

De spelling van het lot.

Over het ontstaan van een nieuwe NTG werkgroep

Frans Goddijn & anderen

fg@fgbbs.iaf.nl

23 Sept 1996

Abstract

Het SPELLINGBESLUIT gepubliceerd in Staatsblad 394 (*Besluit van 19 juni 1996, houdende voorschriften omtrent de schrijfwijze van de Nederlandse taal*) heeft niet alleen gezorgd voor veel gekrakeel tussen neerlandici, maar het vormde ook een buitenkans voor uitgeverijen van werken als HET GROENE BOEKJE (binnen zes maanden 500.000 verkochte exemplaren). Kleinere ondernemingen brachten floppies op de markt met woordenlijsten en algoritmen. Intussen sloeg een groepje mensen (bestaande uit — in alfabetische volgorde — Erick Branderhorst, Erik Frambach, Frans Goddijn, Hans Hagen, Gerard van Nes, Piet Tutelaers, Jacqueline van Weelden en Peter van Zeeland, en terzijde gestaan door Hans Linders) de handen ineen om te komen tot een vrij beschikbare, zo correct mogelijke Nederlandse woordenlijst.¹

Nederlands uit de USA

Nu de nieuwe spelling er is, wil ik graag een nieuwe, correcte Nederlandse woordenlijst om aan mijn spellingchecker te voeren. Ik heb geen nieuwe spellingchecker nodig en al helemaal geen tekstverwerker, ook al adverteren de verkopers ervan ermee dat hun product nu is uitgerust met HET GROENE BOEKJE.

Diverse leveranciers staan klaar om ons onze eigen taal te verkopen: UBIQ BV in Montfoort (0348-475385) heeft een floppy voor krap dertig gulden, echter wie dan denkt een woordenlijst te kopen, krijgt een verrassing: er zitten wel 160.000 woorden op de floppy maar je kunt er niet bij. Sterker, ze kunnen er zelf ook niet bij want, zo verklaart een medewerker van UBIQ BV, “de woordenlijst is gecompriemd en komt van NOVELL in Amerika. Hij is bedoeld voor gebruikers van WORDPERFECT die er hun spellingchecker mee kunnen updaten”.

Zijn we dan *!#@^:=(zover gezakt dat we de woorden van *onze eigen taal* moeten aanschaffen in de Verenigde Staten van Amerika?!

Hans Hagen schreef me: “Alle pogingen van commissies ten spijt, wordt taal gemaakt door de gebruikers. Het is dan ook te gek voor woorden dat diezelfde gebruikers moeten betalen om de door de overheid voorgeschreven woorden te mogen gebruiken.”

INP (Postbus 765, 2270 AP Voorburg) levert ook een oplossing, namelijk een floppy waarmee WP-gebruikers hun bestaande woordenlijst kunnen bijwerken. Niet alle woorden zijn veranderd: er moeten volgens INP ongeveer 1200 woorden weg en er komen zo'n 4800 woorden woorden

bij om de WP-gebruiker in de hedendaagse spelling te laten schrijven.

Verder kan ook bij de STICHTING SOCIALE DATA-BANK NEDERLAND een verbeterde Nederlandse WP-woordenlijst besteld worden. Zij melden dat t.o.v. WORDPERFECT 5.1 ruim 7.000 fouten zijn verbeterd, en bijna 112.000 ontbrekende woorden zijn toegevoegd. Leuk voor de WP-gebruiker, maar ik stel nu eenmaal wat hogere eisen aan mijn zetwerk dan met WP mogelijk is.

Fouten bij WP

Erik Frambach heeft destijds bij het vervaardigen van de woordenlijst in de oude spelling gebruik gemaakt van WORDPERFECT 5.1 ter controle. Maar wat bleek... het was niet zozeer WORDPERFECT waarmee de fouten uit zijn eigen woordenlijst-in-wording werden gehaald:

“Ik ben me daar hardstikke doodgeschrokken van de massa's fouten (in WORDPERFECT). Wat erg dat zoveel mensen dit als referentie gebruiken! Naar aanleiding daarvan heb ik het volgende stukje geschreven in ons facultair computerblaadje:

De niet zo perfecte WordPerfect woordenlijst

Tot voor kort gebruikten velen voor het controleren van de spelling van de Nederlandstalige T_EX-documenten de woordenlijst die bij WORDPERFECT 5.1 wordt geleverd. Deze lijst is echter erg beperkt van omvang en inconsistent wat betreft welke verbuigingen en meervouds- en enkelvoudsvormen zijn opgenomen. Het ergste is echter dat de lijst *honderden* fouten bevat (nee, ik overdrijf niet).

