

Leesindexen berekenen, dat doe je niet zelf, dat laat je doen!

Marc Stevens

Doet u het ook wel eens, beste lezer, 'even' de moeilijkheidsgraad van een voor uw klas aantrekkelijke leestekst berekenen? Want geef toe, de teksten die u voor technisch of begrijpend lezen gebruikt, moeten toch min of meer gegradeerd zijn: wat makkelijker is, komt in principe voor wat moeilijker is. Maar bent u daar bij elke nieuwe tekst die u in uw klas brengt wel zo zeker van?

Of werkt u in uw klas ook met niveaulezen, zodat u goed op de hoogte moet zijn van de leesindex van elke tekst? En gebeurt het dan wel eens dat u 's avonds laat nog aan het rekenen moet slaan?

H

et 'berekenen' van een tekst is een klus. Zelfs om iets simpels als de leesindex A te berekenen (waarmee dan verder het AVI-niveau bepaald wordt, zie bijlage 3), ben je voor een niet eens zo lange tekst al snel een half uurtje kwijt. Vanzelfsprekend geeft die berekening je enig gevoel van zekerheid, maar of dit echt wel rendeert? Bovendien zijn tekstberekeningen niet méér dan hulpmiddelen die slechts enkele indicaties geven. Zo'n leesindex, zoals we de uitkomst van een tekstberekeningsformule wel noemen, zal je immers nooit vertellen of een tekst voor een individuele leerling geschikt is of niet. Dat heeft alles te maken met het feit dat leesindexen de technische elementen van teksten kunnen berekenen (zinslengte, woordlengte, ...), maar bijvoorbeeld niet het gegeven van de graad van vertrouwdheid van de lezer met zijn onderwerp. Een leerling die over zijn hobby leest, zal doorgaans moeilijkere teksten aankunnen dan een leerling die helemaal geen

voorkennis heeft i.v.m. het onderwerp van zijn tekst.

De discussie i.v.m. de bruikbaarheid van leesbaarheidsformules wordt al een tijdje gevoerd. Voorstanders ervan zeggen dat je in een tekst voldoende gegevens kan tellen en berekenen om zijn moeilijkheidsgraad voldoende correct in te schatten. Zij wijzen erop dat je talloze gegevens kan 'vastpakken'.

Er zijn de kwantitatieve tekstvariabelen zoals:

- gemiddelde zinslengte (aantal woorden per zin);
- gemiddelde woordlengte (meestal: aantal lettergrepen per woord);
- aantal herhaalde woorden (uitgebreidheid van de gehanteerde woordenschat);
- aantal ongebruikelijke woorden (vanuit een frequentielijst van woorden voor een bepaalde leeftijd);
- aantal (dubbele) negaties in de tekst;
- aantal genominaliseerde werkwoorden.

Er zijn ook de kwalitatieve tekstvariabelen, zoals:

- aantal keer een 'omgekeerde' woordvolgorde;
- aantal passieve zinnen in de tekst;
- aantal bijzinnen;
- aantal beknopte constructies;
- aantal tussenzinnen en tangconstructies;
- gemiddeld aantal ideeën per zin (hangt samen met aantal bijzinnen);
- hoeveelheid redundantie in het taalsysteem en de informatie die wordt doorgegeven (te veel informatie in een tekst creëert bijvoorbeeld te veel ballast, te weinig informatie vraagt een veel grotere inspanning van de lezer op het vlak van voorspellen en infereren).

Een aantal van deze laatste variabelen laten zich niet netjes tellen; soms moet er zelfs heelwat geïnterpreteerd worden. Dat brengt ons meteen bij de tegenstanders van leesbaarheidsberekening op teksten. Zij stellen namelijk dat in een tekst talloze aspecten niet in formules kunnen worden gedwongen, maar afhankelijk zijn van onderwerp tot onderwerp en van ontvanger tot ontvanger.

Zullen we de technische berekeningen dan maar afschaffen? De gemiddelde leerkracht die voor het vak Nederlands niet over wat extra tijd beschikt, doet dat maar best. Hij of zij werkt best ook maar verder met reeds 'getelde teksten', tenzij het berekenen geen extra belasting zou impliceren. Indien u snel een aantal gegevens over de moeilijkheidsgraad van een tekst zou kunnen te weten komen, blijft het interessant om zulke berekeningen uit te voeren en de resultaten vergelijkend te gebruiken.

voor het zonet geschetste doel kan gebruiken. In wat volgt, wil ik even ingaan op wat dit programma zoal kan en wat ook niet. Vooraf u toch nog even geruststellen: TANSY is een behoorlijk krachtig programma, maar toch is het gratis. Mocht u geïnteresseerd raken, dan vertel ik u graag dat u ongeveer enkel het schijfje en de verzendkosten hoeft te betalen. Het gaat hier dus om een shareware-programma, wat betekent dat het vrij gebruikt en gekopieerd mag worden. Willen we even kijken of er voor u ook een deurtje kan opengaan?

