

Journal of Dutch Linguistics and Literature

Tijdschrift voor Nederlandse Taal- en Letterkunde

TNTL

4

2018 | jaargang 134

LAURENS HAM, LEO LENTZ, HENK PANDER MAAT & FABIAN STOLK*
Zijn romans en kranten sinds 1950 eenvoudiger geworden?

Abstract – Did the language of Dutch novels and newspaper articles become less complex between the 1950s and the 2010s? Based on a corpus of 60 novels, published in three different decades, and 200 newspaper articles, published in two decades, an analysis of stylistic phenomena was conducted, using T-scan, a tool for automatic analysis of Dutch language. The analysis focused on four aspects of complexity: word and sentence complexity, cohesion and personalisation. This analysis showed that while the language of the newspaper articles did change on all levels between 1950 and 2000, the language of the novels did not change substantially. The language of the news articles is more complex than the language of the novels, across different decades. A further analysis of stylistic phenomena in novels revealed some interesting relations between language and content.

1 Inleiding

Wie reflecties verzamelt over de westerse leescultuur van de laatste decennia, ziet dat menige auteur klaagt over versimpeling van het taalgebruik. Zo stelt een redactioneel commentaar van *Trouw* in 2013:

De kunst van het lezen beperkt zich tot de 140 tekens van Twitter en de ‘vind ik leuk’-vignetten van Facebook. De lees- en schrijfvaardigheid van de gemiddelde Nederlander (laag- én hoogopgeleid) wordt dramatisch slechter. Een komma in een zin is een onoverkomelijke barrière, laat staan dat de hindernis van een ingebedde bijzin zou kunnen worden genomen (Anoniem 2013).

Komma's als barrières, ingebedde bijzinnen als hindernissen. Volgens *Trouw* is het met de verwerking van talige complexiteit blijkbaar zo slecht gesteld, dat alles wat de verwerking bemoeilijkt maar beter weggelaten kan worden. Daarom bestaat volgens de auteur sinds enige jaren de tendens om ‘naar beneden’ te schrijven en zo de taal te versimpelen.

De waardering van begrijpelijke taal verschilt per domein. In de wereld van zakelijke communicatie en overheidsvoorlichting wordt het als een deugd gezien om eenvoudig te schrijven. In de literaire wereld ligt dat anders. In literaire recensies vinden we negatieve kwalificaties over romanschrijvers die tegemoet proberen te komen aan vlakke, snelle lezers. Zo oordelen sommige recensenten uiterst negatief over de taal van auteurs als Kader Abdolah (‘ongecomplieerde, eenduidige stijl’, Houwelingen 2017) of Maartje Wortel (‘kleuterverhalen’, Peters 2015).

* Dit onderzoek is mogelijk gemaakt door een Crossing the Borders-beurs van ICON/UiL OTS. We willen Lisa Arnold bedanken voor het selecteren en gebruiksklaar maken van het grootste deel van het romancorpus van dit onderzoek.

Arie Storm spreekt in een kritiek over ‘het stijlmiddel van de korte zinnen’ (Storm 2011) en Jeroen Vullings stelt dat ‘cohorten schrijvers’ vandaag de dag ‘stilistisch onderpresteren’:

Zij schrijven de romans die ik het snelst wegleg, en dat niet alleen omdat er in het verhaal niks van belang passeert, maar vooral omdat ze wel in een mal gegoten lijken te worden: eenvormig, karakterloos, steriel. En dat met een uiterst elementaire woordenschat, waarin het gewone ontstellend regeert, hetgeen de infantilisering van dit massaproduct vervolmaakt (Vullings 2015b).

Met de term ‘massaproduct’ wijst Vullings op een samenhang tussen enerzijds veranderend lees- en koopgedrag en anderzijds het taalgebruik: alleen met een vlakke stijl kan het boek probleemloos als consumptieartikel functioneren in de hedendaagse boekenmarkt. Sommige auteurs melden dat zij zich door uitgevers gedwongen voelen tot een eenvoudige stijl. Zo merkte Rascha Peper op dat haar redacteur haar vermeend ouderwetse taalgebruik probeerde te moderniseren (Peper 2012) en Wessel te Gussinklo klaagde in een interview dat een essay door Meulenhoff vanwege de moeilijkheidsgraad ‘niet marktconform’ werd gevonden (Vullings 2015a). Maxim Februari schreef dat uitgevers ‘een staf van persklaarmakers in dienst [hebben] die rare woorden vervangen door gebruikelijker woorden’, een proces dat hij onder meer in verband bracht met de groeiende macht van het neoliberalisme, het gebruik van programmeertalen, de standaardisatie van bedrijfsmatige en beleidsmatige processen en veranderend mediagebruik (Februari 2016).

In een nog breder kader stelde Alessandro Baricco (2010) dat de cultuurgebruiker van de eenentwintigste eeuw een ‘barbaar’ is: iemand die superieur over de oppervlakte van de samenleving scheert, maar niet in diepgang geïnteresseerd is. Snelheid, daar zou het de barbaar allemaal om gaan. Die snelheid kan volgens Baricco alleen gegarandeerd worden als barrières zoveel mogelijk uit de weg geruimd worden.

Is het inderdaad zo dat komma’s en bijzinnen vandaag in de literatuur vermeden worden, om de haastig lezende en snel afgeleide consument maar te blijven boeien? Korter gesteld: zijn romans de afgelopen decennia eenvoudiger geworden? In dit artikel willen we deze vraag beantwoorden vanuit het perspectief van de taalwetenschap en literatuurwetenschap, waarbij we niet alleen naar de taal van de roman kijken maar ook naar de taal van de krant. Zo kunnen we zien of er sprake is van een genrespecifieke ontwikkeling in de (roman)literatuur of dat er een bredere, algemenere ontwikkeling gaande is in geschreven taalgebruik. We onderzoeken een corpus van romanfragmenten en krantenberichten tussen 1950 en 2014 op stilistische patronen die wijzen op complexiteit en proberen de resultaten van die analyse te interpreteren.

2 Theoretische achtergrond

Het min of meer anekdotische bewijs voor een tendens van ‘naar beneden schrijven’ laat zich aanvullen met een drietal observaties die we putten uit sociologisch, retorisch en stilistisch onderzoek. Respectievelijk gaat het om onderzoek naar hoe we met tijd omgaan, naar de verdeling van verantwoordelijkheden tussen lezer en schrijver, en om internationaal diachroon onderzoek naar talige complexiteit.

2.1 Leesgedrag en tijdsbesteding

Barrico bracht zijn sombere kijk op cultuurverandering direct in verband met de notie snelheid. En de notie *versnelling van de tijd* staat centraal in een invloedrijke studie die de Duitse socioloog Hartmut Rosa (2005) heeft geschreven over de manier waarop de moderne mens met tijd omgaat. Rosa claimt dat er op drie terreinen sprake is van een voortdurende versnelling: (1) in het productieproces van goederen en voedsel; (2) in de organisatie van maatschappelijke processen; en (3) in het privéleven van het individu. Voor de leescultuur is dat laatste domein uiteraard het belangrijkste. In het privé-domein schetst Rosa vier belangrijke aan tijd gerelateerde ontwikkelingen, die alle tot een gevoel van versnelling leiden: (1) we zijn veel activiteiten in een hoger tempo gaan uitvoeren; (2) we pauzeren minder en eten bijvoorbeeld onze boterham achter het beeldscherm op; (3) we voeren verschillende activiteiten tegelijk uit (multi-tasking); (4) en we hebben tijdrovende activiteiten vervangen door andere werkzaamheden die minder tijd kosten.

Op drie van de vier punten zien we een relatie met leesgedrag. We zijn sneller gaan lezen; daar worden tegenwoordig zelfs trainingen voor aangeboden. Tijdens het lezen houden we ons mobieltje in de gaten, kijken we televisie, luisteren we naar muziek, reizen we van A naar B. En ten slotte ‘lezen’ we tegenwoordig ook boeken door naar een van de talloze theatervoorstellingen of verfilmingen te gaan die gebaseerd zijn op het boek dat we dan niet meer hoeven te lezen. In één avond klaar met Griet Op de Beeck of Hendrik Groen. In een tijd van versnelling is lezen een activiteit die het onderspit delft in de concurrentie met andere vormen van vrijetijdsbesteding waarbij meer actie en snelheid ervaren worden, zoals film en televisie. Rosa wijst overigens wel op onderzoek waaruit blijkt dat aan televisiekijken weliswaar meer tijd wordt besteed, maar dat men achteraf aan die activiteit minder waarde toekent dan aan bijvoorbeeld het lezen van een boek. In het algemeen, zegt hij, zijn mensen geneigd meer tijd te besteden aan activiteiten die zij *minder* waardevol achten, vermoedelijk omdat die activiteiten minder cognitieve energie kosten. Dit lijkt tegenstrijdig, maar is volgens Rosa verklaarbaar als we begrijpen dat de mens in zijn stressvolle leven kiest voor vrijetijdsbestedingen die meer passief zijn. In dat licht verbaast het volgens Rosa dan ook niet dat mensen minder televisie kijken en juist meer lezen in periodes waarin zij het werk op afstand houden, zoals tijdens vakantie. Dan hebben zij blijkbaar meer cognitieve energie beschikbaar.

Ook uit Nederlands onderzoek naar tijdsbesteding blijkt dat de leestijd van Nederlanders afneemt. Waar ze in 1955 gemiddeld nog 2,4 uur per week in een boek lasen, was dat in 1975 afgenomen tot 1,6 uur en in 2005 tot 1,3 uur (Leesmonitor). Het lezen van boeken daalde dus al scherp in de periode 1955-1975, maar papieren dag- en nieuwsbladen en tijdschriften raakten vooral tussen 1975 en 2005 in een vrije val. In 1975 las men nog 2,7 uur per week in gedrukte kranten en 1,5 uur in gedrukte tijdschriften, in 2005 was dat allebei praktisch gehalveerd (Huysmans & De Haan 2010: 40). Het meest recente rapport van het CBS (2017) meldt dat de Nederlander nu gemiddeld 37 minuten per dag besteedt aan het lezen van boeken, tijdschriften en kranten. Ook in oplages heeft de printcultuur sterk moeten inleveren: waar in 1995 de totale betaalde printoplage van regionale en landelijke kranten op het hoogtepunt van bijna 4,7 miljoen lag, was dat in 2016 gedaald tot

onder de 2,5 miljoen (Bakker 2017). De boekenmarkt heeft het eveneens niet gemakkelijk sinds de economische crisis van 2008 (KVB Boekwerk 2017).