¹Dit is een bewerking van een artikel uit het septembernummer van Computer Info. De kopij van het artikel is eerst door Erik Frambach voorzien van diverse links naar email-adressen, homepages en ftp-sites en de tekst is later door mij uitgebreid.

- Er staan typfouten in: bv. ‘o’ wordt ‘i’, ‘m’ wordt ‘n’. (Een veel voorkomende fout bij tikwerk, aangezien de ‘o’ en de ‘i’ naast elkaar zitten op het toetsenbord evenals de ‘m’ en de ‘n’).
- Er zijn letters vergeten.
- Er zijn letters omgedraaid.
- Woorden worden ernstig verminkt: ‘belastingruc’ wordt ‘balstingruc’.

Geschrokken van de onbetrouwbaarheid van deze lijst ben ik zelf op zoek gegaan naar betere alternatieven. Uit ‘public domain’ woordenlijsten heb ik een nieuwe, ruim twee keer zo grote lijst samen kunnen stellen die vele malen beter is.

WORDPERFECT Nederland heb ik aangeschreven om te vragen of ze interesse hebben, maar nee, zo schreef Lisa Boerlage, de Teamleider Support van WORDPERFECT: *“Het is bij WordPerfect Nederland en WordPerfect Corporation bekend dat de Nederlandse woordenlijst beperkingen oplevert. Ook zijn wij ons bewust van het feit dat er in de lijst spelfouten voorkomen. Deze oneffenheden zijn door ons verzameld. Wij hopen dan ook in de volgende versie van WORDPERFECT een betere woordenlijst te kunnen leveren.*

Als je een fout vindt kun je die melden. Kennelijk is dit hoe WORDPERFECT denkt de woordenlijst te verbeteren. Maar zo komen we er nooit. De lijst is slecht opgezet en zal niet wezenlijk verbeteren door hier een daar een woord aan te passen.

Vertrouw de Nederlandse spelling van WP niet! Ook niet in volgende ‘verbeterde’ versies. Wie werpt eens een kritische blik op de andere talen?”

Dat schreef Erik Frambach in januari 1993, en toen WP 6.0 uitkwam ging hij eens voorzichtig kijken of er iets verbeterd was. Ja er was iets verbeterd, maar hij kon ook weer moeiteloos massa’s oude fouten aanwijzen.

Overigens is het vaker onthutsend om kritisch te kijken naar woordenlijsten die algemeen als ‘de’ standaard worden gezien, zoals verderop zal blijken. . .

ASCII gewenst

Het is trouwens de vraag of het voldoende is om simpelweg een aantal duizenden woorden uit een oudere woordenlijst te vervangen, zo redeneerde Hans Hagen:

“Onze taal kenmerkt zich door een grote hoeveelheid woorden. Dit is een direct gevolg van het feit dat veel woorden te combineren zijn tot nieuwe. Deze combinaties leiden vervolgens weer tot afspraken over de verbindende letters, zoals een ‘n’ of een ‘s’. Als de volledige woordenlijst van de Nederlandse taal vele meters beslaat en HET GROENE BOEKJE maar enkele centimeters dik is, hoe weet je dan zeker of iets goed geschreven is? De technisch wetenschappelijke wereld heeft daarnaast nog een geheel eigen woordenschat, aangevuld met wederom een veelheid aan combinaties. Gelukkig is deze wereld uitstekend in staat de eigen hulpmiddelen te creëren om teksten op spelling te controleren, maar waar haal je de juiste spelling vandaan?

In de T_EX-wereld bijvoorbeeld, waar een gebruiksvriendelijke interface bij de invoer van tekst veel minder belangrijk wordt gevonden dan de typografische kwaliteit van de uitvoer, wordt hoofdzakelijk gebruik gemaakt van

ASCII teksten. Bij het controleren op de juiste spelling worden eveneens in ASCII vastgelegde woordenlijsten gebruikt. De gecodeerde woordenlijsten van commerciële tekstverwerkers, zoals WP en MS-WORD zijn daarbij onbruikbaar. Commissies die een wijziging in de spelling voorstellen zonder zorg te dragen voor voor ieder toegankelijke bestanden in ASCII formaat, schieten dus deels hun doel voorbij.”