TANSY staat voor tekst-analyse-systeem, en analyseren is wat TANSY inderdaad ook doet. Vooraf moet je als gebruiker de tekst die je wil laten berekenen, ingetypt hebben met een tekstverwerker, en opgeslagen (op harde of zachte schijf) in ASCII- of DOS-formaat. Dat is een formaat dat de meeste tekstverwerkers kennen, zodat ons berekeningsprogramma niet aan elke tekstverwerker moet worden aangepast. In *Wordperfect* bijvoorbeeld kunt u een tekst in ASCII-formaat bewaren met de CTRL+F5-toetsen. Er dient nog bij vermeld dat omwille van de beperkte omgeving waarin werd geprogrammeerd, teksten niet langer kunnen zijn dan 110 regels en/of 1200 woorden. Indien dat toch het geval is, zal TANSY in de tekst knippen en met de eerste 110 regels verderwerken, zodat u toch nog een goede indruk krijgt van de complexiteit en het niveau van de tekst. TANSY is immers in eerste instantie bedoeld om teksten te berekenen waarmee in de klas (in principe van de lagere school) kan worden gewerkt, en niet om complete boeken na te tellen.

Het computerprogramma TANSY

Van toeval gesproken: op mijn computer, een doodgewone PC, draait momenteel het zelfgemaakte programma TANSY dat u

Technische leesindexen

U hoeft TANSY enkel maar te starten, de tekst van uw keuze aan te wijzen (in het hoofdmenu dat dan verschijnt, neemt u keuze 1) en de berekening te starten (in het hoofdmenu: keuze 3). In de volgende

afbeeldingen vindt u resultaten van een berekening die verricht werd op een tekst van Guus Kuijer.

T A N S Y (Tekst analyse systeem) Versie 4.30 van mei '93 Berekening van LEESINDEX A, Y en Z Programma: Marc Stevens	
Hoofdmenu	
<hr/> 1. Tekst inladen 2. Tekst bekijken/wijzigen 3. Berekenen van de leesindex 4. Frequentielijst v/d tekst drukken 5. Uitvallers bekijken 6. Resultaten verwerken 7. Geïntegreerde woordenlijsten bekijken 8. Zelf lijsten aanleggen (op basis van de onderzochte teksten). Functie is UIT 9. Stoppen <hr/>	
Typ uw keuze.	
Ingeladen tekst is: F:\AATEXTS\guus	

Luttele seconden later ziet uw computerscherm er ongeveer zo uit:

T A N S Y (Tekst analyse systeem) Versie 4.30 van mei '93 Berekening van LEESINDEX A, Y en Z Programma: Marc Stevens	
FREQUENT GEBRUIKTE WOORDEN: [bepte] [boes] [denkt] [heel] [lees] [maar] [noma] [niet] [stomme] [stoute] [word]	
Aantal zinnen : 44 Aantal woorden : 244 Aantal verschill. wdn.: 134 Aantal lettergrepen : 288 Woordvariatie : 55 % Schatting abstracte wdn.: 0 % Persoonlijke woorden : 10 % Handicap beginnend lezer : 1.80 Gemiddelde zinslengte : 5.55 Gemiddelde woordlengte : 1.18 Procent lange woorden : 0.00 Steen- vs. Conventwda.: 58% - 42%	Soorten woordbeelden + aantallen + % [nk km : 69] [nm : 86] => 64 % [(M)km : 20] [nm(M) : 9] => 12 % [nmkm : 3] [-nm- : 9] => 5 % [Wdn met aai, ooi... : 1] => 0 % [Meerlettergrep. wdn : 43] => 18 % LEESINDEX A (120>>0) : 105 LEESINDEX Y (150>>>0) : 60 Vermeendelijk AWI-niveau: 3 Geen verdere bijzonderheden.
FREQUENTIE-ANALYSE MOMENT (j/n)	

Even overlopen:

Bovenaan, onder de titel, krijgt u een beeld van de *frequent gebruikte woorden* in de tekst (vak met nummer 1). Dit geeft een indicatie i.v.m. het onderwerp of thema van de tekst.

Daaronder volgen berekeningen. Links in vak 2, een aantal afzonderlijke waarden die door telling of berekening werden verkregen, met name:

- *het aantal zinnen in de tekst;*
- *het aantal woorden dat werd geteld;*
- *het aantal verschillende woorden.* Deze waarde toont de uitgebreidheid van de woordenschat;
- *het aantal lettergrepen in de tekst.* Dat is een belangrijke waarde voor de berekening van leesindex A (zie later);
- *de woordvariatie in procent.* Dit getal geeft weer hoeveel verschillende woorden er worden gehanteerd t.o.v. het totaal aantal getelde woorden.
- *het aantal abstracte woorden in de tekst.* Hierbij dient vermeld dat het om een schatting gaat eerder dan om een sluitende waarde. Het programma is nl. in staat om een heleboel abstracte woorden te herkennen; exhaustief is het helaas op dit vlak niet. In de tekst van Kuijer die we hier bij wijze van voorbeeld berekenden, zitten blijkbaar weinig of geen abstracte woorden, wat de leesbaarheid en de begripelijkheid ten goede komt;
- *het aantal persoonlijke woorden.* Het gaat hier om de persoonlijke en reflexieve voornaamwoorden én een aantal zelfstandige naamwoorden die een concreet persoon aanwijzen, zoals 'moeder', 'vader', 'vriend' ... dus b.v. geen persoonsnamen die een beroep aanduiden. De 19% die we aantreffen in de tekst van Kuijer, is relatief hoog, zodat we mogen concluderen dat het niet om een afstandelijke tekst gaat, maar om wellicht een verhaal met concrete figuren. Het voorkomen van veel van deze woorden maakt de tekst leesbaarder.

