WAT GEBEURT ER IN HET NEDERLANDS?!

OVER TAAL, FREQUENTIE EN VARIATIE

Redactie
Nicoline van der Sijs, Lauren Fonteyn
en Marten van der Meulen

Sterck & De Vreese

© 2021 Nicoline van der Sijs, Lauren Fonteyn
en Marten van der Meulen | uitgeverij Sterck & De Vreese

Omslagontwerp Mijke Wondergem
Boekverzorging Elgraphic

ISBN 978 90 5615 803 3
NUR 620

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of op enige andere manier, zonder voorafgaande schriftelijke toestemming van uitgeverij Sterck & De Vreese, postbus 234, 8400 AE Gorredijk, Nederland –info@sterckendevreese.nl.

De uitgeverij heeft ernaar gestreefd alle copyrights van de in deze uitgave opgenomen illustraties te achterhalen. Aan hen die desondanks menen alsnog rechten te kunnen doen gelden, wordt verzocht contact op te nemen met Sterck & De Vreese.

Sterck & De Vreese is onderdeel van
20 leafdesdichten en in liet fan wanhoop bv

www.sterckendevreese.nl

INHOUDSOPGAVE

Voorwoord 9

Taalverandering 13
Leestekens: onaanzienlijk en ondergewaardeerd
Nicoline van der Sijs 15
Intermezzo Kranten toen, en kranten nu 21
Wat is er gebeurd sinds *luiden loeden* luidde? En waardoor?
Frans Hinskens 23
De kracht van sterke werkwoorden
Isabeau De Smet & Freek Van de Velde 30
Hulpwerkwoorden stapelen: toen en nu
Gerlof Bouma & Evie Coussé 36
Intermezzo Welke woorden verkleinen we het vaakst? 41
Al *dat* wat *wat* werd
Julie Van Ongeval & Rik Vosters 43
De toekomst van *die* en *dat*
Gosse Bouma 51

Grammatica (zinsbouw, woordvorming) 57
Welke werkwoorden willen *aan*?
Timothy Colleman 59
Intermezzo Wat zoeken we, en waar zoeken we naar? 65
De strijd tussen *jij wil* en *hij wilt*
Hans Bennis 67
Intermezzo We verlangen niet noodzakelijk naar wat we verlangen 72
Wat bijwoorden van graad ons leren kunnen
Jack Hoeksema 74

Veelzijdig onzijdig
Jenny Audring 79
Oerend hard en kapot cool
Roeland van Hout 85
Intermezzo Hartstikke, enorm, facking of toch gewoon heel? 90
Venlo, je regent en je bussen rijden niet
Helen de Hoop 92

Uitspraak 99
Wat rijmt er op *smart*?
Marc van Oostendorp 101
Vaak klinkt anders
Mirjam Ernestus 106
Waarom eh... zeggen we eigenlijk zo vaak 'eh'?
Meike de Boer 111
Intermezzo Welke woorden zijn het frequentst in het Nederlands? 116
De /n/ van Aboutaleb
Khalid Mourigh 118

Woorden en namen 123
Frequentie en efficiëntie in je dagelijks leven
Karlien Franco 125
Intermezzo Hoe frequent zijn de populairste woorden? 131
Jeugdjournaal versus Duidelijke Taal
Vincent Vandeghinste & Elke Peters 133
De extreme woorden-schat van de Middelnederlandse literatuur
Mike Kestemont 138
Veel-woorden-drachtig: het koppelteken en de samenstelling
Hans Beelen 145
Telkens weer hetzelfde verhaal
Folgert Karsdorp & Lauren Fonteyn 153
Intermezzo De woorden die we het vaakst gebruiken om 'niets' te zeggen 158
De machtswet van de voornamen
Gerrit Bloothooft 160

Voorschriften, advies en praktijk 167
De canon van taaladvies in de twintigste eeuw
Marten van der Meulen 169

Intermezzo Waar hebben onze kranten het het vaakst over? 174
Dat als in advies en praktijk
Alex Reuneker & Ronny Boogaart 177
Het negatieve effect van herhaling
Dominiek Sandra 183
Verfransing: reëel of imaginair?
Brenda Assendelft & Gijsbert Rutten 189
Intermezzo Het Nederlands van onze wetboeken 194
Moeder, waarom lenen wij?
Eline Zenner & Laura Rosseel 196

