur	modem
bmw	draai	inno	mosseleter
boebel	drieling	integratiedienst	moto
boekenbon	driewieler	inval	motor
boekenverzamelaar	driufuithoek	italiaan	mustang
boel	drink	jobstudent	nacht
boiler	driver	juli	namiddag
bol	drummer	jumbo	nederlander
bon	duik	jumper	nek
boom	duitser	juni	nonkel
boordsteen	duplex	junkie	november
boulette	duvel	kabel	obelisk
brander	e40	kalender	oktober
brief	elektrieker	kameraad	omslag
broekkant	elentrick	kant	onderbodem

DE VERSPREIDING VAN DE -E(N)-UITGANG IN ATTRIBUTIEVE POSITIE

ontspanner	specialist	vrouwenfilm
oom-priester	speech	waal
oprit	spelmaniak	waterkoker
output	sportwinkel	wederdienst
oven	spreker	weerstand
overkant	staanmolen	wekker
overtrek	stapop	wereldkampioen
pamper	sterrenhemel	wereldwinkel
pannenkoek	stoelgang	west-vlaming
papegaai	stofzuiger	wielrenner
paper	stoot	wijsheidstand
paravent	student	windmolen
pasen	sus	winkel
pater	sx	wis-computer
pee	tandarts	woensdag
peter	telefoon	xp
pianist	testrit	zak
piraat	tgV	zaterdag
playmobil-tractor	thalys	zelfscanner
playstation	theatermaker	zetel
poetsdag	tilt	zijtak
pol	tiramisu	zitdag
primé	toer	zitter
primer	toet	zondag
producer	tool	zwart-witfilm
prof	tour	
psychiater	tracks	
psycholoog	tram	
put	trein	
pyjama	trendhopper	
r41	tripel	
radiateur	trolley	
radio	tumulus	
regenput	turbo	
relaxzetel	turbodiesel	
remorque	turk	
rug	uitleg	
rugzak	uitloper	
sanitairman	uitslag	
scanner	uitzichttoren	
schep	vario	
sekspaalbaron	vent	
september	verhuis	
skibril	verjaardag	
slag	verlofdag	
slagroom	vertaalcomputer	
sleutel	vervangwagen	
smos	verzekeraar	
snellunch	video	
snipperdag	viervaksweg	
snoeier	vlieger	
snoeper	voet	
snul	volvo	
spambeweerder	vriend	
spanjaard	vrijdag	

Commuun (119)

agenda	jas	wang
avenue	jeugdherberg	wc
b	kaasschotel	website
babysit	kamerjas	weer
bakkersfiets	knie	wereld
barleys	kreeft	woede
belgacom	kruk	woonkamer
ben	kweddelfiets	worst
benzine	lasagna	zaag
blandijn	living	
bocht	luifel	
boot	maag	
boterham	maat	
café	mail	
camel	mars	
camera	menopauze	
cel	merel	
chipolata	minuut	
curryworst	mountainbike	
d	oester	
dafalgan	officinalis	
damesfiets	pasgeborene	
das	patroonheilige	
demer	pet	
deputé	pitabar	
dinge	pizza	
douche	plant	
droomwereld	plicht	
e-mail	poep	
efteling	post	
eglantier	pull-over	
elf	rechterwang	
elzas	regenjas	
envelop	rennerstaal	
ergopolis	rol	
etap	schommel	
ex	schuld	
fiets	site	
file	sms	
flagiel	staaf	
foto	stem	
gb	tai-chi	
geliefde	teek	
golf	tornado	
gr	trouw	
groep	trui	
helft	twijg	
herk	uil	
hernia	valeriaan	
hoede	vanillepap	
humo	voicemail	
ijzer	vriendenklik	
ijzerbedevaart	vrouwenfiets	
imov	vulgaris	
inferno	wachtlijst	

Feminien (83)

aanraking	recordopkomst
acupunctuur	recuperatie
afrodite	repliek
akoestiek	respons
ambriage	reuma
analogie	rubriek
b-klasse	sabam
bagarre	statistiek
basiskritiek	tactiek
bbc	talenrichting
begeleiding	teevee
boetiek	televisie
brt	theorie
btw	thesis
claus	tirade
commode	trofee
diarree	tut
directory	tuut
elektriciteit	tv
em	ufizi
euthanasie	uia
fotokopie	unief
friteuse	verplaatsing
frituur	videotheek
fysica	voituur
grafiek	vzw
groentemengeling	wiskunde
groepsverplaatsing	zandloperstructuur
handleiding	
intentieverklaring	
java-applet	
kabel-tv	
kav	
kritiek	
krokusvakantie	
limousine	
logopediste	
luchtvervuiling	
luxevoiture	
mam	
mama	
mantis	
mechanica	
miele	
milieubeweging	
moskee	
ms	
mutualiteit	
nmbs	
opstartfase	
overlock	
paniek	
patisserie	
picasso	
pronostiek	