Daaronder dan enkele resultaten van berekeningen:

- *de handicap voor een beginnend lezer.* Het programma gaat ervan uit dat een beginnend lezer die alle 'letters' kan lezen (of eigenlijk: alle grafeem-foneem-correspondenties correct maakt) met woorden van het type MKM (medeklinker-klinker-medeklinker) de minste leesproblemen heeft. Van zodra echter een woord complexer is opgebouwd, bijvoorbeeld met meer medeklinkers vooraan of achteraan, of met meer lettergrepen, slaat ons programma aan het tellen. Elk woord dat afwijkt van de MKM-bouw, heeft dus een zekere handicap die verrekend wordt. Een MKMM-woord (uitgezonderd woorden op -t) scoort 2 handicappunten, een woord met daarin 3 medeklinkers 'pakt' 6 punten enz. Zo ontstaat een waarde. Is die 1, dan betekent het dat er geen technische lees handicaps in de tekst zitten. Boven de 1,5 spreken we van een moeilijke tekst voor beginnende lezers. Boven de 2,5 impliceert dat een tekst in principe niet meer geschikt is voor de lagere school. In de tekst van Kuijer is de waarde 2,11. Dit wijst op een ongeschiktheid voor beginnende lezers.
- *de gemiddelde zinslengte.* Deze waarde geeft weer hoeveel woorden er gemiddeld in een zin zitten. Dus:

$$\text{Gemiddelde zinslengte} = \frac{\text{totaal aantal woorden in de tekst}}{\text{totaal aantal zinnen}}$$

- *de gemiddelde woordlengte* geeft weer hoeveel lettergrepen er gemiddeld in een woord zitten. Dus:

$$\text{Gemiddelde woordlengte} = \frac{\text{totaal aantal lettergrepen}}{\text{totaal aantal woorden}}$$

De laatste twee waarden hebben we nodig om straks (in deze tekst, maar ogenblikkelijk in het programma) de welbekende leesindex A te berekenen. Hier toont zich ook de kracht van de computer: u hoeft zelf niet meer twee of drie bladzijden aan één stuk lettergrepen te tellen. Dat doet de machine voor u.

- Dan volgt de waarde die weergeeft hoeveel procent lange woorden er in de tekst zitten. Het gaat hier om woorden die uit meer dan drie lettergrepen bestaan.
- Tot slot vinden we in deze kolom met afzonderlijke waarden een duo dat weergeeft hoeveel woorden er in de tekst *bouwsteen* zijn, dus semantisch sterk gevuld en hoeveel procent er fungeren als *cement- of mortelwoord*, dus verbanden of verwijzingen uitdrukkelijk (functiewoord).

Al deze waarden zijn interessant genoeg om apart geïnterpreteerd te worden en u een beeld te geven van de tekst die u in handen heeft, maar een aantal ervan wordt ook gebruikt bij verdere berekeningen die ons programma maakt. Vandaar dat we nog even verder overlopen.

Rechts op ons scherm, in vak 3 treffen we een analyse aan die voor de leerkracht van het eerste leerjaar van belang is. Het programma gaat nl. van elk woord na wat de opbouw is in termen van medeklinker en klinker, zodat we snel een beeld krijgen van het aantal moeilijkheden op dat vlak. Uiterst rechts wordt er ook procentueel getoond hoe de moeilijkheden verdeeld zitten.

Tot slot van dit eerste scherm verschijnen er twee technische leesindexen en wordt het vermoedelijke AVI-niveau van de tekst gegeven (vak 4):

- *Leesindex A* wordt algemeen gebruikt om het technische leesniveau te bepalen. Het is een 'klassieker' en afgeleid van een formule die R. Flesch ooit in het leven riep:

$$\text{Leesindex A} = 195 - (2/3 \text{ van de gemiddelde woordlengte} \times 100) - (2 \times \text{gemiddelde zinslengte})$$

Op basis hiervan wordt het AVI-niveau van de tekst vastgesteld. Hoe dat exact gebeurt, leest u in bijlage 3. De waarde van deze leesindex ligt tussen 120 en 0. Een hogere waarde betekent een makkelijker tekst. Doorgaans wordt er gesteld dat een tekst voor het eerste leerjaar zich situeert tussen 120 en 100, maar hij moet bovendien nog aan enkele andere eisen voldoen.