Idiomen, uitdrukkingen en verwerving 203
Vergeetwoordmusea
Sterre Leufkens 205
Intermezzo Welke idiomen horen we het vaakst? 209
De ingewikkelde puzzel die corona heet
Gudrun Reijnierse 211
Thierry doet een Trumpje
Muriel Norde 216
De scheppende kracht van verstrooidheid
Hendrik De Smet 221
Taal- en telhelden
Caitlin Meyer 226
Intermezzo Van de eerste tot de twintigste 231
Waarom herkennen sneller gaat als je jong begonnen bent
Marc Brysbaert 233

Variatie in en om Nederland 239
Gaat de taal in Nederland en Vlaanderen uit elkaar?
Dirk Geeraerts 241
Intermezzo Hoe praten we over koffie in België en Nederland? 246
Groot-Nederlands en 'taalverloedering': een kwestie van geloof
Stefan Grondelaers & Dirk Speelman 248
Gewoonweg verschillend
Ton van der Wouden 255
Oneindige variatie in online taal
Lieke Verheijen 259
Intermezzo Wat vinden we dom in het Nederlands? 265

Is het Limburgs bezig een voornaamwoord kwijt te raken?
Joske Piepers & Ad Backus 267
'Luste gij un quukske jong?' en andere Brabantse memes
Jos Swanenberg 272
Friestalige jongeren op sociale media
Lysbeth Jongbloed 278

Voorbij het Nederlands 283
Not your dad, maar wel je vader
Barend Beekhuizen 285
Hoe automatische vertaling de genderbias van AI (*Artificial Intelligence*) verraadt
Frieda Steurs, Kris Heylen & Vincent Vandeghinste 290
Intermezzo Welke gevoelsmaten gebruiken Nederlanders het vaakst? 296
'Genderneutraal taalgebruik' of 'geschlechtergerechte Sprache'
Matthias Hüning 298
Nederlands en Engels in een veeltalig Suriname
Sita M. Doerga Misier-Patadien 303
Drie keer Nederlands Creools
Cefas van Rossem 307
Intermezzo Vloeken tellen! 314
De inburgering van joodse woorden en uitdrukkingen
Ewoud Sanders 316

VOORWOORD

Taal. Het is zo'n klein woord, dat eigenlijk geen recht doet aan de enorme reikwijdte van het fenomeen. Als we het over taal hebben, dan hebben we het over hoe kinderen taal leren van hun ouders, of over de ontwikkeling van genderneutraal taalgebruik, of de manier waarop we woorden in ons hoofd opslaan, of hoe we nieuwe woorden en zinnen vormen als de wereld om ons heen verandert. Met taal kunnen we droge feiten beschrijven, maar we kunnen ons er ook creatief en origineel in rijm mee uiten, of ons genadeloos boos maken op Twitter. Taal is een duizendkoppig fenomeen, waarvan de ene kop nog interessanter is dan de andere.

Die koppen worden intensief bestudeerd in de taalwetenschap. Dat gebeurt op allerlei manieren. Soms worden mensen in MRI-machines gelegd, om te kijken hoe hun brein zich gedraagt. Soms worden er enquêtes afgenomen met vragen over taalgebruik of over de houding tegenover taalverschijnselen. Soms worden statistische tellingen losgelaten op grote tekstbestanden. Ook worden verschillende van die technieken gecombineerd. En dat allemaal om een beter beeld te krijgen van wat er nu precies echt met taal gebeurt, om vervolgens op zoek te gaan naar terugkerende patronen en sluitende verklaringen voor de wonderlijke verschijnselen die we om ons heen horen en lezen. Taalwetenschappers trekken enthousiast met een vlindernetje naar het platteland, of de stad, of het lab, of de computer, als ware David en Davina Attenboroughs. Op zoek naar taal in het wild, ademloos kijkend naar wat ze allemaal gevangen hebben, en zich verwonderd afvragend: *Wat gebeurt er in het Nederlands?!*