Falende algoritmen

Maar zelfs voor WP-gebruikers is het niet alles goud wat er blinkt. “De spellingchecker van Prisma is voorlopig niet aan te raden”, zo schreef Henk Langerak van het Algemeen Dagblad. In elk geval niet in WP 6.0a en 6.0b. Daarin gooit-ie heel veel woorden weg. Hij herkent veel woorden niet en zeker niet de nieuwe spelling.”

Joost Bloemsma van uitgeverij Het Spectrum reageerde hierop:

“Deze spellingchecker is niet in de handel geweest omdat de tekortkomingen door ons bemerkte werden. Hij was overigens identiek aan die van UBIQ. Deze spellingchecker bevat niet een gecomprimeerde woordenlijst maar een verzameling algoritmen die de spelling moeten controleren, en daar zit nou de kneep: deze nieuwe spelling is zo inconsistent dat ze niet in een verzameling algoritmen te vangen is. De oplossing zou een schier eindeloze lijst van woorden zijn, zie De Nieuwe Spellinggids (Prisma/Van Dale/Wolters) met meer dan 275.000 woorden, zonder verbuigingen en vervoegingen (slechts fl 29,90). Een dergelijke lijst als spellingchecker is echter onaanvaardbaar langzaam.”

Ik heb gevraagd of deze Nieuwe Spellinggids ook op floppy is te verkrijgen, maar dat bleek niet het geval te zijn. Van Dale Lexicografie verwees me naar uitgeverij Het Spectrum waar Joost Bloemsma werkt. Hem heb ik de tip gegeven dat zo’n woordenlijst in samenwerking met een spellingchecker als AMSpell zeker niet *onaanvaardbaar langzaam* is, doordat de woordenlijst voor gebruik wordt gecompileerd en gendexeerd. Het Spectrum zou, met deze spellingchecker, woordenlijst en vooral hun marketingapparaat, eenvoudig een goedkope spellingfloppy op de markt kunnen brengen.

Het Groene Boekje

SDU (070-3789911) is marktleider met HET GROENE BOEKJE (“onvolledig, onsystematisch en gebruiksvriendelijk” aldus Onze Taal), maar de beoogde lancering in januari van de elektronische versie van het Groene Boekje werd uitgesteld tot juni dit jaar. Oorzaak hiervan was het grote aantal verschillen tussen dit boekje en de Dikke Van Dale. Die moesten eerst zoveel mogelijk worden weggevoerd. De kopers van het ELEKTRONISCH GROENE BOEKJE zijn daarmee in het voordeel vergeleken bij de mensen die het papieren boekje kochten. Overzichtelijk is het zeker en systematisch is het ook. Gebruiksvriendelijk zou je het kunnen noemen als je ziet hoe de SDU erin is geslaagd de woorden van onze taal te verpakken in een

user interface als een etalageruit: je kunt de woorden op allerlei manieren bekijken maar je kunt er niet ‘bij’.

Ben je aan het schrijven en twijfel je aan de spelling van een woord, dan kun je het ELEKTRONISCH GROENE BOEKJE starten en je zoekt informatie over het woord. Het ELEKTRONISCH GROENE BOEKJE is echter traag, erg traag: een woord vinden kost enkele tot vele seconden, bijna even lang als het kost om hetzelfde woord in de papieren versie op te zoeken. Wachten of bladeren, het is kiezen tussen twee kwaden. Een heel document op deze manier controleren is onbegonnen werk.

Je kunt met het ELEKTRONISCH GROENE BOEKJE ook dictees maken en met anagrammen goochelen. Ook kun je een verhaal over de nieuwe spellingsregels lezen, maar n ding kun je niet en dat is wat mij betreft het enige van serieus belang: een leesbare lijst van woorden genereren.

En dat terwijl er *gratis* spellingcheckers zijn die beter zijn dan die van WP!

Alternatieve spellingcheckers

AMSpell bijvoorbeeld, geschreven en onderhouden door Arjan Merckens en Erik Frambach, naadloos ingepast in de \LaTeX -structuur, met woordenlijsten voor Nederlands (oude spelling), Engels, Amerikaans, Frans, Duits, Italiaans en Spaans.