Misschien moeten we ons niet laten misleiden door de afname van het lezen op papier, omdat er sinds enkele jaren op grote schaal digitaal gelezen wordt (Wennekers, Van Troost & Wiegman 2016). Het lezen van kranten nam tussen 2005 en 2015 toe van 1,5 naar 1,9 uur, vanwege de brede verbreiding van online (gratis) nieuws (Leesmonitor). Dat neemt niet weg dat er aanwijzingen zijn dat mensen met minder geduld lezen: er worden vermoedelijk minder boeken uitgelezen (KVB Boekwerk 2016) en mensen combineren het lezen vaak met een andere activiteit (Wennekers e.a. 2016). Dat betekent dat ze meerdere teksten afwisselend bekijken, tijdens het lezen iets posten op de sociale media of tussendoor iets online opzoeken (Leesmonitor). Al in 2010 – op een moment dat de sociale media nog veel minder in de haarvaten van de samenleving zaten dan nu het geval is – schreven Huysmans en De Haan: ‘Vermoedelijk leest men ook minder vaak langere periodes aaneengesloten, dus meer fragmentarisch en mogelijk vluchtiger, al ontbreekt voor die laatste indruk empirisch bewijs’ (Huysman & De Haan 2010: 114). De uitgever die de strijd wil aangaan met de snellere en gemakkelijker te consumeren media doet er vermoedelijk dus goed aan niet te veel van de cognitieve energie van de lezer te vragen. En dus wordt de auteur, volgens het genoemde hoofdredactioneel van *Trouw*, gevraagd niet omhoog maar ‘naar beneden’ te schrijven.

2.2 *Verantwoordelijkheid van lezer en schrijver*

Een tweede theoretisch perspectief komt voort uit de zogenaamde contrastieve retorica, een taalwetenschappelijke niche waarin verschillen tussen talen en hun taalgebruiksconventies bestudeerd worden. Hinds (1987) was een van de eerste onderzoekers die stijlverschillen tussen Engelse en Japanse auteurs in verband bracht met een verschil in de verdeling van verantwoordelijkheden tussen schrijver en lezer. In de Engelse cultuur ligt de verantwoordelijkheid voor succesvolle communicatie volgens Hinds in de eerste plaats bij de auteur. Een lezer die de tekst niet begrijpt, heeft als eerste reactie: de auteur heeft het niet goed uitgelegd. Een Japanse lezer zal in een vergelijkbare situatie zichzelf verwijten niet genoeg inzet getoond te hebben. Natuurlijk is dit een clichématige weergave en misschien gaan er ook culturele vooroordelen achter schuil – het essay van Hinds bleef niet zonder weerwoord (vgl. Kubota 1999) – maar dat neemt niet weg dat de verdeling van verantwoordelijkheden tussen schrijver en lezer een interessant begrip is. We vragen ons af of in de afgelopen honderd jaar in de westerse wereld op dit punt misschien een verschuiving plaatsvindt. Het zogeheten ‘naar beneden schrijven’ kan dan gezien worden als een culturele ontwikkeling waarbij de schrijver steeds meer de verantwoordelijkheid draagt voor de geslaagdheid van de communicatie. Hinds legt een expliciet verband tussen de communicatiestijlen van de beide landen en opvattingen over beleefdheid. Hoe zwaarder de normen voor beleefd gedrag de cultuur domineren, des te sterker ligt de verantwoordelijkheid bij de lezer, die zijn best moet doen de complexiteit van de tekst te doorgronden. En in een cultuur waar beleefdheid een veel kleinere rol speelt, mag de lezer hogere eisen stellen aan de schrijver. Hinds spreekt zich niet uit over de historische ontwikkeling in de verantwoordelijkheidsverdeling tussen lezer en schrijver, noch maakt hij

een onderscheid tussen domeinen van communicatie, zoals de literatuur, de pers en de zakelijke communicatie. Maar hij zet deze verdeling wel op de agenda als variabele; hij relateert deze aan machtsverhoudingen tussen tekstproducenten en lezers. Hebben lezers in het westen in de afgelopen decennia meer macht gekregen?

De veeleisende lezer heeft de laatste decennia in elk geval wel een belangrijke partij aan zijn zijde gekregen: de overheid. In een indrukwekkend literatuuroverzicht van de ontwikkelingen in de vs rond de *plain language movement* maakt Schriver (2017) duidelijk dat de overheid tussen 1940 en 2015 op tal van gebieden regels uitvaardigde waarin niet alleen de overheid zelf verplicht werd meer rekening te houden met de lezer, maar ook derden zoals banken, verzekeraars, gezondheidsorganisaties en rechters. In dezelfde periode werd de oorspronkelijke focus op zinslengte en woordcomplexiteit, gesymboliseerd in de invloedrijke Flesch-leesbaarheidsscore, verbreed naar tekstkenmerken die de zinsgrenzen overstijgen, zoals samenhang tussen tekstonderdelen, de kwaliteit van titels en de vormgeving. Ook kwam er een verbreding van genres: niet alleen brieven en circulaire, maar ook websites en formulieren moesten aan een begrijpelijkheidsstandaard voldoen. Ten slotte ging het niet alleen meer om begrijpelijkheid, maar ook om vindbaarheid en gebruiksvriendelijkheid. Interessant zijn de overwegingen die belangrijke actoren in dit proces achteraf formuleren voor hun beleid. Zo zegt de gouverneur van Florida in een interview:

Government can be frustrating enough as it is, and so the opportunity to communicate in language that people can understand and appreciate is very important. When government officials speak to people in government jargon that is difficult to understand, it is not only not appreciated, I think it is unkind (geciteerd in Schriver 2017: 359).

En een van de auteurs van een invloedrijke richtlijn uit 1998 voor financiële communicatie zegt:

Underlying the advice was the idea of transparency and that people had a right to clear communications that affected their lives. We felt it was the responsibility of the writer to make the material clear and understandable, not the responsibility of the reader to ferret out the meaning of a text. This was a radical idea at the time, especially among lawyers and accountants, who fought us every step of the way (geciteerd in Schriver 2017: 358).

Opmerkelijk is dat in deze beide citaten een relatie wordt gelegd met zowel de notie *beleefdheid*, zoals Hinds (1987) ook deed, als de notie *verantwoordelijkheid*, die in het tweede citaat expliciet op de schouders van de auteurs gelegd wordt. Bovendien is er sprake van een *recht op begrijpelijkheid*. Hier lijkt de lezer inderdaad aan macht te winnen en men lijkt dit ook als iets nieuws te ervaren.

Ook in Europa ontstond een wetgevend kader voor de begrijpelijkheid van teksten in diverse domeinen, zoals financiën en gezondheid. Zo moet in Nederland de Autoriteit Financiële Markten sinds 2007 toezien op de begrijpelijkheid, evenwichtigheid en duidelijkheid van informatie over beleggingsproducten, hypotheek en pensioenen. En het College ter Beoordeling van Geneesmiddelen ziet toe op de leesbaarheid van de medische bijsluiter, mede op grond van een Europese richtlijn die in 1992 van kracht werd. In Europa mag een nieuw geneesmiddel alleen toegelaten worden tot de markt als de bijsluiter getest is onder lezers en begrijpelijk is bevonden.

2.3 *Internationaal onderzoek naar complexiteit in kranten*

Dankzij de vorderingen op het gebied van de taaltechnologie kunnen we vragen over talige vereenvoudiging tegenwoordig onderzoeken door grote verzamelingen teksten met leesbaarheidstools te analyseren. Voor het Britse journalistieke domein is dat al eens gedaan door Westin & Geisler (2002). Zij analyseerden 864 redactionele commentaren uit drie Britse dagbladen tussen 1900 en 1993 op vijf verschillende dimensies waarvan er drie in verband gebracht kunnen worden met een alledaagse conversatiestijl, contrasterend met een meer hoog geletterde stijl. Op twee van die drie dimensies constateren zij een ontwikkeling: het taalgebruik wordt minder formeel en minder abstract of passief. Welke tekstkenmerken precies verantwoordelijk zijn voor die ontwikkeling wordt in hun rapportage niet goed duidelijk.

Štajner & Mitkov (2012) maakten zowel een diachrone (jaren dertig, zestig en negentig) als een regionale vergelijking (UK-VS) voor vijftien genres met zowel krantenartikelen als literaire fictie. Zij voerden een analyse uit op vier tekstkenmerken: zinslengte, leesbaarheid (met de Kincaid-formule, een variant op de eerder genoemde Flesch-score), zinscomplexiteit (hoeveelheid deelzinnen) en passiva. Het voordeel van deze analyse is dat preciezer dan bij Westin & Geisler zichtbaar wordt voor welke tekstkenmerken er significante verschillen zijn, maar het nadeel is dat er met drie periodes, vijftien genres en twee regio's verwacht mag worden dat er wel altijd ergens verschillen tevoorschijn komen. Zo wordt de zinslengte in Britse romans in de eerste periode (1930-1960) langer, maar blijft die later stabiel, terwijl de zinslengte in Amerikaanse kranten bij de reportages afneemt in de tweede periode. In dat laatste genre nam de zinscomplexiteit toe in beide regio's. Misschien zijn zulke resultaten dus enigszins toevallig. Een poging tot verklaring bieden de auteurs in elk geval niet. We concluderen dat de ene studie voor het Engelse taalgebied enige aanwijzingen biedt voor een vereenvoudiging in de krantentaal, terwijl de andere studie weinig duidelijkheid biedt.

Bij ons weten zijn wij de eersten die onderzoek doen naar de complexiteit van Nederlandse roman- en krantentaal in diachroon perspectief. Het literaire corpus is ons primaire materiaal: we willen vooral toetsen of de eerder genoemde schrijvers, columnisten en recensenten een punt hebben wanneer ze aannemen dat literaire teksten *stilistisch* steeds eenvoudiger worden, dus op niveaus als zinslengte, bijzinnen, et cetera. Over metaforische of 'literaire' complexiteit spreken we ons niet uit, zoals over het benodigde referentiekader, ingewikkelde motievencomplexen of complexe tijd- en ruimtesprongen. Het zou uiteraard boeiend zijn ook daar een diachrone analyse van te maken – al wordt er voor zover wij weten veel minder geklaagd over een afname van dat type complexiteit in hedendaagse romans.

Het krantencorpus gebruiken we als referentiemateriaal, om te controleren of er algemenere tendensen zijn in de ontwikkeling van de Nederlandse taal. Zo kunnen we nagaan of de eventueel in de romans aanwezige ontwikkeling verklaard zou kunnen worden als exponent van een algemenere taalontwikkeling. In het vervolg bespreken we hoe wij ons corpus met teksten hebben samengesteld, wat we onder eenvoudig taalgebruik verstaan en hoe we de analyse op het materiaal hebben uitgevoerd.