- *Leesindex Y* dan is een index die een gedetailleerder beeld van de tekst geeft, althans als we teksten voor beginnende lezers berekenen op hun technische moeilijkheidsgraad. Het resultaat van de berekening is een waarde tussen 150 en 0. Deze waarde loopt sneller naar beneden dan die van leesindex A. Algemeen genomen zal deze waarde nul worden bij teksten die door een gemiddelde lezer goed gelezen worden aan het einde van het tweede leerjaar. De formule is een tikkeltje moeilijker dan de voorgaande:

$$\text{Leesindex Y} = 254 - (\text{handicap beginnend lezer} \times 100) - \text{gemiddelde zinslengte}$$

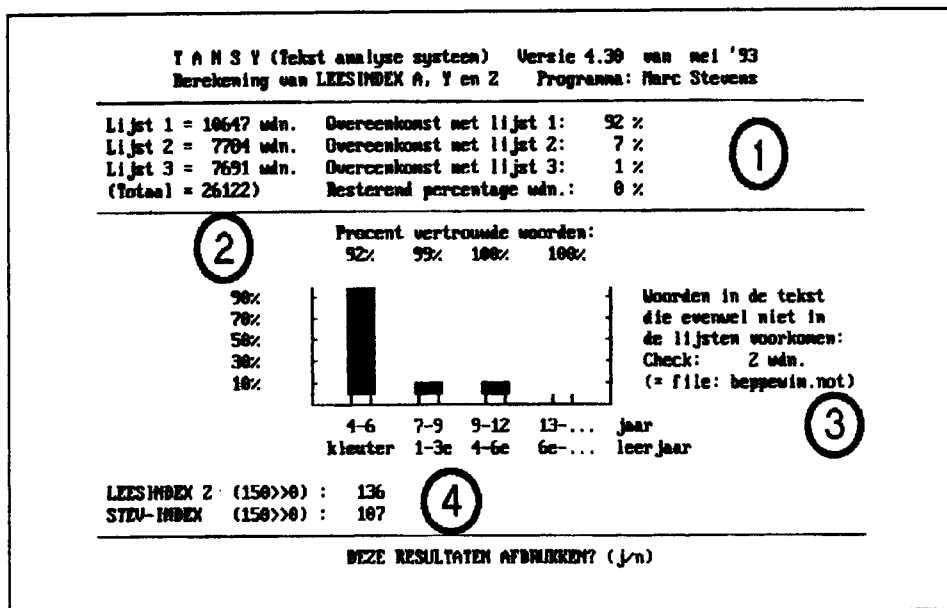
Indexen op basis van woordfrequentie

Tot daar ons eerste cijferscherm. Indien u nu de J-toets indrukt, komen we bij het scherm waarin indexen worden berekend op basis van de frequentie van woorden. In TANSY werden 3 woordenlijsten verwerkt op basis van woordfrequentie. De eerste omvat de meest frequente woorden die door kleuters worden begrepen (ongeveer 3000 basiswoorden), aangevuld met in het Nederlands frequent gebruikte voornamen. De tweede omvat de volgende 2500 meest frequente woorden. Dat zijn er die in principe goed begrepen worden door leerlingen van het eerste tot het derde leerjaar. De

derde en laatste lijst voegt daar weer ongeveer 2500 woorden aan toe. Het gaat om woorden die begrepen worden door vierde- tot zesdeklassers.

Ons computerprogramma vergelijkt elk woord uit de opgegeven tekst met die

woordenlijsten, zodat we te weten komen hoeveel woorden uit de tekst in principe door een bepaalde groep kinderen kunnen worden begrepen. Op ons scherm wordt dit resultaat afgebeeld d.m.v. cijfers en een grafiekje:



In vak 1 ziet u naast het aantal woordvormen in elke lijst het percentage van woorden uit de tekst die een overeenkomstig woord vonden in de lijst. Uiteraard wordt na onderzoek van de drie lijsten ook het resterende percentage vermeld. Het gaat hier om de woorden uit de tekst die niet aangetroffen werden in onze lijsten. Die woorden worden overigens ook weggeschreven naar de schijf (gegevens daarover vindt u op het scherm rechts in het midden bij het grote cijfer 3).

Bij nummer 2 werden dezelfde gegevens in grafiekvorm gepresenteerd. In onze voorbeeldtekst (van Guus Kuijer), blijkt 92% van de woorden tot de gemiddelde woordenschat van 6-jarigen te behoren. Acht procent van de woorden is voor die groep kinderen wellicht niet gekend. Voor de groep tot 9 jaar zijn 96% van de woorden

gekend. Dat is al 4% meer. Vraag is natuurlijk hoe deze getallen genterpreteerd dienen te worden. Veel onderzoek daarover ken ik niet. Wel weet ik dat A. Vermeer stelt dat om een tekst te begrijpen 90% van de woorden gekend dienen te zijn. Zelf zou ik dat percentage iets hoger leggen. In ieder geval lijkt de tekst van Kuijer op het vlak van woordenschat haalbaar voor 7- tot 9-jarigen.

Op basis van de onderzochte overeenkomsten met de woordenlijsten, berekent ons programma *leesindex Z* (links van grote cijfer 4). Het resultaat is weer een waarde tussen 150 en 0. De formule van deze index is nogal complex, daarom nemen we ze hier niet op. Maar het komt er op neer dat er van waarde 150 meer punten worden afgetrokken naarmate we bij woorden komen die gemiddeld door minder kinde-

ren begrepen worden. Dit is uiteraard helemaal geen index meer om via pen en papier te berekenen, maar eentje om de computer tussen een halve en een volledige minuut mee zoet te houden.

