

NL TERM Publicaties

# TiNT dag 2012

Redactie

Els Ruijsendaal, Cornelia Wermuth

Reeks: Terminologie in het Nederlandse Taalgebied 4



ACADEMIA  
PRESS

TiNT dag 2012



NL TERM Publicaties

# TiNT dag 2012

Redactie:  
Els Ruijsendaal, Cornelia Wermuth

Reeks: Terminologie in het Nederlandse Taalgebied 4



ACADEMIA  
PRESS

Reeks *Terminologie in het Nederlandse Taalgebied*, deel 4: **TiNT dag 2012**

Een uitgave van de Vereniging NL Term en Steunpunt Nederlandse Terminologie (SNT)

© Academia Press  
Eekhout 2  
9000 Gent  
Tel. 09/233 80 88 Fax 09/233 14 09  
Info@academiapress.be www.academiapress.be

Vormgeving & zetwerk: bvba Le Pur et l'Impur

Els Ruijsendaal, Cornelia Wermuth (red.)  
TiNT dag 2012

Gent, Academia Press, 2013, VIII + 88p.

ISBN 978 90 382 2253 0  
D/2013/4804/261  
NUR 620  
U 2111

Niets uit deze uitgave mag worden verveelvoudigd en/of vermenigvuldigd door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook, zonder voorafgaande schriftelijke toestemming van de uitgever.

## Woord vooraf

De TiNT dag 2012, alweer de vierde editie, vond plaats in het bekende Meertens Instituut in Amsterdam, waar directeur hoogleraar Hans Bennis onze gastheer was.

Het thema van de dag was “Terminologie en actuele taaltoepassingen”.

Nieuw was, zoals u in deze bundel kunt lezen, dat het thema deels gestalte kreeg in interactieve sessies, zeven stuks van zo’n 10 tot 15 minuten. En die kregen de volle belangstelling van de rondtrekkende aanwezigen, die ondanks het bar en boze winterweer in groten getale waren komen opdagen.

Impressies en verslagen van deze benadering en uitgewerkte artikelen van de erop volgende lezingen vullen deze bundel. Daarin passeren uiteengezette praktijkvormen als die van termen in het Algemeen Nederlands Woordenboek, de termextractor van de Nederlandse Taalunie, de hogeronderwijsterminologie van de universitaire vertaalcentra, de Language Tool Box van KU Leuven enz. de revue.

Een enquête onder de aanwezigen om hun wensen en inzichten mee te kunnen nemen, maakt tot slot zichtbaar waar het de onderzoekers en de gebruikers om te doen is. Men is, na lezing hiervan, weer helemaal bij.

Op de vijfde TiNT dag, ons eerste lustrum(!), slaan wij onze tenten op in Brussel, in het Centrum voor Vaktaal en Communicatie van de Erasmushogeschool (Vrije Universiteit Brussel). En daar hoop ik u wederom te mogen verwelkomen.

Willy Martin  
Voorzitter NL TERM



# Inhoud

<b>Vaktaal in woordenboeken en instituties</b>	1
Tanneke Schoonheim	3
<i>Nederland wordt steeds slimmer – Voormalige vaktaalwoorden en de algemene woordenschat in het Nederlands</i>	
Folkert Zijlstra	15
<i>Het terminologieprogramma van de NAVO</i>	
<b>Terminologie en vertalen</b>	23
Bert Keurentjes	25
<i>Universitaire vertaaldiensten en hogeronderwijs terminologie</i>	
Serge Verlinde & Sara Rymenams	35
<i>De Interactive Language Toolbox, óók voor terminologen</i>	
<b>Terminologie en automatische taaltoepassingen</b>	47
Mart Trautwein	49
<i>Metadatering &amp; automatische classificatie bij de Tweede Kamer</i>	
Oele Koornwinder	55
<i>Terminologie in automatische toepassingen</i>	
<b>Veldenquête terminologie</b>	69
Hennie van der Vliet	71
<i>U over NL TERM – De resultaten van de bezoekersenquête op de TiNT dag 2012</i>	
<b>Medewerkers aan deze bundel</b>	87





# **Vaktaal in woordenboeken en instituties**



Tanneke Schoonheim

(Instituut voor Nederlandse Lexicologie, Leiden)

## **Nederland wordt steeds slimmer**

### **Voormalige vaktaalwoorden en de algemene woordenschat in het Nederlands**

#### ***amenvatting***

Algemene woordenboeken van het Nederlands beschrijven in principe de algemene woordenschat van het Nederlands. Voor de zogenoemde ‘vakwoordenschat’, die specifiek is voor een bepaald vakgebied en daarbinnen ook redelijk exclusief gebruikt wordt, zijn er andere woordenboeken, namelijk vakwoordenboeken. Hoewel over het algemeen redelijk goed is uit te maken wat tot de algemene woordenschat behoort en wat tot de vakwoordenschat gerekend moet worden, zit daartussenin een grijs gebied met woorden die in beweging zijn, en dan vooral woorden die zich van de vakwoordenschat in de richting van de algemene woordenschat bewegen.

Een van de terreinen waarop dit goed zichtbaar is, is het vakgebied van de geneeskunde. Waar vroeger in het dagelijks taalgebruik gesproken werd van *aderverkalking*, *hersenvliesontsteking* en *suikerziekte*, vinden we hiervoor tegenwoordig meer en meer (ook) de termen *arteriosclerose*, *meningitis* en *diabetes*. Deze van oorsprong medische termen lijken zich een plaats te hebben verworven in de algemene woordenschat en dat geldt ook voor een flink aantal van hun collega's: wordt Nederland inderdaad, zoals de reclames van de LOI (Leidse Onderwijsinstellingen) suggereren, steeds slimmer?

Als woorden van de vaktaalwoordenschat in de algemene woordenschat terecht komen, komen ze binnen de reikwijdte van de algemene woordenboeken. Mijn bijdrage gaat over de manier waarop deze woordenboeken hiermee omgaan en wat ze er eventueel nog meer mee zouden kunnen doen.

#### **Inleiding**

In januari 2009 startten de Leidse Onderwijsinstellingen (LOI) met een nieuwe reclamecampagne: *Nederland wordt steeds slimmer*. Een van de eerste commercials van deze campagne gaat over twee verkoopsters in een oliebollen

kraam, van wie de ene een doktersromannetje aan het lezen is. De andere verkoopster helpt een klant. Intussen bekritisieren ze vakkundig de door dokter Greg gevolgde behandelmethode. Andere commercials uit deze reeks gingen onder andere over getatoeëerde sportschoolbezoekers die de filosofie van Kant en Schopenhauer bespreken, een vriendenclubje dat zich met bier en chips voor de televisie installeert om aan het nationaal dictee mee te doen, en voetbalscheidsrechters die na afloop van de wedstrijd de mogelijkheid en onmogelijkheid van de tijdens de wedstrijd naar hun hoofd geslingerde ziektes en aandoeningen bespreken. Een van de meer recente commercials speelt op een vrijgezellenavond, waar het optreden van een mannelijke stripper verandert in een medisch consult als een van de toekijkende dames ontdekt dat er iets ernstig mis is met diens rug. (<http://www.youtube.com/watch?v=t8aHclOHCjs>)

Wat deze commercials voor veel mensen leuk maakt, is dat ze gewone mensen in dagelijkse situaties laten zien, die woorden en begrippen gebruiken uit specifieke vakgebieden. Het is uiteraard de bedoeling van de LOI om mensen ertoe over te halen hun cursussen te gaan volgen, zodat ze ook zo slim worden als de mensen in de commercials en mee kunnen praten over de onderwerpen die erin aan de orde komen.

## **Vakwoordenschat, algemene woordenschat en woordenboeken**

Een aantal woorden en begrippen uit de commercials behoort tot de zogenoemde 'vakwoordenschat'. Deze is specifiek voor een bepaald vakgebied en wordt daarbinnen ook redelijk exclusief gebruikt. De betekenis van de vakwoordenschat wordt beschreven in vakwoordenboeken, bijvoorbeeld dat van Pinkhof voor de medische vakwoordenschat. Als deze woorden en begrippen om de een of andere reden ook buiten het vakgebied in gebruik raken, schuiven ze van de vakwoordenschat op naar de algemene woordenschat. Het kan dan zijn dat ze in een soort van tussengebied blijven hangen, waarbij ze niet meer louter tot de vakwoordenschat behoren, maar ook niet echt tot de algemene woordenschat gerekend kunnen worden; maar het gebeurt ook wel dat ze zo goed ingeburgerd raken bij het grote publiek dat ze van de vakwoordenschat in het hart van de algemene woordenschat terechtkomen. Dat is bijvoorbeeld gebeurd met veel termen uit het vakgebied van de informatica, waar zo goed als iedereen tegenwoordig overweg kan met woorden als *downloaden*, *harde schijf*, *muis* en *USB stick*. Deze woorden moeten dan niet alleen meer in de vakwoordenboeken beschreven worden, maar verdienen ook een plaats in de algemene woordenboeken.

Taal is een dynamisch systeem, waarin voortdurend allerlei veranderingen en verschuivingen plaatsvinden. Dat zijn bijvoorbeeld veranderingen op het gebied van de spelling, de betekenis of de gebruiksmogelijkheden van woorden, verschuivingen naar een ander stijlregister of een ander deel van de woordenschat. Makers van woordenboeken houden deze veranderingen en verschuivingen in de gaten en passen – als het goed is – hun woordenboeken steeds aan op de laatste stand van zaken. Dat kan uiteraard korter of langer duren, al naargelang de frequentie van het verschijnen van nieuwe drukken of updates. Met de komst van de digitale lexicografie zijn makers van woordenboeken niet meer gebonden aan beperkingen in ruimte en in tijd. Onlinewoordenboeken kunnen in principe permanent bijgewerkt en geüpdatet worden, en hebben daarnaast een schier onbeperkte ruimte voor de beschrijving van de woordenschat tot hun beschikking.

De keuze van de in de woordenboeken opgenomen informatie en de manier waarop deze wordt gepresenteerd, maken een woordenboek meer of minder geschikt voor bepaalde groepen gebruikers. Hoewel het woordenboek over het algemeen vooral gebruikt wordt om er de betekenis van een woord in op te zoeken, zijn er ook tal van andere gevallen waarin een woordenboek goed van pas kan komen, bijvoorbeeld bij spellingkwesies en om synoniemen op het spoor te komen. Elk woordenboek heeft zo zijn specifieke groep van gebruikers en elk van deze groepen heeft specifieke wensen voor en verwachtingen van het woordenboek dat hij gebruikt. De tijd en de ruimte die de makers van onlinewoordenboeken ter beschikking staat, zouden misschien dus niet alleen gebruikt moeten worden om de laatste stand van zaken met betrekking tot de woordenschat te beschrijven, maar ook om ervoor te zorgen dat verschillende gebruikersgroepen in het woordenboek precies die informatie krijgen waar ze behoefte aan hebben. Wat zouden bijvoorbeeld vertalers in een hedendaags woordenboek van de algemene Nederlandse woordenschat willen vinden?

### **Diabetes en suikerziekte in verschillende woordenboeken**

Ik wil hier aan de hand van het medische woordpaar *diabetes/suikerziekte* laten zien of, sinds wanneer en hoe precies zij in grofweg de laatste honderd jaar in woordenboeken beschreven zijn. Ik heb daarvoor gebruikgemaakt van het *Geneeskundig Woordenboek* van Pinkhof, het *Woordenboek der Nederlandsche Taal* (WNT), een aantal woordenboeken van Van Dale en het corpus van het *Algemeen Nederlands Woordenboek* (ANW). Natuurlijk zijn deze woorden niet representatief voor de hele medische woordenschat, maar ze geven wel een beeld van de veranderingen en verschuivingen die daarin in de loop van de

tijd hebben plaatsgevonden. Aan de hand van de informatie die in de bestaande, weliswaar inmiddels gedigitaliseerde, maar van oorsprong papieren woordenboeken over deze woorden te vinden is, wil ik daarnaast iets laten zien van de extra mogelijkheden die een van aanvang af digitaal opgezet woordenboek zijn gebruikers zou kunnen bieden.

Suikerziekte is een stofwisselingsziekte waarbij het bloed herhaaldelijk een te hoog suikergehalte heeft als gevolg van een tekort aan insuline. De medische term hiervoor is *diabetes mellitus*, maar de gewone benaming *suikerziekte* is in elk geval al in de negentiende eeuw in gebruik. Het vakwoordenboek van Pinkhof (1998<sup>10</sup>) verwijst bij het trefwoord *suikerziekte* naar *diabetes mellitus*. Bij dit laatste trefwoord wordt aangegeven dat het vaak verkort wordt tot *diabetes* en het wordt gedefinieerd als “stoornis in de koolhydraatstofwisseling t.g.v. verstoorde insulinehuishouding, met als kenmerkende symptomen hyperglykemie en glucosurie”.

In de algemene woordenboeken van het Nederlands komt aanvankelijk alleen de Nederlandse benaming van de ziekte voor. De eerste druk van Van Dale uit 1864 (eigenlijk dan nog Calisch en Calisch) vermeldt het woord als synoniem van het kennelijk in die tijd meer courante woord *suikerpis*, dat omschreven wordt als “eene soort van ziekte”, vergezeld van het label *geneeskunde*. Vanaf het begin van de twintigste eeuw verschuift de hoofddefinitie van *suikerpis* naar *suikerziekte* en in de 5de druk van 1914 wordt het woord verklaard als “eene ziekte waarbij veel suikerdeelen in de urine opgelost zijn”. In diezelfde vijfde druk van 1914 vinden we ook voor het eerst de term *diabetes*, die verklaard wordt met het synoniem ‘suikerziekte’, maar verder niet voorzien wordt van enige vorm van labeling.

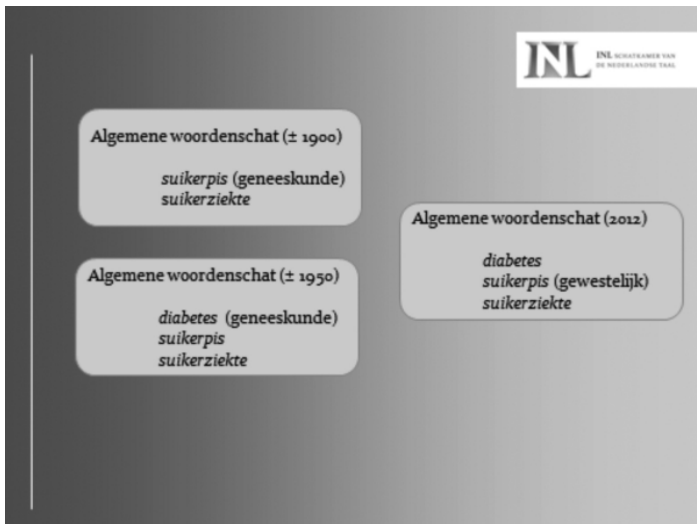
In de loop van de tijd wordt de definitie van *suikerziekte* steeds wat verder bijgeschaafd tot hij in de onlineversie van de huidige (14de) druk uit 2010 2011 geworden is tot “stofwisselingsziekte waarbij het bloed een te hoog gehalte aan suiker (glucose) heeft, als gevolg van een tekort aan insuline”. Als synoniemen worden *diabetes* en *diabetes mellitus* gegeven. *Suikerpis* staat ook nog steeds in het woordenboek, maar wordt nu als gewestelijk gekarakteriseerd. Aan de betekenisomschrijving van *diabetes* is de afgelopen honderd jaar niets veranderd, wel heeft het woord inmiddels gezelschap gekregen van de ingang *diabetes mellitus* die ook van de synoniemdefinitie ‘suikerziekte’ voorzien is.

Het *Woordenboek der Nederlandsche Taal* (WNT) is pas in 1927 toe aan de beschrijving van het woord *suikerziekte*: “Benaming eener stofwisselingsziekte, in hoofdzaak hierin bestaande dat het bloed een te hoog gehalte aan suiker (glucose) bevat, welke ten deele in de urine overgaat, waarin zij gemakkelijk kan worden aangetoond; de wetenschappelijke naam is *diabetes mellitus*.” Voor

het woord *diabetes* is aanvankelijk geen plaats in het woordenboek. Dat wordt pas toegevoegd in 2001 als de *Aanvullingen* op het dan inmiddels voltooide woordenboek verschijnen. Hier worden twee betekenissen onderscheiden: allereerst die van “Ziekelijke toestand gekenmerkt door het regelmatig lozen van abnormale hoeveelheden sterk verdunde urine”, die ter onderscheiding van de volgende betekenis ook wel wordt aangeduid als *diabetes insipidus*, en vervolgens de meer gangbare betekenis “Stoornis in de koolhydraatstofwisseling gekenmerkt door afscheiding van glyucose in de urine en een verhoogd suikergehalte in het bloed (*hyperglycaemie*); suikerziekte”.

In het *Algemeen Nederlands Woordenboek* (ANW), een wetenschappelijk woordenboek van het hedendaags Nederlands, dateren zowel *suikerziekte* als *diabetes* van dit jaar. De definitie van deze woorden is gelijk: “[...] stofwisselingsziekte waarbij het bloed herhaaldelijk een te hoog suikergehalte heeft als gevolg van een tekort aan insuline”, en ze worden in de artikelen dan ook als elkaars synoniem opgevoerd. De woordgroep *diabetes mellitus* is als verbinding bij *diabetes* opgenomen en wordt hetzelfde gedefinieerd als *diabetes* en *suikerziekte*.

In Van Dale Nederlands Duits worden bij het trefwoord *diabetes* de Duitse woorden *Diabetes* en *Zuckerkrankheit* allebei als synoniem opgegeven, met dat laatste woord als ongemarkeerde vorm. Bij *suikerziekte* staan dezelfde woorden en wordt het Duitse *Diabetes* expliciet als vaktaal gelabeld. De Engelse en de Franse woordenboeken van Van Dale kennen allebei slechts één woord voor deze ziekte, het Engelse *diabetes* en het Franse *diabète*. In geen van beide woordenboeken wordt er een label als ‘vaktaal’ aan de trefwoorden of aan de als verklaring opgenomen woorden toegevoegd.





Uit de woordenboeken valt dus op te maken dat *suikerpis* tot ongeveer het begin van de vorige eeuw de algemene benaming voor suikerziekte was. Daarna wordt *suikerziekte* het gewone woord, en vinden we daarnaast de uit de medische vaktaal afkomstige term *diabetes (mellitus)*. De verkorte vorm *diabetes* verschuift in de loop van de tijd zodanig de kant van de algemene woordenschat op, dat deze in de woordenboeken tegenwoordig nauwelijks nog als vaktaal wordt aangemerkt. Verder nemen de Engelse en de Franse woordenboeken dus slechts één woord voor deze ziekte op (*diabetes* en *diabète*), terwijl het Nederlandse (*diabetes* en *suikerziekte*) en het Duitse (*Diabetes* en *Zuckerkrankheit*) er twee hebben, die kennelijk als volledig synoniem worden beschouwd.

### ***Diabetes en suikerziekte in het ANW corpus***

Hoewel vertalers zich bij hun werk vast niet uitsluitend (of misschien zelfs wel helemaal niet) door de hierboven gebruikte woordenboeken laten leiden, kunnen ze in het vertaalproces toch een bepaalde functie hebben. Ze kunnen bijvoorbeeld een rol spelen bij de keuze voor het juiste synoniem. Net zoals *rijwiel* en *fiets* hetzelfde betekenen, maar niet in elke context allebei even goed passen, geldt dat immers mogelijk ook voor *diabetes* en *suikerziekte*. Het is daarom misschien toch aardig om te kijken of woordenboeken hiervoor beter bruikbaar gemaakt kunnen worden. Dit geldt dan naar mijn mening in het bijzonder voor woordenboeken in digitale vorm, die niet gebonden zijn aan beperkingen in tijd en ruimte, en die zo zijn ingericht dat ze flexibel uitbreidbaar zijn en regelmatig geüpdatet kunnen worden.

Een dergelijk woordenboek is het *Algemeen Nederlands Woordenboek* (ANW; <http://anw.inl.nl>), dat momenteel op het Instituut voor Nederlandse Lexicologie gemaakt wordt. Dit is een gestaag groeiend wetenschappelijk woordenboek van het hedendaags Nederlands in Nederland en België, dat min of meer aansluit op het *Woordenboek der Nederlandsche Taal*. De tijdsspanne die het ANW omvat, loopt van ca. 1970 tot 2018, vooralsnog de einddatum van het project. Maar ook na 2018 is het natuurlijk de bedoeling de Nederlandse woordenschat te blijven monitoren en het woordenboek te blijven actualiseren.

Een van de belangrijkste voorwaarden voor een bruikbaar woordenboek van het hedendaags Nederlands is in elk geval een representatief en dynamisch corpus, met een dekkende diversiteit van taalmateriaal, dat voortdurend wordt aangevuld. De gegevens uit een dergelijk corpus kunnen de in het woordenboek beschreven taalfeiten illustreren en ondersteunen. Regelmatig cor

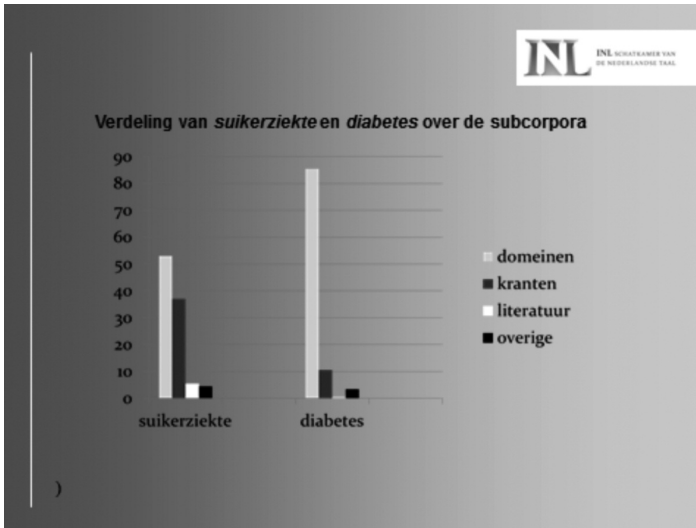
pusanalyse kan ertoe bijdragen dat veranderingen tijdig onderkend worden en dat de informatie in het woordenboek daarop wordt aangepast.