3 Methode

3.1 *De samenstelling van het corpus*

Allereerst bespreken we de verzameling literaire teksten. We hebben handmatig op basis van in PiCarta vermelde teksten een corpus samengesteld van zestig romans, gelijk verdeeld over drie periodes: 1955-1964, 1980-1989 en 2005-2014. Daarbij hebben we geprobeerd de drie deel-steekproeven zo breed mogelijk te maken. Ze bevatten daarom romans die als literair bekend staan (afgaande op een bespreking of vermelding ervan in Brems (2006), Van Boven & Kemperink (2006) of in literaire kritieken). De beperking tot romans en dus de uitsluiting van bijvoorbeeld verhalenbundels vloeit voort uit een pragmatische overweging: dan is het makkelijker het beoogde aantal corpusfragmenten uit integrale, samenhangende teksten te selecteren.

Met het oog op de breedheid van de steekproef hebben we per periode alleen auteurs geselecteerd die maximaal vijftien jaar vóór of die ín de betreffende periode zijn gedebuteerd. Op die manier hoopten we iedere periode zo goed mogelijk vrij te houden van ‘historische’ najleffecten van langlopende oeuvres. Het oeuvre van een (belangrijk) auteur kan immers een lange levenslijn hebben, en zo’n auteur kan de eigen stijl in die periode stevig consolideren. Om te kunnen zien of zich nieuwe ontwikkelingen voordoen in de literaire taal, kiezen we dus per steekproefperiode de werken van ‘nieuwe’ auteurs, zodat we de taal van nieuwe generaties meten. Om te kunnen bezien of niet alleen de taal in het algemeen zou zijn veranderd maar meer specifiek het taalgebruik van literaire schrijvers, zijn de steekproeven wijd uit elkaar genomen en komt iedere auteur in een periode maar éénmaal voor met één roman.

Bij de samenstelling van het corpus hebben we rekening gehouden met de volgende literaire en ideologische afwegingen:

- Noord/Zuid. We hebben gekozen voor literaire teksten van Nederlandse auteurs, hoewel het voor letterkundigen gebruikelijk is de Nederlandstalige literatuur in zijn geheel te beschouwen, dus inclusief de Vlaamse, Surinaamse en Caribische literatuur. In onze analyse wilden we de complexiteit vermijden waar Štajner & Mitkov (2012) onder gebukt gingen, door ook een regionale conditie (met de bijbehorende stilistische variatie) als variabele mee te nemen.
- ‘Hoog’/‘laag’. De selectie is nadrukkelijk niet beperkt tot de canonieke, ‘hoge’ literaire werken. Naast romans van hoog aangeschreven auteurs als Harry Mulisch, Cees Nooteboom, Nelleke Noordervliet, Frans Kellendonk, Kader Abdolah en Franca Treur, bevat het corpus ook proza dat eerder als *middlebrow* bekend staat, van auteurs als Dolf Verroen, Hans Berghuis, Guus Vleugel en Naima el Bezaz.
- Gender en etniciteit. We hebben bij de selectie niet naar een gelijke verdeling van mannen en vrouwen gestreefd, maar naar een verdeling die enigszins in de buurt komt van het aandeel van vrouwelijke auteurs op de boekenmarkt van dat moment. Tot eind jaren negentig was ongeveer 25 procent van de gepubliceerde boeken van een vrouw; pas daarna begon dat aandeel te stijgen (Vos 2008: 56). In onze steekproef valt het aandeel vrouwelijke auteurs overigens iets lager uit, vanwege de digitale beschikbaarheid van

het materiaal en de andere selectiecriteria waarmee we rekening willen houden. Etniciteit speelt in onze corpusafbakening op vergelijkbare wijze een rol: in iedere periode zijn er boeken opgenomen van niet-witte auteurs.

Per boek hebben we 3000 woorden geselecteerd, alleen in hele zinnen; verdeeld over vijftien, gelijkelijk over het gehele boek verspreide fragmenten van elk (ongeveer) 200 woorden.¹ We hebben omwille van de homogeniteit van het materiaal alleen fragmenten met vertellerstekst gekozen; passages met (meer dan 5%) monologen en dialogen zijn achterwege gelaten, omdat we vermoeden dat vertellerstekst wezenlijk andere stilistische kenmerken heeft dan dialogen en monologen. De fragmenten bevatten dus voornamelijk beschrijvingen van situaties, processen en handelingen, waardoor ze goed vergelijkbaar zijn met de teksten die we uit de kranten hebben geselecteerd. Iedere geselecteerde passage moest een aaneengesloten stuk tekst zijn, maar alineagrenzen en witregels mochten daar wel deel van uit maken.

Voor de verzameling van de krantenteksten hanteerden we een vergelijkbare werkwijze. Maar hier kozen we slechts voor de twee periodes aan de uiteinden van het romancorpus: 1950 versus 2002. Net zoals we in het literaire corpus kozen voor zogenaamd hogere en lagere literatuur, werkten we met twee kranten die respectievelijk als 'hoog' en 'laag' gezien kunnen worden: *NRC* (later *NRC Handelsblad*) en *De Telegraaf*.² *De Telegraaf* staat bekend als een populaire krant, *NRC* als een kwaliteitskrant. Dit onderscheid heeft zijn oorsprong al aan het begin van de twintigste eeuw (Kester 2007). In elke krant selecteerden we vijf rubrieken (buitenland, binnenland, politiek, sport en cultuur), en voor elke rubriek kozen we tien stukken tekst van ongeveer 200 woorden. De meeste teksten konden niet toegeschreven worden aan een auteur, waardoor een selectie op basis van gender of etniciteit was uitgesloten. Ook hier kozen we voor berichten zonder citaten; zo vermeden we dus net als bij de romans spreektaal in monologen en dialogen.

3.2 *De analyse van complexiteit*

Voor de analyse is het (computer)programma T-Scan gebruikt (Pander Maat e.a. 2014), dat in de huidige versie ruim 400 tekstkenmerken weergeeft. T-Scan kan niet gebruikt worden om bijvoorbeeld de eerder genoemde literaire complexiteit van een tekst te meten, maar zoals aangegeven gaat het ons om het meten van stilistische complexiteit (zie paragraaf 2.3). We onderzochten de teksten op vier verschillende dimensies van stilistische complexiteit, die van (1) het woord; (2) de zinsbouw; (3) de samenhang; en (4) de mate waarin de tekst verwijst naar concrete personen. We lichten elke dimensie hieronder toe.

1 We hebben met gedigitaliseerde teksten gewerkt. Een gedeelte van het materiaal was aanwezig in een groot corpus van gedigitaliseerde teksten dat al was aangelegd in het kader van onderzoek van Pander Maat. Daarnaast konden we gebruik maken van de grote persoonlijke digitale bibliotheek van neerlandicus Ewoud Sanders. Van de teksten die niet in deze bronnen voorhanden waren, hebben we zelf scans gemaakt.

2 Deze kranten zijn afkomstig uit het VU-DNC-corpus (<https://portal.clarin.inl.nl/vu-dnc/>). We hebben voor 1950 de *Nieuwe Rotterdamse Courant* gebruikt (de fusie met *Algemeen Handelsblad* was toen nog niet doorgevoerd).

Lexicale complexiteit en concreetheid

Woorden kunnen moeilijk zijn omdat ze weinig frequent zijn of omdat ze inhoudelijk complex zijn. We kozen daarom in de analyse voor *woordfrequentie* en *woordconcreetheid* als analysecategorieën (zie verder Pander Maat & Dekker 2016). Meer in het bijzonder is gekeken naar:

- het gemiddelde van de gestandaardiseerde woordfrequentie van inhoudswoorden, gebaseerd op het Subtlex-corpus. Behalve functiewoorden zijn ook namen buiten beschouwing gelaten; hoe lager de woordfrequentie, des te complexer de tekst;
- de proportie zelfstandige naamwoorden die concreet zijn en verwijzen naar personen, planten, dieren, voorwerpen en waarneembare substanties, lichaamsdelen, natuurverschijnselen en gebeurtenissen; hoe concreter de woorden, hoe minder complex de tekst;
- de proportie zelfstandige naamwoorden die algemeen zijn (bijvoorbeeld *idee*, *discussie*); een grotere proportie wijst op een grotere tekstcomplexiteit;
- de proportie adjectieven die concreet zijn, die verwijzen naar emoties en direct waarneembare kenmerken; grotere proportie, geringere tekstcomplexiteit;
- de proportie concrete werkwoorden, die verwijzen naar direct waarneembare situaties en processen; grotere proportie, geringere tekstcomplexiteit;
- de proportie algemene werkwoorden (bijvoorbeeld *vergen*, *argumenteren*); een grotere proportie wijst op een grotere tekstcomplexiteit;
- de dichtheid van zelfstandige naamwoorden die nominalisaties zijn (bijvoorbeeld *noodzakelijkheid*); hoe groter deze dichtheid, des te groter de tekstcomplexiteit.

Op al deze kenmerken zijn substantiële verschillen gevonden tussen Nederlandse tekstgenres (Pander Maat & Dekker 2016). Verder hebben de eerste twee kenmerken aantoonbaar effect op het tekstbegrip (Kleijn 2018).

Syntactische complexiteit

We kozen hier voor maten die de grammaticale omvang van een zin of deelzin betreffen. Op zinsniveau ging het om:

- zinslengte in woorden;
- het aantal persoonsvormen per zin;
- het aantal finiete bijzinnen per zin.

Op het niveau van de deelzin ging het om:

- deelzinslengte, dat wil zeggen het aantal woorden per persoonsvorm;
- het aantal attributieve adjectieven per deelzin;
- het aantal andere bijvoeglijke bepalingen per deelzin;
- het aantal opsommingen per deelzin (dat wil zeggen het aantal keren dat via nevenschikking een extra zinsdeel wordt toegevoegd).

Op al deze kenmerken zijn substantiële verschillen gevonden tussen Nederlandse tekstgenres (Pander Maat 2017); voor elke dimensie geldt dat hoe groter de lengte / het aantal, des te groter de tekstcomplexiteit. Verder hebben met name de deelzinskenmerken aantoonbaar effect op het tekstbegrip (Kleijn 2018). Teksten zijn eenvoudiger als de deelzinnen kort zijn en weinig bepalingen bevatten.

Cohesie

Cohesie is geanalyseerd door te kijken naar woordoverlap. Daar hebben we twee maten voor:

- Measure of Textual Lexical Diversity (MTLD), een soort type-token-ratio die corrigeert voor het effect van tekstlengte;
- argumentoverlap: de proportie inhoudswoorden en voornaamwoorden (gelemmatiseerd) in de tekst die genoemd is in de voorgaande 50 woorden.

Teksten zijn eenvoudiger als een groter aandeel van de woorden in de tekst een herhaling is van eerdere woorden. Dat is verklaarbaar doordat de auteur dan langer spreekt over hetzelfde thema en voor dat thema dezelfde woorden gebruikt.