Frequentie

Een begrip dat van fundamenteel belang is bij het verzamelen van taalfenomenen en zoeken naar verklaringen daarvoor, is *frequentie*. In z'n meest basale vorm betekent frequentie simpelweg dat we tellen hoe vaak iets voorkomt. Dat klinkt nogal

triviaal, maar dat is het totaal niet. Dát we systematisch en op grote schaal talige fenomenen kunnen tellen, is relatief nieuw, en te danken aan het feit dat steeds meer talige gegevens gedigitaliseerd zijn. Hedendaagse teksten van allerlei genres – van kranten via literatuur tot Twitter en appjes – komen digitaal beschikbaar, maar ook oude gegevens, van twaalfde-eeuws Nederlands tot nu. Tegelijkertijd worden er steeds meer (en betere) computerprogramma's ontwikkeld om op zinnige wijze taalverschijnselen in die teksten te tellen.

Mede hierdoor heeft de taalwetenschap de afgelopen decennia alsmaar meer inzicht gekregen in hoe ongelooflijk belangrijk dat ogenschijnlijk simpele gegeven van frequentie is. Zo weten we dat frequentie de verklarende factor is voor allerlei typen taalveranderingen: aan de ene kant vinden bepaalde ontwikkelingen juist wél plaats bij infrequente woorden, maar juist níét bij de hoogst frequente. Aan de andere kant zijn er processen die juist meer voorkomen naarmate een woord of klank vaker voorkomt. Ook heeft frequentie een effect op het leren en doorgeven van de moedertaal. Voorbeelden hiervan zijn in de verschillende stukken in dit boek te vinden.

Hoewel frequentie dus enorm fundamenteel is voor ons begrip van wat er gebeurt in taal, is er aan dit fenomeen nog maar weinig aandacht geschonken in boeken voor een groot publiek. Die lacune willen wij met dit boek vullen.

Variatie

Nu is kijken naar hoe vaak iets voorkomt eigenlijk maar het halve plaatje. Om het hele plaatje te zien moet je vooral ook kijken naar wat er verder óók nog voorkomt, en wat er níét voorkomt. Kortom: als je interesse hebt in frequentie, kijk je al snel naar *variatie*. Ook dat lijkt nogal een simpel begrip: je neemt alle auto's, en dan zie je dat er rode zijn, maar ook witte en blauwe. Zo werkt dat in principe bij taal ook. Sommige mensen zeggen *groter als*, andere *groter dan*; sommigen zeggen *hij wilt naar huus*, anderen *hij wil naar huis*; sommigen zeggen *het mooie meisje dat joeg*, anderen *de mooie meisje die jaagde*. En sommige mensen zouden dat allemaal wel kunnen zeggen, afhankelijk van de situatie waarin ze het Nederlands gebruiken: bij vrienden is het *ne kaffe*, maar op het werk is het gewoon *koffie*. Ook hier geldt dat wie de patronen achter de variatie bestudeert, inzicht krijgt in allerlei aspecten van het menselijk taalvermogen. Zo blijkt variatie vaak samen te hangen met situaties, maar ook met generaties, sociale groepen of geografische gebieden.

Nederlands

Wij mogen ons best gelukkig prijzen met het Nederlands. Die taal is namelijk verhoudingsgewijs heel goed onderzocht en beschreven: er bestaan vele gram-

matica's, woordenboeken, dialectatlassen, uitspraakwoordenboeken en taalgidsen voor een heel divers publiek, er verschijnen vanaf de twintigste eeuw ieder jaar opnieuw een groot aantal proefschriften en wetenschappelijke artikelen, en er is een groot aantal digitale infrastructuren beschikbaar met oude en jonge teksten en geavanceerde zoekapplicaties.