Tot slot berekent het programma nog een vierde leesindex. Dat is, zeg maar, een conglomeraat van indexen Y en Z, samen met nog enkele reeds apart aangeduide waarden, met name het aantal abstracte woorden (meer=moeilijker), het aantal persoonlijke woorden (meer=makkelijker) en de graad van woordvariatie (meer=moeilijker). Dat levert de *stevindex* op, die zich ook weer situeert tussen 150 en 0.

Op dit ogenblik durven we nog niets absoluut zeggen over de waarden die we hier verkrijgen, maar de uitkomsten zijn alleszins wel te gebruiken om teksten mee te vergelijken. De auteur zou het overigens geweldig op prijs stellen mocht u hem de teksten die u in uw klas gebruikt, op een schijfje zetten (met vermelding van de klas en de maand waarin een bepaalde tekst aan bod komt) en toesturen, zodat door de praktische input het programma nog beter op punt kan worden gesteld.

Het programma gebruiken

Waarvoor kan u dit programma dan gebruiken? Dit hangt een beetje af van de klas waarin u lesgeeft. In het algemeen kan u met TANSY op enkele verschillende manieren de moeilijkheidsgraad van een tekst berekenen en die waarden vergelijken met die welke u berekende voor andere teksten, zodat vergelijking van teksten op een min of meer objectieve wijze mogelijk is. In het programma is ook een routine ingebouwd (in het hoofdmenu keuze nummer 6) die de berekende gegevens telkens opslaat, zodat u snel en overzichtelijk resultaten (dus ook teksten) met mekaar kan vergelijken. Werkt u met niveaugroepen in uw klas, dan kan u dit programma gebruiken om teksten

voor bepaalde groepen te kiezen. U hoeft dan niet echt afhankelijk te zijn van een methode waarin reeds berekende teksten op systematische manier mekaar opvolgen. U kan nu zelf bijvoorbeeld aantrekkelijke fragmenten uit de kinderliteratuur halen en daarin na berekening een volgorde aanbrengen.

Prettig is ook dat u met de kleine tekstverwerker die in het programma werd ingebouwd (in het hoofdmenu keuze nummer 2) de mogelijkheid heeft om voor of na berekening van teksten, daarin snel wijzigingen aan te brengen (technisch of inhoudelijk moeilijke woorden verwijderen, zinnen verkorten, enz.) en dan de teksten te herberekenen, zodat die ook beter passen in een lijn van stijgende moeilijkheidsgraad. Het spreekt dat u op dezelfde manier met één tekst aan de slag kan gaan om op basis daarvan een moeilijker en een makkelijker versie te schrijven, zodat u in niveaugroepen kan werken, die uiteindelijk dan toch min of meer dezelfde tekst doorgemaakt zullen hebben.

Taakleerkrachten zullen met dit programma weer op een eigen manier aan de slag kunnen. Het zorgvuldiger ordenen van hun materialen lijkt mij één van de vele mogelijkheden.

Een NT2-leraar zal handig gebruik kunnen maken van de verschillende indexen om de leesbaarheid van teksten voor anderstaligen in te schatten. Daarbij kan hij op basis van de aangegeven frequente woorden bepaalde teksten gebruiken en aanpassen om die woorden te semantiseren en/of consolideren (= woordenschat uitbreiden).

U ziet dat ons programma legio mogelijkheden aanreikt om flexibel met teksten aan de slag te gaan. Wellicht krijgt u er al zin in, daarom toch nog even enkele praktische punten:

- Bij het intikken van een tekst moet u ervoor zorgen dat dat op de gangbare manier gebeurt; dus leestekens vlak na een woord plaatsen en liefst telkens slechts één spatie tussen 2 woorden

laten. Met overdreven veel spaties tussen de woorden wil TANSY wel eens verkeerd tellen.

- De regellengte mag ook niet te groot zijn. Meestal betekent dat dat u niet met een te klein lettertype mag werken. In ieder geval moet u de automatische splitsing aan het einde van de regel uitzetten. TANSY werkt niet correct wanneer er gesplitste woorden in de te berekenen tekst voorkomen.
- Zoals reeds gezegd, dient u er ook op te letten dat u de tekst niet op de normale manier maar in ASCII-formaat naar de schijf wegschrijft.
- Het is goed om weten dat dit programma bij lange na niet perfect is. Bij het tellen van lettergrepen duikt er wel eens een foutje op. U moet ermee rekening houden dat er een kleine afwijking (ongeveer 1%) kan optreden bij de berekening. Maar die fout is zo klein dat u er zich niet druk hoeft over te maken. Wij doen dat wel, en zijn er in de laatste versie in geslaagd om een heleboel afkortingen op correcte manier te verwerken. In een afkorting als 'a.h.w.' ziet de computer immers drie punten staan, zodat hij nogal snel geneigd is om te denken dat het hier om drie zinnen gaat!
- In tegenstelling tot andere doorsnee computerprogramma's is TANSY niet duur. Wil u het ook hebben, dan stuurt u een omslag met daarin 4 postzegels van 15 fr en 2 briefjes van 100 fr. naar *Marc Stevens, Constant Jorislaan 11, 2100 Deurne*. Wie bestelt, wordt wel gevraagd om de teksten die hij/zij ingetikt heeft ook een keertje op schijf terug te zenden aan de auteur, samen met zijn/haar bevindingen i.v.m. het programma. Alleen op die manier is een degelijke verbetering mogelijk, en kan u eigen wensen m.b.t. het programma ook werkelijkheid zien worden. Wie weet wordt het programma na verloop van tijd voor u op maat gesneden.