Hoewel het nog niet helemaal ideaal is, komt het corpus dat gebruikt wordt voor de totstandkoming van het *Algemeen Nederlands Woordenboek*, aardig in de richting van wat we hiervoor nodig hebben. Het is een corpus van ruim 100 miljoen woorden, dat dient als basis voor de bewerking van de woordenboek artikelen van het ANW. Het corpus bevat voor het overgrote deel taalmateriaal uit de periode tussen 1970 en 2005. In het kader van dit onderzoek is het corpus onderverdeeld in drie subcorpora: een subcorpus met krantenmateriaal, waarin ook de neologismenverzameling van het INL is opgenomen (45,5 miljoen woorden), een subcorpus met internetmateriaal van vakdomeinen (32 miljoen woorden) en een subcorpus met literair materiaal (20 miljoen woorden). De resterende 5 miljoen zit in het zogenoemde 'pluscorpus' dat hier verder buiten beschouwing wordt gelaten, omdat het deels krantenmateriaal bevat en deels materiaal dat tot het domeinencorpus gerekend kan worden, maar waarvan helaas lastig vast te stellen is wat precies wat is.

Leek een corpus van honderd miljoen woorden begin deze eeuw nog voldoende voor een goede beschrijving van de kernwoordenschat van het hedendaagse Nederlands, inmiddels komen we meer en meer tot de ontdekking dat een betrouwbaar en representatief corpus eigenlijk nog vele malen groter moet zijn, misschien zelfs wel honderd keer zo groot. In elk geval is duidelijk dat een dergelijk corpus voortdurend moet worden uitgebreid en aanvuld met nieuw materiaal om ervoor te zorgen dat het woordenboek ook daadwerkelijk de hedendaagse situatie beschrijft en niet die van enkele decennia geleden. Zoals alles verouderen ook corpora tegenwoordig snel. Zo bevat het ANW corpus bijvoorbeeld nog overvloedig materiaal uit de tijd waarin guldens en franken een rol speelden, terwijl die inmiddels toch wel als verouderde woorden beschouwd kunnen worden. De euro is daarentegen op het moment in het corpus nog ondervertegenwoordigd omdat het toevoegen van nieuw materiaal pas recent is begonnen.

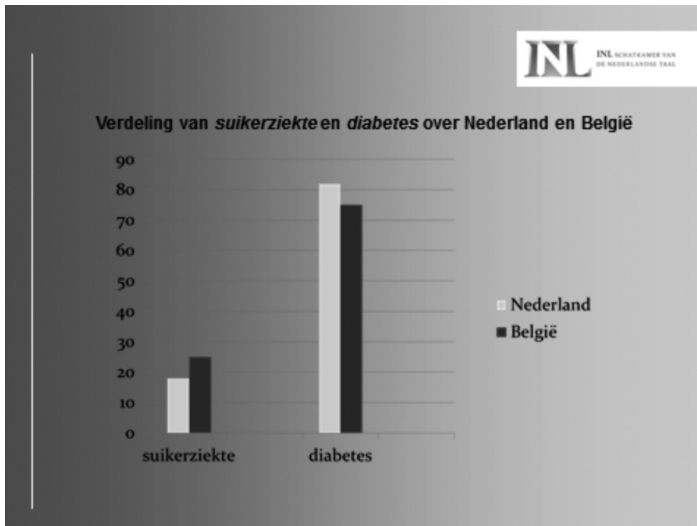
In het ANW corpus komt het woord *diabetes* in totaal 1132 keer voor. De frequentie van *suikerziekte* is veel lager, dat woord is slechts 287 keer aangetroffen. Hieruit zouden we kunnen opmaken dat *diabetes* inmiddels het gangbare woord voor deze ziekte is geworden en dat bij het maken van vertalingen dit woord dus de voorkeur zou moeten genieten. Wanneer we echter kijken naar het voorkomen van de beide woorden in de subcorpora, valt onmiddellijk op dat er een verschil in verdeling is. Van de 1132 voorkomens van *diabetes* bevinden zich er maar liefst 968 in het domeinencorpus, dat is 85,5 procent van het materiaal. Vervolgens vinden we 119 voorkomens in het krantenmateriaal

(10,5 procent) en slechts 6 in het literaire corpus (0,5 procent). De 39 vormen uit het pluscorpus blijven hier zoals gezegd buiten beschouwing.

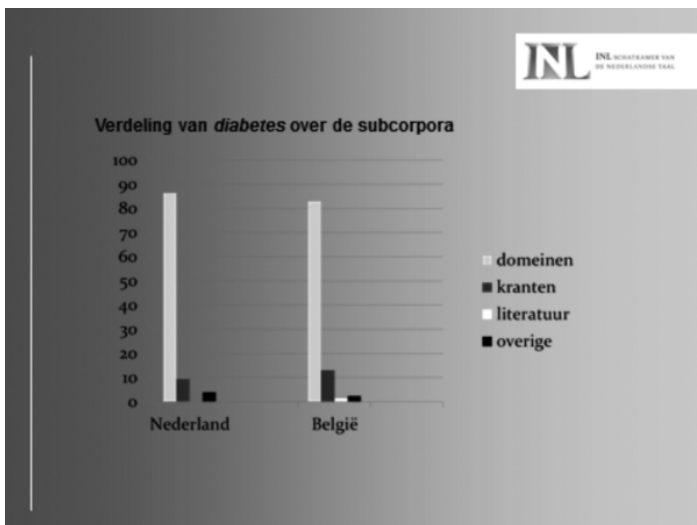


Bij *suikerziekte* zien we heel andere verhoudingen. Hier is weliswaar ook de meerderheid van de voorkomens te vinden in het domeinencorpus (154 stuks, ruim 53 procent), maar de percentages in het krantenmateriaal (107 stuks, ruim 37 procent) en ook in het literaire materiaal (16 stuks, dat is 5,5 procent) geven aan dat dit woord weliswaar significant minder voorkomt dan zijn synoniem *diabetes*, maar kennelijk wel algemener en breder gebruikt wordt. Frequentie zegt dus zeker niet alles over het gebruik van een woord. Ook hier zijn de 10 voorkomens uit het pluscorpus buiten beschouwing gebleven.

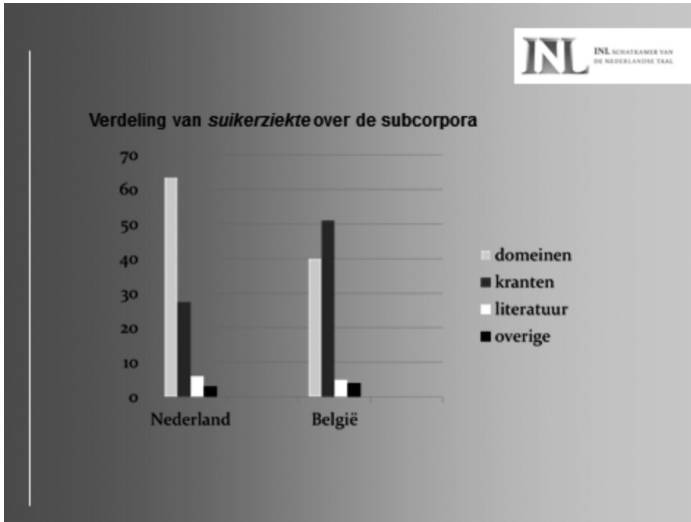
Als we naar de verdeling van beide woorden over de subcorpora kijken, zien we dus dat zowel *diabetes* als *suikerziekte* in alle subcorpora voorkomen. Wel is er een duidelijk verschil in de percentages waarin ze in de verschillen de subcorpora vertegenwoordigd zijn. *Diabetes* is – zoals gezegd – veel frequenter in het domeinencorpus, terwijl *suikerziekte* juist vaker voorkomt in het krantencorpus en het literaire corpus.



Als we het corpus opdelen in een Belgisch Nederlands (BN) en een Nederlands Nederlands (NN) subcorpus en naar de verdeling van de woorden *diabetes* en *suikerziekte* over deze beide subcorpora kijken, zien we dat ze ook hier allebei voorkomen en dat in beide subcorpora *diabetes* het frequentste woord is. Waar de verhouding in het NN subcorpus 82 procent *diabetes* tegen 18 procent *suikerziekte* is, is dat in het BN subcorpus echter 75 tegen 25 procent.



Als we die gegevens combineren en kijken naar de verdeling van het woord *diabetes* over de subcorpora, onderscheiden naar NN en BN herkomst van het materiaal, zien we dat *diabetes* zowel in Nederland als in België in overweldige mate voorkomt in de vakdomeinen, en daarnaast ook wel in de kranten wordt aangetroffen. In de literatuur speelt het woord nauwelijks een rol, wat op zich niet heel verbazingwekkend is.



Voor het woord *suikerziekte* zien we echter een heel ander beeld. Ook dit woord blijkt namelijk behoorlijk vaak in de vakdomeinen voor te komen, in het BN subcorpus in 40 procent van de voorkomens, terwijl dat percentage voor het NN subcorpus zelfs ruim 63 procent is. Een opvallend verschil tussen het NN en het BN materiaal is verder dat *suikerziekte* in Vlaanderen veel vaker in de krant gebruikt wordt, 51 procent van de voorkomens van *suikerziekte* in het BN subcorpus is afkomstig uit krantenmateriaal. In Nederland blijft dit percentage steken op 27,5. In de literatuur blijkt *suikerziekte* zowel in het NN als in het BN materiaal in gelijke mate gebruikt te worden.

## Conclusie

Het zou misschien niet onaardig zijn om gegevens over de frequentie van synoniemen binnen de verschillende subcorpora (Belgisch Nederlands of Nederlands Nederlands, vakdomeinen, literatuur of krantenmateriaal), aan

de bestaande informatie in de woordenboeken toe te voegen. Op die manier wordt immers niet alleen duidelijk welke woorden er beschikbaar zijn voor het vertalen van een woord uit een andere taal, maar ook in welke context ze gebruikt worden en in welke mate.

Op het moment kan deze informatie al wel uit de corpora gehaald worden, maar is ze nog niet beschikbaar in de woordenboeken. Om dat mogelijk te maken en de gegevens daarmee voor een groter publiek beschikbaar te maken, zouden wat aanpassingen verricht moeten worden. Het moet immers een dynamisch systeem zijn, dat telkens weer een representatief beeld van de actuele woordenschat geeft en dat ook bijstelt aan de hand van verschuivingen zoals die van vaktaalwoorden naar woorden uit de algemene woordenschat. De corpora waaruit de gegevens worden gehaald, zouden dus regelmatig moeten worden aangevuld met actueel materiaal uit alle domeinen.

Vanuit het woordenboek zou het vervolgens mogelijk moeten worden gemaakt om op basis van dit materiaal flexibele overzichten te genereren, die voor de verschillende taalvariëteiten van het Nederlands heel precies laten zien welk synoniem in welke omgeving het meest gangbaar is. Op die manier zou er dan een afgewogener keuze gemaakt kunnen worden voor de juiste vertaling van een woord in de context van een medisch handboek, een bijsluiting of een doktersroman.

Het zou volgens mij de moeite waard zijn om een dergelijke voorziening te ontwikkelen, niet alleen voor vertalers, maar ook voor andere taalgerebriekers. Een land dat steeds slimmer wordt, verdient immers woordenboeken die het slimmer maken.

## Geraadpleegde woordenboeken

- Algemeen Nederlands Woordenboek* online (ANW; <http://anw.inl.nl>). Leiden 2009 .
- Woordenboek der Nederlandsche Taal* online (WNT; <http://wnt.inl.nl>). Leiden 2005.
- Pinkhof Geneeskundig Woordenboek*. Tiende, herziene en uitgebreide druk. Houten/Diegem 1998.
- Van Dale, Groot Woordenboek der Nederlandse Taal*. 1ste druk, Tiel [1864] 14de druk, Utrecht, 2005.
- Van Dale Groot woordenboek Nederlands Duits*, Utrecht/Antwerpen [1986].
- Van Dale Groot woordenboek Duits Nederlands*, Utrecht/Antwerpen [1983].
- Van Dale Groot woordenboek Nederlands Engels*, Utrecht/Antwerpen [1986].
- Van Dale Groot woordenboek Engels Nederlands*, Utrecht/Antwerpen [1984].
- Van Dale Groot woordenboek Nederlands Frans*, Utrecht/Antwerpen [1985].
- Van Dale Groot woordenboek Frans Nederlands*, Utrecht/Antwerpen [1984].



# Het terminologieprogramma van de NAVO

## 1. Inleiding

Ook de NAVO doet aan terminologiebeheer. De belangrijkste doelstelling daarvan is normalisatie. Dat is voor de NAVO belangrijk om 'interoperabiliteit' te bereiken. Binnen de NAVO worden drie niveaus van interoperabiliteit onderscheiden: compatibiliteit, uitwisselbaarheid en gemeenschappelijkheid. Onder **compatibiliteit** verstaat men bij de NAVO de geschiktheid van producten, processen of diensten om, onder bepaalde voorwaarden, gezamenlijk te worden gebruikt zonder onaanvaardbare interactie te veroorzaken.

Onder **uitwisselbaarheid** verstaat men bij de NAVO dat bepaalde producten, processen of diensten gelijkaardige eigenschappen hebben en daarom, indien nodig, elkaar kunnen vervangen. En onder **gemeenschappelijkheid** verstaat men bij de NAVO de staat die wordt bereikt wanneer dezelfde doctrine, procedure of apparatuur wordt gebruikt.

Deze normalisatie van terminologie maakt deel uit van de veel bredere inspanningen van de NAVO op dit gebied. De operationele doelstellingen van de NAVO kunnen slechts gehaald worden als de interoperabiliteit – het werken met verschillende systemen, eenheden, landen – gewaarborgd is op alle ondersteunde niveaus, te weten:

*C3: communication, command and control*  
terminologie  
doctrine  
tactiek  
logistiek  
training  
materieel.





Figuur 1: Niveaus van interoperabiliteit

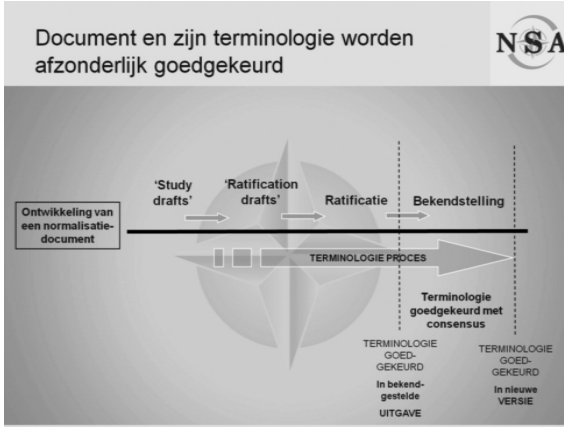
Initiatieven tot normalisatie komen zowel ‘van bovenaf’ (opgelegd vanuit de politiek en/of de NAVO bevelsstructuur), als van ‘onderop’ (vooral: zogenoemde *lessons learned* van operaties).

## 2. Het NAVO terminologieprogramma

Al sinds het begin van de NAVO hebben veel groepen van specialisten hun eigen terminologie op hun eigen manier vastgelegd in allerlei woordenlijsten, glossaria etc. Zo ontstonden terminologische bronnen bij onder andere *Logistics*, *Air and Missile Defence*, *Engineers*, *Civil Emergency Planning* en *Fire Support*, elk met hun eigen manier om de relaties te definiëren tussen verschillende concepten en termen.

In de *NATO Policy for Standardization*, die in het jaar 2000 werd goedgekeurd door de Noord Atlantische Raad, werd voor het eerst een directe verwijzing gemaakt naar terminologie in het verband van de normalisatieactiviteiten van de NAVO. Dit beleidsdocument stelde dat NAVO documenten “must use NATO Agreed terminology” (zie het Noord Atlantisch verdrag s.v.). Deze zin leidde tot invoering van het *NATO Terminology Programme* in de gehele NAVO. De invoering van dit programma heeft geleid tot uniforme regels voor terminologie in de gehele NAVO organisatie, met als gevolg dat de terminologie die wordt ontwikkeld en goedgekeurd via dit programma uiteindelijk de status *NATO Agreed* krijgt. Dit betekent dat iedereen die een NAVO document schrijft, hierbij goedgekeurde terminologie moet gebruiken. De auteur is

thus verantwoordelijk voor zijn terminologie. De terminologie wordt echter via een aparte procedure goedgekeurd, los van de goedkeuring van het document. Figuur 2 toont de verschillende stappen van het terminologieproces.



**Figuur 2: Besluitvormingsproces voor terminologie**

De verschillende stappen in het terminologieproces, zoals in figuur 2 gevisualiseerd, kunnen als volgt beschreven worden.

1. **Vorbereiding en indienen**

Controleer *NATO Agreed* terminologie in het *NATO Terminology Management System* (NTMS): zijn er termen/definities die moeten worden toegevoegd, gewijzigd of verwijderd? Maak op basis hiervan de benodigde voorstellen voor het toevoegen, veranderen of verwijderen en pas de zogenaamde *Guidance* toe. Dit is een document dat gebaseerd is op de ISO normen voor terminologie en beschrijft aan welke eisen een term moet voldoen, hoe een definitie moet worden geschreven enz.

2. **Kwaliteitsbewaking**

Dialogo tussen experts en het *NATO Terminology Office* (NTO).

3. **Goedkeuring**

Wanneer de experts het eens zijn over de inhoud en het NTO akkoord gaat met de vorm, kan de terminologie worden goedgekeurd. Indien zij is goedgekeurd met consensus, wordt de terminologie *NATO Agreed* terminologie en wordt zij verplicht in de gehele NAVO.

#### 4. Bekendstelling

Bekendstelling geschiedt door wijziging in de database (toevoeging wijziging verwijdering). *NATO Agreed* terminologie is verplicht in de gehele NAVO.

Het besluitvormingsproces voor de terminologie is een separaat proces, parallel aan het besluitvormingsproces voor het brondocument. Het moge duidelijk zijn, dat het terminologieproces eerder of later afgerond kan zijn dan dat voor het brondocument.

Afhankelijk van het moment van goedkeuring wordt de terminologie opgenomen in het nieuwe document of in een nieuwe goed te keuren uitgave van het document (als de terminologie wordt goedgekeurd samen met of vóór het document zelf), of (wanneer de terminologie pas wordt goedgekeurd na het document zelf) in een nieuwe versie, die geen nieuwe goedkeuring behoeft.

Hierna zal het terminologieprogramma van de NAVO worden voorgesteld aan de hand van de bovenstaande vier stappen in de goedkeuring van terminologie.

### 3. De verschillende stappen in het NAVO terminologieprogramma

#### Stap 1: Voorbereiden en indienen van terminologie

De eerste stap in de cyclus vraagt om raadpleging van het *NATO Terminology Management System* (NTMS), een terminologiedatabank, die is opgezet als een woordenboek. Hieronder zijn twee schermfoto's van het NTMS. Figuur 3 geeft een indruk van de zoekfaciliteiten, figuur 4 toont een deel van de uitkomsten van een zoekvraag naar de term *target*.




Figuur 3.

LIST OF DEFINITIONS			
ENGLISH	STATUS	Abbr. FRENCH	STATUS Abbr.
<b>target</b>	NA 30 Jun 2003		
1.Chapter: 723.07.15 The electrode in a camera tube that is scanned by the electron beam.			
Source(s): ACOMP-01(B);			
<b>target</b>	NA 9 Jul 1997	<b>objectif</b>	NA 1 Mars 1973 Obj
1.A geographical area, complex, installation or specified unit(s) planned for capture, neutralization or destruction by military forces. <i>Related term(s): objective area.</i> Responsible group(s): MC, MCTC; Source(s): AAP-06;		1.But géographique d'une opération; par exemple : un mouvement caractéristique du terrain dont la conquête, l'occupation ou la défense est essentielle à la manoeuvre du chef.	
<b>target</b>	NA 9 Jul 1997	<b>objectif</b>	NA 9 Juil 1997 Obj
2.In intelligence usage, a country, area, installation, agency or person against which intelligence activities are directed.  Responsible group(s): MC, MCTC; Source(s): AAP-06;		2.Zone géographique, complexe, installation ou unité(s) désignée(s) dont la saisie, la neutralisation ou la destruction par des forces militaires est planifiée. <i>Terme(s) connexe(s): zone de l'objectif.</i>	
<b>target</b>	NA 9 Jul 1997	<b>objectif</b>	NA 9 Juil 1997 Obj
3.In artillery, an area designated and numbered for future firing.		3.Pays, zone, installation, organisme ou individu contre lesquels des activités de	

Figuur 4.


In het NATO Terminology Programme wordt terminologie benaderd op een conceptuele manier, zoals ook door ISO wordt aanbevolen. In figuur 5 staan enkele voorbeelden van concepten die op verschillende manieren in woorden, afkortingen of zelfs pictogrammen kunnen worden geconcretiseerd. Alle concretisering van een concept zullen onder dezelfde noemer verzameld worden. Deze benadering is in principe taalonafhankelijk en is daarom een ideaal uitgangspunt voor de NAVO als tweetalige organisatie.

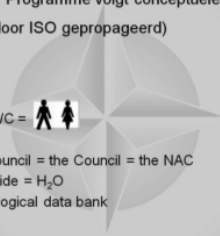


**Vergeet termen! Denk eerst aan 'concepten'!!**

NATO Terminology Programme volgt conceptuele benadering van terminologie (o.a. door ISO gepropageerd)

'Designations':

- toilet = restroom = WC = 
- close = shut
- the North Atlantic Council = the Council = the NAC
- water = hydrogen oxide = H<sub>2</sub>O
- term bank = terminological data bank



Figuur 5.

In mijn presentaties voor diverse groepen leg ik deze benadering nader uit. Ik vertel hun dat het vaak zinvol is om de diverse begrippen of concepten die aan bod komen in een document te ordenen in een begrippendiagram. Een dergelijk schema kan namelijk behulpzaam zijn voor het maken van definities. Af en toe hoor ik dat deze activiteit op zichzelf alleen al als zinvol en verhelderend wordt ervaren.

De definities mogen slechts uit één zin bestaan en mogen alléén die informatie bevatten die helpt het begrip te omschrijven en af te bakenen. De definities bevatten zodoende geen encyclopedische informatie, doctrinaire verklaringen, informatie over procedures etc.

