

De ontwikkeling van een woordenlijst academisch Nederlands

MADELEINE HULSEN, EDITH SCHOUTEN,

HARALD BAAYEN EN ROB SCHREUDER

Jaarlijks gaan er honderden buitenlandse studenten in Nederland studeren aan een hogeschool of universiteit. Gezien de steeds verdergaande internationalisering is de verwachting dat dit aantal in de komende jaren verder zal toenemen. De studenten dienen over een goede (academische) woordenschat te beschikken. In dit artikel wordt de ontwikkeling beschreven van een woordenlijst schriftelijk academisch taalgebruik. Het doel van de woordenlijst is de anderstalige (aspirant-)student te helpen bij het lezen en begrijpen van academische teksten. De lijst verschijnt in een idioomboek waarin ook een groot aantal woordenschatoefeningen is opgenomen. De lijst kan dienen als naslagwerk, maar vooral ook als uitgangspunt om de woordenschat uit te breiden en te trainen.

In deze bijdrage wordt ingegaan op het hoe en waarom van de woordenlijst academisch Nederlands. Bovendien wordt verslag gedaan van het corpusonderzoek waarop de lijst is gebaseerd en de stappen die zijn genomen om tot de definitieve lijst te komen. Tenslotte wordt de opzet en inhoud van het idioomboek nader toegelicht.¹

Achtergronden

Het leren van vocabulaire is een essentieel onderdeel van het verwerven van een vreemde taal. De laatste jaren staat woordenschatontwikkeling en -training volop in de belangstelling. Wat verstaan we precies onder academische woordenschat? Academische woordenschat is die specifieke set van woorden die niet frequent voorkomt in algemeen en alledaags (schriftelijk) taalgebruik, maar wel frequent in een grote verscheidenheid aan academische teksten. Het gaat hierbij uitdrukkelijk niet om vakjargon of puur specialistische termen die alleen worden gebruikt in één bepaald onderzoeksgebied, maar om woorden die men veelvuldig tegenkomt in academische teksten; woorden die het onderwerp van de teksten ondersteunen, maar niet verwant zijn aan het onderwerp zelf.

In de jaren tachtig van de vorige eeuw is door Paul Nation in Nieuw-Zeeland al een dergelijke woordenlijst voor het Engels opgesteld. Engelse academische woorden die in deze 'University Word List' voorkomen, zijn bijvoorbeeld *analyze*, *conclude* en *principle*, maar de lijst bevat ook meer specialistische woorden als *x-ray* en (enkele) namen van

vakgebieden als *philosophy* en *geography* (Xue & Nation 1984). De 'University Word List' is opgesteld door een viertal eerdere woordenlijsten te combineren. De belangrijkste bronnen waren een woordenlijst voor toekomstige niet-Engelstalige studenten opgesteld door de New Zealand Council for Educational Research (Campion & Elley 1970) en de 'American University Word List' (Praninskas 1972). Voor deze lijsten waren onder meer woorden uit college teksten en studieboeken voor eerstejaars universitaire studenten gebruikt.

Meer recentelijk heeft Coxhead (2000) een nieuwe academische woordenlijst voor het Engels ontwikkeld, de 'New Academic Word List', die uit 570 woordfamilies bestaat en op een corpus van 3,5 miljoen woorden is gebaseerd. Deze lijst is op een meer systematische en geautomatiseerde wijze tot stand gekomen, doordat de teksten in elektronische vorm zijn verzameld en geanalyseerd. Het corpus van Coxhead bestaat uit 414 academische teksten, geschreven door meer dan 400 auteurs. In het corpus zijn artikelen uit wetenschappelijke tijdschriften, wetenschappelijke internetpublicaties, complete studieboeken, hoofdstukken daaruit en andersoortige academische teksten opgenomen. De academische disciplines die in het corpus vertegenwoordigd zijn betreffen letteren en sociale wetenschappen, beleidswetenschappen, rechten en de bètawetenschappen. Uit een vergelijking van de 'University Word List' van Xue & Nation en de 'Academic Word List' van Coxhead bleek dat de laatste, ondanks een geringer aantal woorden, een hogere dekkingsgraad heeft en een groter aantal disciplines bestrijkt.