Er wordt, met andere woorden, enorm veel onderzoek uitgevoerd naar het Nederlands. Onderzoek waar soms wel over wordt bericht, in de krant of op de radio, maar waar best méér over bericht kan worden. Want een behoorlijk deel van het onderzoek blijft besloten binnen wetenschappelijke tijdschriften, die bovendien deels slechts tegen betaling kunnen worden geraadpleegd. Ook dat is een reden om dit boek samen te stellen: ons doel was wetenschappelijke inzichten over de rol van frequentie in een groot aantal taalverschijnselen voor een breed publiek toegankelijk te maken. Op die manier snijdt het mes aan twee kanten: de taalverschijnselen die in deze bundel ter sprake komen, hebben een gemene deler, namelijk de rol van frequentie, en door frequentie als uitgangspunt te nemen, sluiten de stukken automatisch aan bij de meest recente methodes en inzichten binnen de taalkunde. Op die manier geeft de bundel een breed perspectief op het veelzijdige onderzoek dat wordt uitgevoerd naar de Nederlandse taal.

De directe aanleiding voor het samenstellen van de bundel was het feit dat Nicoline van der Sijs op 1 augustus 2021 bij de Radboud Universiteit met emeritaat gaat. Gerenommeerde en aanstormende onderzoekers hebben een bijdrage aan de bundel geleverd. Die bijdragen zijn inhoudelijk geordend. Zo zijn er hoofdstukken met stukken over taalverandering, over grammatica, uitspraak, woorden en namen, taalvoorschriften en taaladvies, idiomen en taalverwerving, en tot slot over variatie in en om Nederland, met aandacht voor streektalen en Nederlandse variëteiten als het Belgisch-Nederlands en Surinaams-Nederlands. Verder zijn er in ieder hoofdstuk ook 'intermezzi' toegevoegd: deze bevatten veel of juist weinig voorkomende woorden of uitdrukkingen in een bepaald genre of op een bepaald terrein, met korte toelichtingen. De intermezzi illustreren het begrip frequentie op een concrete wijze, als aanvulling op de meer verhalende populair-wetenschappelijke bijdragen in de bundel.

Methode

Er is een sterke neiging in wetenschapsnieuws en populair-wetenschappelijke berichten om te focussen op *resultaten*: Wát gebeurt er? Wat zien we, wat betekent dat? Maar de laatste jaren klinkt de roep steeds luider om ook aandacht te besteden aan het *hoe* van wetenschap, om mensen meer inzicht te geven in wat er allemaal achter de schermen gebeurt. Zulke vragen worden bijvoorbeeld gesteld

op de website *Neerlandistiek,* of in de blogserie *Methodologie op Maandag* van oudheidkundige Jona Lendering. Maar ze komen ook aan de orde in dit boek. Ieder hoofdstuk bevat een kader waarin nadere toelichting wordt gegeven op de methode of de data en dataverzameling. Nuttig, maar ook leuk op zich: zo leren de lezers bijvoorbeeld welke tekstverzamelingen er beschikbaar zijn en hoe ze daar zelf mee aan de slag kunnen. Ieder hoofdstuk wordt bovendien afgesloten met een korte alinea 'Verder lezen?', waarin relevante achtergrondliteratuur op een rijtje is gezet.

Er gebeurt veel in het Nederlands!

In de krant is regelmatig te lezen dat het vak Nederlands op de scholen en de studie Nederlands op de universiteit onder druk staan. Deze bundel laat echter zien dat er ondertussen enorm veel spannende dingen gebeuren *in* het Nederlands en in het onderzoek *naar* het Nederlands. Voor ons is dat allemaal erg inspirerend, en wij hopen dat dat ook voor de lezers geldt.