### *Stap 2: Kwaliteitsbewaking*

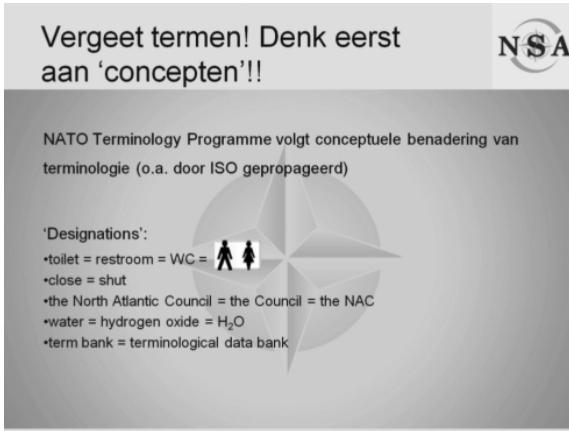
De terminologievoorstellen worden ingediend bij het NATO Terminology Office (NTO) in Brussel. In het NTO controleren wij of het voorstel correct is ingediend als toevoegings, wijzigings of verwijderingsvoorstel. Vooral toevoegingen moeten echte toevoegingen zijn, d.w.z. toevoegingen van nieuwe concepten, omdat het NTMS anders verder vervuild zou raken. Verder controleren we uiteraard of het voorstel in de beide officiële talen (Engels en Frans) is ingediend en zorgen we desnoods voor een vertaling. Ten slotte gaan we na of het voorstel niet in tegenspraak is met andere *NATO Agreed* terminologie. Uiteraard wordt ook bekeken of de bovengenoemde *Guidance* correct en volledig werd toegepast. Tot slot wordt nog gecontroleerd of de voorgestelde terminologie in overeenstemming is met eventueel terminologie van gezaghebbende civiele organisaties en dergelijke.

Deze kwaliteitsbewaking is in essentie een dialoog tussen het NTO en de auteur van de voorstellen. De dialoog is meestal een uitwisseling van e-mails, maar soms proberen we ook in een vergadering door de enorme rijstebrij van voorstellen heen te komen. Dit gaat doorgaans veel sneller en levert voor beide partijen een meerwaarde op.

### *Stap 3 en 4: Goedkeuring en bekendstelling*

Als de experts het eens zijn met de inhoud van het voorstel voor nieuwe terminologie en het *NATO Terminology Office* met de vorm, zijn we tot een consensus gekomen en kunnen de voorstellen officieel worden goedgekeurd op het naast hogere niveau. Als ook daar de voorstellen met consensus zijn goed

gekeurd, is de terminologie *NATO Agreed* en als zodanig verplicht binnen de gehele NAVO.



Figuur 6.

Als de voorstellen met consensus zijn goedgekeurd, is de terminologie *NATO Agreed* en als zodanig verplicht binnen de gehele NAVO. De officiële bekendstelling van de *NATO Agreed terminology* gebeurt door het actualiseren van het onder 1. besproken *NATO Terminology Management System* (NTMS).

#### 4. Tot slot

Tot slot nog een paar woorden over NTMS, de termenbank van de NAVO. NTMS zal worden vervangen door een nieuwe termenbank, *NATOTerm TermOTAN*. We verwachten in de loop van 2014 te kunnen lanceren. Het zal technisch ook mogelijk zijn om andere talen dan het Frans en Engels in deze termenbank op te nemen. De verantwoordelijkheid daarvoor moet worden genomen door het land / de landen waar de desbetreffende taal de officiële taal is. Of er van die mogelijkheid ook wat betreft het Nederlands gebruik gemaakt zal worden, betwijfel ik echter.



# **Terminologie en vertalen**





Bert Keurentjes

## Universitaire vertaaldiensten en hogeronderwijsterminologie

### 1 Inleiding

Het merendeel van de Nederlandse en Vlaamse universiteiten heeft een talen centrum, dat zich vaak vooral richt op het verzorgen van taalcursussen. Een aantal universitaire talencentra heeft daarnaast een vertaal- en redactiedienst. De omvang, opzet en doelgroep van deze dienst is bij elk talencentrum weer anders, maar wat ze gemeen hebben, is dat ze zich voornamelijk bezighouden met het vertalen van Nederlandse teksten in het Engels. Tot voor kort werd door elk van deze universitaire vertaaldiensten op geheel eigen wijze omgegaan met (Engelse) hogeronderwijsterminologie, wat tot uitdrukking komt in het veelvoud aan universiteitsspecifieke termenlijsten die op het internet te vinden zijn. Dankzij verscheidene initiatieven begint er nu echter verandering te komen in deze situatie waarbij elke universitaire vertaaldienst het wiel opnieuw moet uitvinden.

Op nationaal niveau is er de themagroep Vertalen van de vereniging van Nederlandstalige Universitaire Talencentra in Nederland en België, die gedurende de afgelopen jaren jaar zeer actief is geweest in het uitwisselen van kennis, expertise en ervaringen over alles wat te maken heeft met universitaire vertaaldiensten in het algemeen en hogeronderwijsterminologie in het bijzonder. De uitwisseling van terminologie gebeurt vooral via de door Taalcentrum VU opgezette *Higher Education Terms Wiki*. Daarnaast is er onder de vlag van CercleS (Confédération Européenne des Centres de Langues de l'Enseignement Supérieur), een European Translation Focus Group in het leven geroepen. Op korte termijn is het belangrijkste doel van deze groep om een partnerschap op vertaal- en terminologiegebied aan te gaan met het EU project TUNING Educational Structures in Europe.

In dit artikel wil ik een beeld schetsen van de werkzaamheden van universitaire talencentra in het algemeen en universitaire vertaal- en redactiediensten in het bijzonder. Hierbij ga ik met name in op de manier waarop deze diensten omgaan met hogeronderwijsterminologie en bekijk welke ontwikkelingen er op dit moment gaande zijn op dit gebied op nationaal én internationaal niveau.

## 2 Universitaire talencentra en vertaaldiensten

Naar een voorbeeld van een universitair talencentrum hoef ik niet lang te zoeken, daar ik zelf als vertaalcoördinator werkzaam ben bij een dergelijk centrum. En wel bij Radboud in'to Languages, het expertisecentrum voor taal en communicatie van de Radboud Universiteit Nijmegen. Radboud in'to Languages is onderdeel van de Faculteit der Letteren en zoals de meeste universitaire talencentra richt het zich vooral op het verzorgen van taaltrainingen, die in 17 talen worden aangeboden. Andere belangrijke diensten van ons centrum zijn trainingen op het gebied van communicatie en nascholing. In totaal volgen jaarlijks circa 5000 personen een van de circa 400 trainingen bij Radboud in'to Languages.<sup>1</sup>

Vanaf 2011 werkt Radboud in'to Languages samen met het talencentrum van Wageningen UR, dat voorheen bekend stond als Language Services, maar recent werd omgedoopt tot Wageningen in'to Languages. Evenals de Nijmeegse partner houdt Wageningen in'to Languages, dat onderdeel is van het Facilitair Bedrijf van de universiteit, zich voornamelijk bezig met het geven van taaltrainingen.

Zowel Radboud in'to Languages als Wageningen in'to Languages verzort naast trainingen ook vertaal en redactiewerk. Samen handelden de universitaire vertaaldiensten van deze partnerinstituten in 2012 een kleine 1000 opdrachten af. Bij negen van de tien vertaalopdrachten was Engels de taal waarin vertaald werd. Dit kan worden verklaard door het feit dat beide vertaaldiensten voornamelijk opdrachten verzorgen voor Nederlandse medewerkers van de universiteit, die hun teksten vaak in het Nederlands schrijven en deze dan in het Engels vertaald willen hebben, zodat ook buitenlandse medewerkers of studenten van de informatie kennis kunnen nemen. Denk hierbij aan een brief over het aankomende jaargesprek, een universiteitsspecifieke regeling of een facultaire intranetpagina. Vanzelfsprekend schrijven Nederlandse medewerkers ook wel eens teksten in het Engels, die ze vervolgens door een *native speaker* met relevante ervaring en expertise willen laten redigeren. Ook dit kan natuurlijk bij een universitaire vertaal en redactiedienst. Vaak gaat het dan om wetenschappelijke artikelen die een medewerker van de universiteit (of van het universitair medisch centrum) wil gaan indienen bij een *journal*.

## 3 De NUT en de themagroep Vertalen

De universitaire talencentra in het Nederlandse taalgebied komen samen in de NUT, de vereniging van Nederlandstalige Universitaire Talencentra in

Nederland en België. Deze in 1993 opgerichte vereniging telt 23 leden, 5 Belgische en 18 Nederlandse, en één geassocieerd lid in Suriname. Lidmaatschap is niet voorbehouden aan talencentra van universiteiten; op dit moment zijn er ook vier talencentra van hogescholen aangesloten bij de NUT.<sup>2</sup>

De NUT is opgedeeld in een aantal themagroepen, waaronder de themagroep Vertalen. Deze werd in 2004 opgericht en is na een periode van inactiviteit in 2011 weer nieuw leven ingeblazen. De groep heeft een harde kern van circa acht leden, allen vertegenwoordigers van universitaire vertaal- en redactiediensten.

Naast het digitaal uitwisselen van kennis en hulpmiddelen is de voor naamste ambitie van de NUT themagroep Vertalen het nationaal en internationaal invloed uitoefenen op besluitvorming over zaken die met vertalen in het algemeen en (hogeronderwijs)terminologie in het bijzonder te maken hebben. Aan het realiseren van deze ambitie wordt hard gewerkt, onder andere door het verkennen van een samenwerkingsverband met NL Term en de sterke vertegenwoordiging van de themagroep binnen de European Translation Focus Group (TFG), waarover later meer.

## 4 Hogeronderwijsterminologie

### 4.1 Wat is hogeronderwijsterminologie?

Een belangrijke voorwaarde voor het runnen van een succesvolle universitaire vertaaldienst is de consciëntieuze omgang met terminologie in het algemeen en terminologie op het gebied van hoger onderwijs in het bijzonder. Als vertaler *by training and trade* ben ik eerst en vooral een pragmaticus en hecht ik weinig waarde aan een exacte definitie van het begrip 'hogeronderwijs terminologie'. Waar het mij in mijn dagelijkse praktijk als vertaler, proeflezer en coördinator vooral om gaat, is dat terminologie in onze vertalingen op een logische en consistente wijze toegepast wordt. De Nederlandse universiteit met al zijn verschillende procedures en afdelingen is immers al ingewikkeld genoeg, zeker voor een buitenlandse medewerker of student. Het is dan wel prettig als een bepaalde procedure of afdeling in alle Engelstalige communicatie van de universiteit hetzelfde genoemd wordt. Wat overigens ook geldt voor de Nederlandse communicatie, maar daar hebben universitaire vertaaldiensten over het algemeen een stuk minder invloed op, daar wij een Nederlandse tekst vaak pas in handen krijgen als deze al gefinaliseerd – en soms zelfs al gepubliceerd – is.

De meest voor de hand liggende voorbeelden van hogeronderwijsterminologie zijn termen als 'hoogleraar', 'promovendus' en 'hoorcollege'. De Engelse vertaling van dit soort termen, die al sinds jaar en dag gebezigd worden aan de universiteit, is over het algemeen weinig problematisch. Een hoogleraar zal zich in het Engels *professor* noemen, of *full professor* in gevallen waar in hij zich wenst te onderscheiden van *assistant professors* ('universitair docenten') en *associate professors* ('universitair hoofddocenten'). Voor 'promovendus' bestaan er weliswaar meerdere vertalingen (*PhD student*, *PhD candidate*, *junior researcher*, *doctoral candidate*, etc.), maar omdat deze allemaal de lading dekken is het hierbij simpelweg een kwestie van één Engelse term kiezen en je daar (hopelijk) zo veel mogelijk aan houden. Wat voor 'hoorcollege' dan weer in mindere mate geldt: er zullen in Nederland weinig universiteiten zijn die voor een andere Engelse vertaling dan *lecture* kiezen.

Lastiger wordt het op het moment dat er nieuwe onderwijstermen worden bedacht die uitsluitend van toepassing zijn op de Nederlandse situatie, zoals 'bindend studieadvies'. Tot op zekere hoogte lijken de universitaire vertaaldiensten in Nederland een consensus te hebben bereikt over de vertaling hiervan: *binding recommendation regarding the continuation of studies*. Het nadeel van deze Engelse term is natuurlijk dat hij nogal lang is, wat de meeste vertalers zal bewegen om na één keer voluit schrijven te kiezen voor een verkorte vorm. Maar welke? *Binding study recommendation* lijkt voor de hand te liggen, maar er zijn ook vertaaldiensten die naar analogie van het Nederlands kiezen voor *binding study advice*, waarbij een voordeel is dat de afkorting BSA in de Engelse vertaling kan worden gehandhaafd. Ook andere vertalingen zijn natuurlijk mogelijk, zoals afbeelding 3 even verderop in dit artikel aantoont.

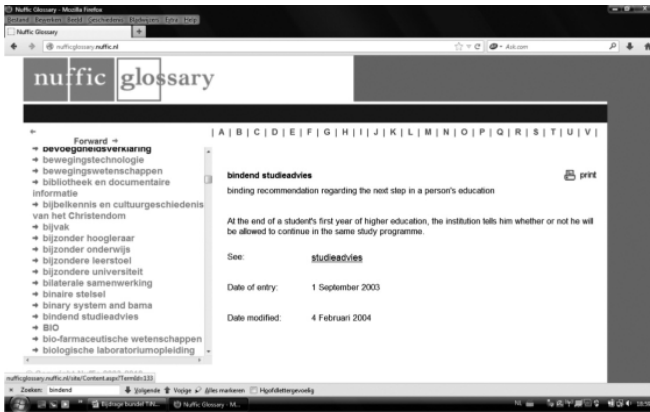
Een laatste voorbeeld dat ik hier wil noemen om te laten zien dat hogeronderwijstermen soms verrassend ingewikkeld kunnen zijn, is de op het oog eenvoudige term 'vak'. Anders dan bij 'bindend studieadvies' beginnen de problemen hierbij al in het Nederlands, omdat de term 'vak' aan veel Nederlandse universiteiten vaak naast synoniemen als 'cursus' en 'module' wordt gebruikt. Niet zelden gebeurt dit binnen één en dezelfde tekst, wat voor een Nederlandse lezer – die waarschijnlijk de typisch Nederlandse voorliefde voor synoniemen deelt met de schrijver – wellicht geen onoverkomelijk probleem is. Maar dit stelt een vertaler wel voor een aantal lastige keuzes, te beginnen met: welke van de gangbare Engelse vertalingen voor 'vak' (*course unit*, *module*, *course*, *subject* etc.) kies ik? En als ik eenmaal een vertaling gekozen heb, moet ik deze dan ook gebruiken voor de Nederlandse synoniemen 'cursus' en 'module' of moet ik het Nederlandse origineel volgen en variëren? Hetgeen extra lastig wordt als de gekozen vertaling voor 'vak' *module* is, omdat het nogal vreemd zou zijn om 'vak' te vertalen als *module* en 'module' als *subject*

of *course*. Even daargelaten dat de wettelijk verplichte Engelse term *course unit* is (waarover later meer), is mijn advies aan vertalers: gebruik een en dezelfde vertaling voor ‘vak’, ‘cursus’ en ‘module’ om verwarring bij de lezer te vermijden. Mits je er zeker van bent dat de schrijver van de Nederlandse tekst deze termen daadwerkelijk als synoniemen bedoelt, natuurlijk.

## 4.2 Terminologielijsten en databases op nationaal niveau

De vorige paragraaf heeft hopelijk aangetoond dat het vertalen van hogeronderwijs termen geen eenvoudige zaak is en bij voorkeur zo consciëntieus en consistent mogelijk moet gebeuren. Een van de eerste pogingen om tot een consistente tweetalige lijst van hogeronderwijs terminologie te komen, werd ondernomen door de Nuffic, de Nederlandse organisatie voor internationale samenwerking in het hoger onderwijs. In 1989 gaf de Nuffic voor het eerst de *Woordenlijsten Nederlands Engels voor hoger onderwijs* uit.<sup>3</sup> Deze papieren uitgave mondde uiteindelijk uit in de publicatie van de online *Nuffic Higher Education Glossary* in 2003.

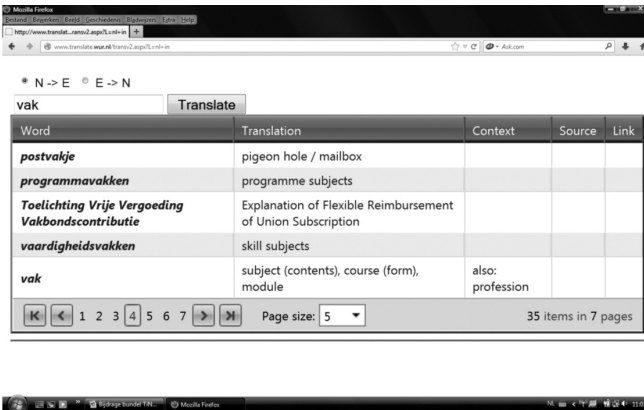
Dat deze termenlijst niet bedoeld is om bepaalde termen op te leggen maar puur als hulpmiddel om onderwijs termen te pláátsen, blijkt wel uit de omschrijving van de lijst (of beter gezegd: de inhoud daarvan) op de site van de Nuffic als “Dutch English terminology for describing higher education in the Netherlands”.<sup>4</sup> En ook de vertaling van een term als ‘bindend studieadvies’ en de verklaring daarbij verraadt het descriptieve karakter van de lijst (zie afbeelding 1).



**Afbeelding 1: Engelse vertaling (*binding recommendation regarding the next step in a person's education*) en definitie (*At the end of a student's first year of higher education, the institution tells him whether or not he will be allowed to continue in the same study programme*) van 'bindend studieadvies' in de Nuffic Glossary**

Een voorbeeld van een meer prescriptieve lijst is de officiële vertaallijst van de Radboud Universiteit Nijmegen. Deze door het International Office beheerde lijst bevat circa 600 veelgebruikte universitaire termen in het Nederlands en het Engels, die goedgekeurd zijn door het college van bestuur. Ten tijde van het schrijven van dit artikel is deze lijst uitsluitend toegankelijk voor medewerkers van de Radboud Universiteit, maar het kan goed zijn dat dit op korte termijn zal veranderen. Er is namelijk sprake van dat het beheer van de lijst in handen zal komen van Radboud in'to Languages, wat ten gevolge zou kunnen hebben dat de lijst voor iedereen toegankelijk wordt gemaakt.

Ook Wageningen UR heeft een eigen tweetalige lijst<sup>5</sup> met termen die binnen de instelling veel worden gebruikt. De huidige lijst stamt in beginsel uit februari 2010, maar wordt regelmatig bijgewerkt. Met meer dan 3000 termen is deze lijst een stuk uitgebreider dan bijvoorbeeld de Nijmeegse lijst, wat vooral komt doordat er destijds voor is gekozen ook algemene HRM gerelateerde termen zoals 'functie' en 'contract' op te nemen, alsmede termen uit het Life Sciences domein zoals 'veenweide' en 'plantenafweer'. Dit heeft als gevolg dat als men bijvoorbeeld op zoek is naar de vertaling van 'vak' en deze term intypt in het zoekbalkje van de database waarin de lijst gegoten is, men zeven pagina's met hits krijgt en er dus even moet worden doorgelinkt om te vinden waar men naar op zoek is (zie afbeelding 2). Voordeel hierbij is natuurlijk wel dat men zo in één moeite door kennis kan nemen van de vertaling van aanverwante termen als 'programmavakken' en 'vaardigheidsvakken'.



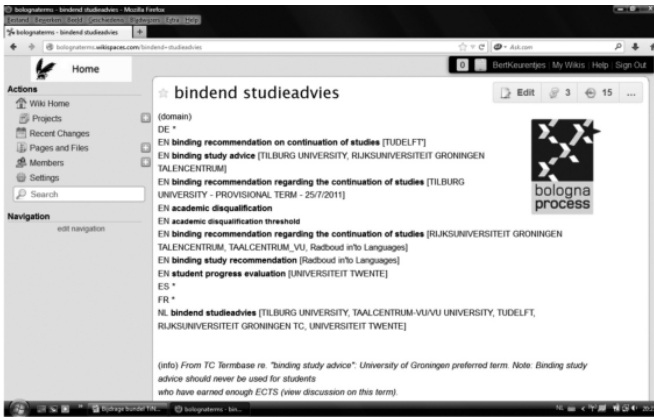
Afbeelding 2: Resultaat van zoekactie 'vak' in de terminologielijst van Wageningen UR

Waar de Nuffic Glossary een beschrijving geeft van hogeronderwijs termen die worden gebruikt, en de lijsten van Radboud in'to Languages en Wageningen in'to Languages voorschrijven welke termen aan de respectieve universiteiten móeten worden gebruikt, doet de terminologiedatabase van de Universitaire Vertaal- en Correctiedienst van de Rijksuniversiteit Groningen een poging om deze uiteinden van het spectrum met elkaar te verenigen. Op de bijbehorende webpagina wordt de volgende achtergrondinformatie gegeven bij de database:

De database bevat vertalingen van termen die binnen de Rijksuniversiteit Groningen worden gebruikt op het gebied van onderwijs en HRM, aangevuld met meer algemene onderwijs termen. Het College van Bestuur van de RUG bepaalt de standaardtermen die binnen de universiteit worden gebruikt, en hun vertalingen. Bij alle termen in de database staat duidelijk aangegeven of het gaat om een terminologiekeuze die is goedgekeurd door het College van Bestuur van de RUG, of dat het een term uit een andere betrouwbare bron betreft.<sup>6</sup>

De in 2009 door Taalcentrum VU opgezette Higher Education Terms Wiki<sup>7</sup> brengt bovengenoemde termenlijsten samen en biedt gelegenheid om termen te bediscussieren. De Wiki bevat op het moment van schrijven meer dan 7000 termen en heeft 146 leden.