Voor het Nederlands is er tot nu toe geen woordenlijst opgesteld met een selectie van woorden die de academische woordenschat vertegenwoordigen. Een relevant onderzoek is dat van Hazenberg (1994) naar de wenselijke en feitelijke receptieve woordenschat van niet-Nederlandstalige studenten. In dit proefschrift stond de vraag centraal hoeveel

en welke Nederlandse woorden (aspirant-)studenten receptief dienen te kennen om Nederlandse studieteksten te kunnen lezen. Het gaat hierbij echter om de totale receptieve woordenschat, waarbij Van Dales *Basiswoordenboek Nederlands* als uitgangspunt is genomen. Aan de hand van het basiswoordenboek en een vergelijking met het (Nederlandse) INL-corpus van de CELEX-database bestaande uit 42.380.000 woordtekens is een lijst samengesteld van 23.550 woorden (zie paragraaf 2 voor uitleg over het INL-corpus). Uit het onderzoek bleek dat (aspirant-)studenten die de ongeveer 11.123 meest frequente woorden uit de woordenlijst van Hazenberg receptief kenden, een willekeurige Nederlandse tekst redelijk konden begrijpen. Aspirantstudenten met een Germaanse moedertaal hadden een grotere receptieve woordenschat dan anderstalige studenten met een niet-Germaanse moedertaal. Zoals gezegd spitste het onderzoek van Hazenberg zich echter niet toe op specifiek academische woorden, maar ging het om de totale receptieve woordenschat.

In het algemeen wordt er bij cursussen Nederlands voor buitenlandse (aspirant-)studenten geen of weinig expliciete aandacht besteed aan het leren van woorden die frequent voorkomen in academische en/of studieteksten. Uit onderzoek blijkt echter dat studenten over het algemeen meer moeite hebben met het leren van academisch vocabulaire dan met de algemene, basiswoordenschat, omdat academische woorden minder frequent voorkomen (Coxhead 2000). Voor het succesvol kunnen volgen van een studie aan een hbo- of universitaire instelling is het dus zinvol aandacht te besteden aan de verwerving van academische woorden, te meer omdat deze in de betreffende taalomgeving wél frequent worden gebruikt.

De op te stellen woordenlijst kan gebruikt worden als richtlijn voor de ondersteunende (academische) woordenschat die (buitenland-

se) studenten dienen te hebben bij aanvang van een studie aan een Nederlandse universiteit. De lijst is, zoals gezegd, vooral bedoeld voor buitenlandse of uitwisselingsstudenten (in hbo of wo) die geheel of gedeeltelijk hun opleiding in het Nederlandse taalgebied volgen, maar ook Nederlandse studenten, leerlingen in de Tweede Fase van het voortgezet onderwijs en beleidsmakers kunnen profijt hebben van de lijst.

De samenstelling van het corpus

Het corpus van Nederlandse academische en wetenschappelijke teksten werd aangelegd met behulp van twee bronnen:

Een verzameling van Nederlandse academische teksten, zoals syllabi, cursusboeken, handleidingen, collegesamenvattingen, die werden gebruikt in colleges voor eerste- en tweedejaars studenten aan de Katholieke Universiteit Nijmegen (KUN).

Een verzameling van (populair-)wetenschappelijke artikelen die verschenen zijn in de wetenschapsbijlagen van *de Volkskrant* en *NRC Handelsblad* tussen 1991 en 2000.

Alle teksten werden verkregen in digitale vorm om automatische en systematische frequentietellingen te kunnen realiseren.

Voor de verzameling van academische teksten werd docenten van alle faculteiten van de KUN² gevraagd schriftelijk cursusmateriaal ter beschikking te stellen (in digitale vorm), dat bedoeld was voor eerste- en tweedejaars studenten. De reden waarom we specifiek materiaal voor deze groep studenten zochten, was dat er voor ouderejaars studenten veelal meer specialistische teksten zullen worden gebruikt en ook vaker gebruik zal worden gemaakt van niet-Nederlandstalig materiaal. Verder gingen we ervanuit dat syllabi en cursusboeken die voor eerste- of tweedejaars studenten worden geschreven het beste het algemene, ondersteunende academische