Persoonsdichtheid

Wij hebben dit kenmerk gedefinieerd als het aantal woorden dat verwijst naar personen. Daarbij gaat het om:

- de dichtheid van persoonlijke (incl. bezittelijke) voornaamwoorden;
- de dichtheid van nomina verwijzend naar personen;
- en de dichtheid van persoonsnamen;
- de totale dichtheid van persoonsverwijzingen.

Vaak is verondersteld (Quereshi 1991) dat teksten waarin veel mensen voorkomen eenvoudiger zijn, of minimaal eenvoudiger lijken.

Hoe verloopt de analyse nadat deze kenmerken zijn bepaald? T-scan produceert een dataset die ingelezen kan worden in een programma voor statistische analyse, zoals SPSS. De analyse-categorieën die we hierboven bespraken, worden in deze dataset op verschillende manieren weergegeven. Soms zijn het absolute aantallen, zoals een gemiddeld aantal letters per woord of aantal woorden per zin. Andere kenmerken worden als proportie weergegeven van een deelverzameling van de tekstwoorden, zoals de proportie concrete werkwoorden op het totaal aantal werkwoorden. Wanneer alle tekstwoorden worden gebruikt bij het bepalen van de frequentie van een verschijnsel, is er sprake van een dichtheid; dat is dan bijvoorbeeld het aantal persoonlijke voornaamwoorden per 1000 woorden. Een analyse kan uitgevoerd worden op het niveau van woorden, (deel)zinnen, alinea's, het volledige fragment, het genre (zoals sportberichten) en de roman als geheel. Voor de verschillende analysevragen werken we in het vervolg op twee niveaus: bij de nieuwsberichten gaat het om de tekst als geheel, bij de romans zijn er twee niveaus: dat van het tekstfragment en dat van alle tekstfragmenten per roman tezamen.

4 Zijn er verschillen tussen kranten en romans?

Voordat we romans en kranten afzonderlijk beschouwen, vergelijken we beide genres wat betreft complexiteit. Daarbij kijken we ook naar eventuele verschillen tussen de krantenrubrieken, om te zien in hoeverre het krantengere heterogeen is. We hebben in deze analyse de romandata versimpeld, door elke roman als één analyse-eenheid te beschouwen, gedefinieerd als het gemiddelde over de 15 fragmenten. Om de tijdsfactor constant te houden, beperken we ons in deze paragraaf tot de romans uit 1955-1964 en 2005-2014. Op die manier komen we tot 240 ob-

servaties: tweemaal 100 krantenberichten (twintig per rubriek) en tweemaal 20 romans. We splitsen hieronder de resultaten uit naar periode, maar in de bespreking gaan we nog niet in op periodeverschillen. We beperken ons tot acht van de 21 complexiteitskenmerken, waar we interessante verschillen vonden.

Qua lexicale complexiteit zijn de genreverschillen groot. Wat betreft woordfrequentie laat Tabel 1 zien dat de krantenberichten in het algemeen lager scoren (en dus minder frequente woorden bevatten) dan de romans ($F_{[1,236]}=99.00$, $p<.001$, $\eta^2=.296$). In kranten valt op dat in de sportberichten bekendere woorden gebruikt worden dan in andere berichtsoorten, maar nog lang niet zo bekend als die in romans. Wat betreft concrete nomina scoren kranten ook lager dan romans ($F_{[1,236]}=67.18$, $p<.001$, $\eta^2=.222$), zie Tabel 2. In de kranten zijn de politieke berichten veruit het meest abstract. Verder komen in de kranten veel minder concrete adjectieven voor ($F_{[1,236]}=102.72$, $p<.001$, $\eta^2=.303$), zie Tabel 3. Soortgelijke verschillen tussen kranten en romans zijn gerapporteerd in Pander Maat & Dekker (2016). Onze conclusie op woordniveau is dat kranten complexer zijn dan romans.

TABEL 1 Genre- en periodeverschillen wat betreft de woordfrequentie van de inhoudswoorden in de tekst*

Periode	Kranten politiek	Kranten cultuur	Kranten binnenland	Kranten buitenland	Kranten sport	Kranten totaal	Romans
1950/55	4,06 (0,18)	4,02 (0,17)	4,06 (0,30)	4,09 (0,15)	4,24 (0,12)	4,09 (0,20)	4,50 (0,21)
2002/ 2005	4,20 (0,19)	4,16 (0,31)	4,20 (0,25)	4,16 (0,18)	4,30 (0,22)	4,20 (0,24)	4,54 (0,18)
Totaal	4,13 (0,20)	4,09 (0,25)	4,13 (0,28)	4,12 (0,17)	4,27 (0,18)	4,15 (0,23)	4,52 (0,19)

* De woordfrequenties zijn gebaseerd op een corpus met ondertitels bij tv-programma's. Ze zijn uitgedrukt als de logaritme met grondtal 10 van de gemiddelde woordfrequentie per miljard woorden. Dus een woordfrequentie van 4 betekent dat de inhoudswoorden in een tekst gemiddeld 10.000 keer voorkomen per 1.000.000.000 woorden.

TABEL 2 Genre- en periodeverschillen wat betreft proportie concrete zelfstandige naamwoorden op alle zelfstandige naamwoorden per tekst

Periode	Kranten politiek	Kranten cultuur	Kranten binnenland	Kranten buitenland	Kranten sport	Kranten totaal	Romans
1950/55	0,22 (0,13)	0,39 (0,08)	0,29 (0,12)	0,29 (0,15)	0,32 (0,10)	0,30 (0,13)	0,45 (0,05)
2002/ 2005	0,26 (0,10)	0,35 (0,10)	0,33 (0,10)	0,29 (0,12)	0,35 (0,11)	0,32 (0,11)	0,48 (0,07)
Totaal	0,24 (0,11)	0,37 (0,09)	0,31 (0,11)	0,29 (0,14)	0,33 (0,10)	0,31 (0,12)	0,47 (0,06)

TABEL 3 Genre- en periodeverschillen wat betreft proportie concrete adjectieven op alle adjectieven per tekst

Periode	Kranten politiek	Kranten cultuur	Kranten binnen-land	Kranten buiten-land	Kranten sport	Kranten totaal	Romans
1950/55	0,04 (0,04)	0,11 (0,07)	0,07 (0,07)	0,05 (0,07)	0,09 (0,06)	0,07 (0,07)	0,19 (0,05)
2002/ 2005	0,02 (0,03)	0,09 (0,06)	0,04 (0,05)	0,07 (0,10)	0,06 (0,04)	0,06 (0,06)	0,16 (0,04)
Totaal	0,03 (0,03)	0,10 (0,07)	0,05 (0,06)	0,06 (0,08)	0,08 (0,05)	0,06 (0,07)	0,17 (0,04)

Op syntactisch niveau zijn de verschillen tussen krant en roman minder groot. De zinslengte in kranten is groter dan die in romans ($F[1,236]=17.02, p<.001, \eta^2=.067$). Weer zijn de politieke berichten het meest complex. Waaraan worden de extra woorden besteed? Het blijkt dat het verschil niet zozeer afkomstig is van bijzinnen of nevengeschikte deelzinnen, maar van bepalingen binnen deelzinnen.

Met name het verschil in bijvoeglijke bepalingen is groot. Zo komen attributieve adjectieven in kranten veel meer voor dan in romans ($F[1,236]=78.53, p<.001, \eta^2=.250$), zie Tabel 4. Hierboven zagen we al dat die krantenadjectieven vaker abstract zijn dan de romanadjectieven. Maar attributieve adjectieven voegen niet veel woorden aan de zin toe. Dat geldt wel voor andere bijvoeglijke bepalingen, bijvoorbeeld in de vorm van voorzetselgroepen. Het blijkt dat de ‘overige’ bijvoeglijke bepalingen in kranten veel vaker voorkomen ($F[1,236]=98.43, p<.001, \eta^2=.294$), zie Tabel 5. Bijvoeglijke bepalingen komen in alle rubrieken voor en dienen in kranten vaak om verwijzingen naar entiteiten te preciseren (vgl. *minister Koenders van Buitenlandse Zaken*). Dat soort preciseringen is in romans waarschijnlijk veel minder vaak nodig. Hoe dan ook, bijvoeglijke bepalingen zorgen voor een hogere informatiedichtheid binnen de deelzin. Onze conclusie is dus dat ook op zinsniveau de krant complexer is dan de roman.

TABEL 4 Genre- en periodeverschillen wat betreft zinslengte in woorden

Periode	Kranten politiek	Kranten cultuur	Kranten binnen-land	Kranten buiten-land	Kranten sport	Kranten totaal	Romans
1950/55	29,68 (9,16)	25,44 (6,81)	21,40 (4,79)	23,60 (6,07)	24,70 (6,44)	24,96 (7,21)	20,53 (5,43)
2002/ 2005	20,69 (3,92)	19,94 (4,00)	18,74 (4,51)	18,72 (4,48)	17,68 (2,97)	19,16 (4,08)	15,54 (2,62)
Totaal	25,18 (8,32)	22,69 (6,18)	20,07 (4,78)	21,16 (5,81)	21,19 (6,09)	22,06 (6,53)	18,04 (4,92)

TABEL 5 Genre- en periodeverschillen wat betreft aantal attributieve adjectieven per deelzin

Periode	Kranten politiek	Kranten cultuur	Kranten binnenland	Kranten buitenland	Kranten sport	Kranten totaal	Romans
1950/55	0,63 (0,24)	0,68 (0,22)	0,56 (0,22)	0,64 (0,17)	0,55 (0,15)	0,61 (0,20)	0,34 (0,10)
2002/ 2005	0,52 (0,14)	0,61 (0,24)	0,50 (0,14)	0,62 (0,17)	0,48 (0,18)	0,55 (0,18)	0,27 (0,06)
Totaal	0,57 (0,20)	0,64 (0,23)	0,53 (0,18)	0,63 (0,17)	0,51 (0,17)	0,58 (0,20)	0,30 (0,09)

TABEL 6 Genre- en periodeverschillen wat betreft aantal andere bijvoeglijke bepalingen per deelzin

Periode	Kranten politiek	Kranten cultuur	Kranten binnenland	Kranten buitenland	Kranten sport	Kranten totaal	Romans
1950/55	1,45 (0,26)	1,48 (0,43)	1,30 (0,53)	1,47 (0,71)	1,29 (0,35)	1,40 (0,49)	0,61 (0,18)
2002/ 2005	1,14 (0,36)	1,21 (0,42)	1,25 (0,44)	1,19 (0,34)	1,28 (0,51)	1,22 (0,40)	0,58 (0,13)
Totaal	1,30 (0,35)	1,35 (0,44)	1,28 (0,48)	1,33 (0,57)	1,29 (0,43)	1,32 (0,46)	0,60 (0,16)

TABEL 7 Genre- en periodeverschillen wat betreft argumentoverlap. Het percentage inhoudswoorden in de tekst dat voorkomt in de 50 eraan voorafgaande woorden

Periode	Kranten politiek	Kranten cultuur	Kranten binnenland	Kranten buitenland	Kranten sport	Kranten totaal	Romans
1950/55	5,69 (2,21)	6,15 (2,47)	6,18 (1,96)	5,88 (2,05)	6,99 (2,52)	6,18 (2,25)	10,30 (1,28)
2002/ 2005	6,24 (2,40)	5,48 (3,31)	6,81 (2,31)	5,94 (2,19)	6,42 (2,29)	6,18 (2,52)	10,35 (1,86)
Totaal	5,96 (2,29)	5,82 (2,90)	6,49 (2,13)	5,91 (2,09)	6,71 (2,39)	6,18 (2,38)	10,33 (1,57)

We kunnen de informatiedichtheid ook op hoger niveau bekijken, namelijk met de argumentoverlap. Die maat verwijst naar het aantal inhoudswoorden dat ook voorkomt in de voorgaande 50 woorden. Hoe lager dat aantal, hoe informatiever de tekst is. Kranten blijken veel informatiever te zijn dan romans ($F[1,236]=110.46, p<.001, \eta^2=.319$). In kranten worden 40% minder vaak woorden herhaald uit de voorgaande 50 woorden dan in romans. Dit bevestigt het beeld dat krantenberichten meer nieuwe referenten introduceren, die een meer

uitgebreide formulering behoeven. Ook op het niveau van de samenhang, met name de informatiedichtheid, is de krant dus complexer dan de roman.