Bij PRAGMA (038-4229775) heeft men een spellingchecker TEXSPELL gemaakt die gratis mag worden verspreid. Zelf gebruiken ze deze samen met de editor TEXEDIT, maar hij is ook afzonderlijk te gebruiken. Ik citeer Hans Hagen van PRAGMA:

“Met TEXSPELL kunnen teksten worden gecontroleerd op foutief gespelde woorden. Men kan daarbij werken met verschillende lijsten, bijvoorbeeld een die grotendeels overeenkomt met HET GROENE BOEKJE (ruim 100.000 woorden). Bij het ontwikkelen van cursusmateriaal gebruiken we echter meestal een zelf samengestelde lijst van ruim 30.000 woorden, die slechts gedeeltelijk overlapt met de groene lijst. Er zijn dus net zoveel boekjes als kleuren.

In TEXEDIT kan men real-time een tekst op de juiste spelling controleren. De juist gespelde woorden worden groen weergegeven, de foute woorden rood en woorden beneden een bepaalde omvang, bijvoorbeeld 4 karakters, wit. Het blijkt dat een op deze manier gekleurde file snel ‘in orde’ te brengen is, niet in de laatste plaats omdat accenten automatisch worden geplaatst. Bovendien worden bepaalde categorieën woorden, zoals passieve werkwoorden, geel of cyaan weergegeven, zodat men snel een ‘beeld’ van de stijl krijgt. Om dit alles te kunnen ondersteunen, moeten we kunnen beschikken over goede, al dan niet zelf samengestelde, lijsten van woorden.”

Het WORDS-L team

Wat zou er nu meer voor de hand liggen dan een nieuwe, correcte woordenlijst te maken waar iedere tekstverwerker en iedere spellingchecker op ieder computersysteem in ieder land mee uit de voeten kan?

Deze vraag leidde tot het ontstaan van een WORDS-L-team (WLT) van mensen die vanuit hun woon- en werkplaatsen verspreid over Nederland werken aan een grote, correcte en bovenal kostenloze Nederlandse woordenlijst voor iedere gebruiker.

Eerst was het Gerard van Nes die de zomerse nacht voor zijn vertrek naar Florida offerde aan het programmeren in PERL van een utility, `match.pl` waarmee een bestaande woordenlijst kan worden geschoond en aangevuld volgens een aantal nieuwe spellingregels. Een slim PERL-script. Waar ‘uiering’ stond, ook als onderdeel van een woord, werd er het nieuwe ‘uienring’ van gemaakt. Wel moest het resultaat nog met de hand worden nabewerkt, want ook het woord ‘sluiering’ bijvoorbeeld werd voorzien van een extra letter. Zo bleken er meer *woordgrapjes* te zijn die moesten worden gecontroleerd. Gelukkig schreef het PERL-script van Gerard een verslagje van alles wat het had gedaan, zodat de correcties gemakkelijk konden worden uitgevoerd. Intussen was de discussie over woordenlijsten van TEX-NL verhuisd naar WORDS-L, een eigen mailinglijst vanaf FGBBS.

Van welke woordenlijst kon eigenlijk het beste worden uitgegaan? Het simpelst zou het natuurlijk zijn als iemand alle woorden uit het ELEKTRONISCH GROENE BOEKJE zou halen. Hans Linders van de TU EINDHOVEN liet zien dat dit zonder al te veel omhaal mogelijk is, maar de woordenlijst die hieruit rolde kan niet zomaar worden verspreid. Het was dankzij Linders wel mogelijk om duizenden woorden in HET GROENE BOEKJE op te zoeken zonder dat er dagen voor moest worden gebladerd of gewacht.

Copyright?

Op het Internet zijn hier en daar Nederlandse woordenlijsten te vinden, zelfs exemplaren van HET GROENE BOEKJE uit 1954 en 1990 (het eerste is nog ‘gered’ uit de stroken geponst papier waarop destijds de data werden bewaard), bijvoorbeeld op `gopher://olt.et.tudelft.nl:72/1/words/` van de TU DELFT en bij `ftp://ftp.nl.net/pub/textproc/dictionaries/` van NLNET.