Toch even relativieren

Eigenlijk is TANSY bedoeld voor teksten waarmee in de lagere school wordt gewerkt, maar ik ken al gebruikers die ook stiekem de leesindexen berekenen op hun eigen teksten; kwestie van eens nagaan hoe makkelijk of moeilijk men zelf schrijft. En dat mag best hoor, zolang dit programma voor u geen absolute maat gaat neerzetten i.v.m. de moeilijkheidsgraad van een tekst. De waarschuwing die ik uitsprak aan het begin van mijn betoog, blijft dus gelden: gebruik het programma om teksten te vergelijken of om een vermoeden i.v.m. de moeilijkheid van een tekst bevestigd te zien; gebruik het a.u.b. niet om zelf niet meer te moeten lezen. De interesse van uw leerlingen en uw eigen aanvoelen daarvan bepalen nog steeds in grote mate of een tekst zinvol en waardevol is voor uw klas. Dat is immers een beslissing die geen enkel computerprogramma in uw plaats kan nemen! Overigens dient vermeld dat dit programma niet op 100% wetenschappelijke manier tot stand is gekomen, maar via het concretiseren van een aantal hypothesen en het woordtelwerk van anderen (zie daarvoor bibliografie). Volledig sluitend zijn de berekeningen (nog) niet. Dat hoeven ze wat mij betreft ook niet te worden. De zinvolheid en aantrekkelijkheid van het leven worden nu eenmaal niet door computers berekend! Wat kost een avontuur uit een spannend boek? En hoeveel punten krijgt een verhaal waarin een kind ontdekt dat het niet alleen staat met een probleem en waarin het een oplossing voor dat probleem aangereikt krijgt?

Het was niet mijn eerste bedoeling, beste lezer, maar ik merk dat u nu toch eventjes met de mond vol tanden zit. Het programma blijft evenwel bestelbaar.

*Marc Stevens
Constant Jorislaan 11
2100 Deurne*

Bibliografische springplankjes

Coenen, M. & A. Vermeer: **Nederlandse woordenschat allochtone kinderen**. Zwijssen, 1988.

Damhuis, R. e.a.: **Woordenlijst voor 4- tot 6-jarigen. Een streeflijst voor kleuters**. Rotterdam: Projectbureau OVB.

Jaspaert, K. & K. Van Gorp: **Woordenlijst voor de lagere school** (voorlopige versie). Leuven: Steunpunt NT2, 1990.

Kohnstamm, G. A. e.a.: **Nieuwe streeflijst woordenschat voor 6-jarigen**. Lisse: Swets & Zeitlinger, 1981.

Koning, L.: **Speciale leesbegeleiding**. Den Bosch: Malmberg.

Struiksma, A.J.C., A. van der Leij & J.P.M. Veijs: **Diagnostiek van technisch lezen en aanvankelijk spellen**. Amsterdam: VU Uitgeverij, 1991 (4e druk).

Van den Berg, R.M. & H.G. Te Lintelo: **Individualisering van het leesonderwijs**. 's-Hertogenbosch: Katholiek Pedagogisch Centrum, 1977.

Van Hauwermeiren, P.: **Het leesbaarheidsonderzoek**. Groningen: H.D. Tjeenk Willink, 1975.

Wieringa, C.J. & J.F. Sas: **Leesmoeilijkheden, diagnose en behandeling**. Groningen: Wolters-Noordhoff, 1990 (2e druk).

BIJLAGEN

In de eerste twee bijlagen vindt u twee berekeningsresultaten zoals die uit de printer tevoorschijn komen wanneer het programma zijn werk gedaan heeft. Het eerste betreft het voorbeeld dat we in het artikel gebruikten, nl. de berekening op basis van een tekst van Guus Kuijer (bijlage 1). Het tweede voorbeeld toont berekeningsresultaten van een tekst van Roald Dahl (bijlage 2). Het is wellicht interessant beide ook even met mekaar te vergelijken. Uit een heleboel gegevens komt nl. naar voor dat de tekst van Dahl moeilijker is dan die van Kuijer.

Bijlage 1

TEXT ANALYSE SYSTEEM
Berekening van LEXINDEX A, Y en Z Programma: Marc Stevens
versie 4.30 --- van mei 1993

blad 1

ANALYSE VAN DE TEKST: [F:\ARTKST\gsm]

De tekst begint als volgt:

Pasje.

Modelief heeft een klein passje gekregen. So houdt het ...
mar het is heel wild. Met wil steeds naar beneden. Maar ...
spelen. ...