Afbeelding 3: Pagina voor ‘bindend studieadvies’ in de Higher Education Terms Wiki

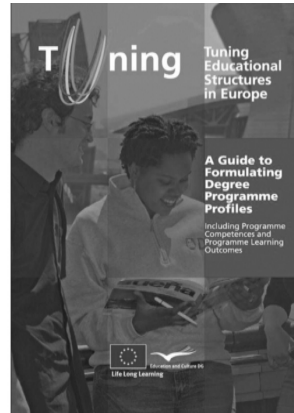
Na een demonstratie tijdens de 2010 editie van de jaarlijkse terminologiedag, georganiseerd voor en door vertalers in het hoger onderwijs, kende de Wiki een korte opleving maar inmiddels is het alweer geruime tijd te stil op het discussieforum. Aangezien het een zeer handig hulpmiddel betreft bij het vertalen van teksten op het gebied van hoger onderwijs en lidmaatschap bovendien kosteloos is, roep ik geïnteresseerden hierbij op zich aan te melden voor de Wiki.<sup>8</sup>

## 5 Hogeronderwijsterminologie op Europees niveau

De NUT kent als nationale overkoepelende organisatie van universitaire talencentra in CercleS zijn Europese tegenhanger. Deze in 1991 opgerichte European Confederation of Language Centres in Higher Education<sup>9</sup> telt inmiddels 290 leden, waaronder dus ook de talencentra die zijn aangesloten bij de NUT.

Ook CercleS kent een aantal themagroepen, waaronder de European Translation Focus Group (TFG), die ongeveer 15 actieve leden heeft. De oprichtingsbijeenkomst van deze Europese vertaalthemagroep vond plaats in november 2011 in Madrid en een jaar later werd in Londen bepaald dat het belangrijkste doel van de TFG het aangaan van een partnerschap met TUNING Educational Structures in Europe<sup>10</sup> is. Dit EU project werd in 2000 opgezet om uitvoering te geven aan het Bolognaproces, waarin wordt gestreefd naar de tot standbrenging van één Europese ruimte voor hoger onderwijs. Dit heeft onder andere geresulteerd in de invoering van de bachelor masterstructuur.

De wettelijk bindende besluiten die TUNING neemt, worden vastgelegd in publicaties die regelmatig online en in gedrukte vorm verschijnen. Deze publicaties worden in het Engels opgesteld en vervolgens in alle officiële talen van de EU (minus het Nederlands) vertaald. En nu het mandaat en het budget van TUNING binnenkort naar alle waarschijnlijkheid met tien jaar verlengd worden en het project ook steeds meer wereldwijde impact begint te krijgen, verschijnen de publicaties ook steeds vaker in niet Europese talen.



Afbeelding 4: Een tweetal TUNING-publicaties

De TFG heeft opgemerkt dat de vaststelling en vertaling van hogeronderwijs termen namens TUNING niet altijd worden verzorgd door taalprofessionals die daartoe het best gekwalificeerd zijn, omdat zij niet noodzakelijkerwijs voldoende vertrouwd zijn met de bron- en doeltaal en de onderwijssystemen in de verschillende Europese landen. Dit heeft tot gevolg dat de (vertaling van) onderwijstermen niet altijd optimaal is; de keuze voor het onhandige *course unit* als officieel Engels equivalent van 'vak'<sup>11</sup> is hier een voorbeeld van. De universitaire vertaaldiensten, aangesloten bij de TFG, willen graag als partner van TUNING fungeren en zo meepraten over de Engelstalige terminologie die vastgesteld wordt en de vertaling hiervan.

Een dergelijk partnerschap zou onder andere bestaan uit het delen en bediscussiëren van termen. Dit zou bijvoorbeeld kunnen gebeuren via een Wiki naar model van de Higher Education Term Wiki. De officieel goedgekeurde terminologie zou vervolgens kunnen worden opgenomen in de terminologiedatabase van de Universitaire Vertaal- en Correctiedienst van de Rijksuniversiteit Groningen, mits er budget beschikbaar wordt gesteld om de database uit te breiden met andere talen.

De projectleiders van TUNING hebben al aangegeven oren te hebben naar een partnerschap met de TFG. Op dit moment is de TFG daarom bezig een lijst van contactpersonen per EU land samen te stellen, die vervolgens gekoppeld kunnen worden aan de vertegenwoordigers van TUNING in het zelfde land. Hopelijk is dit een eerste stap op weg naar een partnerschap dat de kwaliteit en consistentie van de hogeronderwijsterminologie in Europa en de rest van de wereld ten goede zal komen.

## Noten

- <sup>1</sup> [http://www.ru.nl/radboudintolanguages/overons/facts figures/](http://www.ru.nl/radboudintolanguages/overons/facts%20figures/)
- <sup>2</sup> <http://www.nut-talen.eu/documenten/fact%20sheet%20NUT.pdf>
- <sup>3</sup> Warman, M. (1989). *Woordenlijsten (Nederlands Engels) voor hoger onderwijs = Terminology (Dutch English) for higher education in the Netherlands*. Den Haag: Nuffic.
- <sup>4</sup> <http://nufficglossary.nuffic.nl/>
- <sup>5</sup> [http://www.wageningenur.nl/nl/Onderwijs Opleidingen/Wageningen into Languages/Faciliteiten Wageningen into Languages/Vertaalservice/Terminologie.htm](http://www.wageningenur.nl/nl/Onderwijs%20Opleidingen/Wageningen%20into%20Languages/Faciliteiten%20Wageningen%20into%20Languages/Vertaalservice/Terminologie.htm)
- <sup>6</sup> [http://www.rug.nl/science and society/language centre/translation correction and editing/rug terminology database](http://www.rug.nl/science%20and%20society/language%20centre/translation%20correction%20and%20editing/rug%20terminology%20database)
- <sup>7</sup> <http://bolognaterms.wikispaces.com/>
- <sup>8</sup> Aanmelden kan door naar <http://bolognaterms.wikispaces.com/> te gaan en daar te klikken op *Join now* (links bovenin het scherm).
- <sup>9</sup> <http://www.cercles.org/main.html>
- <sup>10</sup> <http://www.unideusto.org/tuning/>
- <sup>11</sup> Aangezien de besluiten van TUNING wettelijk bindend zijn, betekent dit strikt genomen dat alle universitaire vertaaldiensten in Nederland 'vak' móeten vertalen als *course unit*. In de praktijk ziet men echter dat het college van bestuur van een universiteit vaak het laatste woord heeft in de keuze van de Engelse onderwijstermen die aan de universiteit in kwestie worden gebruikt.

Serge Verlinde  
Sara Rymenams

## **De Interactive Language Toolbox, óók voor terminologen**

### **1. Inleiding**

Een bijdrage leveren op een terminologiedag zonder over terminologie te spreken lijkt paradoxaal. Terminologie in gebruik staat echter niet los van de 'algemene' taal en bij de redactie van een 'technische' tekst, rechtstreeks in L1 of als resultaat van een L1 L2 vertaling, speelt die 'algemene' taal hoe dan ook een belangrijke rol. Het raadplegen van diverse lexicografische bronnen, meer en meer via het internet, is bij dit proces dan ook onvermijdelijk. Door het overaanbod dreigt men echter te vervallen in de 'Paradox of choice – Why more is less', naar de titel van een boek van de Amerikaanse psycholoog Barry Schwartz (2004).

De *Interactive Language Toolbox* (<http://ilt.kuleuven.be/inlato>) wil daar om een gebruiksvriendelijke schrijfhulp aanbieden voor het Engels, het Frans en het Nederlands, bestaande uit een *predictive writing aid* en een *corrective writing aid* (Ziyuan, 2012). De toepassing is in eerste instantie bedoeld voor vreemdetaallearners en studenten in opleiding, maar telt zeker ook tal van interessante functies voor meer gevorderde gebruikers.

### **2. De Interactive Language Toolbox als predictive writing aid**

Het *predictive writing aid* luik is opgevat als een lexicale *hotspot* met snelle en gebruiksvriendelijke toegang tot een aantal geselecteerde bronnen op het internet en dit voor drie talen: het Engels, het Frans en het Nederlands (afb. 1).<sup>1</sup>

**Interactive Language Toolbox**  
(Almost) everything you always wanted to know about... texts, words and word combinations in Dutch, English and French

**Take a tour**  
• search for words and word combinations: click here  
• text revision: click here

**Interface**  
• in English  
• en français  
• in het Nederlands

**FAQ**  
click here

**More... (French only)**  
• Discover vocabulary  
• Exercises

**Tip**  
• Words combine together. Discover how.  
• Learn how to find good translations.

**Personal database**  
click here

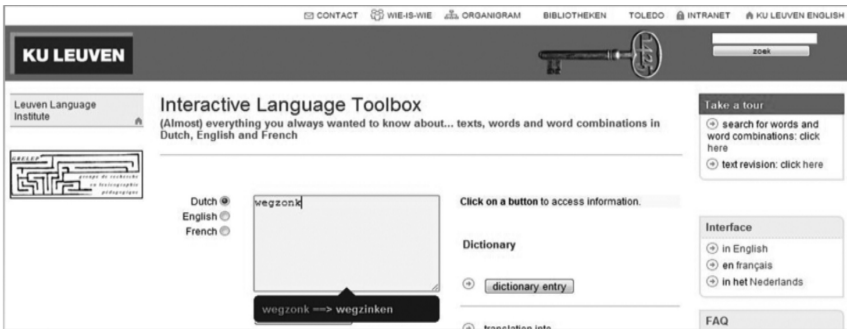
**Accented letters and special characters**  
á é í ó ú  
à è ì ò ù  
â ê ô  
â é î ó ú  
ç ß å ä -

**Try it yourself. copy (ctrl/cmd+c) and paste (ctrl/cmd+v)**

- word
  - prijs price prix
  - omwille\_van in\_order\_to parce\_que
- word combination
  - misleidende reclame misleading advertising publicit  trompeuse
- search for an appropriate word
  - een \* idee an \* idea [ under construction ]
- search for a synonym in a word combination
  - [ under construction ] -great price [ under construction ]
- compare two words or word combinations
  - server opstellen/server opzetten cut prices/reduce prices belle journ e/beau jour
- text
  - Hiermee wil Foucault aantonen dat de moderne maatschappij eugeneticamechanismen toelaat, waaronder het klonen van embryo's die voldoen aan gewenste eigenschappen. In de paragraaf van Foucault werd beschreven dat de mens vastgehecht is aan onzichtbare machtsstructuren, namelijk instituties die de mens disciplineren. Een discipline brengt individualiteit tot stand en bezit het volgend genetisch kenmerk: het controleren van de evolutie van de lichamen door heen de tijd. Een discipline

Afbeelding 1: Homepage *Interactive Language Toolbox*

De toepassing biedt zowel de mogelijkheid om te zoeken naar informatie over woorden en woordcombinaties, als om de frequentie van alternatieven te vergelijken of het gebruik van woorden en woordcombinaties te controleren. Bij het zoeken naar informatie over woorden verschijnen in een pop up venster de woorden uit onze database waarvan de beginletters overeenkomen met de al ingetikte letters (incrementele zoekfunctie, afb. 2). Dit heeft voornamelijk voordelen bij de herkenning van onregelmatige vormen door vreemde taalleiders.



Afbeelding 2: Incrementele zoekfunctie

Afhankelijk van de zoekopdracht verschijnt rechts van het zoekveld ook een reeks knoppen die toegang verlenen tot de beschikbare bronnen (afb. 3).

Dutch   
English   
French

prix

clear search form

reload page

Click on a button to access information.

Dictionary

- dictionary entry

---

translation into

- Dutch
- English

---

meaning

- definition
- words with same meaning
- words with opposite meaning
- words with more general meaning
- words with more specific meaning
- related words

word combinations:

*raise a question, a key question*  
word + ...

- verb
- noun
- adjective
- adverb
- preposition
- all words

word forms : *think/taught, foot/feet*

gender : *le/la, de/het*

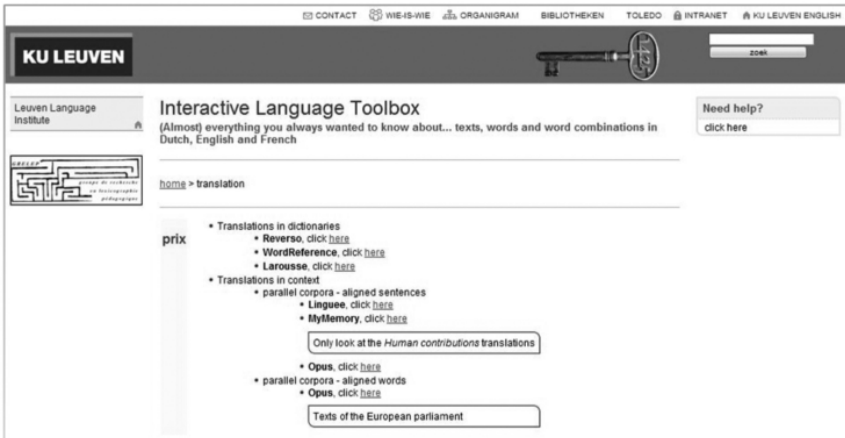
syntax - how to form a sentence:  
*compare with, to something?*

examples (from the web, from literature, ...)

Afbeelding 3: Beschikbare informatie voor een zoekopdracht op een woord

Het aantal beschikbare knoppen is uiteraard het meest uitgebreid wanneer de zoekopdracht op een woord slaat. De knoppen leiden niet alleen naar de web sites met volledige woordenboekartikels, maar ook naar sites die zeer specifieke informatie leveren. In deze gevallen belicht men enkel de informatie uit één specifieke rubriek van een klassiek woordenboek. Uit onze logbestanden is gebleken dat 90% van de gebruikers met die beperkte informatie genoegen neemt en geen verdere, meer algemene informatie meer wenst (Verlinde 2010).

Als de gebruiker doorklikt, verschijnt een scherm met gegevens die wij halen uit onze databases en/of snelkoppelingen naar externe sites. De gebruiker maakt een keuze aan de hand van het menu. Om die keuze eventueel wat te sturen of om de lezing van de externe pagina te vergemakkelijken, geven we daar waar nodig bijkomende informatie over deze pagina's (afb. 4).



Afbeelding 4: Menukeuze bij vertaling van het Frans naar het Engels

De snelkoppelingen dienen om de externe pagina's met het antwoord op de vraag te openen, dus zonder nogmaals de zoekopdracht te moeten invoeren. In tegenstelling tot andere lexicale sites, die informatie van verschillende bronnen verzamelen (<http://www.thefreedictionary.com>), behoudt de *Interactive Language Toolbox* zo de layout van de externe websites. De selectie van bronnen onderscheidt de *Interactive Language Toolbox* dan weer van sites die voornamelijk bronnen opsommen (<http://onlook.com>).



### 3. De *Interactive Language Toolbox* als predictive writing aid voor vertalingen

Traditioneel worden vertaalwoordenboeken gebruikt om teksten te schrijven in een vreemde taal. Iedereen kent de belangrijkste gebreken van dergelijke woordenboeken, zoals de afwezigheid van context, de vermelding van laagfrequente woordcombinaties of zinsneden, en de vaak (te) lange artikels voor courante woorden. De *Interactive Language Toolbox* biedt hier een interessant alternatief door de verwijzing naar parallele corpora die gealigneerd zijn op het niveau van de zin. Dergelijke sites dekken bovendien een veel ruimer deel van het taalgebruik en zijn dus rijker dan vertaalwoordenboeken. Een nadeel van deze sites is dat niet elke gevonden vertaling eenduidig is. De gebruiker wordt dan ook sterk aangeraden om de gevonden informatie te controleren door gebruik te maken van de snelkoppelingen naar Google News en Google Scholar, afhankelijk van het onderwerp en de context (afb. 5).

The image shows a screenshot of a web interface with three distinct sections, each enclosed in a rounded rectangular box. The first section is titled 'Tip!' and contains the text: 'Try to translate sequences of words. If you find a translation for such a sequence of words, it will be more reliable than the combination of translations of single words.' The second section is titled 'Check the context!' and contains the text: 'If you found a translation, check its use in context. No "" needed for a search on multiple words.' Below this text is a form with the label 'check occurrence of' followed by a text input field. Underneath the input field, the word 'on' is followed by a bulleted list of search options: 'Google News (newspaper texts written by native speakers of the language: click here)', 'Google Scholar (texts written by scholars and scientists): click here', and 'the web:'. Under 'the web:', there is a sub-bullet: 'Google: click here'. The third section is titled 'No translation found?' and contains a bulleted list of suggestions: 'Submit your problem to a translation forum' (with sub-bullets 'WordReference.com' and 'en-anglais.com'), 'If the word refers to an object:' (with sub-bullets 'Try to find sites of companies that manufacture or sell this type of object.' and 'Check whether their website is available in different languages.'), and 'If the word is a technical term:' (with sub-bullet 'IATE database (European Union): click here').

Afbeelding 5: Controlefunctie voor gevonden vertaling

Beide bronnen zijn verkozen boven de standaard Google zoekmachine, omdat de teksten in principe geschreven zijn door native speakers (Google News) of van een hoger en meestal betrouwbaarder niveau (Google Scholar) zijn.

Naast de traditionele zoekopdrachten op basis van een woord of een woordcombinatie, biedt de *Interactive Language Toolbox* ook de mogelijkheid om onbekende woorden op te zoeken, zonder een beroep te doen op de vertaling. Eén voorbeeld hiervan is het onderstaande resultaat, verkregen op basis van de eenvoudige zoekopdracht 'a \* idea', verstuurd naar de *Netspeak* website (afb. 6).

**Netspeak** One word leads to another.

an ? idea i x Q

an excellent idea	98,000	21.4%	+
an interesting idea	97,000	21.2%	+
an original idea	40,000	8.7%	+
an old idea	28,000	6.3%	+
an awesome idea	17,000	3.7%	—

What an awesome idea. ... photo ...: Am I the only one? What an awesome idea - Turkish media right now in one image. I said no more photos! ㄹ

SBB.connect. an awesome idea! BUT... Last December, the Swiss national railway company SBB/CFR launched a new mobile application to connect rail ... ㄹ

Saw this on a friends FB wall. What an awesome idea. 2 days ago · 43,027 views · 10.23 GB bandwidth. View comments on reddit · Download full resolution ... ㄹ

It's an awesome idea. rich, creative, colorful. Pages. blog · gentle magic yoga · gentle magic massage · lust for life coaching · my body, my dear friend · art ... ㄹ

PINK Pink Parties - Coming out parties are an awesome idea. 207 likes · 0 talking about this. ㄹ

Funny pictures about What an awesome idea.... Oh, and cool pics about What an awesome idea.... Also, What an awesome idea... photos. ㄹ

an innovative idea	16,000	3.6%	+
an abstract idea	12,000	2.6%	+
an amazing idea	10,000	2.3%	+
an intriguing idea	8,800	1.9%	+
an accurate idea	8,100	1.8%	+
an important idea	8,100	1.8%	+
an initial idea	6,000	1.3%	+
an approximate idea	5,000	1.1%	+
an attractive idea	4,800	1.1%	+
an ingenious idea	4,300	0.9%	+
an inspired idea	3,800	0.8%	+
an overall idea	3,800	0.8%	+
an adequate idea	3,800	0.8%	+
an absurd idea	3,500	0.8%	+
an exciting idea	3,400	0.8%	+
an awful idea	3,400	0.7%	+

Afbeelding 6: Woordcombinaties op de *Netspeak*-website

In een aantal gevallen zijn snelkoppelingen technisch niet realiseerbaar en wordt er enkel verwezen naar de homepage. Het is namelijk belangrijk dat de gebruiker toegang krijgt tot de beste bronnen die op internet te vinden zijn.

#### **4. De *Interactive Language Toolbox als corrective writing aid***

De *corrective writing aid* (beschikbaar voor het Frans en in ontwikkeling voor het Engels en het Nederlands) focust op de revisie van een tekst door de gebruiker. De toepassing scant een tekst op een systematische en geautomatiseerde manier op zoek naar elementen, patronen of sequenties die problematisch zijn of die vaak aanleiding geven tot fouten. Vergeleken met de spelling en grammaticacontrole van een tekstverwerker geeft deze toepassing minder rechtstreekse verbeteringen van fouten, maar trekt ze wel de aandacht van de gebruiker voor een aantal specifieke elementen of passages uit zijn tekst die vaak problemen veroorzaken.

Voor het Frans wordt hierbij vooral gekeken naar de vormcorrectheid vanwege het niveau van de universiteitsstudenten, onze primaire doelgroep; voor het Engels en het Nederlands zal vooral aandacht besteed worden aan stijl en registerkenmerken. Voor elk van de drie talen blijft ook de correctie van lexicale fouten een bijzonder grote uitdaging (Wanner, Verlinde en Alonso Ramos (te verschijnen)).

Voor het Frans verloopt de controle in zes stappen, die deels thematisch bedoeld zijn. Zo worden achtereenvolgens de spelling, de overeenkomst onderwerp – werkwoord, vaak voorkomende zinsconstructiefouten en specifieke problemen rond het substantief en het werkwoord bekeken. De laatste stap behandelt een aantal lexicale items (keuze van voorzetsel en woordcombinaties). Bij elk van de stappen worden de problemen in de tekst geïdentificeerd (linkerkolom) en vindt de gebruiker feedback in de rechterkolom (afb. 7).

Check 1 of 6: words (spelling, morphology, ...)

---

Possible errors found for these items: click on the link(s)

**Les mots**

- [Orthographe \(spelling\), morphologie et influences de l'anglais et du néerlandais](#)

Check 2 of 6

[Start check 2 of 6](#)

Orthographe, morphologie et influences de l'anglais et du néerlandais

**exemple** = correct this word with the information on the right.

**export** = check for possible errors for this word using the information on the right.

Il donne l'**exemple** de l'**export**.

CORRECT YOUR TEXT

• Influences de l'anglais ou du néerlandais

<b>exemple</b>	an example	= un exemple
----------------	------------	--------------

CHECK, IF NECESSARY, CORRECT YOUR TEXT

• Influences de l'anglais ou du néerlandais

<b>export</b>	de export	= l'exportation sauf l'import-export (commerce de produits importés et exportés, DAFA p.311) MA/S exportateur, exportatrice voir aussi: un pays exportateur (exportland), une entreprise exportatrice
---------------	-----------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

**Afbeelding 7: Spellingcontrole voor het Frans**

Zoals men op het voorgaande schermafdruk kan zien, benaderen we de correctie en de feedback vanuit een didactisch standpunt en proberen we te wijzen op de bron van de fout.