De vraag is echter of die woordenlijsten ‘vrij’ zijn. Erick Branderhorst, lid van het WORDS-L-team hierover:

“Internet en free software zijn onlosmakelijk met elkaar verbonden. Velen zijn bekend met bijvoorbeeld de GNU, XFREE, \TeX en talloze andere soortgelijke projecten. Het vrij ter beschikking stellen van software houdt echter niet op bij het aanbieden van een bestandje op een server en het in een Nieuwsgroep bekend maken van dit bestandje. Het is wel degelijk belangrijk dat er duidelijk bij vermeld wordt wie het wel en wie het niet mag gebruiken en onder welke voorwaarden.

Van alle vrij beschikbaar gestelde lijsten van Nederlandse woorden hebben we geen enkele lijst kunnen vinden waarvoor dit duidelijk was omschreven. De herkomst van de lijsten is vaak duister en niemand heeft klaarblijkelijk de moeite genomen om voorwaarden voor het gebruik afdoende te definiëren. Wij willen voor eens

en voor altijd een einde aan deze situatie maken en een pakket met Nederlandse woordenlijsten, afbreekpatronen en aanverwante zaken vormen en dit distribueren onder de voorwaarden zoals gesteld in de GENERAL PUBLIC LICENSE II (afgekort als GPL, er is ook een Nederlandse versie). Vrij vertaald behelst de GPL het volgende: je mag alles doen met dit pakket wat je wilt, als je er maar voor zorgt dat je alles wat jij kreeg ook weer beschikbaar stelt aan anderen.”

We kozen voor het gebruik van een woordenlijst die niet van de redactie van HET GROENE BOEKJE afkomstig was. Erik Frambach had al eens een Nederlandse woordenlijst gecompileerd voor het gebruik bij AMSpell. Het script van Gerard van Nes genereerde een nieuwe lijst van 220.000 woorden die we `n1963.txt` hebben genoemd.

ISO LATIN1 & Piet Tutelaers

Vanaf dit punt werd Piet Tutelaers de spil waarom het project draaide. Aanvankelijk om de nodige afbreekpatronen te genereren, maar ook om de kwaliteit van de verkregen woordenlijst te verbeteren. Het was zijn idee om te werken aan een universeel masterbestand in het ISO LATIN1 formaat. Van daaruit kan voor elke denkbare karakterset een vertaling worden gemaakt, dus ook naar ASCII. Voor deze vertaalslag leverde Piet het PERL-script `lat2ansi.pl`

Piet schreef tijdens het werk aan de lijst het volgende:

“Het bestand `n1963.txt` heb ik omgezet naar ISO LATIN1 en vervolgens geverifieerd met mijn benadering van het ELEKTRONISCH GROENE BOEKJE (EGB96). In mijn EGB96 benadering (laten we dit gemakshalve maar even E-G-B96 dopen omdat hierin ook de afbreekplaatsten staan) ontbreken werkwoordsvormen als werkt en werk-ten, echter niet werk-te en ge-werkt. Ook meervoudsvormen als kaart-jes (wel kaart-je) en de overtreffende trappen kaler, kaalst (wel kaal). Ik heb geprobeerd op basis van eenvoudige heuristieken de woordenlijst te screenen met mijn PERL-script `verifieer`. Het resultaat is `spell-nl-iso.txt`² Hierin worden de woorden voorafgegaan door codes die aangeven wat er nog met een bepaald woord moet gebeuren:

```
=<woord>
  <woord> identiek aan E-G-B96 (zonder
    afbreekstreepjes)
-<woord>:<woord1>
  <woord> moet vervangen worden
    door <woord1>.
-<woord>:<woord1>|<woord2>
  <woord> moet vervangen worden door
    <woord1> en/of <woord2>
?<woord>
  <woord> met twijfelachtige spelling.
    Bijvoorbeeld 'kapotmaakt'
    omdat dit los wordt
    geschreven 'kapot maakt'.
#<woord>
  <woord> ontbreekt in E-G-B96
    terwijl verwante
    woordvormen wel voorkomen
    (werkten, kaalst, etc.)
+<woord>
  <woord> ontbreekt in E-G-B96.
    Het lijkt mij verstandig
    deze woorden na te lopen
    in de Dikke van Dale.
```

Mijn ideaal zou zijn een groot ISO-LATIN1 woordenbestand met correcte woorden, dus inclusief hoofdletters, koppeltekens (streepje), apostrof en *met* afbreekplaatsen (klein streepje).