Aantal zinnen: 51
Aantal woorden: 345
Aantal verschillende wds.: 165
Aantal lettergrepen: 452

Woordgrepen: modelief/klinter opaanvolgingen:

Woordvorm: Aantal wds. + % Voorbeelden:

- M / m : 90 (-26%)
- Hm : 122 (-35%)
- (H)hm : 35 (-10%)
- Hm(H) : 4 (-1%)
- Hmhm : 10 (-3%)
- -hm- : 1 (-0%)
- -Hm- : 4 (-1%)
- maele.gre.: 79 (-23%)

Waardvariatie = 48.00 %
Persoonlijke woorden = 10.00 %
Handloop aanvankelijk lezer = 2.11 (Handloop van 303 punten)
Schattig abstracte woorden = 0.00 (0 woorden)
Stem- vs. Consonanten = 55% - 45%

Gemiddelde zinlengte: 6.76
Gemiddelde woordlengte: 1.21
Procent lange woorden: 1.16

LEXINDEX A = (120>>0) 94
Verwachtelijk AVI-niveau: 5
LEXINDEX Y = (150>>0) 36

Geen verdere bijzonderheden.

WOORDEN MET EEN SPYFVUEL TECHNISCHE MOEILIKHEID:

PROFESSE VOORNAME:
[haar] [hoort] [heel] [meer] [modelief] [meer] [niet] [pasje] [veel] [vegt] [zijn]

TEXT ANALYSE SYSTEEM
Berekening van LEXINDEX A, Y en Z Programma: Marc Stevens
versie 4.30 --- van mei 1993

blad 2

ANALYSE VAN DE TEKST: [F:\ARTKST\gsm]

GEGEVENS N.A.V. OVERZICHT MET WOORDENLIJSTEN:

Overeenkomst met lijst 1 (kleuter) = 91.52 % (151 woorden)
Overeenkomst met lijst 2 (lagere klassen) = 4.24 % (7 woorden)
Overeenkomst met lijst 3 (hogere klassen) = 0.61 % (1 woorden)
Niet-herkende woorden van de tekst = 3.64 % (6 woorden)

OVERZICHT:

Loefijde	4-6	7-9	9-12	13>>
Leerjaar (België)	kleuter	1-3e	4-6e	secund.
Groep (Nederland)	1-2	3-5	6-8	hoger
Vertroude	v	v	v	v
4 gekende woorden	92%	96%	97%	100%

LEXINDEX X = (150>>0) 124
STY-INDEX = (150>>0) 110

SAMENVATTING & CONCLUSIES:

LESTERINDEX:

A	Y	Z	STY-INDEX
(van 120 tot 0)	(150>0 - maal)	(van 150 tot 0)	(van 150 tot 0)
94	36	124	110

GEKENDHEID + VERKLEINING OP EEN SCHAA VAN 120 TOT 0:

GEKENDHEID VAN A EN Z (NAAR VAN 120 TOT 0 !)	STY-INDEX (van 120 tot 0)
94	88

- Deze tekst lijkt geschikt voor het (2e-) 3e leerj. (B) / groep 5 (W).

Bijlage 2

TEXT ANALYSE SYSTEEM
Berekening van LEXINDEX A, Y en Z Programma: Marc Stevens
versie 4.30 --- van mei 1993

blad 1

ANALYSE VAN DE TEKST: [F:\ARTKST\acc1001]

De tekst begint als volgt:

Laister den naar," sei Bertonia. "Bistoren heeft de Bu ...
[Joupe, Julius Betsuinkel, hetrept op het eten van Eng ...
de schrijfties. Se pakte hem op bij zijn era en elingd ...
open voor naar buiten. Gue klas is op de eerste verdie ...
Julius Betsuinkel) over de tuin heen seilen als een fris ...

Aantal zinnen: 95
Aantal woorden: 950
Aantal verschillende wds.: 400
Aantal lettergrepen: 1443

Woordgrepen: modelief/klinter opaanvolgingen:

Woordvorm: Aantal wds. + % Voorbeelden:

- M / m : 260 (-26%)
- Hm : 217 (-20%)
- (H)hm : 56 (-6%)
- Hm(H) : 21 (-2%)
- Hmhm : 10 (-1%)
- -hm- : 15 (-2%)
- -Hm- : 11 (-1%)
- maele.gre.: 336 (-35%)

Waardvariatie = 42.00 %
Persoonlijke woorden = 13.00 %
Handloop aanvankelijk lezer = 2.00 (Handloop van 1322 punten)
Schattig abstracte woorden = 0.53 (5 woorden)
Stem- vs. Consonanten = 52% - 48%

Gemiddelde zinlengte: 10.00
Gemiddelde woordlengte: 1.52
Procent lange woorden: 1.38

LEXINDEX A = (120>>0) 74
Verwachtelijk AVI-niveau: 9
LEXINDEX Y = (150>>0) 0

Geen verdere bijzonderheden.