Bij een andere stap (vaak voorkomende zinsconstructiefouten) wordt het probleem gekaderd en moet de gebruiker de feedback toepassen op zijn zin (afb. 8).

**Construction des phrases**

[1] Check, if necessary, correct the marked sentence(s) with the information on the right.

[2] A cause de la crise, ils vont pas en vacances.

**CHECK, IF NECESSARY, CORRECT YOUR TEXT**

• [3] Phrase avec négation incomplète (**pas, personne, rien, ... sans ne**)?

A cause de la crise, ils vont pas en vacances.

En français, la négation comporte (au moins) deux éléments : **ne ... pas/plus/jamais/rien/aucun(e)/personne/...**

ex. Il **ne** danse **pas**.

Elle **ne** joue **plus** aux cartes.

Je **ne** bois **jamais** de café avant d'aller dormir.

Nous n'avons **rien** compris.

Elle n'a **aucun** échec.


Je n'ai vu **personne**.

Je **ne** suis **pas** heureux ni malheureux.

Cela **ne** l'intéresse **guère**.

Cela **ne** me gêne **nullement/aucunement**.

Je **ne** le vois **nulle part**.

 Si **rien, personne, aucune** sont employés comme sujet en début de phrase, il faut également les faire suivre par un **ne** de la négation:

ex. **Rien ne** l'arrête.

**Personne ne** l'a reconnu.

**Aucune** photo n'est réussie.

**Remarque:**  
En français parlé, on a tendance à omettre le **ne** de la négation, surtout dans la construction **ne ... pas**:

ex. C'est **pas** vrai.

Je sais **pas**.

Afbeelding 8: Grammaticacontrole voor het Frans

Bij de *Interactive Language Toolbox* volstaat het dus niet om een gegeven suggestie door een muisklik eenvoudig goed te keuren, zoals dit in Word gebeurt, maar moet de gebruiker zijn tekst effectief aanpassen. Hij kan dit wel recht streeks op het scherm doen. De versie van de tekst die in de toepassing verschijnt, is namelijk editeerbaar. Correcties worden zo van de ene stap naar de andere meegenomen. Na de laatste stap kan de gebruiker zijn laatste versie dan eenvoudig kopiëren en plakken in een Word document bijvoorbeeld. Op die laatste pagina verschijnt ook een overzicht van alle problemen die de toepassing geregistreerd heeft.

Een effectanalyse, uitgevoerd bij een grote groep universiteitsstudenten van het eerste jaar, waarvan de resultaten nog gepubliceerd moeten worden, heeft aangetoond dat deze *corrective writing aid* wel degelijk een statistisch sig

nificante impact heeft op de kwaliteit van een tekst. Of deze tool ook daadwerkelijk het leerrendement bevordert, is minder duidelijk. In sommige gevallen, voornamelijk bij elementgerichte correctie en feedback, lijkt dit zo te zijn; bij andere, zeker wanneer het meer om interpretatie gaat, is dit veel minder duidelijk.

## Besluit

De *Interactive Language Toolbox* is een poging om *more is less* om te buigen in *less is more*. Het luik *predictive writing aid* heeft tot doel om de gebruiker zo efficiënt mogelijk te leiden naar de voor hem meest relevante informatie. De tool probeert daarbij zo laagdrempelig mogelijk te zijn en kwaliteitsvolle bronnen aan te bieden. Bij de *corrective writing aid* staat de didactische aanpak voorop en worden items gecontroleerd die vaak aan de bron liggen van fouten, maar veelal onvoldoende aan bod komen in de beschikbare spelling- en grammaticaconcontrole.

Ondanks alle pogingen om deze tools zo transparant mogelijk te maken, is een goede begeleiding van de gebruiker noodzakelijk en moet hij leren ze te gebruiken.

De toepassing is en blijft in ontwikkeling, zowel voor wat betreft het opnemen van nieuwe referentiebronnen als voor het verfijnen van de revisietool, waar in wellicht nog inzichten uit het veld van de NLP alsook massa's lexicale gegevens uit bestaande databanken (bijv. informatie over woordcombinaties) geïntegreerd moeten worden.

## Literatuur

Verlinde, S. en J. Binon (2010). "Monitoring Dictionary Use in the Electronic Age". In: A. Dykstra, T. Schoonheim (eds.). *Proceedings of the XIV Euralex International Congress. Leeuwarden, 6-10 July 2010*. Leeuwarden: Fryske Akademy.

Wanner, L., S. Verlinde en M. Alonso Ramos (te verschijnen). *Writing assistants and automatic lexical error correction: word combinatorics*.

*Netspeak*: <http://www.netspeak.org/>

## Noot

<sup>1</sup> Op de volgende schermafdrucken is de interface te zien in het Engels. Men kan echter ook voor het Frans of het Nederlands kiezen.



# **Terminologie en automatische taaltoepassingen**





# Metadatering & automatische classificatie bij de Tweede Kamer

## 1. Inleiding

De waardevolste goederen van een kennisintensieve organisatie zijn inzicht, kennis, informatie en gegevens. De dragers van inzicht en kennis zijn de medewerkers van de kennisorganisatie: aan de ene kant mensen die op basis van accurate kennis kunnen anticiperen op toekomstige situaties, aan de andere kant mensen die juiste verbanden kunnen leggen tussen beschikbare informatie ten behoeve van die kennis. Beide zijn van vitaal belang. Om die reden besteden kennisintensieve organisaties veel aandacht aan het opbouwen van informatie door het ontsluiten van de beschikbare gegevens.

Het spreekt voor zich dat de keten van informatie, kennis en inzicht leunt op de kwaliteit van de ontsluiting van de gegevens. Fouten en/of onvolledigheid in de ontsluiting leiden tot gebrekkige kennis en falend inzicht.

De Tweede Kamer der Staten Generaal is een dergelijke kennisintensieve organisatie. Dagelijks worden in het parlementaire proces honderden stukken geproduceerd over de meest uiteenlopende onderwerpen. De kern hiervan vormen de parlementaire stukken (Kamervragen, moties, handelingen e.d.), naast publieke informatie, zoals pers en tijdschriftartikelen. De Dienst Informatievoorziening ontsluit deze documenten middels kwalitatief hoogwaardige metadata ten behoeve van niet alleen de eigen organisatie, maar juist ook van het algemene publiek en de journalistiek. Dit alles ter versterking van het democratisch proces.

Knelpunt voor de Tweede Kamer der Staten Generaal is de frictie tussen snelheid en kwaliteit. Het ontsluiten van documenten met kwalitatief hoogwaardige metadata vergt veel doorlooptijd. Hoewel de parlementaire informatie goed toegankelijk gemaakt wordt, duurt het lang voor de informatie beschikbaar is.

Op zoek naar een automatisch hulpmiddel waarmee de informatiespecialisten in kortere tijd hun werkzaamheden kunnen uitvoeren, kwam de Tweede Kamer uit bij GridLine, specialist op het gebied van taal en zoekoplossingen voor het Nederlands. De taal en zoekoplossingen van GridLine helpen organisaties efficiënter en beter te werken met ongestructureerde informa-

tie: documenten, brieven, websites. Met *GridWalker T* biedt GridLine de Tweede Kamer een trefzeker systeem voor automatische classificatie van documenten. *GridWalker T* kent automatisch trefwoorden toe die de gebruiker alleen nog hoeft te controleren.

## **2. *GridWalker T* en automatische classificatie**

De metadatering van parlementaire stukken is een arbeidsintensieve taak. Stukken moeten worden ingedeeld naar ministerie en/of onderwerp; functie en/of politieke partij van betrokkenen vastgesteld; titels, citeertitels, abstracts opgesteld; dossiers gevormd; trefwoorden toegekend etc. *GridWalker T* is een hulpmiddel voor informatiespecialisten bij deze metadatering. *GridWalker T* kent automatisch trefwoorden toe op basis van de tekstuele inhoud. De informatiespecialist kan zich concentreren op de overige zaken en hoeft alleen nog de voorgestelde trefwoorden te controleren. *GridWalker T* is gebouwd met de open source zoekmachine Lucene, waarop extra Nederlandse taaltechnologie uit GridLines TaalServer is toegepast. De TaalServer voegt modules toe voor herkenning van de Nederlandse morfologie, samenstellingen, woordgroepen, spelling, tikfouten, OCR fouten (fouten in de optische tekenherkenning), synoniemen, afkortingen, naamherkenning en zo nog meer. Om die reden alleen al was *GridWalker T* een logische keuze voor de Tweede Kamer.

Het doel van de automatische classificatie is tweeledig. Het hoofddoel is een versnelling van de ontsluiting. De onderwerpen die in parlementaire stukken aan bod komen, zijn divers.

Om al deze onderwerpen met trefwoorden te kunnen ontsluiten, heeft de Tweede Kamer een uitgebreide thesaurus opgesteld: de Parlementsthesaurus bevat meer dan 7.500 termen en is meer dan 10 niveaus diep. De informatiespecialist heeft bij de ontsluiting van stukken met een of meer trefwoorden dus de keus uit vele termen uit een complexe structuur.

Automatische classificatie geeft trefwoordsuggesties zodat de informatiespecialist uit een klein aantal gerichte suggesties de gewenste trefwoorden kan kiezen. Het arbeidsintensieve onderdeel waarbij de informatiespecialist moet kiezen uit duizenden diverse termen, is vervangen door de keuze uit een handvol gerichte suggesties.

Daarnaast zorgt automatische classificatie voor eenduidigheid. Waar binnen de ene groep van informatiespecialisten richtlijnen zijn opgesteld over de wijze van toekenning, zijn die niet gedeeld met andere groepen. Ook heeft niet iedere groep evenveel kennis over hetzelfde deel van de thesaurus.

Automatische classificatie geeft een eenduidige classificatie. Hoe *GridWalker T* een tekst classificeert, is gebruikersonafhankelijk.

Natuurlijk verandert met de tijd ook de toepassing van richtlijnen, onder andere door verloop van medewerkers en veranderende omstandigheden. Automatische classificatie borgt de kennis waarmee geclassificeerd wordt. De automatische classificatie kan met de tijd opnieuw getraind, afgestemd en ingericht worden om aan veranderende omstandigheden tegemoet te komen. Vanaf dat moment is de kennis geborgd en beschikbaar voor de toekomst.

### **3. Inrichting automatische classificatie**

Automatische classificatie van documenten staat niet op zichzelf, net zo min als zoeken of andere toepassingen van taaltechnologie. Om documenten automatisch te kunnen classificeren zijn koppelingen met bestaande systemen nodig, binnen een bestaande IT infrastructuur en niet in de laatste plaats binnen bestaande werkprocessen. Automatische classificatie vergt specialisten in taaltechnologie, inrichting van automatische classificatie IT specialisten. GridLine heeft juist daarom beide groepen in dienst: geen taaltechnologie zonder IT.

Zoals alle producten van GridLine is *GridWalker T* gebaseerd op een open architectuur: koppelingen via open standaarden zoals XML, REST, SOAP. Een open architectuur maakt het mogelijk een nieuwe service op te nemen in de IT infrastructuur zonder invloed op overige systemen. Voor de Tweede Kamer classificeert *GridWalker T* meer dan 700.000 parlementaire stukken tegen een parlementsthesaurus van meer dan 7.500 termen. Alleen de invoerapplicatie waarin de metadata worden geregistreerd, is bekend met *GridWalker T*. De applicatie krijgt nu direct geverifieerde trefwoorden uit *GridWalker T*. Alle overige systemen in de IT infrastructuur zijn ongewijzigd gebleven. Voor die systemen lijkt niets te zijn veranderd.

De introductie van een nieuwe applicatie mag de bestaande werkprocessen niet verstoren. Juist deze Open Architectuur zorgt ervoor dat bestaande werkprocessen behouden kunnen blijven. Omdat de onderliggende systemen niet of nauwelijks gewijzigd hoeven te worden, kunnen de bestaande werkprocessen behouden blijven. Daar staat tegenover dat de introductie van automatische classificatie en de open architectuur wel een versimpeling van de systemen kan betekenen. Werkprocessen die in het verleden opgelegd zijn vanuit bestaande systemen, kunnen nu tegen het licht gehouden en opnieuw ingericht worden.

Een voorbeeld als de automatische classificatie voor de Tweede Kamer laat zien hoe belangrijk dit aspect is. Aanvankelijk werd gevraagd naar automatische classificatie van één gegevensbron met parlementaire stukken. Gedurende dat traject werd al gevraagd naar automatische classificatie van een tweede en een derde gegevensbron, met andersoortige stukken. Op korte termijn zal nu een datawarehouse (datapakhuis) ingericht worden, waarin alle gegevensbronnen met parlementaire stukken beschikbaar komen. Dankzij de open architectuur van *GridWalker T* hebben deze veranderingen geringe invloed op de automatische classificatie.

#### **4. Voorwaarden voor succes**

Om automatische classificatie tot een succes te maken moet aan een aantal voorwaarden worden voldaan. Ten eerste moet de klant geloven in de mogelijkheden van taaltechnologie. Automatische classificatie vergt training en tuning. De eerste resultaten zijn doorgaans matig. Evaluatie van de resultaten is belangrijk. Indien er een 'gulden standaard' beschikbaar is waarlangs geëvalueerd kan worden, is automatische evaluatie mogelijk. Omdat met automatische classificatie ook eenduidiger classificeren wordt nagestreefd, ontbreekt vaak zo'n gulden standaard. De classificaties uit het verleden zijn immers niet per se representatief voor de nieuwe classificaties. Dientengevolge is een handmatige evaluatie door informatiespecialisten nodig. Zij kunnen inhoudelijk beoordelen of de automatische classificatie goede of zelfs betere trefwoorden heeft gegeven.

Ten tweede, is zoals hierboven al is beschreven, taaltechnologie zonder IT niet mogelijk. De leverancier van de automatische classificatie moet in staat zijn de applicatie in te passen in de bestaande infrastructuur. Dat vraagt om specialistische kennis op het gebied van IT architectuur, naast die van taaltechnologie, en een open architectuur.

Ten derde zullen gebruikers de automatische classificatie alleen dan accepteren indien de applicatie aan hun behoeften aangepast kan worden. Een nieuwe applicatie mag aan de ene kant de bestaande werkprocessen niet omgooien en moet aan de andere kant verbeteringen in de werkprocessen kunnen ondersteunen.

En ten laatste, maar minstens even belangrijk, moet de automatische classificatie ondersteund worden door kennis van de taal. De automatische classificatie moet dankzij taaltechnologie in staat zijn vervoegingen, spelfouten, tikfouten, OCR fouten, samenstellingen, synoniemen, namen e.d. te herkennen.

## 5. Toekomst

Mijn professionele carrière begon in 1995 met de bouw van een zoekmachine. Voor die toepassing gebruikten wij specifieke apparatuur van een parallelle computer op krachtstroom. Zoekoplossingen waren een noviteit en alleen al vanwege de kosten schaars. Sindsdien hebben de kosten van de apparatuur en de rekenkracht een dergelijke vlucht gemaakt dat alleen daarom al tegenwoordig de taaltechnologische mogelijkheden zoveel groter zijn. Die groei zal alleen maar verder doorzetten.

Ik verwacht dan ook dat de toepassingen van automatische classificatie in de nabije toekomst breder en breder zullen worden. Automatische classificatie is allang niet meer voorbehouden aan organisaties met de omvang van de Tweede Kamer. Het is slechts een kwestie van tijd dat de organisaties zich dat ook zelf realiseren.



# Terminologie in automatische taaltoepassingen

## 1. Introductie

Wie pakt er nog een woordenboek bij twijfel over de schrijfwijze, betekenis of vertaling van een moeilijk woord of een vakterm? En wie maakt nog gebruik van een trefwoordenregister om informatie op te zoeken in een boek of document? Bij velen hebben deze papieren hulpmiddelen al lang plaatsgemaakt voor snellere digitale alternatieven. Maar zelfs het raadplegen van een online woordenboek wordt in deze digitale tijd al gauw als oponthoud ervaren.

De toekomst is aan slimme applicaties, die professionals op hun wenken bedienen zonder dat ze hun *workflow*, hun werkstroom, hoeven te onderbreken: de tekstschrijver wil tijdens het schrijven van een brief of beleidstekst met een druk op de knop feedback ontvangen op spelfouten, naamgebruik, (vak)termkeuze en leesbaarheid, zonder dat hij allerlei huisstijlgidsen moet doorbladeren. En als de tekstschrijver een thema wil bespreken waarvoor hij specifieke informatie uit een bestaand document nodig heeft, verwacht hij dat de zoekmachine deze tekstpassage direct voor hem kan terugvinden aan de hand van een kenmerkende vakterm, zelfs als die vakterm er niet in voorkomt. Voor een deel is dit al bestaande praktijk.

Dit soort toepassingen vereist de inzet van geavanceerde taaltechnologie, die automatisch kan worden afgestemd op het informatiedomein waar de toepassing voor bedoeld is. Hoe beter de beschikbare terminologische bronnen, des te beter het resultaat. Hier ligt dan ook een veelbelovend toepassingsgebied van de terminologie.

Traditioneel is terminologie het vakgebied dat zich bezighoudt met de woordenschat in de professionele en semiprofessionele praktijk. Vóór het digitale tijdperk was het nut van de terminologie vooral gelegen in de mogelijkheid om via systematisch opgezette termenlijsten en bijbehorende betekenisdefinities een bijdrage te leveren aan eenduidige communicatie tussen professionals binnen een organisatie of vakgebied, ook buiten het eigen taalgebied (via terminologische vertaallijsten, bijvoorbeeld in Europees verband). Daarbuiten bleef de praktische toepassing van de opgebouwde termenlijsten meestal beperkt tot het aanleggen van trefwoordregisters voor de ontsluiting van studietoelagen of bibliotheken. In het laatste geval was meestal sprake van een



taxonomie van vaktermen, of soms zelfs van een thesaurus (netwerk van systematisch gerelateerde vaktermen met hyperoniemen, hyponiemen, synoniemen, antoniemen en overige relaties).

Met de komst van het digitale tijdperk heeft de terminologie talloze nieuwe toepassingsmogelijkheden gekregen, al zijn veel bedrijven en organisaties hier nog niet goed mee bekend. Vooral bij domeinspecifieke toepassingen van taaltechnologie is een aanzienlijke kwaliteitswinst mogelijk als gebruik kan worden gemaakt van systematisch opgezette termenlijsten of thesauri. Denk bijvoorbeeld aan zoekmachines (al dan niet thesaurusgebaseerd), systemen voor automatische documentclassificatie, voorleessystemen, informatiebeheer, verwijsystemen (bijv. kenniskaarten), automatische spellingcontrole en systemen voor woordkeuzeadvies.



## 2. Terminologietoepassingen bij GridLine

Als grootste aanbieder van zoek- en taaltoepassingen voor het Nederlands heeft GridLine veel ervaring met de hier genoemde toepassingen. Bij tal van projecten is inderdaad een kwaliteitssprong gemaakt door de inzet van bestaande of nieuw afgeleide termenlijsten. Voor domeinen waar nog geen (goede) termenlijst voorhanden is, biedt GridLine de mogelijkheid om “met een schone lei” een nieuwe termenlijst te creëren door middel van automatische termextractie. Zie de volgende paragraaf voor nadere uitleg.

De eindgebruiker krijgt vervolgens de mogelijkheid om de ruwe resultaatlijst te beoordelen en op te schonen door de woorden die geen term zijn eruit te halen, en de overgenomen termen nader te classificeren en van extra kenmerken te voorzien. Dit proces van termselectie en verrijking ondersteunt GridLine met een gespecialiseerde beheerapplicatie voor woordenlijsten, genaamd Woordenlijstbeheer.

Woordenlijstbeheer is ook geschikt voor diverse andere beheertoepassingen rondom terminologie. Hieronder volgt een opsomming van de diverse mogelijkheden:

termsselectie als onderdeel van een termextractietraject (hierboven al genoemd);  
 het inventariseren, vergelijken en toekennen van betekenisdefinities;  
 het bekijken en aanpassen van concepten en conceptrelaties in een thesaurus;  
 het beheren, combineren en uitbreiden van termenbanken (zowel eentalig als meertalig);  
 het beheer van Klinkende Taalwoordenlijsten voor standaardisering of vereenvoudiging van moeilijke woorden, vaktermen en organisatieterminen.

Deze toepassingen worden beschreven in paragraaf 5: *Toepassingen van Wordenlijstbeheer*. Om dit goed te kunnen plaatsen is het praktisch om eerst wat uitleg te geven over automatische termextractie; dit gebeurt in paragraaf 3 en 4. Tot slot wordt in paragraaf 6 een applicatie beschreven die automatische afstemming mogelijk maakt van de in dit artikel geïntroduceerde taaltoepassingen: de Kwaliteitsbox. Tussendoor wordt ook een aantal praktijktoepassingen beschreven, steeds na de algemene uitleg over het product of de productfunctie waar ze op gebaseerd zijn.

### 3 Termextractie met GridWalker X

Speciaal voor klanten die een representatieve termenlijst willen opbouwen via automatische termextractie, heeft GridLine de applicatie GridWalker X ontwikkeld. Op een taalkundig ontlede tekst kan GridWalker X een keur aan extractiemethodes toepassen om een termenlijst af te leiden. Voor de taalkundige vooranalyse maakt deze applicatie gebruik van de GridLine TaalServer; dit is een webservice (autonome applicatie die door andere applicaties in het net werk kan worden aangeropen voor simpele of ingewikkelde analyseopdrachten) die volautomatisch allerlei soorten taalkundige analyses kan uitvoeren op Nederlandstalige teksten, zoals tokenizing (opdelen in woorden en zinnen), POS tagging (bepalen van grammaticale woordsoorten), lemmatisering (bepaling van het grondwoord), compound analyse (opdeling van samenstellingen), chunking (bepaling van woordgroepen), Named Entity Recognition (naamsherkenning) en constructieanalyse.

Voor het opbouwen van een allround inzetbare termenlijst levert corpuscontrastering de beste resultaten op. Hierbij wordt een door de klant of branche aan te leveren corpus van vakdocumenten vergeleken met een algemeen refe

rentiecorpus. Welk referentiecorpus het beste werkt, hangt van de toepassing af.