Dit bestand zou diverse behoeften kunnen dekken, zoals het genereren van afbreekpatronen voor T_EX en spellingscontrole. Onze eerste taak is echter te zorgen voor een nauwkeurige woordenlijst. Hoe pakken we dat aan? Van de Dikke Van Dale bestaat een CD-ROM versie. Kunnen we een programma maken dat onze woorden uit de categorie + hiermee vergelijkt? Of moeten we dit zelf gaan doen? De categorie woorden met # en ? wil ik wel voor mijn rekening nemen evenals het toevoegen van de afbreekstreepjes. Ook moeten we ons afvragen of ‘werken’ naast ‘werkte’ wel nodig is.”

Het onvermijdelijke handwerk

Een aantal van deze vragen is nog onbeantwoord. Leden van het WLT controleerden elk een deel van de ongeveer 50.000 woorden waarvan niet automatisch kon worden bepaald dat ze correct zijn. Inmiddels was het besef doorgedrongen dat je een woordenlijst niet ‘even’ maakt of corrigeert, ook al heb je computers tot je beschikking, plus slimme mensen om ze te bedienen. . .

Bij het nakijken van de 50.000 te controleren woorden was te zien dat ergens in de voorgeschiedenis ervan een programmeur bezig was geweest om woorden bij te maken. Er stonden woorden als ‘complexmeest’ en ‘beduudmeest’. Kennelijk had zijn programma de overtreffende trap in *complex*, *complexer*, *meest complex* verkeerd geïnterpreteerd. Piet Tutelaers wist de etymologie van deze fout te traceren:

“Ongetwijfeld heeft iemand deze per computer gefabriceerd uit GB54, de woordenlijst van HET GROENE BOEKJE uit 1954. Hierin komen woorden voor als:

```
110000000#bed*uusd#bed*uusde#bed*
uusder#meest bed*uusd#06
```

De superlatief van beduud is dus niet beduudst maar ‘meest beduud’. Alleen welke snoodaard heeft dit omgedraaid en de spatie weggelaten?”

We zullen het nooit weten, maar het betekende wel dat we alle woorden eindigend op “meest” konden verwijderen (behalve “allermest”).

In de datafile van HET GROENE BOEKJE uit 1954 stond trouwens

```
910000000#boeddh*istisch#boeddh*istische#
boeddh*istisch#boeddh*istisch#76
```

Hieruit blijkt dat het woord ‘boeddhistisch’ destijds door de beugel kon. Onzin natuurlijk en in HET GROENE BOEKJE van vandaag staat dat ook niet meer. Woorden als deze maakten het des te verstandiger dat we een aantal woorden met de hand naliepen of, beter gezegd, met de hand aan het toetsenbord en met de woordenboeken naast het scherm.

Het woordenlijst-checken (Jacoline en ik ‘deden’ meerdere happen van ongeveer 6000 woorden per keer) werkte

²Inmiddels zijn er nieuwe versies verschenen, zie de verwijzingen onderaan dit artikel.

bij ons op een bepaalde manier verslavend. Na een uurtje merkten we dat we eindelijk de kop eraf hadden, aan het schuifblokje dat rechts in het scherm van de editor aangeeft waar je ongeveer zit in het hele bestand. De eerste tijd, als het blokje aan het begin blijft staan net alsof er nog *niets* is gepresteerd, is het moeilijkst. Daarna was het zaak af en toe te pauzeren en vooral niet door te willen prakken als de letters gaan dansen en we niet meer lekker op de stoelen konden zitten. Vervolgens, wanneer we eenmaal over de helft heen waren, leek het wel alsof de zwaarte-kracht meehielp en zakten we in een paar uur (koffie en thee tussendoor) naar het einde, tot het begeerde <EOF> teken ineens van onderaf het scherm opschoof. . .

Dwaas, dwazer, dwaast?

We ontdekten overigens ook dwaasheden in de zogenaamde ‘standaardwerken’ waarin we doorlopend naar wijsheid zochten bij twijfelgevallen. Onze Dikke van Dale gaf nota bene aan dat je *onzichtbaar*, *onzichtbaarder*, *onzichtbaarst* kunt schrijven! Belachelijk. Je bent zichtbaar, slecht zichtbaar of onzichtbaar, maar als niemand je meer kan zien, kun je niet nog wat onzichtbaarder worden, toch? De Dikke kan dunner, is onze conclusie.

Zelf hebben we vast en zeker nog wat fouten laten zitten. Een eerste, van mijn eigen hand, is onlangs ontdekt: ik had het woord ‘aangeft’ laten staan wat onzin is. De trema moet eraf!