Ten slotte kijken we naar de persoonsdichtheid van krantenberichten versus romans. We zouden dit kenmerk kunnen opvatten als een specifieke vorm van concreetheid. Er blijken in romans ongeveer 50% meer verwijzingen naar personen voor te komen dan in kranten ($F[1,236]=56.93$, $p<.001$, $\eta^2=.194$).

TABEL 8 Genre- en periodeverschillen wat betreft het percentage tekstwoorden dat naar een persoon verwijst: persoonlijke en bezittelijke voornaamwoorden, nomina die naar personen verwijzen en persoonsnamen

Periode	Kranten politiek	Kranten cultuur	Kranten binnenland	Kranten buitenland	Kranten sport	Kranten totaal	Romans
1950/55	6,49 (3,35)	10,21 (3,07)	6,93 (3,83)	9,08 (3,82)	9,93 (3,06)	8,53 (3,71)	12,53 (1,82)
2002/ 2005	9,54 (3,39)	9,15 (4,80)	8,80 (3,74)	7,73 (2,49)	9,67 (3,53)	8,98 (3,59)	13,93 (1,84)
Totaal	8,02 (3,59)	9,61 (4,01)	7,87 (3,84)	8,41 (2,18)	9,80 (3,26)	8,75 (3,65)	13,23 (1,94)

We concluderen dat romans qua taalgebruik aanzienlijker eenvoudiger lijken te zijn dan krantenberichten wat betreft de bekendheid van de woorden, de abstractheid van de gebruikte concepten, de hoeveelheid informatie per naamwoordgroep, de bijvoeglijke bepalingen, de zinslengte en het aantal persoonsverwijzingen.

5 Wordt het taalgebruik in kranten eenvoudiger?

Er zijn verschillen tussen romans en kranten, maar zien we die verschillen ook in de tijd? We bespreken die vraag eerst voor de kranten en daarna voor de romans. Bij de analyse van de kranten gaan we ervan uit dat vrijwel elk fragment door een andere auteur is geschreven. We hebben daar geen zekerheid over, omdat meestal geen auteursnaam bij het fragment vermeld wordt. Daarom werken we uitsluitend met de variabelen periode, krant en rubriek. Er zijn verschillen tussen rubrieken, zo hebben we hierboven gezien, maar daarop gaan we hieronder niet in. De factor rubriek is alleen als controlevariabele in de variantieanalyses opgenomen. In deze sectie staat immers de vraag centraal of het taalgebruik in de twee periodes eenvoudiger is geworden.

Het antwoord op die vraag is bevestigend. De kranten van de jaren 1950 en 2002 verschillen aanzienlijk in complexiteit (zie Tabel 9). Zo zien we in de kranten verschillen per periode wat betreft woordfrequentie, algemene nomina en nominalisaties. Het verschil in woordfrequentie zou misschien kunnen worden toegeschreven aan het feit dat onze woordfrequentielijst stamt uit de eenentwintigste eeuw en niet uit 1950. Maar de verschillen betreffen ook een semantische eigenschap van de zelfstandige naamwoorden, namelijk hun abstractheid. We zagen geen verschillen op concrete adjectieven en algemene werkwoorden.

TABEL 9 Periode- en krantenverschillen in het krantencorpus; alleen de variabelen met significante verschillen zijn weergegeven. * = $p < 0,05$, ** = $p < 0,01$

	F periode/ krant	F inter- actie	Gem. 1950 NRC (SE)	Gem. 2002 NRC (SE)	Gem. 1950Tele-graaf (SE)	Gem. 2002Telegraaf (SE)	Effect-grootte (eta ²)
<i>Lexicale kenmerken</i>							
Woord- frequentie	$F^p[1,180]=$ 12,89***	$F[1,180]=$ 4,53*	4,11 (0,31)	4,15 (0,31)	4,08 (0,31)	4,25 (0,31)	periode 0,067
Concrete nomina	$F^k[1,180]=$ 8,51**		0,28 (0,02)	0,29 (0,02)	0,32 (0,02)	0,34 (0,02)	krant 0,045
Algemene nomina	$F^p[1,180]=$ 13,60***		0,14 (0,01)	0,11 (0,01)	0,13 (0,01)	0,10 (0,01)	periode 0,070
Concrete werkwoorden		$F[1,180]=$ 4,98*	0,022 (0,01)	0,037 (0,01)	0,042 (0,01)	0,034 (0,01)	
<i>Zinsbouwkenmerken</i>							
Zinslengte	$F^p[1,180]=$ 54,05***		25,44 (0,79)	18,64 (0,79)	24,49 (0,79)	19,67 (0,79)	periode 0,231
Finiete bijzin- nen per zin	$F^p[1,180]=$ 16,65***		0,84 (0,05)	0,63 (0,79)	0,88 (0,79)	0,69 (0,79)	periode 0,085
Nevenge- schikte deel- zinnen per zin	$F^p[1,180]=$ 12,17***		1,29 (0,03)	1,18 (0,03)	1,13 (0,03)	1,13 (0,03)	periode 0,061
Deelszins- lengte	$F^p[1,180]=$ 16,83***		11,88 (0,27)	10,55 (0,27)	11,63 (0,27)	10,74 (0,27)	periode 0,085
Attributieve adjectieven/ deelzin	$F^p[1,180]=$ 6,03*		0,61 (0,03)	0,54 (0,03)	0,61 (0,03)	0,55 (0,03)	periode 0,032
Andere bijv. bepalingen/ deelzin	$F^p[1,180]=$ 8,26**		1,43 (0,07)	1,26 (0,07)	1,38 (0,07)	1,17 (0,07)	periode 0,044
Opsommin- gen per deelzin	$F^p[1,180]=$ 21,66***		0,32 (0,02)	0,21 (0,02)	0,30 (0,02)	0,21 (0,02)	periode 0,107
Maximale afhankelijk- heidslengte	$F^p[1,180]=$ 29,26***		9,77 (0,30)	8,85 (0,30)	9,80 (0,30)	9,53 (0,30)	periode 0,140
<i>Cohesie</i>							
mtld	$F^k[1,180]=$ 4,97*		106,9 (3,92)	114,9 (3,92)	120,0 (3,92)	119,2 (3,92)	krant 0,027
<i>Persoonlijke stijl</i>							
Persoonlijke nomina	$F^k[1,180]=$ 5,87*		0,21 (0,01)	0,22 (0,01)	0,24 (0,01)	0,25 (0,01)	krant 0,032

Wat betreft de syntactische complexiteit zijn er periodeverschillen op alle geanalyseerde variabelen. De lengte van de zinnen in 1950 is bij beide kranten groter dan in 2002. Dat komt niet alleen door meer nevenschikkingen van deelzinnen, maar ook door een groter aantal bijzinnen. We mogen stellen dat bijzinnen meer complexiteit bevatten voor de verwerker dan nevensgeschikte deelzinnen. Maar er zijn bij kranten ook periodeverschillen op het lagere niveau van de deelzin, al zijn deze verschillen wat kleiner. Rond elke persoonsvorm staan in 1950 meer woorden, die met name afkomstig lijken van bijvoeglijke bepalingen, en in mindere mate van opsommingen van zinsdelen.

Er zijn geen periodeverschillen wat betreft de samenhang en de persoonlijke stijl.

Zien we verschillen in de tijd voor *NRC (Handelsblad)* en *De Telegraaf* afzonderlijk? Ja, en daarbij wordt hier en daar de verwachting bevestigd dat berichten in een populair dagblad minder complex zijn dan in een 'kwaliteitskrant'. Hoewel de krantenverschillen kleiner zijn dan de periodeverschillen, vinden we dat *NRC (Handelsblad)* minder concreet en persoonlijk is in zijn nomina dan *De Telegraaf*; daarnaast heeft *NRC* een lagere MTLD dan *De Telegraaf*, wat wil zeggen dat er minder woorden herhaald worden. Wat betreft woordfrequentie en concreetheid van werkwoorden vinden we interacties die een gemengd beeld laten zien; we gaan daarop niet verder in.

6 Wordt het taalgebruik in romans eenvoudiger?

We richten ons nu op periodeverschillen in de romans. In de romandata zijn drie niveaus te onderscheiden: periode (3), roman/auteur (20 per periode) en fragment (15 per roman). Bij de analyse van periodeverschillen moeten we dus rekening houden met afhankelijkheden in de observaties: er zijn immers telkens 15 observaties gedaan in dezelfde roman. Als we dat over het hoofd zien, doen we net alsof we 900 romans hebben onderzocht en onderschatten we de standaardfouten rondom de waargenomen waarden. Om recht te doen aan die afhankelijkheden, voeren we een multilevelanalyse uit.³ Die analyse heeft als bijkomend voordeel dat we kunnen nagaan welk deel van de totale variantie is toe te schrijven aan verschillen tussen individuele romans (en niet tussen periode). Een maat voor dat deel is de 'intra-class-correlation' (ICC). Hoe hoger de ICC, hoe sterker een bepaald kenmerk hoort bij een bepaalde roman. Zie Tabel 10 voor de resultaten.

Uit onze analyse blijkt dat het taalgebruik in romans op vier kenmerken verandert in de geanalyseerde periodes: (1) zinslengte en (2) het aantal nevensgeschikte deelzinnen per zin; (3) het aantal attributieve adjectieven per deelzin; en (4) de proportie van persoonlijke nomina.⁴ We vonden geen verschillen op de lexicale kenmerken en op samenhang.