vocabulaire weerspiegelen dat nodig is om een opleiding te volgen aan een universitaire of hbo-instelling. We hebben getracht een zo groot mogelijke spreiding te krijgen over de faculteiten en studierichtingen. Helaas was het niet voor alle disciplines mogelijk geschikte digitale teksten te vinden (bijvoorbeeld voor geschiedenis, economie en psychologie). Dit is voor een groot deel te verklaren door het feit dat (natuurlijk) niet alle docenten gebruik maken van eigen lesmateriaal, maar bijvoorbeeld van (Engelstalige) studieboeken of artikelen. In het academische corpus zijn de natuurwetenschappen enigszins oververtegenwoordigd. Het zou kunnen dat bij deze studierichtingen het gebruik van Nederlandse teksten vaker voorkomt of dat de teksten vaker digitaal beschikbaar zijn.

Het resulterende corpus van academische teksten bestond uit 40 teksten, geschreven door 37 verschillende auteurs (zie Tabel 1).

Discipline	Aantal teksten	Auteurs
Letteren		
Taalwetenschappen	3	2
Bedrijfscommunicatie	1	1
Filosofie	2	2
Rechten	5	4
Sociale wetenschappen		
Antropologie	1	1
Pedagogiek	1	5
Medische wetenschappen		
Tandheelkunde	5	4
Medicijnen en gezondheidswetenschappen	4	4
Bètawetenschappen		
Biologie	4	2
Wiskunde	5	3
Informatica	1	1
Natuurkunde	2	2
Scheikunde	6	6
Totaal	40	37

Tabel 1: Academische disciplines vertegenwoordigd in het corpus van academische teksten, onderverdeeld naar aantal teksten en auteurs.

De teksten varieerden in lengte van 316 tot 71.866 woorden. In totaal bestond het corpus van academische teksten uit 1.410.871 woorden.

Het corpus van artikelen uit de wetenschapsbijlagen³ bestond in totaal uit 4.712 artikelen. De artikelen uit *de Volkskrant* (n = 2472) werden gepubliceerd tussen 1997 en 2000 en die uit *NRC Handelsblad* tussen 1991 en 2000 (n = 2240). Hoewel teksten in wetenschapsbijlagen in zekere mate zullen afwijken van academische teksten voor eerste- en tweedejaars studenten, hebben we gemeend dat ze toch een zeer nuttige aanvulling kunnen zijn voor de samenstelling van een algemene, ondersteunende academische woordenlijst. We zijn er daarbij vanuit gegaan dat de moeilijkheidsgraad en de mate van specialisatie van de wetenschapsbijlagen ongeveer gelijk is aan die van de academische teksten voor eerste- en tweedejaars studenten. Door de statistische bewerking, die hieronder wordt beschreven, kon worden voorkomen dat woorden die meer tot het alledaagse Nederlands behoren niet werden geselecteerd. Door de aard en de grootte van het corpus - alle krantenartikelen stonden achter elkaar in één bestand - was het niet mogelijk de spreiding in de lengte van de teksten aan te geven en een onderscheid te maken tussen academische disciplines en auteurs.

Het totale corpus bestond uit meer dan vijf miljoen woorden, of tokens (zie Tabel 2). Gemiddeld bestonden de academische teksten uit 35.272 woorden. De wetenschappelijke artikelen waren korter van lengte: de artikelen uit de wetenschapsbijlage van *de Volkskrant* waren gemiddeld 725 woorden lang en die uit *NRC Handelsblad* gemiddeld 865 woorden. Het aantal types, dat wil zeggen het aantal verschillende woorden in het corpus, was 172.233.