Helemaal ‘af’ zal de woordenlijst niet komen, want de Nederlandse taal wordt gemaakt door de Nederlanders. En die raken nooit uitgepraat. Dit wordt prachtig gellustreerd door Wim Danils en Felix van de Laar, de auteurs van het boek SPELLINGCHAOS. Zij schreven een verslag van hun naspeuringen naar de verwickelingen die hebben geleid tot de nieuwe spelling (de officile WOORDENLIJST). Het is ronduit onthutsend maar vaak ook vermakelijk om te lezen hoe de diverse hotemetoten van gewichtige instituten langs elkaar heen blunderen. Het blijkt dat de hoogste instanties op het gebied van de nieuwe spelling het op belangrijke punten met elkaar oneens zijn. In plaats van elkaars kritiek te verwerken in verbeteringen, bestrijdt men de kritiek en dat levert tot standaard verheven dwaasheden op (bij het schrijven van het volgende heb ik gegevens en zinnen ontleend aan het boek van Danils en Van de Laar):

- in de WOORDENLIJST zou de *dubbelspelling* nu verdwenen moeten zijn, zodat alleen *akkoord* nog goed is, en *accoord* niet meer. Maar: msli is gebleven naast muesli, shockeren naast choqueren, clerus naast klerikaal, vacant naast vakantie, sex-appeal naast seksshop en ga zo maar door.
- een nieuwe spellingregel zegt dat namen van tijdperken met een hoofdletter moeten, maar dat wordt officieel vergeten bij de oudheid, de bronstijd, de ijstijd en de steentijd.
- willekeur: bij de nieuwe spellingregels staat dat het woord *pistool* moet worden afgebroken als pi-stool,

maar het is wel waterpis-tool, laserpis-tool en laspis-tool!

Een stereo-installatie heeft een streepje na stereo, maar videoinstallatie niet.

- de officile WOORDENLIJST bevat tientallen spel- zet en/of drukfouten. Bij herdrukken van HET GROENE BOEKJE is een aantal daarvan stilzwijgend verbeterd en ook is er op een gegeven moment een velletje met *errata* bijgevoegd, maar nog altijd blijven er veel bekende fouten zitten (lijkt dit niet erg op de eerder gemelde houding van WP?).
- hoofdletters: sommige volkeren verdienen er wel een (de Batavier, de Saks) maar andere niet (de azteek, de zoeloe en de bosjesman). Bij de ziektes moet pfeiffer het zonder hoofdletter stellen, Alzheimer heeft hem wel gekregen. Het lijkt wel een lintjesregen! De oscar, bekende filmprijs, heeft geen hoofdletter, tenzij hij wordt uitgereikt: Oscaruitreiking.
- de tussen-n zorgt voor veel hilariteit. ‘Pannenkoek’ is de beroemdste geworden, maar het is erg gesteld met de belastingen. ‘Aangiftebiljet’ hoort zonder tussen-n maar hij moet wel staan in ‘belastingaangiftenbiljet’. Oorzaak is dat de samensteller van de officile woordenlijst — INSTITUUT VOOR NEDERLANDSE LEXICOLOGIE, INL samen met de TAALUNIE — is uitgegaan van een verzameling documenten die eerst als standaard is aangenomen en waaruit vervolgens pas de woorden zijn gesorteerd. De samensteller nam zich voor onder geen beding af te wijken van deze standaard-input en verdedigt deze dan ook te vuur en te zwaard tegen elke inmenging van buitenaf. . . *tenzij* de samensteller vindt dat er nog iets moet worden veranderd, bijvoorbeeld toen deze ontdekte dat wel het woord *bondsdagpresident* was opgenomen, maar niet het woord *bondsdag-president*. . .

Het stof, dat de introductie van de nieuwe spelling deed opwaaien, is nog niet gaan liggen. Het ANP hanteert de nieuwe spelling zoveel mogelijk, voor zover als ze het tenminste kunnen bijhouden. De Vlaamse krant DE STANDAARD vaart zijn eigen koers en geeft daarmee de toon aan voor veel andere kranten. NRC HANDELSBLAD heeft besloten het woord *product* te blijven spellen als *produkt*.