3 Dat doen we met periode als 'fixed' factor en roman als 'random' factor. De vraag is dan telkens of de analyse met de factor periode een beter model oplevert dan het model met alleen de random factor roman.

4 We hebben deze analyse gecontroleerd met een analyse op geaggregeerde scores per roman; de dataset werd dus op die manier ingekrompen van bijna 900 naar 60 cases. We vonden significante verschillen voor dezelfde vier factoren.

TABEL 10 Intra-class correlaties en periodeverschillen in het romancorpus op basis van een analyse per fragment; * = $p < .05$, ** = $p < .01$

	ICC factor roman	F voor factor periode (indien significant)	Gem. 1955 (SE)	Gem. 1980 (SE)	Gem. 2005 (SE)	Effect-grootte (R^2)
<i>Syntactische kenmerken</i>						
Zinslengte in woorden	,258	$F[2,56.94]=4,49^*$	20,53 ^a (1,18)	18,41 (1,18)	15,54 ^{b*} (1,18)	0,041 (0,136) ⁺
Aantal nevenschikte deelzinnen per zin	,157	$F[2,56.36]=8,49^{**}$	1,55 ^a (0,62)	1,34 (0,62)	1,20 ^{b**} (0,62)	0,044 (0,228) ⁺
Aantal attributieve adjectieven per deelzin	,162	$F[2,57.17]=3,85^*$	0,338 ^a (0,18)	0,300 (0,18)	0,267 ^{b*} (0,18)	0,025 (0,120) ⁺
<i>Persoonlijke stijl</i>						
Persoonlijke nomina (dichtheid)	,197	$F[2,57.03]=3,40^*$	0,152 ^a	0,177	0,195 ^{b*}	0,026 (0,107)

Toelichting bij de superscripten: een gemiddelde met superscript a verschilt significant van een gemiddelde met superscript b. Een gemiddelde zonder superscript verschilt niet van de andere gemiddelden. Bij de effectgrootte geeft een + de waarde voor een unilevel-analyse met een gemiddelde per roman.

Tabel 10 laat die verschillen zien plus de effectgroottes.⁵ Nergens wordt meer dan 5% van de variantie verklaard door deze vier kenmerken.⁶ De gemiddelden lopen geleidelijk op in de drie periodes; dat wil zeggen dat de trend doorloopt en geen breuk vertoont. De paarsgewijze verschillen tussen de drie periodes onderling zijn getest met een Bonferroni-post hoc-analyse. Het resultaat daarvan is in Tabel 10 aangegeven door superscripten bij de gemiddelden. We vinden alleen significante verschillen tussen de twee uiteinden van de ontwikkeling (1955 en 2005).

Onze conclusie over het taalgebruik in de romans is dat tussen 1955 en 2005 de zinnen iets minder complex worden: ze worden korter, bevatten minder nevenschikte deelzinnen (maar niet minder bijzinnen) en minder attributieve adjectieven. Maar er waren ook syntactische kenmerken waarop we geen verschillen vonden: finiete bijzinnen per zin, deelzinslengte, andere bepalingen per deelzin, opsommingen per deelzin en de maximale afhankelijkheidslengte. Daarnaast worden de teksten iets persoonlijker: ze bevatten immers meer persoonlijke naamwoorden. Deze verschillen zijn significant maar substantieel kleiner dan de ver-

5 Voor de multilevelanalyse zijn de effectgroottes berekend door het kwadraat te nemen van de correlatie tussen enerzijds de feitelijke scores en anderzijds de predicties van de modellen met alleen de fixed factor.

6 Effectgroottes in multilevelanalyses zijn altijd klein omdat veel variantie op het onderste niveau (dat van de fragmenten per roman) onverklaard blijft. Daarom melden we ook de effectgroottes voor de analyses van de geaggregeerde scores per roman. Die zijn tussen haakjes vermeld. Deze effectgroottes zijn hoger, waarbij opvalt dat dit vooral geldt voor het verschil in aantallen nevenschikte deelzinnen.

schillen in de kranten. Er zijn immers minder kenmerken waarop we verschillen vonden en de kenmerken waarop we wel verschillen vonden zijn minder bepalend voor de begrijpelijkheid van de tekst.

7 Een nadere blik op het romancorpus

We hebben gezien dat er weliswaar enkele verschillen zijn in het taalgebruik van romans tussen de periodes 1955-2005, maar de verschillen tussen romans binnen een periode lijken substantiëler. In het vervolg willen we aan de hand van voorbeelden laten zien dat die geringe verschillen tussen periodes slechts gedeeltelijk te interpreteren zijn als een bewijs voor de stelling dat het taalgebruik in romans in de loop van de tijd steeds eenvoudiger wordt. We verkennen een aantal patronen in het gebruik van enkele stilistische verschijnselen, die misschien niet zozeer met de periode verbonden zijn maar met andere literaire stijlverschijnselen. De fragmenten zijn gekozen op grond van opvallende scores in de resultaten van de analyse met T-scan. Ze lijken interessant voor verder onderzoek.

In de oudere romanfragmenten vonden we gemiddeld langere zinnen. Dat leek het gevolg van nevenschikking van deelzinnen. In Josepha Mendels' *Heimwee naar Haarlem* staat een fragment met lange zinnen en een hoog aantal persoonsvormen per zin.

Het water onder de planken van de kettingbrug staat hoog en als Roberta er vlug overheen loopt en het hout onder haar voeten kleppert, spat het tot om haar oren. Ze droogt zich af met haar zakdoek en kleedt zich uit, ze draagt nu alleen nog maar een hemd en broekje, niemand ziet haar, de kinderen zijn op school en volwassenen komen hier weinig. Ze wil nog eens proberen of ze er niet goed overheen kan komen. Want waarom zal zij niet die uitzondering zijn waarover juffrouw Bloemstra zo veel jaren geleden al gesproken heeft? Iemand moet dat toch worden. Maar ze is vlug moe en geeft het op, ze is niet meer gewend zich in te spannen en gaat in het gras zitten. Ze heeft honger, het is misschien al theetijd, thee in haar roze kopje, een roze servet op haar knieën (Mendels 1983 [1958]: 40).

Het fragment illustreert dat de langere zinnen in de eerste periode veelal nevenschikkend van aard zijn. We betwijfelen of de tekst op deze manier veel complexer wordt: een nevenschikking met komma's stelt weinig eisen aan de lezer. Het lijkt erop dat auteurs in 1955 komma's gebruiken om gerelateerde predicaties in los verband aaneen te rijgen. Die gedachte wordt ondersteund door Parkes; hij laat zien dat het afbakenen van de zinsgrens al heel lang aan verandering onderhevig is (Parkes 2016).

In de drie fragmenten die het hoogst scoren op nevenschikking lijkt iets anders aan de hand te zijn. Een fragment van één zeer lange zin uit Mulisch' *Het stenen bruidsbed* beschrijft de traumatische herinnering aan een gebeurtenis in de Tweede Wereldoorlog: het bombardement op Dresden (Mulisch 1968 [1959]: 190-191). Mulisch lijkt deze stijl vooral te gebruiken om de heftigheid van de oorlogsherinnering van een tegelijk surrealistische en lyrische kleur te voorzien. *Ik Jan Cremer* (1964) en Jeroen Brouwers' *Bezonen rood* (1981: 107-108) gebruiken eveneens lange zinnen vol nevenschikte bijzinnen om een oorlogssituatie te beschrijven.

Elektriciteit werd vervangen door carbid, aardappels werden in de schil gegeten, porselein knalde kapot in bombardementen, de boeren hadden vreten genoeg, de burens zaten in Drente op de hei, de moffen zongen Erika en in beide kampen zong men Lili Marlene, de Amerikanen vielen als mussen van het Normandische dak, de Nederlandse vloot werd in de zeebodem gestampt, onze grammofoonplaten vielen kapot, brandbommen vielen bij de fietsenmaker in huis, op het ziekenhuis werden grote rooie kruisen geschilderd, kolen worden gegapt, moffen worden verzopen, mitrailleurs ratelen, (geen Nederlandse), de *Telegraaf* blijft de hele oorlog door uitkomen, in de katholieke kerken wordt gebeden om de Russen te laten komen, zes miljoen joden worden vermoord (waaronder Anne Frank) en veel joden vluchten naar Amerika [...] (Cremer 1976 [1964]: 9).

Hier lijkt de aaneenschakeling van nevenschikkingen geen symptoom van een schrijfstijl in een bepaalde periode, maar een bewust ingezet stilistisch middel om een effect van chaos en verwarring bij de lezer te bewerkstelligen. Hoewel de afzonderlijke zinsdelen niet erg complex zijn, doet het geheel wel een groter beroep op het uithoudingsvermogen van de lezer.

Een ander patroon dat samenhangt met zinslengte, zien we in het gebruik van attributieve adjectieven. Ook hierbij is het de vraag of dat de teksten uit die periode extra complex maakt: attributieve adjectieven maken de tekst wat informatiedichter en de zin iets langer, maar ze lijken niet bijzonder lastig te verwerken voor een lezer. Toch is het interessant om wat langer bij dit verschijnsel stil te staan, in het bijzonder om te bepalen of een groot aantal adjectieven misschien vooral voorkomt in een bepaald type fragmenten. Neem Hans Berghuis' *Drie vrouwen*:

Stel je deze *kleine* geschiedenis voor: de zon stond vanmiddag laag boven de waaierpalmes aan de *zuidwestelijke* kim. Geen vleugje wind speelde in de *groene* bamboehagen aan de rand van het dorp. Ik stond op de veranda van de *geïsoleerde* barak. In het *natte* rijstveld tussen de *verre* palmen en de *nabije* bamboestruiken kwebbelden vreedzaam een menigte eenden. Zij liepen elkaar na, wroetten in het *ondiepe* water en scheurden rijsthalmen los uit de *weke* bodem. De rijst heeft gebloeid en over tien dagen zullen onze mensen met de oogst beginnen. Twee *kleine bruine* vrouwen stapten vanaf de *smalle* strook *droge* grond tussen de velden de *natte* akker in. Met *lange* stokken trachtten zij de eenden bijeen te drijven. Maar de *weerspannige* vogels ontweken de *langzame* bewegingen van de *Tonkinese* boerinnen [...] (Berghuis 1956: 110).⁷

Het lijkt erop dat zo'n hoge dichtheid aan adjectieven vooral te vinden is in passages die een omgeving beschrijven: we zagen dat bijvoorbeeld in een aantal opvallend adjectief-dichte passages in romans uit de periode 1955 (naast Berghuis' genoemde boek bijvoorbeeld Oscar Timmers' *Landklimaat* (Timmers 1957) en Tip Maruggs *Weekendpelgrimage* (Marugg 1958)), maar ook in romans uit de andere twee periodes (bijvoorbeeld Marion Bloems *Geen gewoon Indisch meisje* (Bloem 1983) en Nelleke Noordervliets *Millemorti* (Noordervliet 1989)). Opvallend is dat het daarbij vaak om niet-Nederlandse, of preciezer om niet-westerse werelden gaat. Dat zou een aanwijzing kunnen zijn dat auteurs de neiging hebben niet-westerse of koloniale werelden met veel stilistische omhaal te beschrijven, zoals al eerder is gesuggereerd (Meijer 1995: 129-130).