Om het aantal en de frequentie van de woorden in het corpus te bepalen werden alle tekstbestanden omgezet tot platte ASCII-bestanden en alle woorden en letters werden omgezet in kleine letters. Accenten, trema's, interpunctie en eventuele opmaakcodes werden verwijderd. Woorden met nummers werden ook verwijderd. De woord- en frequentietellingen werden gebaseerd op de overblijvende reeksen van letters. Voor elk woord werd de woordfrequentie bepaald binnen de tekst en binnen het totale corpus. De volgende stap was het vergelijken van de frequentie van de woorden in ons Nederlandse corpus met de frequentie van de woorden in het INL-corpus van CELEX (Centrum voor Lexicale Informatie), dat 42 miljoen hedendaagse Nederlandse woorden bevat (Baayen, Piepenbrock & Gulikers 1995). Het INL-corpus is samengesteld uit 930 volledige boeken (ongeveer 30% fictie en 70% non-fictie), die zijn geschreven tussen 1970 en 1988. De woorden in het corpus kunnen worden gekarakteriseerd als leesmateriaal dat gevonden kan worden in een openbare bibliotheek voor volwassenen (Hazenberg 1994, 37). Kinderboeken, studieboeken, specialistische literatuur, kranten en tijdschriften zijn niet vertegenwoordigd in het corpus.

Door het academische corpus, dat uiter-

Corpus	Aantal teksten	Totaal aantal woorden
Academische studieteksten	40	1.410.871
Wetenschappelijke kranten-artikelen		
<i>de Volkskrant</i>	2.472	1.793.163
<i>NRC Handelsblad</i>	2.240	1.938.173
Totaal	4.752	5.142.207

Tabel 2: Aantal woorden in de corpora.

aard ook algemene niet-academische woorden bevat en het INL-corpus, waarin academisch vocabulaire in een veel mindere mate is vertegenwoordigd, met elkaar te vergelijken kan bepaald worden of een woord beschouwd kan worden als 'academisch'. Om deze vergelijking tot stand te brengen, werd voor elk woord een zogeheten entropiewaarde berekend, die de mate van onzekerheid uitdrukt dat een woord in het academische corpus dezelfde frequentie heeft als hetzelfde woord in CELEX. De entropiewaarde kan elke waarde innemen tussen nul en één. Een entropiewaarde van nul houdt in dat er met absolute zekerheid kan worden gezegd dat het woord alleen in het academische corpus voorkomt, terwijl een entropiewaarde van één aangeeft dat er geen of heel weinig verschil is tussen de frequenties van een woord in het academische corpus en het INL-corpus van de CELEX-database. Kortom, woorden waarvan de entropiewaarde dicht bij nul ligt hebben een grotere kans als academisch beschouwd te kunnen worden. Woorden met een entropiewaarde in de buurt van één zijn hoogstwaarschijnlijk algemene, alledaagse Nederlandse woorden.

De ontwikkeling van de woordenlijst academisch Nederlands

De ontwikkeling van de academische woordenlijst vond plaats in verschillende stappen en omvatte zowel objectieve als subjectieve procedures. Het was enerzijds niet mogelijk, maar anderzijds ook niet gewenst om de selectie van de woorden voor de academische woordenlijst geheel op 'objectieve' wijze plaats te laten vinden. Kort samengevat werden de woorden geselecteerd op basis van de volgende criteria, die achtereenvolgens werden toegepast:

1. De woorden die opgenomen worden in de woordenlijst academisch Nederlands komen frequenter voor in academische teksten dan

in algemeen alledaags Nederlands.

2. Een woord moet minstens vijf keer voorkomen in het academische corpus.
3. Een woord moet geselecteerd worden als academisch door minimaal de helft van een groep onafhankelijke beoordelaars.
4. Een woord wordt opgenomen in de uiteindelijke lijst als deze op basis van ervaring en intuïtie (ook) relevant wordt geacht door de onderzoekers.⁴

Vaak wordt ook de spreiding van een woord over de verschillende teksten in het corpus als selectiecriteria gebruikt bij het maken van woorden- of frequentielijsten (zie onder andere Hazenberg 1994, Coxhead 2000). Dit stuitte in ons geval echter op praktische bezwaren, omdat de wetenschapsbijlagen niet verder uit te splitsen waren in deelcorpora.

Wat hielden de wél gebruikte selectiecriteria precies in? Om het aantal tokens te reduceren werden uit het corpus die woorden geselecteerd die meer dan vijf keer voorkwamen en die daarnaast een entropiewaarde van minder dan 0,80 hadden. Deze entropiewaarde lijkt hoog, maar uit eerste analyses bleek dat er anders te veel relevante woorden uit de lijst zouden worden gefilterd. Het frequentiecriteria werd gebruikt om de kans te verkleinen dat het een vakspecifiek woord betrof. Daarom werd besloten dat een woord minstens vijf keer in het totale corpus voor moest komen. Deze twee stappen leverde een ruwe lijst op van 21.045 woorden.