Lauren Fonteyn, Marten van der Meulen en Nicoline van der Sijs

TAALVERANDERING

Leestekens: onaanzienlijk en ondergewaardeerd

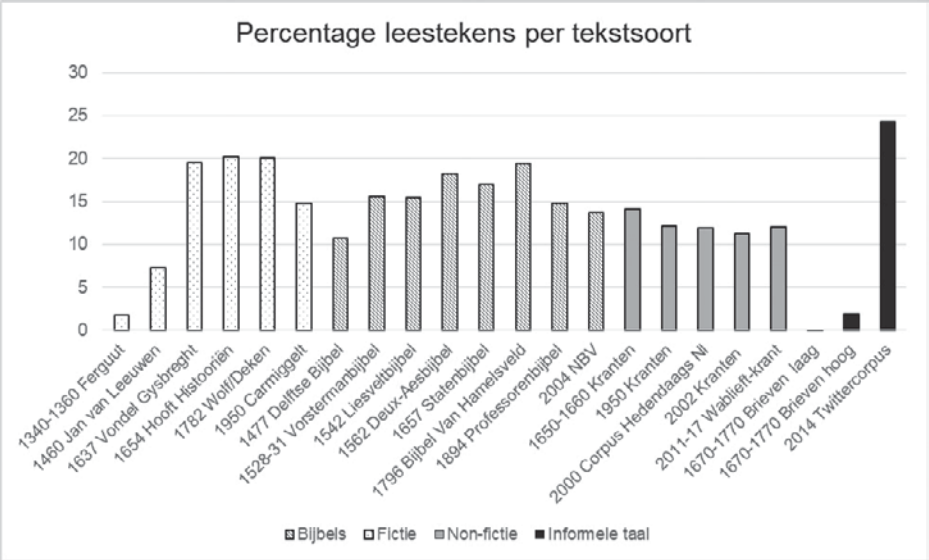
Nicoline van der Sijs
Instituut voor de Nederlandse Taal en Radboud Universiteit

Taalkundigen bestuderen taalverandering door te kijken naar het veranderende gebruik van woorden en zinnen in teksten. De onbeduidende vlekjes in die teksten, de leestekens, zijn in de Nederlandse context tot nu toe grotendeels verwaarloosd. Dit ondanks het feit dat in een gemiddelde moderne tekst de komma vaker voorkomt dan het lidwoord *de*. Die leestekens geven aan waar een langere of kortere pauze moet worden ingelast in het (hardop) lezen – vandaar de naam *leestekens*. De stilzwijgende veronderstelling is dat die leestekens irrelevant zijn voor taalveranderingsonderzoek. Maar is die veronderstelling wel terecht? Is het gebruik van leestekens in Nederlandse teksten constant of verandert het in de loop van de tijd? En als het verandert, wat betekent dat dan? Laten we voor de beantwoording van deze vragen eens kijken hoe leestekens door de tijd heen zijn gebruikt. Omdat er geen geschikt tekstcorpus beschikbaar is (zie het kader) heb ik voor dit leestekenonderzoek enkele omvangrijke en betrouwbare teksten of tekstcollecties geselecteerd uit verschillende perioden en van verschillende genres: bijbeltaal, fictie, kranten, en informele taal zoals genoteerd in brieven en tweets.

Het percentage leestekens

Figuur 1 toont een diachroon overzicht van het percentage leestekens (punt, komma, dubbelepunt, puntkomma, vraagteken en uitroepteken bij elkaar opgeteld) ten opzichte van het totale aantal woorden in de tekst, voor de vier genoemde genres.

De figuur laat een duidelijke trend zien. In de veertiende eeuw (Ferguut) ligt het percentage leestekens in de tekst heel laag. In fictie en in het bijbelcorpus stijgt het percentage sterk in de vijftiende en zestiende eeuw, vertoont een hoogtepunt in zeventiende- en achttiende-eeuwse literatuur, en daalt weer in de loop van de twintigste eeuw. Daarbij is het interessant dat het bijbelcorpus, waarin



Figuur 1. *Percentage leestekens t.o.v. het totale aantal woorden in verschillende genres en tijden*

het toch telkens om de vertaling van dezelfde brontekst gaat, dezelfde evolutie laat zien – dat bewijst wel dat leestekens een eigen regelsysteem kennen, dat net als de grammatica in de loop van de tijd verandert. In kranten, die zijn gericht op een algemeen publiek, is het percentage leestekens door de tijd vrij constant en wat lager dan in fictie of bijbeltaal, maar het is goed denkbaar dat dit beeld moet worden bijgesteld als er ook een gecorrigeerd krantencorpus beschikbaar komt van de periode tussen 1650 en 1950. In informele brieven worden nauwelijks leestekens gebruikt, maar hoe hoger opgeleid de schrijver, hoe meer leestekens hij of zij gebruikt.