Bij de toepassing Klinkende Taal (adviesstelsel voor het meten en verbeteren van de leesbaarheid van een tekst) wordt bijvoorbeeld gebruik gemaakt van een deelverzameling van het Corpus Gesproken Nederlands, omdat dit een goede benadering geeft van de algemeen bekende woordenschat en dus de kans verhoogt dat er bij de contrastering termen worden gevonden die gewone mensen als moeilijk ervaren. Voor het aanleggen van een representatieve termenlijst (inclusief eenvoudige vaktermen) kan weer beter gebruik worden gemaakt van een zo algemeen mogelijk geschreven corpus. Dit geeft een beter beeld van de termfrequenties in het meer formele taalgebruik, zodat een betere maat wordt verkregen voor het vaststellen van de gewoonte van vaktermen in de algemene taal.

De via GridWalker X verkregen termenlijsten kunnen pas in gebruik worden genomen als ze door deskundige redacteuren (vaktaalspecialisten binnen de organisatie) zijn verwerkt. Dit is een van de toepassingen die wordt ondersteund door Woordenlijstbeheer.

#### **4. Praktijkttoepassing: vaktermen verzamelen met de Termtreffer**

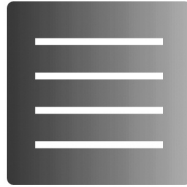
De Termtreffer is een eindgebruikersapplicatie voor termextractie en terminologiebeheer die GridLine voor de Nederlandse Taalunie heeft ontwikkeld. De Termtreffer biedt een allround oplossing voor onder meer vertalers en terminologen die regelmatig vaktermenlijsten nodig hebben voor hun werkzaamheden. Het handmatig aanmaken van een dergelijke lijst door de analyse van bedrijfsdocumenten of door expert interviews is erg duur en tijdrovend. Dit gaat veel efficiënter met de Termtreffer. Hiermee kan de gebruiker automatisch domeinspecifieke termen en trefwoorden verzamelen vanuit documentcollecties.

Desgewenst kan de gebruiker de verzamelde termen semi automatisch expanderen (woordvormen toevoegen) en grammaticaal verrijken door integratie van informatie uit bestaande referentiebestanden, zoals digitale woordenboeken. Onderliggend maakt de Termtreffer gebruik van de automatische termextractiemethodes van GridWalker X (zie introductie).

De gebruiker kan kiezen uit een groot aantal standaardmethodes, zowel statistische als regelgebaseerde. Het is ook mogelijk om concordanties aan te maken. Het gebruik van de Termtreffer vereist wel het nodige inzicht in de werking en de specifieke voor- en nadelen van de beschikbare extractiemethodes. Bovendien zal altijd een redactioneel validatieproces nodig zijn, dat effi

ciënter verloopt als de gekozen extractiemethode goed aansluit bij de doelstelling van de termenlijst. Begeleiding door een expert zal dan ook bijdragen aan de bruikbaarheid van de extractieresultaten.

De afgeleide termenbanken kunnen eenvoudig worden doorzocht, gecombineerd en onderhouden met een geavanceerde omgeving voor woordenlijstbeheer. De gebruiker kan ook termen in context aanstrepen. De functionaliteit van deze module is gebaseerd op de module Woordenlijstbeheer.



WOORDENLIJSTBEHEER

## 5. Toepassingen van Woordenlijstbeheer

Als specialist in zoek- en taaltoepassingen voor het Nederlands voert GridLine vaak opdrachten uit voor levering waarbij een representatieve termenlijst nodig is om een acceptabel kwaliteitsniveau te bereiken. Dit speelt met name bij de domeinspecifieke inzet van zoek- en taaltoepassingen, wat bij veruit de meeste leveringen het geval is. Zoals is toegelicht in paragraaf 3, Termextractie met GridWalker X, kan GridLine de opdrachtgever hierin faciliteren door een termenlijst op te stellen via automatische termextractie.

Omdat terminologie geen statische zaak is, maar zich voortdurend verder ontwikkelt, is het essentieel dat een dergelijke termenlijst (of een complexere variant ervan, bijvoorbeeld een thesaurus) na aanmaak onderhouden wordt. Hiervoor heeft GridLine een speciale applicatie ontwikkeld: Woordenlijstbeheer. Hieronder worden vijf actuele toepassingen van Woordenlijstbeheer beschreven, in een aantal gevallen wordt de beschrijving gecombineerd met een korte schets van een of meer praktijktoepassingen.



## 5.1 Toepassing 1: Klinkende Taal

Klinkende Taal helpt professionals met het schrijven van begrijpelijke teksten. Zij kunnen met dit programma snel nagaan of hun tekst niet te moeilijk is voor de doelgroep. Ook vertelt Klinkende Taal wat de tekstschrijver hieraan kan doen. Zo markeert het moeilijke woorden en/of vaktermen (met verbeter suggesties), lange zinnen en alinea's en allerlei lastige zinsconstructies. Hiernaast kan het ook spelfouten aanstrepen. Klinkende Taal is onder meer leverbaar als plug in in Microsoft Office Word.

Met Woordenlijstbeheer kunnen klanten de voor hun organisatie ingestelde woordenlijsten bekijken, wijzigen en aanvullen. Zo kunnen zij Klinkende Taal beter laten aansluiten op de voertaal binnen hun organisatie. Dit gaat gemakkelijker als de organisatie al over eigen standaardlijsten met te vermijden woorden en de bijbehorende voorkeustermen of namen beschikt.

Er zijn beheermodules beschikbaar voor de woordenlijsten van de modules Dure Woorden, Vaktermen, Organisatienamen en Spellingcontrole. Verder is er een module waarmee de beheerder nieuwe woorden en woordalternatieven kan toevoegen, bijvoorbeeld voor woorden en vaktermen die door gebruikers van de Word plug in zijn aangemeld.

Voor de modules Dure Woorden en Vaktermen zijn ook standaardlijsten beschikbaar. De standaardlijst voor Dure Woorden is voor alle klanten hetzelfde. Bij Vaktermen zijn meerdere branchelijsten beschikbaar, waaronder Overheidsterminologie, Farmaceutische terminologie, Ziekenhuisterminologie, GGD terminologie en Woningcorporatie terminologie. Er komen regelmatig nieuwe standaardlijsten bij. Deze lijsten ontstaan in nauwe samenwerking met de klanten die erom gevraagd hebben.

Met Woordenlijstbeheer kunnen klanten de lijsten voorbereiden (zie toepassing 2) en na oplevering aanvullen en onderhouden. Woordenlijstbeheer biedt ook mogelijkheden voor het beheer van woordenlijsten in andere talen en voor meertalige termenlijsten. Dit is bijvoorbeeld nuttig voor Klinkende Taal toepassingen bij tweetalige organisaties of de ondersteuning van tekstschrijvers met een anderstalige achtergrond.

### 5.1.1 *Praktijkvoorbeeld: begrijpelijke geneesmiddelinformatie*

Stichting Health Base schrijft begrijpelijke geneesmiddelinformatie voor de patiënt (GIP). Deze informatie is een aanvulling op de bijsluiters van de fabrikant. De bijsluiters van de fabrikant is een wettelijk document en bevat daar

door (te) veel informatie.

De wettelijke bijsluiter is volgens patiënten vaak niet leesbaar en niet toepasbaar. Het gevaar is dat patiënten hierdoor medicijnen niet goed gebruiken. De GIP moet daarom leesbaar en toepasbaar zijn. Dit bereiken we door Klinkende Taal in te zetten. Het meet het taalniveau en helpt je om dit taalgebruik te vereenvoudigen.

Een extra druk op de leesbaarheid van de GIP komt door gebruik van medische termen. Met de aan Klinkende Taal gekoppelde vaktermbeheermodule lukt het om de medische termen zoveel mogelijk te vervangen. Dit levert een nog leesbaarder tekst op voor de patiënt, waardoor het medicijngebruik verbetert.

### 5.1.2 *Praktijkvoorbeeld: begrijpelijke GGD terminologie*

De Gemeentelijke Geneeskundige Dienst (GGD) Gelre IJssel gebruikt vaak geschreven tekst om onze doelgroepen te bereiken. Dat zijn uitnodigingen voor onderzoeken, brieven en folders, maar ook via de website richten wij ons tot de mensen. Door heldere taal te gebruiken is meteen duidelijk waar het over gaat. Ouders begrijpen waarom, waar en wanneer hun kind voor een onderzoek opgeroepen wordt, publiciteitscampagnes leveren meer effect op en rapporten zijn prettiger te lezen en te begrijpen.

Bij de GGD werken veel hoogopgeleide mensen: academici en onderzoekers. Voor hen is het vaak lastig om in eenvoudige taal hun boodschap over te brengen. Klinkende Taal helpt hierbij, onder meer door moeilijke GGD termen aan te strepen. De GGD vaktermenlijst is tot stand gekomen door automatische termextractie toe te passen op de website van GGD Gelre IJssel. Dit onderdeel is uitgevoerd door GridLine. Vervolgens heeft de communicatie afdeling van GGD Gelre IJssel de lijst redactioneel doorgenomen m.b.v. Woordenlijstbeheer en de bruikbare termen onderverdeeld in verboden en toelaatbare vaktermen.

### 5.1.3 *Praktijkvoorbeeld: het Taalweb Nederlands Friese schrijfhulp voor vaktermen*

In de loop der jaren is er een zekere norm voor het geschreven Fries ontstaan. Die norm is niet officieel en vaak nogal los, zelfs in de woordenboeken van het Fries. Om de taalnorm te versterken wordt er voor het eerst in de geschiedenis van het Fries gewerkt aan een tegenhanger van het Groene Boekje: de Standertwurldlist Frysk. De Standertwurldlist is de basis voor het online Taalweb dat bij de Fryske Akademy ontwikkeld wordt.

Het Taalweb bestaat uit een bijzondere spellingcontrole, een automati

sche vertaler Nederlands – Fries en vice versa, en een woordenboekportaal. Een van de onderliggende doelen van het taalweb is om het schrijven van ambtelijk Fries te bevorderen. Door het Taalweb wordt gedemonstreerd hoe terminologie en taaltechnologie kunnen worden aaneengesmeed tot een handige schrijfhulp.

## 5.2 Toepassing 2: termen selecteren en verrijken

GridLine biedt organisaties de mogelijkheid om snel termenlijsten op te bouwen door middel van automatische termextractie. Hiervoor heeft GridLine GridWalker X ontwikkeld. GridWalker X kan een ruwe termenlijst (of namenlijst) afleiden uit een representatieve verzameling documenten of webpagina's van het werkterrein waarvoor de terminologie in kaart moet worden gebracht. In een dergelijke ruwe termenlijst zitten doorgaans enige woorden die ten onrechte de status vakterm hebben gekregen. Deze kunnen er alleen uit worden gehaald door menselijke redacteurs. Die redacteurs zijn ook nodig om spelvorm en grammaticale kenmerken te controleren en om synoniemen en betekenisdefinities toe te kennen.

Als de woordenlijst bestemd is voor Klinkende Taal, moet de redacteur voor alle geselecteerde termen aangeven of het om moeilijke (verboden of toelaatbare) woorden gaat, en bij elk moeilijk woord een of meer geschikte alternatieven opgeven. In het geval van organisatienamen moet de redacteur bij elke naamvorm aangeven wat de juiste schrijfwijze is. De applicatie Klinkende Taal gebruikt deze informatie om moeilijke woorden of fout geschreven namen aan te strepen en bruikbare alternatieven voor te stellen.

Woordenlijstbeheer heeft een speciale werkomgeving om dit redactie proces te faciliteren: de termselectieomgeving. Voor Klinkende Taallijsten werkt het als volgt: een redacteur (of groep redacteurs) loopt alle woorden langs. Voor iedere termkandidaat geeft de redacteur een gestandaardiseerde beoordeling, bijvoorbeeld *geen vakterm* (eventueel met de opmerking "geen goed woord" of "fout gespeld"), *verboden vakterm* ("vakterm die beter vermeden kan worden"), *toelaatbare vakterm* ("vakterm die de eerste keer van uitleg moet worden voorzien") of *vrij te gebruiken vakterm* ("woord dat erkend wordt als vakterm, maar dat zonder verdere uitleg gebruikt kan worden").

Wanneer alle redacteurs klaar zijn met hun beoordelingstaak, berekent de applicatie automatisch wat de gemeenschappelijke voorkeuren zijn. De resulterende lijst bevat per woord een overzicht van welke beoordelingen de redacteurs gegeven hebben, en welke woorden na weging van deze oordelen in

aanmerking komen voor opname in de eigen Klinkende Taallijst. Hier kan vervolgens een test mee worden gedaan in een proefomgeving met de aangepaste versie van de Klinkende Taalcontrole. Met Woordenlijstbeheer kunnen de redacteuren alle woorden uit deze selectie eenvoudig van alternatieven voorzien, desnoods voorzien van commentaar voor andere redacteuren. Zo kunnen zij efficiënt een complete organisatielijst opbouwen.

Bij de opbouw van een onderwerpenlijst voor documentclassificatie kan de redacteur volstaan met het selecteren van geschikte onderwerpsnamen en het controleren van de spelling. Zo heeft elke woordenlijsttoepassing haar eigen beoordelingscategorieën.

### 5.2.1 *Praktijkvoorbeeld: begrijpelijke overheidsterminologie*

Begin 2011 heeft het Algemeen Management Team van de gemeente 's Hertogenbosch bepaald dat publieksteksten voortaan in Klare Taal de deur uitgaan. Het gaat om brieven en e mails aan inwoners en teksten op de gemeentepagina's in de Bossche Omroep en op de website. Taalniveau B1 is daarbij de norm. Ter ondersteuning van dit besluit kregen alle medewerkers de Klinkende Taal Word plug in. In 's Hertogenbosch is de plug in uitgebreid met extra taalmodules voor het nakijken van teksten op het gebruik van (on)heldere overheidsterminologie en spelfouten (met verbeteringsuggesties).

De module overheidsterminologie is in nauwe samenwerking met de gemeente 's Hertogenbosch en het taaladviesbureau Texplain tot stand gekomen. Als eerste stap heeft de gemeente een representatief taalcorpus aangeleverd, namelijk een selectie van haar website. GridLine heeft hier automatische termextractie op toegepast. Vervolgens hebben enkele communicatieadviseurs de ruwe termenlijst doorgenomen in de termselectiemodule van Woordenlijstbeheer en onderscheid gemaakt tussen niet termen, vrij te gebruiken termen, toelaatbare termen en verboden termen. Alle verboden termen en toelaatbare termen zijn bovendien van direct bruikbare woordalternatieven of uitleg voorzien.

Deze lijst is vervolgens als optioneel selecteerbare standaardlijst toegevoegd aan de Klinkende Taal Server, met de gemeente 's Hertogenbosch als eerste gebruiker. Tal van andere overheidsorganisaties zijn inmiddels gevolgd, zowel lokaal als op ministerieniveau. Hierbij heeft elke organisatie de mogelijkheid individuele aanpassingen door te voeren in de standaardlijst.



### 5.3 Thesaurusbeheer

Woordenlijstbeheer biedt gebruikers van GridWalker T (en andere thesaurus browsers) de mogelijkheid om thesaurusbeheer uit te voeren. Een thesaurus is simpel gezegd een woordenlijst waarbij de woorden door hun betekenis onderling gerelateerd zijn. Dit kunnen verschillende relaties zijn, bijv. synoniem (straat = weg), of hyperoniem hyponiemen (fruit > banaan).

In het thesaurusbeheer moeten gebruikers relaties tussen termen kunnen zien. Dus van een woord moeten bijvoorbeeld alle synoniemen opgevraagd kunnen worden. Gebruikers moeten ook relaties tussen woorden kunnen aangeven. Woordenlijstbeheer faciliteert dit en biedt de thesaurusbeheerder een extra overzicht door de thesaurus te visualiseren met behulp van de grafische thesaurusweergave in GridWalker T. In die weergave kan de gebruiker gemakkelijk van term naar term browsen en alle bijbehorende eigenschappen, metadata en relaties bekijken. Deze toepassing van Woordenlijstbeheer kan eenvoudig worden geïntegreerd in websites en zoekmachines die gebruikmaken van de GridWalker T thesaurusbrowser.

#### 5.3.1 *Praktijkvoorbeeld: de Aquo objectcatalogus van Informatiehuis Water.*

Het gaat om een publieke website met zoek- en beheeromgeving van de Aquo objectcatalogus van het Informatiehuis Water. Dit is een uitgebreide thesaurus van water- en waterschapstermen die gebruikt wordt voor de onderlinge afstemming van het terminologiegebruik bij de waterschappen. Zie: <http://aquo.gridwalker.nl/>

### 5.4 Woordbetekenissen bepalen

Bij het aanmaken van termenlijsten is het vaak lastig om de definitie van de opgenomen termen te bepalen. Woordenlijstbeheer kan dit proces op de volgende manieren ondersteunen.

Concordanties aanmaken: inventarisatie van alle zinnen waarin de term voorkomt, desgewenst onderverdeeld in equivalentieklassen op basis van abstracte zinspatronen (analyse van grammaticale en lexicale kenmerken, bijvoorbeeld alle zinnen waarin term X deel uitmaakt van een predicatie over object X, ofwel: een werkwoord in combinatie met een lijdend voorwerp X).

Alle bestaande definities opsporen door raadpleging woordenboeken/indexen en door opsporing van definities in lopende tekst met behulp

van signaalwoorden.

Definities aanmaken door middel van informatie extractie (opsporing van specifieke grammaticale patronen voor de aanmaak van definities in termen van thesaurusrelaties, bijvoorbeeld van het type A is een B, A is een deel van B, etc.).

Bij elk van deze analysemethodes biedt Woordenlijstbeheer de mogelijkheid om de term waarvoor de analyse wordt uitgevoerd direct van een bestaande of nieuw ingevoerde definitie te voorzien, en te koppelen aan de onderliggende voorbeelden. Desgewenst kan de gebruiker ook kiezen voor een definitietem plate die gebaseerd is op een ISO norm of metadata systeem (bv. Dublin core).

## 5.5 Automatische verrijking van woordenlijsten

Woordenlijstbeheer kan worden ingezet voor het beheer van complete terminologiebanken en de verrijking (zowel semi automatisch als redactioneel) van de hierin opgenomen vaktermen (inclusief namen). Woordenlijstbeheer biedt bijvoorbeeld ondersteuning bij de controle en aanvulling van woordvormvarianten bij een gegeven term, zowel ten aanzien van de spelling (toelaatbare spellingvarianten; gebruik streepjes en hoofdletters; wel/niet afkorten) als van de grammaticale woordvormen. Hierbij worden woorden in de lijst automatisch geanalyseerd door de Taalserver. Bij het aanvullen van de grammaticale woordvormen zijn twee mogelijkheden:

### 1. Automatisch lemma's bepalen

Voor ieder woord in de woordenlijst (standaard of eigen) bepaalt de Taalserver automatisch wat het lemma is (basisvorm, 'bananen' 'banaan'). Dit lemma wordt het lemma van het woord, het originele woord wordt als variant (verbuigingsvorm) opgeslagen in de kenmerken van het lemma.

### 2. Automatisch verbuigingsvormen bepalen

Voor ieder woord in de woordenlijst (standaard of eigen) bepaalt de Taalserver automatisch alle verbuigingsvormen ('banaan' 'banaan', 'bananen', 'banaantje', 'banaantjes'). Deze worden als varianten opgeslagen in de kenmerken van het lemma.

Bij de opbouw van meertalige woordenlijsten kan Woordenlijstbeheer ook automatisch vertalingen en de bijbehorende woordkenmerken toevoegen

door koppeling van bestaande vertaalwoordenboeken. Deze functies zijn beschikbaar in de Termtreffer (zie sectie 4).

## 6 KwaliteitsBox

De KwaliteitsBox is GridLines gereedschapskist voor het automatisch evalueren van Nederlandstalige zoek-, taal-, en classificatiesystemen. De Kwaliteitsbox maakt een hoger niveau van kwaliteitsmanagement mogelijk voor producten als Klinkende Taal (automatisch schrijfadvisen en termstandaardisatie), GridWalker ES (Enterprise Search), GridWalker T (documentclassificatie en thesaurusbeheer) en de GridLine Taalserver (analyse van documenten). De overeenkomst tussen deze producten is dat ze gebruikmaken van taaltechnologie, dus dat er een proces is rondom taaldata en vakterminologie.

Omdat deze producten heel verschillende toepassingen ondersteunen, zijn er uiteenlopende evaluatiemethodes nodig. Zo moet bij zoekmachines als GridWalker ES een ranking geëvalueerd worden. Bij GridWalker T moet vastgesteld worden of het systeem documenten juist classificeert. En bij Klinkende Taal moet worden gecontroleerd of verbeteringen in de automatische herkenning van de taalkenmerken geen ongewenste verschuivingen oplevert in het berekende taalniveau: producten die bij veel klanten in gebruik zijn, moeten immers een continue kwaliteit leveren, ook na upgrades van de Taalserver en de producten die hier gebruik van maken.

De Kwaliteitsbox evalueert de kwaliteit van de output door deze te toetsen aan een gold standard, een gulden standaard, dat wil zeggen een dataset die door menselijke redacteuren is samengesteld en gevalideerd. Deze toetsing levert informatie op over de dekkingsgraad en de precisie van het geteste product. Zo kan tijdens de ontwikkeling van een nieuwe productmodule gecontroleerd worden of dit tot betere prestaties leidt op de gold standard. Hiernaast kan de Kwaliteitsbox worden gebruikt voor de validatie van zoek-, taal- en classificatiesystemen op nieuwe, organisatiespecifieke taaldomeinen, met specifieke vakterminologie en/of huisstijl voorschriften. Dit vereist dat de klant een gold standard aanlevert voor zijn eigen toepassingsdomein.

Bij een zoekmachinetoepassing is bijvoorbeeld een lijst nodig met exact gedefinieerde zoekopdrachten en de hierbij verwachte resultaten (= documenten die hoog gerankt moeten worden). Bij een classificatietoepassing is een lijst van documenten nodig met de hierbij verwachte trefwoorden uit het toe te passen trefwoordsysteem. En Klinkende Taal kan desgewenst op specifieke genres afgestemd worden door een specifieke verzameling tekstvoorbeelden aan te leveren met leesbaarheidsoordelen van redacteuren of lezers.