Frequentie en entropie zijn belangrijke criteria voor woordselectie, maar ze zeggen niet alles. Deze twee criteria kunnen niet voorkomen dat er woorden worden geselecteerd die weliswaar relatief vaak voorkomen in academische teksten, maar die intuïtief minder belangrijk of relevant zijn voor leerders van het Nederlands of voor de academische omgeving. Tevens kunnen hierdoor woorden worden geselecteerd die té vakspecifiek zijn, wat gezien de doelstellingen van het onderzoek evenmin gewenst is. Daarom maakten de twee 'subjectieve' criteria de selectie van woorden voor de woordenlijst academisch Nederlands niet alleen mogelijk, maar ook gewenst.

tieve' stappen of criteria een belangrijk deel uit van het selectieproces.

Bijna 40% van de woorden uit de totale lijst van 21.045 woorden werd beoordeeld door 12 onafhankelijke beoordelaars, aan wie de achtergrond en het doel van het onderzoek was uitgelegd alsmede de definitie van academisch vocabulaire. De beoordelaars werden geïnstrueerd die woorden aan te kruisen die zij als academisch beschouwden. Twijfelgevallen konden zij markeren door middel van een vraagteken. Alle beoordelaars hadden een universitaire opleiding voltooid en waren dus bekend met academische literatuur. Een aantal beoordelaars was docent NT2 aan hoogopgeleiden. De oordelen van de beoordelaars werden gecontroleerd door de onderzoekers, zodat eventuele bias in de resultaten eruit gehaald kon worden.

Oorspronkelijk was het de bedoeling dat de beoordelaars de volledige lijst van geselecteerde woorden zouden beoordelen. Helaas bleek dit door technische en organisatorische problemen niet mogelijk. Uiteindelijk hebben de 12 beoordelaars 8.134 woorden beoordeeld uit de ruwe lijst, zo'n 39% van de totale lijst. De resterende lijst van academische woorden werd beoordeeld en gecontroleerd door drie onderzoekers, allen academisch opgeleid en zeer ervaren met NT2-onderwijs.

De procedure voor de woorden die door de beoordelaars aangekruist waren als academisch, was als volgt. Alleen woorden die als academisch waren gemarkeerd bij zes of meer beoordelaars werden opgenomen in de uiteindelijke lijst. Op basis van intuïtie werd daarnaast gekeken of er geen 'onterechte afvallers' zaten bij de lijst van woorden die door minder dan zes beoordelaars waren gemarkeerd als academisch. Als gevolg hiervan werd een klein aantal woorden aan de lijst toegevoegd, ondanks het feit dat ze door minder dan de helft van de beoordelaars was gemarkeerd als academisch (bijvoorbeeld *aantasten*, *afronden*, *ondergeschikt*). Als een woord

echter door minder dan drie beoordelaars was gemarkeerd, werd het niet alsnog toegevoegd. Verder werd ook gecontroleerd of er geen 'onterechte blijvers' zaten in de lijst van woorden die wel door zes of meer beoordelaars was gemarkeerd. In een aantal gevallen werd besloten een woord niet op te nemen terwijl meer dan zes beoordelaars toch vonden dat het een academisch woord betrof. In het algemeen ging het hierbij om vakjargon of om woorden die te specifiek voor één bepaalde discipline werden geacht (bijvoorbeeld *differentiaal*, *transcriberen*), of te algemeen (*inhoudsopgave*, *docent*, *tentamen*).

Tot slot werd de lijst aangevuld met een klein aantal woorden die niet uit het totale corpus was geselecteerd op basis van de twee criteria frequentie en entropie, bijvoorbeeld omdat ze ook een hoge frequentie in het alledaagse Nederlands hebben, maar met een iets andere betekenis. Ook werd een klein aantal extra woorden toegevoegd die gevonden waren in recente wetenschappelijke krantenartikelen, andere academische teksten of op basis van intuïtie.