WOORDEN MET EEN SPYFVUEL TECHNISCHE MOEILIKHEID:

PROFESSE VOORNAME:
[aande] [inistrenk] [haar] [bertonia]

TEXT ANALYSE SYSTEEM
Berekening van LEXINDEX A, Y en Z Programma: Marc Stevens
versie 4.30 --- van mei 1993

blad 2

ANALYSE VAN DE TEKST: [F:\ARTKST\acc1001]

GEGEVENS N.A.V. OVERZICHT MET WOORDENLIJSTEN:

Overeenkomst met lijst 1 (kleuter) = 73.50 % (294 woorden)
Overeenkomst met lijst 2 (lagere klassen) = 16.75 % (67 woorden)
Overeenkomst met lijst 3 (hogere klassen) = 6.75 % (27 woorden)
Niet-herkende woorden van de tekst = 3.00 % (12 woorden)

OVERZICHT:

Loefijde	4-6	7-9	9-12	13>>
Leerjaar (België)	kleuter	1-3e	4-6e	secund.
Groep (Nederland)	1-2	3-5	6-8	hoger
Vertroude	v	v	v	v
4 gekende woorden	74%	91%	96%	100%

LEXINDEX X = (150>>0) 94
STY-INDEX = (150>>0) 80

SAMENVATTING & CONCLUSIES:

LESTERINDEX:

A	Y	Z	STY-INDEX
(van 120 tot 0)	(150>0 - maal)	(van 150 tot 0)	(van 150 tot 0)
74	0	94	80

GEKENDHEID + VERKLEINING OP EEN SCHAA VAN 120 TOT 0:

GEKENDHEID VAN A EN Z (NAAR VAN 120 TOT 0 !)	STY-INDEX (van 120 tot 0)
74	64

Bijlage 3: de AVI-niveaus

We onderscheiden 9 AVI-leesniveaus. Die werden aanvankelijk bedacht door het KPC (Katholiek Pedagogisch Centrum) van 's- Hertogenbosch. Bepalend zijn: woordbouw, woordlengte, zinslengte en bij de eerste niveaus ook lay-out. Het lijkt erop dat dit centrum van plan is de normen voor de eerste vier niveaus lichtjes te wijzigen, wellicht in de richting van de indeling zoals die voorgesteld wordt door L. Koning (Map: Speciale leesbegeleiding), die een veel genuanceerdere invulling suggereert.

Niveau 1 (eerste leerjaar)

- éénlettergrepige woorden van de typen: KM, MK, MKM, Mkm+t;
- de basiswoorden uit een leesmethode;
- korte zinnen; één zin per regel;
- geen hoofdletters;
- leesindex A: 110 en hoger.

Niveau 2 (eerste leerjaar)

- éénlettergrepige woorden, ook met 'd' en 'b' aan het eind;
- persoonsvorm-t na twee medeklinkers (uitgezonderd als de tweede een 'd' is!);
- 'ng' of 'nk' aan het eind van een woord;
- dubbele medeklinker aan het begin van een woord (MMK...-woord);
- tweelettergrepige woorden zonder spellingsmoeilijkheden (je zegt wat je ziet);
- verkleinwoorden;
- korte zinnen; één zin per regel;
- eventueel hoofdletters;
- leesindex A: 110 en hoger.

Niveau 3 (soms eerste leerjaar, niet voor alle kinderen)

- tweelettergrepige woorden met klinker in open lettergreep (tafel, komen, na, nu...);
- persoonsvorm-t achter een stam die eindigt op 'd' (houdt, landt,...);
- woorden met drie medeklinkers aan het begin of vier medeklinkers aan het eind;
- drie- en meerlettergrepige woorden zonder spellingsmoeilijkheden (boerderij ...);
- woorden met voorvoegsels ge-, be-, ver-, (want een nieuw probleem: beken);
- korte zinnen die wel kunnen overlopen op de volgende regel;
- hoofdletters;
- leesindex A: 109 - 100.

Niveau 4

- woorden eindigend op -lijk, -ig, enz. (met doffe klank);
- moeilijkere drie- en meerlettergrepige woorden (muzikanten,...);
- korte hoofdzin/bijzin-combinaties;
- niet te lange, overlopende zinnen;
- leesindex A: 109 - 100.

Niveau 5 (eind tweede, begin derde leerjaar)

- leesindex A: 99 - 94;
- gemiddelde woordlengte van 1,27 lettergrepen (min 1,23 / max 1,30);
- gemiddelde zinslengte van 7 woorden.

Niveau 6

- leesindex A: 93 - 89;
- gemiddelde woordlengte van 1,32 lettergrepen (min 1,29 / max 1,35);
- gemiddelde zinslengte van 8 woorden.

Niveau 7

- leesindex A: 88 - 84;
- gemiddelde woordlengte van 1,36 lettergrepen (min 1,34 / max 1,39);
- gemiddelde zinslengte van 9 woorden.

Niveau 8

- leesindex A: 83 - 79;
- gemiddelde woordlengte van 1,41 lettergrepen (min 1,38 / max 1,44);
- gemiddelde zinslengte van 10 woorden.

Niveau 9

- leesindex A: 78 - 74;
- gemiddelde woordlengte van 1,46 lettergrepen (min 1,43 / max 1,48);
- gemiddelde zinslengte van 11 woorden.