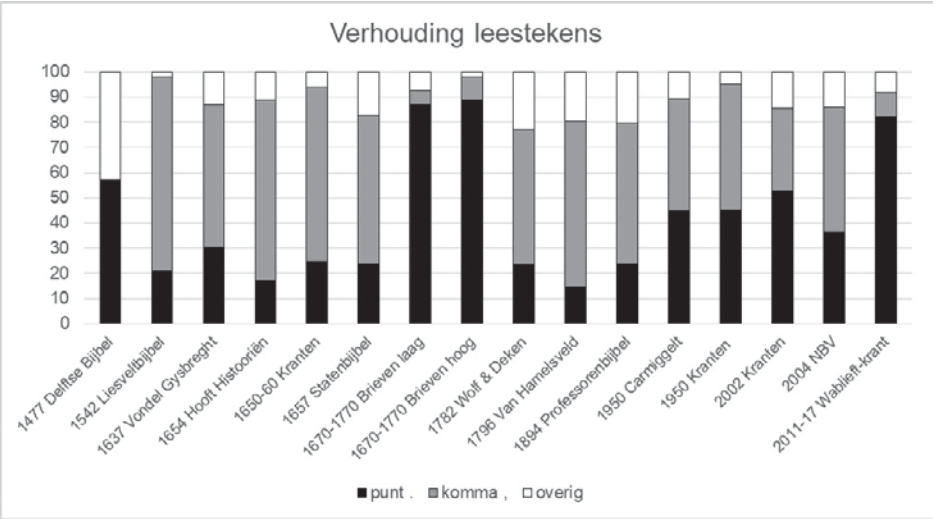
Terwijl in alle genres de trend is dat het percentage leestekens in de modernste tijd afneemt, zien we precies het tegenovergestelde in het Twittercorpus: daarin komen verreweg de meeste leestekens voor, meer dan in enige eerdere periode. Dat heeft te maken zowel met de inhoud van de berichten, die vaak informeel en persoonlijk van aard zijn, als met de vorm: een tweet bevat maximaal 280 tekens. Twitteraars compenseren de beperkte lengte door het toevoegen van leestekens, emoticons en smileys.

De fluctuering van leestekens blijkt samen te vallen met veranderingen in het taalgebruik. De veranderingen in de vijftiende en zestiende eeuw reflecteren de invloed van drukkers na de uitvinding van de boekdrukkunst: die kozen voor een vastere spelling, vermeden streekgebonden woorden en verbindingen, en

deelden zinnen met behulp van leestekens op in kleinere eenheden. In de zeventiende eeuw kreeg de standaardtaal vorm. In die periode gold een lange, meanderende volzin met veel bijzinnen en bijstellingen als buitengewoon welluidend en navolgenswaard, want zo schreven de Romeinen. Dat verklaart de toename van leestekens in literaire en religieuze teksten in die periode. Interessant genoeg geldt dat voor álle literatuur: niet alleen voor de berucht ingewikkelde *Histooriën* van Hooft, maar ook voor het toneelstuk *Gysbreght van Aemstel* van Vondel, dat was bedoeld om op te voeren voor het publiek.

Onderlinge verdeling van de leestekens

Hoe zit het met de onderlinge verdeling van de leestekens door de tijd heen, vertonen ook die een evolutie? Ja, zo blijkt uit Figuur 2. Voor de overzichtelijkheid is de grafiek beperkt tot een verdeling in punten, komma's en overige leestekens. In de oudste fictieteksten uit de veertiende en vijftiende eeuw, die sowieso slechts weinig leestekens bevatten, komt maar één leesteken voor, namelijk de punt. Die teksten zijn weggelaten uit de figuur.



Figuur 2. *Onderlinge verdeling van de leestekens door de tijd heen*

De vijftiende-eeuwse tekst bevat alleen punten en dubbelepunten. In de zestiende-eeuwse teksten rukt de komma op, en neemt direct de eerste plaats in: 77 procent van de leestekens in de Liesveltbijbel is een komma.

Vanaf de zeventiende eeuw bestaat er een delicaat evenwicht tussen de verschillende leestekens, die ieder een eigen functie krijgen. In brieven overheerst