Om de lijst zo compleet en uitgebreid mogelijk te maken hebben we ook tegengestelde betekenissen - indien relevant - toegevoegd van woorden die in het corpus voorkwamen. In de meeste voorkomende gevallen kwam alleen de 'negatieve' vorm voor in het corpus, in welk geval de 'positieve', ongemarkeerde vorm werd toegevoegd (bijvoorbeeld *ongeldig-geldig*, *incompleet-compleet*). Als alleen de positieve vorm in het corpus voorkwam, werd echter niet automatisch de negatieve, gemarkeerde vorm toegevoegd (*opvallend* kwam bijvoorbeeld voor in het corpus, maar *onopvallend* niet en is ook niet toegevoegd).

De uiteindelijke lijst bevat niet alleen de 'stamvorm' van een woord, maar zo veel mogelijk van dezelfde stamvorm het substantief, de infinitief en het adjectief. Daarnaast bevat de lijst ook een aantal bijwoorden, voegwoorden en voorzetsels (bijvoorbeeld *aldus*, *circa*, *dank-*

zij). Samenstellingen werden alleen in de lijst opgenomen als ze vaak voorkomen in deze vorm in wetenschappelijk taalgebruik (bijvoorbeeld *deelverzameling, controlegroep, proefpersoon*). De uiteindelijke lijst bestaat uit 1.484 woorden.

Welke woorden uit de woordenlijst academisch Nederlands komen nu het meest voor in het academische corpus? Om dit te bepalen hebben we naast de academische woordenlijst, waarin de woorden op alfabetische volgorde staan vermeld, een lijst gemaakt van de honderd meest frequente woorden. De drie meest frequente woorden in het academische corpus zijn achtereenvolgens: *wetenschap* (18.073 maal), *artikel* (15.842 maal) en *onderzoek* (7.100 maal). Bij het opstellen van de lijst met de honderd meest frequente woorden zijn de in het corpus aangetroffen woordvormen bij elkaar opgeteld, maar niet over de woordsoorten heen. Bijvoorbeeld de frequenties van de substantieven *wetenschap* en *wetenschappen* zijn bij elkaar opgeteld en onder het substantief vermeld, maar het adjectief *wetenschappelijk* heeft een aparte ingang. De frequentie daarvan is de som van het aantal keer dat de vormen

wetenschappelijk en *wetenschappelijke* in het corpus voorkomen (in totaal 3.871 maal).

Toepassingen

Op basis van de woordenlijst academisch Nederlands wordt door medewerkers van het Universitair Talencentrum Nijmegen een idioomboek ontwikkeld waarin zowel de woordenlijst als idioom oefeningen zijn opgenomen (Giezenaar & Schouten in ontwikkeling). Het boek richt zich op hoogopgeleide anderstaligen met een uitgebreide basiskennis Nederlands en is geschikt voor studie- en beroepsdoeleinden. De woordenlijst academisch Nederlands is in alfabetische volgorde in zijn geheel opgenomen in het idioomboek, waarbij voor elk woord een vertaling in het Engels, Frans en Duits is gegeven. Bovendien is er ruimte opengelaten voor het invullen van een vertaling in de eigen taal.

Daarnaast is ook de lijst opgenomen van de honderd meest frequente academische woorden uit het corpus. In het oefengedeelte van het idioomboek komt de woordenschat

aandachtspunt – afbeelden		
Nederlands	English	Français
A andachtspunt, het	point of interest	singularité
aandragen	put forward	avancer
aaneengesloten	uninterrupted	ininterrompu
aangeven	indicate	indiquer
aanleren	learn/teach	apprendre
aanleveren	supply	fournir
aanmerken	comment	observer
aanname, de	assumption	hypothèse
aannemelijk	plausible	plausible
aanpak, de	approach	approche
aanpakken	deal with	aborder
aanpassen	adapt	adapter
aanpassing, de	adaptation	adaptation

Deel van de woordenlijst academisch Nederlands uit 'Wijze woorden' (Giezenaar & Schouten, in ontwikkeling).

uit de lijst aan bod in een breed scala aan opdrachten, gerangschikt van meer gesloten naar meer open. In het eerste deel wordt de woordvorming en betekenis geoefend van achtereenvolgens substantieven, verba, adjectieven en adverbia, gekoppeld aan het oefenen met de woorden in een betekenisvolle context.