7 De cursiveringen in dit fragment en de volgende zijn door ons aangebracht.

Een juist heel zuinig gebruik van attributieve adjectieven treffen we aan in passages die handelingen of gedachtestromen beschrijven. Het volgende fragment van Vonne van der Meer (*De reis naar het kind*) is daar een voorbeeld van:

Ze wist niet of Max op dat moment de beslissing genomen had, of later toen hij alleen was en zijn vaders woorden door zijn hoofd maalden. Misschien had hij wel helemaal geen beslissing over Pablo genomen. Had hij alleen maar gedacht: zoals mijn vader wil ik niet worden.

Ze durfde niet op een verzoening te hopen. Toen Pablo 's avonds uit de badkamer riep dat hij hulp nodig had en Max als eerste opstond, was ze achter hem aan naar boven gelopen. Het idee dat de twee mannen, samen, zonder haar, in een ruimte zouden vertoeven maakte haar onrustig. Wat ze door de *half openstaande* badkamerdeur zag, deed haar blozen en haar argwaan verdwijnen. Ze zag Max een gebaar maken dat ze vrienden weleens had zien maken als ze hun kind uit bad haalden. Een gebaar dat haar altijd had getroffen als allertzorgzaamst. Ze zag hoe hij eerst zijn handen door het *warme* badwater haalde voor hij ze naar Pablo uitstrekke (Van der Meer 1989: 142-143).

Het patroon dat we hier vermoeden is dat een intensief gebruik van deze adjectieven samengaat met een specifieke functie: de lezers een beeld tonen van een omgeving waar ze niet mee vertrouwd zijn. Terwijl de afwezigheid van die adjectieven duidt op een heel andere functie: de lezers meenemen in de gedachtegang van een personage. De vermeende 'kaalheid' van het Nederlandse proza, onder meer opgemerkt door een internationale literatuurcriticus als Parks (2011), wordt wellicht deels opgewekt of versterkt door de neiging om scènes die zich afspelen in Nederland in een sober idioom te vatten. In ons corpus zagen we dat in een aantal romans die in een typisch Nederlandse context spelen, zoals *Een winterreis* (Brakman 1961) en *Boven is het stil* (Bakker 2006).

Het derde patroon heeft te maken met de verwijzingen naar personen. Vooral de nominale verwijzingen nemen toe in 2005. Wat het effect is van die toename, is moeilijk te bepalen. Uiteraard komen persoonlijke nomina vooral veel voor in situaties waarin mensen en menselijke verhoudingen beschreven worden. Een grote dichtheid aan zulke nomina vinden we bijvoorbeeld in dit fragment uit Kader Abdolahs *De boodschapper*:

Als ik vragen stelde over de kindertijd van Mohammad, begon men vaak over een ontmoeting tussen de jonge Mohammad en de oude *priester* Bahira.

Het is eigenlijk een verhaal dat ooit door de *aanhangers* van Mohammad is verteld en dat hier en daar is aangedikt.

Maar de *vijanden* van Mohammad en vooral de *joden* noemen het een leugen, een hersenschim bedoeld om Mohammad op een lijn te plaatsen met de oude *profeten* Jezus, Mozes en Ibrahim.

De *joden* mogen gelijk hebben en het verhaal kan verzonnen zijn, maar de *priester* Bahira en zijn klooster waren geen verzinsels. [...]

Onderweg trof ik een oude karavanserai naast een groot kruispunt, waar *handelaren* uit andere steden met hun kamelen langskwamen om hun goederen te verkopen (Abdolah 2008: 28).

Het wemelt hier van de persoonlijke nomina ('priester', 'aanhangers', 'vijanden',

‘joden’, ‘handelaren’), die vooral gebruikt worden om een historische context te schetsen. De dichtheid aan namen en nomina maakt de tekst hier niet per se persoonlijker, eerder iets lastiger te begrijpen: de lezer moet, denken we, enige moeite doen om de verhouding tussen de personen te begrijpen. Dat zit anders bij het volgende fragment van Boudewijn Büch (*De kleine blonde dood*): ook dat scoort hoog op persoonlijke nomina, maar omdat die veelal in een overzichtelijke opsomming bij elkaar staan of verwijzen naar zeer concrete personen (‘vader’, ‘burgemeester’, ‘genodigden’), zal de lezer weinig moeite hebben de tekst te volgen:

Totdat hij verscheen met zijn korps, was de optocht een dorps zootje. De *welpen* en *verkenners* slenterden, MIOD liep uit de pas en DoMuToJe speelde vals. De *jongens* en *meisjes* van de turnvereniging waren zo gebiologeerd door het gezelschap dat ze uit de rij liepen en bijna tegen de marmeren leeuwen die voor het bordes stonden op botsten. [...] Mijn *vader* gaf weer een schreeuw en het gezelschap keerde zich naar de *burgemeester* en zijn *genodigden* (Büch 2002 [1985]: 66-67).

De ‘kleuring’ van persoonlijke nomina kan ingrijpend verschillen, zo tonen de twee geciteerde fragmenten aan: woorden als ‘vader’ en ‘burgemeester’ zijn vrij neutraal, terwijl de woordgroep ‘de vijanden van Mohammad en vooral de joden’ sterk ideologisch gekleurd is. Soms kan die kleuring iets zeggen over de ideologische visie van de verteller of de (impliciete) auteur van het boek, soms ook over de visie van het personage. Zo bezieet een personage uit P.F. Thoméses *Vladiwostok!* zijn omgeving vooral met gebruik van seksistisch en beledigend gekleurde nomina: ‘fucker’, ‘rukkers’, ‘smeerkezen’ (Thomése 2007: 193).

8 Conclusie en discussie

Het taalgebruik in kranten is in allerlei opzichten complexer dan de taal van de roman. Romans bevatten eenvoudiger en concretere woorden, meer verwijzingen naar personen, de teksten zijn minder informatiedicht, zowel binnen de zinnen als over de zinsgrenzen heen. Er is een duidelijke verandering zichtbaar in het taalgebruik van de kranten in de tweede helft van de twintigste eeuw. Het taalgebruik wordt minder complex: een breed scala van kenmerken verandert, zowel op het niveau van de woorden als op dat van de zinnen. In romans is de afname in complexiteit veel minder duidelijk. De verschillen die we daar vinden, lijken niet van grote invloed op de complexiteit. De grotere hoeveelheid nevenschikte deelzinnen maakt oudere teksten niet erg moeilijker. Het toevoegen van adjectieven is ook geen aanslag op het werkgeheugen van de lezer. En het toevoegen van persoonlijke naamwoorden maakt een tekst niet veel makkelijker, aangezien het totale aantal verwijzingen naar personen toch al hoog was.

Dit wil niet zeggen dat deze verschillen literair oninteressant zijn. Met name het nevenschikken van deelzinnen en het gebruik van adjectieven kunnen we zien als literaire procedés waarvan het gebruik kennelijk is afgenomen. Vooral de afname in het gebruik van adjectieven is interessant, omdat dit stijlmiddel vooral lijkt te worden gebruikt in (natuur)beschrijvingen en in niet-westerse, althans niet-Nederlandse contexten. Dat zou te maken kunnen hebben met een achteraf bezien niet gebalanceerde corpuselectie van romans uit de vroegste periode: wellicht

hebben wij precies die romans eruit gepikt die in zo'n niet-Nederlandse context spelen. Maar het is ook denkbaar dat het beschrijven van niet-Nederlandse werelden zeldzamer wordt in de moderne(re) roman. Onderzoek van Van der Deijl e.a. (2016) heeft aangetoond dat een groot corpus Nederlandse romans van rond 2013 weinig divers is op het gebied van de representatie van personages: 'Dominant identities are the male, young, highly educated and Western ones. Less prevailing in recent Dutch literature are elder, lowly educated, female and/or non-Western individuals' (Van der Deijl 2016: 37). De onderzoekers vermoeden dat hedendaagse auteurs romanwerelden lijken te representeren die dichtbij hun eigen leefwereld liggen.

We kunnen dus, afgaande op onze analyse, niet stellen dat het taalgebruik in romans eenvoudiger geworden is, terwijl dat voor kranten duidelijk wél het geval is. Dat recensenten en cultuurcritici klagen over de stilistische armoede van de Nederlandse literatuur heeft misschien meer te maken met ergernis over een vernauwende thematische horizon in de Nederlandse literatuur dan met een versimpeling van het taalgebruik. Wellicht is er sprake van een versimpeling in andere literaire middelen dan de stilistische complexiteit die wij hebben gemeten – waarbij dan wel opmerkelijk is dat het juist de stijl is waarover in hedendaagse cultuurkritische reflecties geklaagd wordt.

Voor zover wij kunnen nagaan, is dit de eerste systematische analyse van verschillen in taalgebruik tussen genres en periodes met behulp van taaltechnologische instrumenten. Wij denken dat dit een veelbelovende lijn van onderzoek zou kunnen worden. In de eerste plaats omdat op deze manier onderzoek naar taalverandering goed langs strakke methodische lijnen uitgevoerd kan worden, waarbij gecontroleerd wordt voor toevallige verschillen tussen fragmenten binnen een grotere tekst, voor verschillen tussen auteurs en voor verschillen tussen genres. In de tweede plaats lijkt genrevergelijking op een bredere schaal veelbelovend. Zo zouden we op eenzelfde manier onderzoek kunnen uitvoeren naar verschillen in taalgebruik tussen subgenres, zoals bijvoorbeeld de landelijke journalistiek en de regionale journalistiek, of tussen de stijlen van een auteur binnen een oeuvre (de jonge en de oudere Hermans). In de derde plaats biedt deze analysemethode de kans om binnen een genre, zoals de literaire roman, opvallende stilistische patronen zichtbaar te maken en die vervolgens met meer heuristische methoden te interpreteren. Is er inderdaad een relatie tussen een intensiever gebruik van adjectieven en de beschrijving van koloniale landschappen? Is de techniek van de aaneenschakelende nevenschikkingen inderdaad verbonden met de expressie van ingrijpende ervaringen die verwarring bij personages reflecteren? Zijn de verschillen tussen dialogen en beschrijvende fragmenten in literaire teksten terug te vinden in de verschillen tussen spreektaal en schrijftaal buiten de roman? Ten slotte biedt deze methode een nieuw perspectief op het onderzoek naar regionale en sociale taalvariatie. Wat zijn de kenmerkende verschillen tussen de moderne Vlaamse en de Nederlandse roman of tussen het literaire werk van nieuwe Nederlanders en hun moedertalige collega's, tussen streekromans en de werken van nationale prijswinnaars in het literaire circuit? Studenten en onderzoekers worden van harte uitgenodigd om met het hier beschreven instrumentarium zulke vragen op te pakken.