Voor elk van bovengenoemde toepassingen zijn al evaluatiemodules beschikbaar. De volgende stap is om deze modules samen te brengen in een systeem voor *continuous monitoring*. In *continuous monitoring* wordt een verandering aan een product automatisch getest, zodat iedere nieuwe versie van het product ook een verbetering is. Dit vereist de ontwikkeling van de juiste, instelbare metrieken voor elk type applicatie en anderzijds het ontwikkelen van een generiek protocol waarmee verschillende applicaties hun uitvoer aan de KwaliteitsBox kunnen communiceren. Het uiteindelijke doel is om de Kwaliteitsbox naadloos te laten aansluiten op het ontwikkelproces van de bestaande GridLineproducten, zodat de automatische kwaliteitscontrole er standaard deel van uitmaakt.

### **6.1 Praktijktoepassing: workbench voor terminologische classificatiesystemen**

Informatiediensten gebruiken vaak geavanceerde trefwoordsystemen voor de ontsluiting van hun boeken en documenten. Het toekennen van de trefwoorden was lange tijd handwerk, maar tegenwoordig kan dit ook automatisch. Om na te gaan in hoeverre een classificatiesysteem de juiste termen toekent aan de tekstdocumenten is het van belang om de classificaties te evalueren. Omdat handmatige evaluatie tijdrovend is, heeft GridLine i.s.m. Christophe Deloo (stagiair TU Delft) een workbench, een werkbank, ontwikkeld voor het evalueren van classificatiesystemen. De workbench omvat zowel traditionele binaire evaluatietechnieken als een techniek die classificaties beoordeelt aan de hand van conceptafstanden, gedefinieerd door een thesaurus.

De evaluatiemethode werkt als volgt. Stel dat het classificatiesysteem de term 'energiebron' toegekend heeft in plaats van 'windenergie'. Traditionele technieken zouden deze toekenning simpelweg beoordelen als fout. De techniek gebaseerd op conceptafstanden zal de toekenning juist als vrij goed beoordelen gezien de grote mate van gerelateerdheid tussen de termen. Deze methode is toegepast bij de evaluatie van GridLines systeem voor document classificatie bij de Tweede Kamer. Dit systeem zorgt voor automatische trefwoordtoekenning aan Tweede Kamerstukken met termen, afkomstig van de Parlementsthesaurus.<sup>1</sup> Bij de evaluatie is nagegaan in hoeverre de automatisch toegekende termen verwant waren aan de termen die vermeld staan in de *gold standard*.

## 7. Tot slot

Dit artikel is geen losstaand verhaal, maar is gebaseerd op de presentaties die samen de interactieve sessie vormden van de TiNT dag 2012. Deze presentaties waren opgezet als korte demonstraties van producten en bijbehorende klanttoepassingen, waarbij het publiek de presentatie kon sturen via vragen. De bedoeling is de gepresenteerde onderwerpen in werking en achtergrond te schetsen en toe te lichten. Het gaat specifiek om de volgende presentaties

Woordenlijstbeheer Oele Koornwinder (GridLine);

Aquo objectcatalogus: praktijktoepassing van Woordenlijstbeheer bij Informatiehuis Water, Oele Koornwinder (GridLine)

Overheidsterminologie: praktijktoepassing van Woordenlijstbeheer en Klinkende Taal bij gemeente 's Hertogenbosch, Martijn van der Wind (Texplain)

GGD terminologie: praktijktoepassing van Woordenlijstbeheer en Klinkende Taal bij GGD Gelre IJssel, Erna de Vries (GGD) / Oele Koornwinder (GridLine) Terminologie in medische bijsluiters: praktijktoepassing van Woordenlijstbeheer en Klinkende Taal bij Health Base, Channah Zijlstra (HealthBase);

Het Frysk Nederlandse Taalnet: praktijktoepassing van Klinkende Taal en Woordenlijstbeheer bij de Fryske Akademy, Anne Dykstra (Fryske Akademie) / Dennis de Vries (GridLine);

Automatische termextractie met GridWalker X; demonstratie met behulp van de NTU Termtreffer, Dennis de Vries (GridLine);

Praktijktoepassing van de Kwaliteitsbox: Workbench voor evaluatie classificatiesystemen, Christophe Deloo (TU Delft/GridLine).

Gezien de belangstelling van het publiek is dit een goede manier gebleken om iedereen die direct bij terminologie en automatisering betrokken is, professionele toelichting te geven.

## Noten

<sup>1</sup> Zie het artikel over metadatering bij de Tweede Kamer van Mart Trautwein voor een nadere toelichting.

# **Veldenquête terminologie**



Hennie van der Vliet

## **U over NL Term**

### **De resultaten van de bezoekersenquête op de TiNT dag 2012**

De vereniging NL TERM is een belangenvereniging die zich wil inzetten voor mensen die in een professionele omgeving met vaktaal omgaan. Dat roept meteen de vraag op wie die professionals dan zijn, welke hun belangen zijn en hoe die belangen behartigd kunnen worden. Om daarachter te komen houdt de vereniging een enquête onder ruim twintig mensen die volgens de bestuursleden op professionele wijze met terminologie omgaan. De resultaten worden nu bestudeerd en zullen binnenkort gepubliceerd worden.

Daarnaast zijn op TiNT dag 2012 de bezoekers bevraged. De resultaten van deze bevraging worden in deze bijdrage gepresenteerd.

#### **1. NL Term zoekt haar achterban**

De professionals waarop NL TERM zich richt, zijn in eerste instantie de experts uit de vakgebieden. Dat zijn in principe alle mensen die een vaktaal spreken: artsen, metaalbewerkers, sportleraren, noem maar op. De vaktaal is hun gereedschap en de leden van NL TERM vinden het belangrijk dat die vaktaal een goed stuk Nederlandstalig gereedschap is. In de praktijk zal slechts een deel van die mensen zich bewust zijn van het belang van de vaktaal die zij spreken. Anderen, zoals vakdocenten, houden zich wel bewust met die vaktaal bezig. Al die mensen behoren tot de doelgroep van NL TERM. In tweede instantie zijn de professionals mensen die zich beroepsmatig met communicatie bezighouden en daarbij met vaktaal in aanraking komen. Dat zijn bijvoorbeeld journalisten, voorlichters en vertalers, maar ook informatici die informatie uit teksten of andere tekstuele data halen of op andere wijze taal digitaal verwerken. Uiteindelijk zijn er dus zeer, zeer veel gebruikers van vaktaal.

Het zal duidelijk zijn dat de verschillende groepen professionals op heel verschillende wijze met hun vaktaal omgaan. Voor de expert is de vaktaal als de schoenen dat zijn voor een wandelaar: in het beste geval ben je er geen moment van bewust. Dat ligt anders voor de vertalers en voorlichters. Zij zijn zich zeer bewust van hun taalgebruik. De informatici zijn zich ook zeer bewust van de taal, maar op een specifieke, technische manier. Kortom, er zijn



misschien wel net zo veel manieren om met vaktaal om te gaan als er professionele gebruikers van vaktaal zijn.

Als er zoveel gebruikers van vaktaal en zoveel manieren van gebruik zijn, wat zijn dan de belangen van de gebruikers en hoe ga je die behartigen? NL TERM heeft zich sinds haar oprichting vooral op de vertalers gericht. De oprichters en bestuursleden zijn goed bekend met de vertaalwereld: veel bestuursleden van NL TERM komen uit de vertaalwereld of zijn op een andere manier in lexicale verschillen tussen talen geïnteresseerd. Vertalers zijn de gebruikers die de meeste problemen ondervinden bij het werken met vaktaal en de bestuursleden zijn zich dan ook zeer bewust van de problemen die vertalers met vaktaal ondervinden.

Hoewel met name de laatste jaren nadrukkelijk ook naar bestuursleden uit andere doelgroepen is gezocht en die ook zijn gevonden, blijft de vertaalwereld in de terminologie en in NL TERM toch dominant aanwezig. Dat is niet erg, de vertaalwereld toont ook de meeste interesse in de vereniging. Toch wil de vereniging nadrukkelijk haar blik verruimen. De doelstelling van de vereniging is uiteindelijk de belangen te behartigen van alle gebruikers van vaktaal en dat zijn er véél meer dan alleen de vertalers.

Het is moeilijk een zeer grote en diverse doelgroep te bedienen die je deels nauwelijks kent. Om de diversiteit van de doelgroep recht te doen heeft de vereniging haar doelstellingen in vier tamelijk algemene punten uitgewerkt. NL TERM staat voor de bevordering van:

- de samenwerking binnen de terminologie als vakgebied;
- de bekendheid en het belang van de terminologie in het Nederlandse taalgebied;
- de deskundigheid op het gebied van het omgaan met vaktaal;
- onderzoek van en de ontwikkeling in de terminologie.

De achterban van de vereniging is potentieel zeer groot maar het grootste deel is, zoals gezegd, moeilijk te bereiken. Bovendien is de vereniging goed bekend met de vertalers, maar veel minder met de andere doelgroepen. Voeg daarbij dat het vakgebied van de terminologie zelfs in de taalwetenschap al tamelijk onbekend is, laat staan daarbuiten, en de problemen van NL TERM zijn duidelijk. De doelgroep is groot, moeilijk te bereiken en men weet van het grootste deel niet precies waar hun behoeftes liggen.

Het besef dat de vereniging maar moeizaam toegang heeft tot de zeer diverse doelgroep en bovendien niet goed in beeld heeft welke vragen en wensen er bij de doelgroep leven, deed de vereniging ertoe besluiten een onder

zoek te starten naar de doelgroep en de wensen van de doelgroep. Er is voor gekozen om met een enquête het veld te ondervragen. De vereniging maakte een selectie van ruim 20 professionals uit zo divers mogelijke disciplines. De ondervraagden zijn door de bestuursleden geselecteerde professionals, die op een bewuste, doelgerichte manier met terminologie bezig zijn; ze gelden als goede voorbeelden, of anders op zijn minst als bovenmodaal geïnteresseerden. Voor alle ondervraagden geldt dat de vereniging heel graag wil weten hoe zij met terminologie omgaan, welke problemen zij daarbij ervaren, of en hoe die problemen worden opgelost en wat voor rol NL TERM (en misschien anderen, zoals het Steunpunt Nederlandstalige Terminologie en de Nederlandse Taalunie) voor hen zouden kunnen spelen. Een drietal enquêtes is vervolgens op pad gaan.

Doordat op deze wijze toch weer professionals ondervraagd gaan worden die bij minimaal één lid van de werkgroep bekend zijn, zal de enquête weliswaar inzicht bieden in de werkwijze in een groot aantal domeinen en vakgebieden, maar zal ze geen nieuwe inzichten bieden in de wensen en noden van ons onbekende mensen in onbekende vakgebieden. Daarom wordt over een tweede ronde nagedacht. De veldenquête is inmiddels uitgevoerd en de resultaten worden verwerkt. Volgens plan zal daarover gerapporteerd worden op de TiNT dag 2013.

## 2. De vragen aan de deelnemers aan de TiNT dag 2012

De jaarlijkse TiNT dag is voor NL TERM een goede gelegenheid om contact met haar achterban te onderhouden. Natuurlijk is de vereniging ook zeer geïnteresseerd in wat de deelnemers aan TiNT dagen van haar functioneren vinden. Juist de mensen die zich voor een TiNT dag inschrijven, hebben interesse in het onderwerp, koesteren wellicht bepaalde verwachtingen van het evenement en hebben positieve of negatieve ervaringen met terminologie in hun beroepssituatie. Daarom was ter gelegenheid van TiNT dag 2012 een meervragenformulier ontworpen om door de bezoekers te laten invullen. De deelnemers kregen de vragenlijst bij binnenkomst uitgereikt. Op de presentatie van de veldenquête werd ook de vragenlijst kort toegelicht en werd iedereen verzocht een ingevuld exemplaar in te leveren. Anonimiteit was uiteraard gewaarborgd.

De vragen op het formulier zijn inhoudelijk een bewerking van de open vragen die in de veldenquête gesteld zijn. In alle gevallen konden per vraag meerdere opties tegelijk worden aangekruist en soms werd als laatste optie de gelegenheid geboden zelf een antwoord te formuleren (*anders, nl: ...*). De vra

genlijst bevat achttien vragen, onderverdeeld in vijf onderwerpen. De onderwerpen zijn:

1. terminologie en terminologiebeheer: uw situatie
2. terminologie en terminologiebeheer: evaluatie
3. de Nederlandse Taalunie
4. NL TERM
5. het Steunpunt voor Nederlandstalige Terminologie.

Het eerste onderwerp betreft de doelgroep waartoe de bezoekers behoren, en de context waarin zij werken: wat voor professional is men, is men speciaal op het Nederlands gericht of ruimer georiënteerd, wordt er aan terminologiebeheer gedaan en wordt de terminologie serieus genomen? De antwoorden op die vragen moeten een beeld geven van de manier waarop de bezoekers van de TiNT dag met terminologie in aanraking komen en van het belang dat daarbij aan terminologie wordt gehecht.

### **1. Terminologie en terminologiebeheer: uw situatie**

Bent u betrokken bij terminologie in de rol van:

- a) overheid
- b) instelling, bedrijf
- c) vertaler
- d) ICT'er
- e) wetenschapper
- f) anders, nl.: ...

Verloopt de vaktalige communicatie in uw geval:

- a) uitsluitend in het Nederlands?
- b) meertalig, waaronder Nederlands?
- c) niet Nederlandstalig?

Stelt u, of uw organisatie, zich actief op als het om terminologie gaat, is terminologie een belangrijk onderwerp?

- a) Ja, terminologie wordt zeer serieus genomen.
- b) Ja, er is aandacht voor terminologie.
- c) Er is weinig aandacht voor terminologie.
- d) Er is geen aandacht voor terminologie.

Wordt er in uw werksituatie aan terminologiebeheer gedaan?

- a) We beheren zelf termenlijsten, thesauri, termenbestanden.
- b) We betrekken termenlijsten, thesauri, termenbestanden van elders.
- c) We zoeken termen en dergelijke naar behoefte, verder geen actief beheer.
- d) We maken ons niet druk om terminologie.

De vragen over het tweede onderwerp zijn bedoeld om erachter te komen hoe tevreden de bezoekers zijn met de situatie die in de eerste serie vragen geschetst is, en of er specifieke plannen, wensen of vragen zijn met betrekking tot de terminologie in de praktijk. De antwoorden op de eerste drie vragen van deze serie geven de vereniging een beeld van de manier waarop de deelnemers de omgang met vaktaal in de praktijk ervaren. De twee laatste vragen zijn directe vragen naar de wensen van de deelnemer rond informatie of voorlichting over terminologie: is daar behoefte aan, waar moet dat dan over gaan en op welke manier moet aan die wensen worden voldaan. Bij elkaar moeten de antwoorden in deze serie vragen de vereniging duidelijk maken op welke manier zij zich nuttig kan maken voor haar achterban.

## 2. Terminologie en terminologiebeheer: evaluatie

Vindt u zelf dat er in uw organisatie voldoende aandacht is voor terminologie?

- a) te veel aandacht
- b) voldoende aandacht
- c) te weinig aandacht
- d) er is geen of nauwelijks aandacht voor (gelukkig / helaas).

Zijn er problemen en knelpunten of juist positieve ervaringen?

- a) de aanwezige competentie  
positief / negatief
- b) de manier van werken (terminologiebeheer)  
positief / negatief
- c) de gebruikte software  
positief / negatief

Zijn er plannen?

- a) de competentie vergroten
- b) de manier van werken
- c) de gebruikte software
- d) anders, nl: ...

Heeft u rond elektronisch beheer, software etc. specifieke behoeften:

- a) voor het systematisch verzamelen van termen of vertalingen?
- b) voor het systematisch opslaan en beheren van terminologie?
- c) anders: ...

Heeft u specifieke behoeften aan informatie over terminologiebeheer?

- a) workshop over:
- b) themadagen, congressen over:
- c) informatie op internet over:
- d) anders, nl: ...

De derde, vierde en vijfde serie vragen zijn gericht op de spelers in het veld: de vereniging, de NTU en het Steunpunt. De vereniging wil graag weten of de deelnemers de spelers kennen, wat zij van die spelers verwachten en of zij tevreden zijn met de manier waarop die opereren. De rolverdeling tussen NL TERM, NTU en SNT is ingewikkeld; de drie zijn moeilijk uit elkaar te houden, de afkortingen maken het er niet duidelijker op en bovendien onderhouden de drie elk een eigen website. In het geval van de NTU heet die website dan ook nog niet gewoon *NTU*, maar *Taalunieversum* en de naam van de website van het SNT heeft zelfs geen enkele relatie met het Steunpunt zelf, die websi te heet *NedTerm*. Natuurlijk zijn de vereniging en beide andere spelers zeer geïnteresseerd in de oordelen van de deelnemers.

### 3. De Nederlandse Taalunie (NTU)

Kent u de NTU?

- a) nee
- b) ja, maar niet in verband met terminologie
- c) ja, ook in verband met hun inspanningen voor Nederlandstalige terminologie.

Heeft de overheid (NTU) een taak met betrekking tot de Nederlandstalige terminologie?

- a) nee
- b) ja, overheidsterminologie vastleggen en beschikbaar stellen
- c) ja, gebruikers van Nederlandstalige terminologie ondersteunen en informeren
- d) anders, nl: ...

Bent u tevreden over terminologiebeleid van de overheid?

- a) geen mening
- b) ja
- c) nee
- d) min of meer: ...

#### 4. NL TERM

Kent u NL TERM?

- a) nee
- b) ja, maar ik weet niet wat ze doen
- c) ja, en ik weet (min of meer) wat ze doen.

Wat kan NL TERM (als belangenvereniging voor wie aan terminologie doet) voor u doen?

- a) niets
- b) weet ik niet
- c) de belangen van leden behartigen bij de overheid
- d) leden met elkaar in contact brengen
- e) informatie bieden via workshops, themadagen, congressen
- f) anders, nl: ...

Bent u tevreden over NL TERM als belangenvereniging?

- a) geen mening
- b) ja
- c) nee
- d) min of meer: ...

#### 5. Het Steunpunt Nederlandstalige Terminologie (SNT)

Kent u het SNT?

- a) nee
- b) ja, vaag
- c) ja, tamelijk goed.

Wat kan het SNT als Steunpunt voor u doen?

- a) niets
- b) weet ik niet
- c) termenlijsten maken en publiceren
- d) informatie over terminologie en terminologiebeheer leveren:  
www.NedTerm.nl / loket / workshops / nieuwsbrieven
- e) anders, nl: ...

Bent u tevreden over het SNT als steunpunt voor wie met Nederlands talige terminologie werkt?

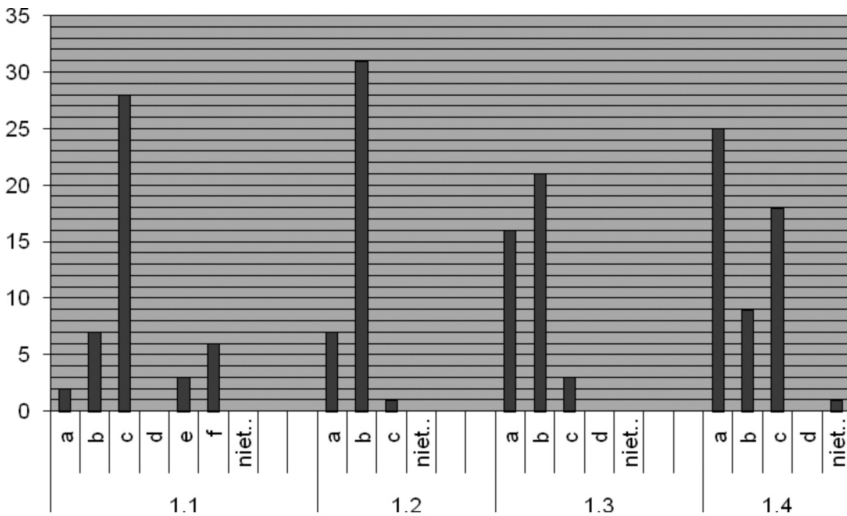
- a) geen mening
- b) ja
- c) nee
- d) min of meer.

### **3. De resultaten**

Op de bewuste TiNT dag waren ongeveer 65 bezoekers aanwezig. Daarvan leverden 38 bezoekers een enquêteformulier in. We hebben de antwoorden in tabellen in kaart gebracht. Hieronder vindt u de vijf tabellen, voor elk onderwerp één tabel. In elke tabel is per vraag het aantal antwoorden gepresenteerd dat de bezoekers kozen. Daarbij moet er rekening mee worden gehouden dat er per vraag meer dan één antwoord gegeven mocht worden en dat de deelnemers niet verplicht waren op alles antwoord te geven. Daardoor is de som van het aantal antwoorden zelden of nooit precies 38. Eventuele toelichting en andere bijschriften bij de antwoorden zijn niet in tabellen weer te geven, maar natuurlijk wel bijzonder interessant. Die worden meegenomen bij de bespreking van de resultaten onder de tabel.

#### **3.1 Uw situatie**

In tabel 1 staan de antwoorden weergegeven op de vragen betreffende het eerste onderwerp, de terminologische stand van zaken in de werkomgeving van de bezoekers van de TiNT dag. Alle tabellen zijn op dezelfde manier opgebouwd, op de staande as staan de aantallen antwoorden, op de liggende as staat het nummer van de vraag en de letter van het antwoord. De inhoud van de vragen ziet u hierboven in paragraaf 2.