In het tweede deel wordt de woordenschat geoefend op basis van authentieke en bewerkte teksten waarbij verschillende thema's aan bod komen. Te denken valt aan het beschrijven van gegevens (grafieken en tabellen) of het verwoorden van conclusies met als doel de woorden uiteindelijk actief te kunnen toepassen in eigen teksten. Aangezien het de bedoeling is dat het boek zelfstandig kan worden doorgewerkt, is er bij alle oefeningen een sleutel. Dit neemt niet weg dat het boek ook in de les kan worden gebruikt door docenten NT2 die studenten voorbereiden op een studie aan een instelling voor hoger onderwijs.

NOTEN

1. Onze dank gaat uit naar Bert Weltens, die de ontwikkeling van de woordenlijst academisch Nederlands heeft geïnitieerd en naar Liesbet Korebrits voor haar waardevolle commentaar op eerdere versies van dit artikel.
2. De KUN heeft acht faculteiten (Theologie, Filosofie, Letteren, Rechtsgeleerdheid, Sociale wetenschappen, Managementwetenschappen, Natuurwetenschappen, wiskunde en informatica, Medische wetenschappen) en meer dan 13.000 studenten in 45 studierichtingen.

3. De auteurs danken de Nederlandse Persdatabank voor het ter beschikking stellen van de bestanden met de artikelen uit de wetenschapbijlagen.
4. Zie Hazenberg (1994) voor een opsomming van de nadelen van het enkel en alleen selecteren op basis van frequentie.

LITERATUUR

- Baayen, H., Piepenbrock, R. & Gulikers, L. (1995). *The CELEX lexical database*. (cd-rom). Philadelphia, PA.: Linguistic Data Consortium, University of Pennsylvania.
- Campion, M. & Elley, W. (1971). *An Academic Vocabulary List*. Wellington: NZCER.
- Coxhead, A. (2000). A New Academic Word List. *Tesol Quarterly*, 34, 2, 213-238.
- Giezenaar, G. & Schouten, E. (in ontwikkeling). *Wijze woorden. Woordenlijst Academisch Nederlands met idioom oefeningen*. Amsterdam: Intertaal.
- Hazenberg, S. (1994). *Een keur van woorden. De wenselijke en feitelijke receptieve woordenschat van anderstalige studenten*. Proefschrift, Vrije Universiteit.
- Hulstijn, J. & Hazenberg, S. (1992). *Woorden op zicht. Wordselectie ten behoeve van het NT2-onderwijs*. *Levende Talen*, 467, 2-7.
- Praninskas, J. (1972). *American University Word List*. London: Longman.
- Xue, G. & Nation, P. (1984). A University Word List. *Language Learning and Communication*, 3, 2, 215-228.

Opdracht 21

A Zet de adjectieven op de juiste plaats in de vaste combinaties. Verander indien nodig de vorm.

haaks + aselekt + gepaard + ~~doorlaggend~~ + naadloos + afgerond
 + baanbrekend + onomstotelijk + afwijzend + beproefd + gebaat

Voorbeeld:

van doorlaggend	belang zijn
1. een	geheel vormen
2. ergens	tegenover staan
3. een	steekproef
4.	onderzoek
5. een	methode
6.	gaan met
7.	zijn bij
8.	staan op
9.	aansluiten bij
10.	vaststaan

B Vul bovenstaande vaste combinaties in.

- De commissie heeft de plannen voorgelegd aan de directie, maar die
er jammer genoeg nogal
- Het dat het
verplicht dragen van autogordels het aantal fatale verkeersongevallen drastisch heeft
verminderd.
- De drie delen van het standaardwerk over taalkunde samen een
.....
- Een onderzoeksteam van het Academisch Ziekenhuis heeft een prestigieuze prijs gekregen
omdat ze hebben gedaan naar astma bij
jonge kinderen.
- Zijn ideeën over de ontwikkelingspsychologie zijn controversieel. Ze
..... de gangbare theorieën op dit gebied.
- De conclusies uit de tweede serie experimenten
..... het eerste deel van het onderzoek. Dit stemt de
onderzoekers tot tevredenheid.
- We spreken van een wanneer de proefpersonen
willekeurig worden gekozen.