Bibliografie

- Anoniem 2013 – ‘De versimpeling van onze taal’. In: *Trouw*, 19 januari 2013; via <http://www.trouw.nl/tr/nl/5009/Archief/article/detail/3379710/2013/01/19/De-versimpeling-van-onze-taal.dhtml>.
- Abdolah 2008 – K. Abdolah, *De boodschapper. Een vertelling*. Breda: De Geus, 2008.
- Bakker 2006 – G. Bakker, *Boven is het stil*. Amsterdam: Cossee, 2006.
- Bakker 2017 – P. Bakker, ‘Oplage 2016: Telegraaf grote verliezer, Trouw en FD bijna stabiel’. In: *Stimuleringsfonds voor de Journalistiek*, 30 maart 2017; via <https://www.svdj.nl/de-stand-van-de-nieuwsmedia/oplage-2016-telegraaf-grote-verliezer-trouw-en-fd-bijna-stabiel/>.
- Baricco 2010 – A. Baricco, *De barbaren*. Vertaald door Manon Smits. Amsterdam: De Bezige Bij, 2010 [2006].
- Berghuis 1956 – H. Berghuis, *Drie vrouwen. Roman*. Leiden: A.W. Sijthoff, 1956.
- Bloem 1983 – M. Bloem, *Geen gewoon Indisch meisje*. Haarlem: In de Knipscheer, 1983.
- Van Boven & Kemperink 2006 – E. van Boven & M. Kemperink, *Literatuur van de Moderne Tijd. Nederlandse en Vlaamse letterkunde in de 19^e en 20^e eeuw*. Bussum: Coutinho, 2006.
- Brakman 1961 – W. Brakman, *Een winterreis*. Amsterdam: Querido, 1961.
- Brems 2006 – H. Brems, *Altijd weer vogels die nesten beginnen. Geschiedenis van de Nederlandse literatuur 1945-2005*. Amsterdam: Bert Bakker, 2006.
- Brouwers 1981 – J. Brouwers, *Bezonnen rood*. Amsterdam: De Arbeiderspers, 1981.
- Büch 2002 – B. Büch, *De kleine blonde dood. Roman*. Utrecht/Amsterdam/Antwerpen: De Arbeiderspers, 2002 [1985].
- CBS 2017 – CBS, *De sociale staat van Nederland 2017*. Den Haag, 2017.
- Cremer 1976 – J. Cremer, *Ik Jan Cremer*. Amsterdam: De Bezige Bij, 1976 [1964].
- Van der Deijl e.a. 2016 – L. van der Deijl e.a., ‘Mapping the Demographic Landscape of Characters in Recent Dutch Prose: A Quantitative Approach to Literary Representation’. In: *Journal of Dutch Literature* 7 (2016) 1, p. 20-42.
- Februari 2016 – M. Februari, ‘Neoliberalisme bedreigt de veelvormigheid van taal’. In: *NRC Handelsblad*, 7 juni 2016.
- Hinds 1987 – J. Hinds, ‘Reader versus writer responsibility: A new typology’. In: U. Connor & R.B. Kaplan (ed.), *Writing across languages: Analysis of L2 text*. Reading: Addison-Wesley, 1987, p. 141-152.
- Van Houwelingen 2017 – B. van Houwelingen, ‘Tam verslag van emotionele reis’. In: *de Volkskrant*, 4 november 2017.
- Huysmans & De Haan 2010 – F. Huysmans & J. de Haan, *Alle kanalen staan open: De digitalisering van mediagebruik*. Sociaal en Cultureel Planbureau, 2010.
- Kester 2007 – B. Kester, ‘Windows on the world: The presentation of international news, 1880-1980’. In: M. Boersma (ed.), *Form and Style in Journalism. European newspapers and the representation of news, 1880-2005*. Leuven: Peeters Publishers, 2007, p. 115-132.
- Kleijn 2018 – S. Kleijn, *Clozing in on readability: How linguistic features affect and predict text comprehension and on-line processing*. Proefschrift Universiteit Utrecht, 2018.
- Kubota 1999 – R. Kubota, ‘Japanese Culture Constructed by Discourses. Implications for Applied Linguistics Research and ELT’. In: *Tesol Quarterly* 33 (1999) 1, p. 9-35.
- kvb Boekwerk 2017 – ‘Monitor 2017 – de markt’. In: kvb Boekwerk, z.d.; via: <https://www.kvbboekwerk.nl/monitor/markt/monitor-2017>.
- kvb Boekwerk – ‘Monitor 2016 – Lezers’. In: kvb Boekwerk, z.d.; via: <https://www.kvbboekwerk.nl/monitor/lezers/monitor-2016>.
- Leesmonitor – Leesmonitor, ‘Leestijd’. In: leesmonitor.nu, z.d.; via <https://www.leesmonitor.nu/nl/leestijd#de-leestijd-daalt-behalve-van-online-nieuws>.
- Marugg 1958 – T. Marugg, *Weekendpelgrimage*. Amsterdam: De Bezige Bij, 1958.
- Van der Meer 1989 – V. van der Meer, *De reis naar het kind. Een vertelling*. Amsterdam: De Bezige Bij, 1989.
- Mendels 1983 – J. Mendels, *Heimwee naar Haarlem. Roman*. Amsterdam: Meulenhoff, 1983 [1958].
- Meijer 1995 – M. Meijer, ‘Witheid in de literaire verbeelding. *Rubber* (1931) van Madelon Székely-Lulofs in het licht van *Playing in the Dark* (1992) van Toni Morrison’. In: *Forum der Letteren* 36 (1995), p. 121-139.
- Mulisch 1968 – H. Mulisch, *Het stenen bruidsbed*. Amsterdam: De Bezige Bij, 1968 [1959].
- Noordervliet 1989 – N. Noordervliet, *Millemorti. Roman*. Amsterdam: Meulenhoff, 1989.

- Pander Maat e.a. 2014 – H. Pander Maat e.a., ‘T-Scan: a new tool for analyzing Dutch text’. In: *Computational Linguistics in The Netherlands Journal* 4 (2014), p. 53-74.
- Pander Maat & Dekker 2016 – H. Pander Maat & N. Dekker, ‘Tekstgenres analyseren op lexicale complexiteit met T-Scan’. In: *Tijdschrift voor Taalbeheersing* 38 (2016) 3, p. 263-304.
- Pander Maat 2017 – H. Pander Maat, ‘Zinslengte en zinscomplexiteit. Een corpusbenadering’. In: *Tijdschrift voor Taalbeheersing* 39 (2017) 3, p. 297-328.
- Parks 2011 – T. Parks, ‘The Dutch are coming. An essay on Dutch fiction and the position of Dutch literature in the process of globalization’. In: *New York Review of Books*, 29 november 2011.
- Parkes 2016 – M. Parkes, *Pause and Effect: an Introduction to the History of Punctuation in the West*. Routledge, 2016.
- Peper 2012 – R. Peper, ‘Ben je mal!’ In: R. Peper, *Fantoompoezen*. Amsterdam/Antwerpen: Querido, 2012, p. 76-79.
- Peters 2015 – A. Peters, ‘Kinderlijke stijloefeningen vermoeien de lezer. Vermoeiende stijloefeningen’. In: *de Volkskrant*, 28 november 2015; via <http://www.volkskrant.nl/recensies/kinderlijke-stijloefeningen-vermoeien-de-lezer~a4195684/>.
- Quereshi 1991 – M.Y. Quereshi, ‘Logical versus empirical readability and human interest scores: Refinement and update’. In: *Psychological reports* 68 (1991) 3, p. 1035-1040.
- Rosa 2005 – H. Rosa, *Beschleunigung. Die Veränderung der Zeitstrukturen in der Moderne*. Frankfurt Am Main: Suhrkamp Verlag, 2005.
- Schriver 2017 – K. Schriver, ‘Plain language gains momentum’. In: *IEEE Transactions on Professional Communication* 60 (2017) 4, p. 343-383.
- Štajner 2012 – S. Štajner & R. Mitkov, ‘Using comparable corpora to track diachronic and synchronic changes in lexical density and lexical richness’. In: *The 5th Workshop on Building and using Comparable Corpora* 17 (2012), p. 88-97.
- Storm 2011 – A. Storm, ‘In dit proza komt geen leven’. In: *Het Parool*, 2 maart 2011.
- Thomése 2008 – P.F. Thomése, *Vladivostok!* Amsterdam/Antwerpen: Contact, 2008.
- Timmers 1957 – O. Timmers, *Landklimaat*. Den Haag: Boucher, 1957.
- Vos 2008 – L.C.V. Vos, *Uitzondering op de regel. De positie van vrouwelijke auteurs in het naoorlogse Nederlandse literaire veld*. Onuitgegeven proefschrift Rijksuniversiteit Groningen, 2008.
- Vullings 2015a – J. Vullings, ‘Nog een halve eeuw en dan zijn we verdwenen’. In: *Vrij Nederland*, 22 juli 2015.
- Vullings 2015b – J. Vullings, ‘Pleidooi voor weerbaarheid’. In: *Vrij Nederland*, 19 december 2015.
- Wennekers, Van Troost & Wiegman 2016 – A.M. Wennekers, M.M. van Troost & P.R. Wiegman, *Media:Tijd* 2015. Amsterdam/Den Haag: NLO, SKO, BRO en SCP, 2016.
- Wennekers e.a. 2016 – A.M. Wennekers e.a., ‘Multitasking met media’. In: *Media: tekst in kaart*, 2016.
- Westin 2002 – I. Westin & C. Geisler, ‘A multi-dimensional study of diachronic variation in British newspaper editorials’. In: *Iceme Journal* 26 (2002), p. 133-152.

Correspondentieadres van de auteurs

Universiteit Utrecht
 Trans 10
 3512 JK Utrecht
 L.J.Ham@uu.nl

Er is een duidelijke verandering zichtbaar in het taalgebruik van de kranten in de tweede helft van de twintigste eeuw. Het taalgebruik wordt minder complex: een breed scala van kenmerken verandert, zowel op het niveau van de woorden als op dat van de zinnen. In romans is de afname in complexiteit veel minder duidelijk. De verschillen die we daar vinden, lijken niet van grote invloed op de complexiteit. [...] Dat recensenten en cultuurcritici klagen over de stilistische armoede van de Nederlandse literatuur heeft misschien meer te maken met ergernis over een vernauwende thematische horizon in de Nederlandse literatuur dan met een versimpeling van het taalgebruik.