Slotsom

Mijn conclusie: nu de oude spelling wel door iedereen van gezag is verlaten, zijn er vele, soms kostbare en tijdrovende publicaties te koop die beloven een leidraad te geven. Geen enkel bedrijf biedt een leesbare en betrouwbare woordenlijst van enige omvang aan waarmee de gebruiker zelf een spellingchecker kan uitrusten. Het WORDS-L-team is bezig hiervoor een uitstekend alternatief aan te bieden. Inmiddels is de WORDS-L-mailinglijst³ vervangen

³Om een indruk te geven van het drukke en dagelijkse overleg: vanaf 11 juli tot 14 september werden er in WORDS-L 386 berichten geschreven.

door een sneller alternatief bij de TU EINDHOVEN. Door een spelfout (!) is deze per ongeluk SPELING⁴ gaan heten.

De bestanden en programma's die we gebruiken, alsmede de resulterende woordenlijsten, zijn te vinden op het Internet en op FGBBS. Gebruikers van *middle-of-the-road* tekstverwerkers als WP en MS-WORD willen we zelfs zoveel mogelijk tegemoet komen door het beschikbaar stellen van plug-and-play modules waarin onze lijst is verwerkt (bij de laatste versie van WP lijkt dat moeilijker en voor MS-WORD het lastigste, maar Hans Linders is hiervoor aan het werk gegaan). Ook is een van de leden van het WLT begonnen met het programmeren van een spellingchecker die nauwer samen zou moeten gaan werken met verschillende tekstverwerkers.

Diverse referenties op het Internet

Let wel: deze adressen en filenamen zijn in ontwikkeling, en daardoor aan verandering onderhevig.

- <http://www.pi.net/~fg/words.htm>
HTML-versie van dit artikel, met vele links naar relevante sites
- <ftp://ftp.tue.nl/pub/tex/GB95/>
Bestanden van het WORDS-L project
- <ftp://ftp.iaehv.nl/pub/users/branderh/words-l/>
Een *nix verzamelpackage, neergezet door Erick Branderhorst (branderh@debian.iaehv.nl)
- <ftp://ftp.tue.nl/pub/tex/GB95/spell-nl>
SPELL-NL bestanden van het WORDS-L project
In deze directory staan:
 - <ftp://ftp.tue.nl/pub/tex/GB95/spell-nl/LEESMIJ>
LEESMIJ, een toelichting
 - <ftp://ftp.tue.nl/pub/tex/GB95/spell-nl/COPYING.nl>
COPYING.nl, over copyrights

- <ftp://ftp.tue.nl/pub/tex/GB95/spell-nl/spell-nl-ansi.v3>
spell-nl-ansi.v3, woordenlijst versie 3, ANSI formaat (voor DOS / OS/2, WIN PC's)
- <ftp://ftp.tue.nl/pub/tex/GB95/spell-nl/spell-nl-ansi.zip>
ZIP versie spell-nl-ansi.v3, woordenlijst versie 3, ANSI formaat (voor DOS / OS/2, WIN PC's)
- <ftp://ftp.tue.nl/pub/tex/GB95/spell-nl/spell-nl-iso.v3>
spell-nl-iso.v3, woordenlijst versie 3, ISO-LATIN formaat
- <ftp://ftp.tue.nl/pub/tex/GB95/spell-nl/spell-nl-iso.zip>
ZIP versie spell-nl-iso.v3, woordenlijst versie 3, ISO-LATIN formaat
- <http://net.info.nl/ci/>
Computer Info
- <ftp://ftp.iaehv.nl/pub/users/branderh/words-l/gpl-nl.html>
html versie van de nederlandse public license

Enkele schriftelijke publicaties

- SPELLINGCHAOS, door Wim Danils en Felix van de Laar, Uitgeverij Scheffers, Utrecht, 1996, ISBN 90.5546.048.6
- WAT VERANDERT ER NU EIGENLIJK?, door Nel Korstanje en Jan Heerze, Walvaboek, Laren Gld, 1995, ISBN 90.6675.589 X
- VAN DALE EN DE NIEUWE SPELLING, van Dale Lexicografie, Utrecht / Antwerpen 1996, ISBN 90.6648.998.7 (met een beetje geluk gratis *ritselen* bij de betere boekhandel)
- SPELLINGBESLUIT, SDU uitgevers, Den Haag en Standaard Uitgeverij, Antwerpen, 1996 (regels en bijna 16.000 probleemwoorden, in de aanbieding voor fl 9,90)

⁴153 berichten in de eerste maand.