Tabel 1: Terminologie en terminologiebeheer: uw situatie

Uit de antwoorden op de eerste vraag blijkt dat de vertalers inderdaad in de meerderheid waren op de TiNT dag 2012. We tellen 28 vertalers. Daarvan werken er sommigen bij de overheid, een instelling of een vertaalbureau, maar velen werken zelfstandig. De meesten die voor *anders, nl: ...* (antwoord f) zijn overigens ook bij de vertaalwereld betrokken, zoals docenten aan vertaalhogescholen. Men werkt in grote meerderheid meertalig, waarbij het Nederlands een van de talen is. Een enkeling werkt niet Nederlandstalig. De organisaties waarvoor de bezoekers werken, hebben in overgrote meerderheid aandacht voor de terminologie. Slechts een drietal organisaties heeft weinig aandacht en geen enkele respondent meldt dat zijn organisatie geen aandacht aan terminologie schenkt. De laatste vraag in deze sectie gaat over terminologiebeheer. De meeste bevrageden melden dat zij zelf termenlijsten, thesauri of andere termbestanden maken. Een kleine 20 respondenten zoeken termen naar behoefte maar doet niet op een actieve manier aan beheer. Slechts negen respondenten vullen in dat zij hun termen van elders betrekken, de overigen verzamelen dus op actieve wijze hun terminologie. Er is niemand die helemaal niet aan terminologie doet en één bezoeker vult deze vraag niet in.

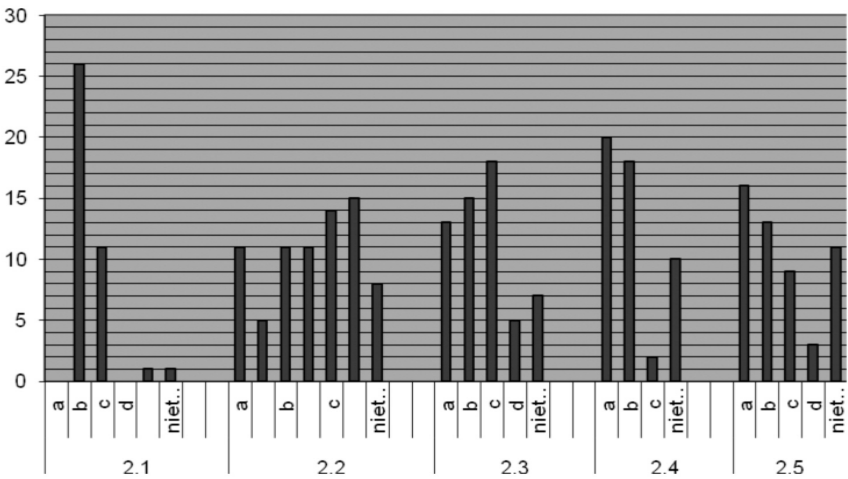
Omdat sommige deelnemers meer dan één antwoord hebben gegeven en sommige combinaties van antwoorden tegenstrijdig zijn, mogen we concluderen dat het overgrote deel van het TiNT publiek zich in de dagelijkse praktijk met vertalen bezighoudt. Men doet actief aan terminologie, door gaans in een meertalige omgeving. Dat iedereen zich wel op de een of andere



manier met terminologie bezighoudt, en de meesten ook heel actief, heeft ons verrast. Het publiek van de TiNT dagen bestaat uit ervaren professionals.

### 3.2 Evaluatie

Het tweede onderwerp is de evaluatie van de in 3.1 geschetste situatie. Hieronder zijn de scores in tabel 2 weergegeven. De scores per vraag worden daaronder besproken.



Tabel 2: Terminologie en terminologiebeheer: evaluatie

Een ruime meerderheid van 27 respondenten vindt dat de eigen organisatie voldoende aandacht voor terminologie heeft. Elf respondenten vinden de aandacht toch onder de maat. Eén deelnemer meldt dat zijn organisatie niet of nauwelijks aan terminologie doet en een andere deelnemer onthoudt zich van antwoord. In combinatie met het beeld uit de eerste sectie, waar bleek dat er zoveel aandacht voor terminologie is, mogen we uit het feit dat toch nog elf mensen de aandacht voor terminologie onvoldoende vinden, concluderen dat de het publiek van deze TiNT dag de terminologie zeer serieus neemt.

De tweede vraag betreft de competentie, de manier van werken en de gebruikte software. Per vraag zijn nu twee balken in de grafiek weergegeven; de eerste balk betreft positieve ervaringen, de tweede betreft negatieve ervaringen. Het aantal positieve ervaringen ontloopt elkaar niet veel, het lijkt informatie

ver de positieve en negatieve ervaringen per vraag te beschouwen. Dan blijkt dat de software de meeste emoties oproept, zowel positief als negatief, waar bij de negatieve iets in de meerderheid zijn. De manier van werken wordt even vaak positief als negatief ervaren. Over de aanwezige competentie zijn elf respondenten positief en vijf negatief. Acht respondenten vullen de vraag niet in.

De antwoorden op vraag drie, naar de plannen voor de toekomst, zijn goed te duiden op basis van de antwoorden op de vorige vraag. De meeste respondenten hebben plannen op het gebied van de software, maar ook de manier van werken en het vergroten van de competentie worden genoemd. Respondenten die optie *d.* kozen, deden dat in de regel om aan te geven dat er geen plannen voorhanden zijn.

Met vraag vier wilden we te weten komen welke specifieke behoeften de respondenten hebben op het gebied van elektronische ondersteuning. Zowel het verzamelen als het beheren van terminologie scoren hier hoog met respectievelijk 20 en 18 respondenten. Slechts twee respondenten hebben geen specifieke wensen en tien laten deze vraag liggen. Dat betekent dat 12 respondenten zowel behoefte heeft aan systematisch elektronisch beheer als aan het elektronisch verzamelen van termen.

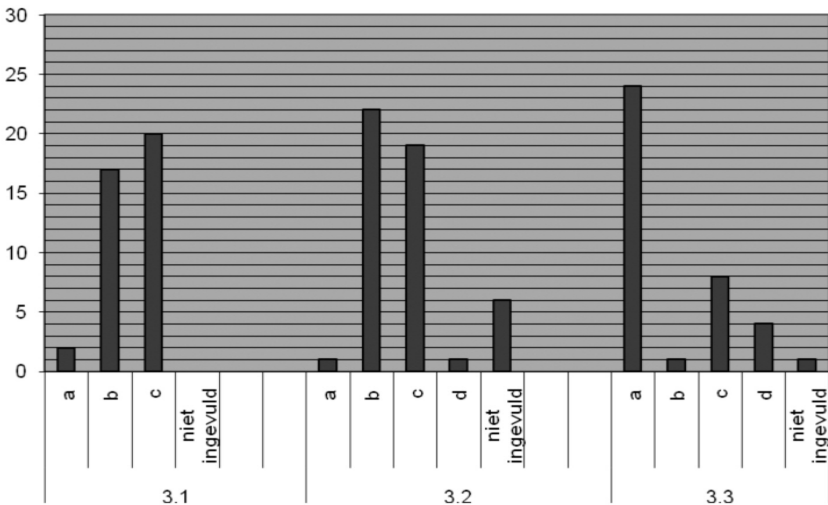
Hoe kan NL TERM de doelgroep het best van kennis en informatie voorzien? De beste manier is een *workshop*, dan *themadagen* en ten slotte *informatie op internet*. Dat is verrassend, juist het raadplegen van internet is gemakkelijk, goedkoop en vrijblijvend. Opnieuw blijkt dat de bezoekers van TiNT dagen de terminologie zeer serieus nemen. Wel opvallend is dat liefst elf respondenten niets invullen.

De bezoekers van de TiNT dag 2012 zijn duidelijk geïnteresseerd in modern terminologiebeheer. De belangstelling is breed, maar vooral elektronische *tools* en werkmethoden staan in de belangstelling. NL TERM zou er goed aan doen workshops te organiseren. Themadagen en congressen worden ook op prijs gesteld. Verschillende respondenten schreven ook een persoonlijke noot betreffende knelpunten en/of wensen voor workshops. Die waren divers, maar kwamen toch steeds weer neer op het organiseren van elektronisch terminologiebeheer en het vinden van termen. Enkelens meldden op zoek te zijn naar eenvoudige en doeltreffende software, iemand noemde als grootste probleem het gebrek aan tijd om aan terminologiebeheer te doen, een ander wil wel met *tools* gaan werken, maar wil de zaken niet grootschalig gaan aanpakken. Praktische toepasbaarheid en eenvoud zijn belangrijke factoren.

### 3.3 De NTU

NL TERM werkt nauw samen met het Steunpunt Nederlandstalige Terminologie en heeft heel goede contacten met de Nederlandse Taalunie, die zich namens de overheid met terminologie bezighoudt. De onderlinge relaties tussen deze drie spelers en hun specifieke taken zijn lang niet voor iedereen duidelijk. Het wordt de doelgroep ook niet gemakkelijk gemaakt: in paragraaf 2 bespraken we al de naamgeving van de verschillende websites. Reden genoeg om te weten of de bezoekers van de TiNT dag de spelers in het veld kennen. Daarbij is NL TERM natuurlijk ook zeer geïnteresseerd in de waardering voor haar werk.

We bespreken achtereenvolgens de NTU, NL TERM en het Steunpunt.

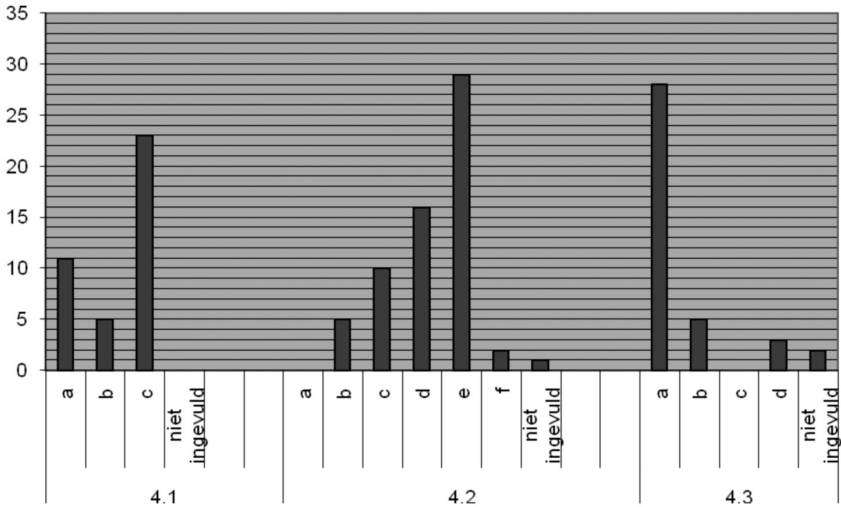


Tabel 3: De NTU

Vrijwel iedereen kent de NTU, slechts twee respondenten bekennen van niet. Een kleine meerderheid weet ook dat de NTU zich met Nederlandstalige terminologie bezighoudt. Slechts een enkeling ziet geen taak voor de overheid. De overheid moet terminologie verzamelen en vastleggen en daarnaast gebruikers van terminologie ondersteunen en informeren. Maar is men ook tevreden over de manier waarop de NTU opereert? De meesten hebben geen mening en acht respondenten zijn ontevreden. Vier respondenten zijn min of meer tevreden en slechts één respondent geeft de NTU een voldoende. Dat kan duidelijk beter.

### 3.4 NL TERM

We stelden de bezoekers min of meer dezelfde vragen over NL TERM als over de NTU. In tabel 4 zijn de resultaten gepresenteerd.



Tabel 4: NL-TERM

NL TERM en de NTU zijn onvergelykbare grootheden en daarom niet goed onderling te vergelijken. De NTU is als verdrags- en beleidsorganisatie bij veel meer betrokken dan alleen terminologie, bijvoorbeeld bij de altijd geruchtmakende spellingshervormingen. Het is dan ook begrijpelijk dat NL TERM minder goed scoort op bekendheid dan de NTU. Toch is de score van NL TERM op bekendheid wat teleurstellend, aangezien NL TERM een thuiswedstrijd speelt. Vooral dat elf mensen de vereniging niet kennen (dat was vóór de TiNT dag, mogen we hopen ...) is opmerkelijk. Wat wel optimistisch stemt is dat een ruime meerderheid NL TERM goed zegt te kennen.

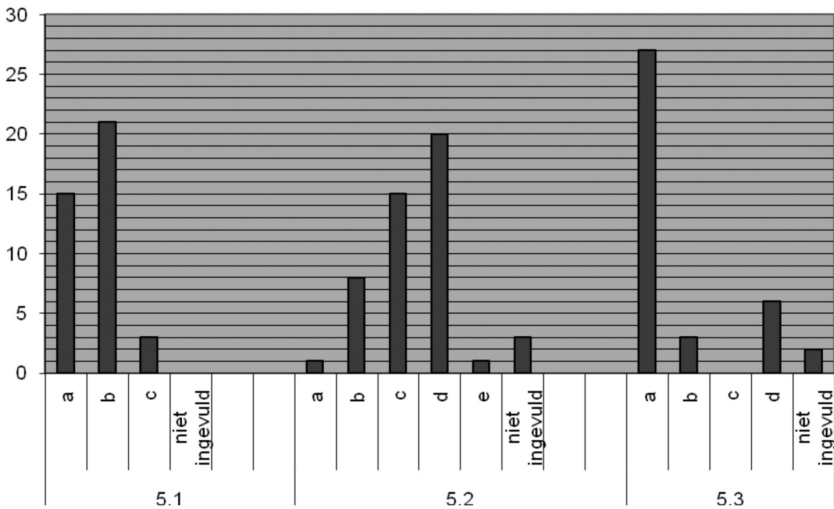
De tweede vraag handelt over wat NL TERM voor de leden zou kunnen doen. De positieve antwoorden (*c*, *d* en *e*) zijn ver in de meerderheid. Opvallend is dat 29 van de 38 respondenten behoefte heeft aan informatie via workshops, themadagen en congressen. Een vijftal respondenten weet niet precies wat NL TERM voor hen zou kunnen doen. Toch hebben ze de weg naar de TiNT dag kunnen vinden. Twee respondenten schreven bij de vragen dat NL TERM een forum of ander platform voor terminologiebeheer zou moeten

bieden. Het Steunpunt heeft een dergelijk platform, maar daar wordt weinig gebruik van gemaakt.

De tevredenheid met het werk van de vereniging is hoog in vergelijking met de score van de NTU (op terminologisch gebied). De meeste respondenten hebben geen mening, vergelijkbaar met de score van de NTU. Er is echter niemand ronduit ontevreden met betrekking tot de prestaties van NL TERM en het aantal tevreden respondenten is wat hoger dan bij de NTU. De conclusie kan zijn dat maar weinigen NL TERM kennen, maar die weinigen waarderen de vereniging doorgaans wel.

### 3.5 Steunpunt

Ten slotte is het publiek gevraagd zijn mening te geven over het Steunpunt Nederlandstalige Terminologie. Het Steunpunt is door de NTU opgericht en werkt zeer nauw samen met NL TERM.



Tabel 5: Het Steunpunt

Op de vraag naar de bekendheid met het Steunpunt lijkt de score enigszins bij die van de NTU en NL TERM achter te blijven. Wat dat betreft zijn de eerste vragen van sectie 4 en 5 wat onhandig gesteld, waardoor ze moeilijk vergelijkbaar zijn. We mogen ervan uitgaan dat respondenten de spelers wel kennen,

maar niet zo heel precies weten hoe de rolverdeling is. Die respondenten zullen voor antwoord *c* kiezen bij de vraag over NL TERM (*ik weet (min of meer) wat ze doen*). Antwoord *b* is namelijk *ja, maar ik weet niet wat ze doen*. Diezelfde respondenten kiezen bij de eerste vraag over het Steunpunt wellicht eerder voor *b* (*ja, vaag*) dan voor *c* (*ja, tamelijk goed*). Gaan we daarvan uit, dan tellen we 23 mensen die (min of meer) weten wat NL TERM doet, tegenover 24 mensen die dat van het Steunpunt min of meer weten ( $b + c$ ). Dat zijn dus vergelijkbare aantallen. Een reden tot zorg is dat vijftien respondenten het Steunpunt niet kennen.

De belangrijkste taak voor het Steunpunt is, evenals voor NL TERM, de doelgroep informeren over terminologie. In de antwoorden waaruit de respondenten dit keer konden kiezen, zijn de themadagen en de congressen afwezig. Die antwoorden waren wel heel populair bij de corresponderende vraag over NL TERM. Daarvoor in de plaats noemen we nu het maken en publiceren van termenlijsten en het leveren van (digitale) informatie. Ook voor deze zaken is veel interesse, maar die is beduidend lager dan die voor de workshops die bij NL TERM genoemd zijn. We zagen dit opmerkelijk resultaat ook al in sectie twee; hoewel de verwachting is dat vooral vertalers graag kant en klare termenlijsten willen downloaden, kiest men juist voor workshops en themadagen boven digitale informatie of een passieve aanpak.

Ten slotte valt bij vraag drie opnieuw op dat veel respondenten geen mening hebben over het functioneren van, ditmaal, het Steunpunt. Drie respondenten zijn tevreden, zes zijn dat min of meer. Dat is vergelijkbaar met wat we bij NL TERM zagen.

We moeten concluderen dat het Steunpunt te weinig zichtbaar is. De verschillen met NL TERM zijn klein, maar NL TERM komt er iets positiever vanaf. Dat is niet ernstig, NL TERM en het Steunpunt werken zo nauw samen dat het verschil vooral aan de wat onhandige vraagstelling kan worden toegerekend. Opnieuw blijkt dat de doelgroep op een serieuze manier met terminologie bezig is; intensieve workshops worden verkozen boven goedkope en gemakkelijke, maar vrijblijvende digitale informatie.

#### 4. Tot slot

We hebben nadrukkelijk niet voor statistische analyse van de resultaten gekozen. De hoeveelheid materiaal is daarvoor te beperkt en de doelstelling van de enquête is daarvoor ook minder geschikt. De doelstelling is een indruk te krijgen van wie de TiNT dag bezoekt en wat de bezoeker van de vereniging wil.

Het kan wel interessant zijn nader onderzoek te doen naar de samenhang van de antwoorden. Zijn vooral de vertalers voorstander van praktisch en eenvoudig terminologiebeheer? Kennen de docenten het Steunpunt? Dat soort onderlinge afhankelijkheden zijn (nog) niet in kaart gebracht.

Samenvattend kunnen we stellen dat een groot deel van het publiek van de TiNT dag 2012 vertaler is. Het publiek neemt de terminologie zeer serieus en werkt op professionele wijze. Desalniettemin zijn er volop plannen voor een nog betere aanpak. Wat de plannen voor de toekomst betreft, scoort het inzetten van digitale *tools* het hoogst, maar men ziet kennelijk in dat nieuwe *tools* om een ruimere eigen competentie en een andere werkmethode vragen. Dat is voor de vereniging en voor het Steunpunt op zich niet onverwacht, maar dat deze positieve houding zo algemeen en overduidelijk gedragen wordt, overtreft toch de verwachting. De beste manier om zich te blijven ontwikkelen op het gebied van de omgang met terminologie is volgens de bezoeker van de TiNT dag niet het internet, maar persoonlijk bezoek aan workshops en thema dagen. Ook hier geldt dat de uitslag niet onverwacht is, maar de overtuiging waarmee dat uit de cijfers blijkt toch wel.

Het is de vereniging nog niet gelukt een breed publiek voor de TiNT dagen te interesseren. De vereniging – en ook het Steunpunt – zijn veel te weinig zichtbaar bij de doelgroep. Dat blijkt ook weer als twee respondenten nadrukkelijk om een platform vragen, terwijl het Steunpunt een dergelijk platform onderhoudt. Maar overduidelijk is gebleken waar dit deel van de doelgroep, de vertalers, behoefte aan heeft. NL TERM en het Steunpunt kunnen aan het werk!

*Met dank aan G. Scholten voor de presentatie van de resultaten.*

## **Geraadpleegde literatuur**

W. Martin (2009), Waar moet dat naartoe. Toekomstperspectieven voor de terminologie in Vlaanderen en Nederland. In: *TiNT dag 2009*, 1 12.

## Medewerkers aan deze bundel

Drs. Bert Keurentjes, Radboud in'to Languages, Radboud Universiteit, Nijmegen,  
[vertaalservice@into.ru.nl](mailto:vertaalservice@into.ru.nl)

Dr. Oele Koorwinder, GridLine, Amsterdam,  
[oele@gridline.nl](mailto:oele@gridline.nl)

Lic. Sara Rymenams, Departement ECHO, Katholieke Hogeschool Leuven,  
[sara.rymenams@khleuven.be](mailto:sara.rymenams@khleuven.be)

Dr. Tanneke H. Schoonheim, Instituut voor Nederlandse Lexicologie, Leiden,  
[tanneke.schoonheim@inl.nl](mailto:tanneke.schoonheim@inl.nl)

Dr. Mart Trautwein, GridLine B.V., Amsterdam,  
[trautwein@gridline.nl](mailto:trautwein@gridline.nl)

Prof. dr. Serge Verlinde, Instituut voor Levende Talen (ILT), KU Leuven,  
[serge.verlinde@ilt.kuleuven.be](mailto:serge.verlinde@ilt.kuleuven.be)

Dr. Hennie van der Vliet, Vrije Universiteit, Amsterdam – Steunpunt  
Nederlandstalige Terminologie,  
[h.d.vander.vliet@vu.nl](mailto:h.d.vander.vliet@vu.nl) / [steunpunt@vu.nl](mailto:steunpunt@vu.nl)

Folkert Zijlstra, NATO Terminology Office, NATO Terminology Programme,  
[zijlstra.folkert@nsa.nato.int](mailto:zijlstra.folkert@nsa.nato.int)

### Redactie

Prof. dr. Els Ruijsendaal, BeNeLux Universitair Centrum,  
[els@ruijsendaal.nl](mailto:els@ruijsendaal.nl)

Dr. Cornelia Wermuth, KU Leuven/Subfaculteit Taal en Communicatie,  
[cornelia.wermuth@arts.kuleuven.be](mailto:cornelia.wermuth@arts.kuleuven.be)





Een uitgave van de Vereniging NL Term en Steunpunt Nederlandse Terminologie (SNT)

Dit boek bevat de lezingen van de TiNT dag 2012. Het is een initiatief van NL Term (Vereniging voor Nederlandstalige Terminologie) en Steunpunt Nederlandse Terminologie (SNT). De redactie van deze bundel berust bij Els Ruijsendaal en Cornelia Wermuth.

Aan deze bundel leverden de volgende auteurs een bijdrage: Bert Keurentjes, Oele Koornwinder, Sara Rymenams, Tanneke Schoonheim, Mart Trautwein, Serge Verlinde, Hennie van der Vliet en Folkert Zijlstra.

