

Leesbaarheid van teksten

Wanneer is een tekst goed of slecht leesbaar? Daarover bestaan verschillende meningen. Onderzoek ernaar heeft ook niet een eenduidig antwoord opgeleverd. Het heeft geleid tot heel verschillende wiskundige formules. Deze worden gebruikt om snel een indruk te kunnen krijgen of teksten redelijk leesbaar zijn.

Eén van deze leesbaarheidsformules (namelijk de formule van Flesch) wordt gebruikt in sommige versies van het tekstverwerkingsprogramma Microsoft Word. Je kunt in een filmpje zien hoe dat bij een tekst werkt en welke resultaten je krijgt.

 Start het filmpje *LEESBAARHEID_WORD.EXE*.

Met de knoppenbalk (zie figuur 3) kun je het filmpje terug laten gaan, laten stoppen of verder gaan en afsluiten.

figuur 3



In het filmpje voert iemand een aantal opdrachten uit om van een bestand (*FRANK-JANSEN-LEESBAARHEID.DOC*) de leesbaarheidsindex *FreS* te bepalen. Dit is een afkorting voor Flesch Reading Ease Score.

figuur 4



Statistiek Leesbaarheid			
Aantallen			
Tekens	2060	Korte zinnen	12
Lettergrepen	692	Lange zinnen	1
Woorden	379	Eenvoudige zinnen	16
Zinnen	22	Lange woorden	55
Alinea's	7		
Gemiddelden			
Lettergrepen per woord			1,82
Woorden per zin			17,22
Zinnen per alinea			3,14
Leesbaarheid			
Uw Flesch Reading Ease score (100 = zeer eenvoudig)			35
Passieve vorm (% van werkwoordelijke constructies)			4
Zinscomplexiteit (100 = zeer ingewikkeld)			27
Lexicale complexiteit (100 = zeer ingewikkeld)			28

Word maakt bij het document een rapport (zie figuur 4). De Flesch Reading Ease Score (*FreS*) is gelijk aan 35.

In het filmpje zag je ook dat er een helppagina is die uitleg geeft over de *FreS*. In figuur 5 zie je een deel van deze helppagina bij de leesbaarheidsstatistiek weergegeven.

figuur 5

Leesbaarheid

Het dialoogvenster **Statistiek Leesbaarheid** bevat een evaluatie van de leesbaarheid van uw document. De statistische gegevens kunt u als richtlijn gebruiken. U moet de statistieken zeker niet zien als een oordeel of rapportcijfer.

De volgende criteria voor leesbaarheid worden gehanteerd:

♦ Flesch Reading Ease Score

Een hoog cijfer betekent een goed leesbare manier van schrijven. Een laag cijfer betekent een ingewikkelde manier van schrijven. Toelichting:

Score	Interpretatie
90-100	Zeer eenvoudig
80-90	Eenvoudig
70-80	Tamelijk eenvoudig
60-70	Standaard
50-60	Tamelijk moeilijk
30-50	Moeilijk
0-30	Zeer moeilijk

De formule voor de Flesch Reading Ease Score is:

$$206,835 - (1,015 \times \text{gemiddeld aantal woorden/zin}) - (84,6 \times \text{gemiddeld aantal lettergrepen/woord}) = \text{Score}$$

Door in het laatste scherm op het kruisje van de knoppenbalk (zie figuur 6) naast de *i* te klikken, sluit je het filmpje.

figuur 6



 Sluit het filmpje af.

Complex wiskunde A1 vwo 2005-I

havovwo.nl

Als formule voor de Flesch Reading Ease Score nemen we:

$$FreS = 206,835 - (84,6 \times \text{Woordlengte}) - (1,015 \times \text{Zinslengte})$$

of

$$FreS = 206,835 - 84,6 \cdot W - 1,015 \cdot Z$$

Hierin geldt:

- W (= woordlengte) is het gemiddelde aantal lettergrepen per woord;
- Z (= zinslengte) is het gemiddelde aantal woorden per zin.

Op deze $FreS$ en figuur 5 komen we bij vraag 19 terug.

- 3p **14** Leg uit hoe je aan deze formule van $FreS$ kunt zien dat een hoge $FreS$ op een eenvoudig te lezen tekst duidt.

Reading-age

Op het internet wordt software aangeboden waarmee de zogeheten reading-age (leesleeftijd) van Engelse teksten bepaald kan worden. De reading-age geeft aan op welke leeftijd de gemiddelde lezer 50% correcte antwoorden geeft op een test over begrijpend lezen.

Een hoge reading-age betekent dat de tekst moeilijk te begrijpen is. Men moet immers een wat hogere leeftijd hebben om vragen over de tekst te kunnen beantwoorden.

Ondanks dat leesbaarheidsformules beperkt bruikbaar zijn, worden ze toch veel gebruikt door schrijvers van handleidingen van technische apparaten, politieke pamfletten, studieboeken, enzovoorts.

Er zijn verschillende formules om van Engelse teksten de reading-age te berekenen. In het bestand READING-AGE.XLS zijn er drie verschillende gegeven: de FOG-formule, de FK-formule en de SMOG-formule. De reading-age wordt daar op drie verschillende manieren berekend met behulp van één of meer van de begrippen:

- W = woordlengte,
- Z = zinslengte en
- L = gemiddeld aantal *lange woorden* per zin; een lang woord is een woord van drie of meer lettergrepen.



Open het bestand READING-AGE.XLS

In het bestand READING-AGE.XLS blad 1 kun je de cellen D7 (W), D8 (L) en D9 (Z) invullen. In de cellen D14, D15 en D16 worden dan de reading-ages volgens de FOG-formule, de FK-formule en de SMOG-formule berekend.

- 3p **15** Onderzoek van welke variabele(n) W , L en Z de SMOG-reading-age afhankelijk is. Licht je werkwijze toe.

De FOG-formule is: $\text{reading-age} = \left(\frac{100 \cdot L}{Z} + Z\right) \cdot 0,4 + 5$.

Iemand doet over de FOG-reading-age voor teksten met $L = 2$ de volgende bewering:

Als we in zo'n tekst (met $L = 2$) bij de bestaande zinnen woorden toevoegen met minder dan drie lettergrepen, blijft het aantal lange woorden en dus L gelijk, maar krijgen we wel langere zinnen (Z is groter geworden). Je mag verwachten dat de nieuwe tekst hierdoor moeilijker leesbaar wordt en dat de FOG-reading-age toeneemt.

- 5p **16** Toon met getallenvoorbeelden aan dat dit niet altijd het geval is.

De reading-age volgens de FK-formule wordt berekend met een formule van de vorm:

$$\text{reading-age} = a \cdot W + b \cdot Z + c.$$

De waarden van a , b en c zijn respectievelijk ongeveer 12, 0,4 en -11 .

In READING-AGE.XLS blad 1 wordt de FK-formule gebruikt met de precieze waarden van a , b en c .

- 5p **17** Bereken met behulp van 'handige' keuzes in READING-AGE.XLS blad 1 de waarden van a , b en c in twee decimalen nauwkeurig.

Vaak zullen we bij het analyseren van een tekst niet direct kunnen beschikken over W , L en Z . Deze gemiddeldes zullen berekend worden aan de hand van tekstkenmerken die we wel direct kunnen tellen, namelijk: aantal zinnen, aantal woorden, aantal woorden van 1 lettergreep, aantal lange woorden en het totaal aantal lettergrepen.

In het bestand READING-AGE.XLS blad 2 zie je in de cellen D7 tot en met D10 waarden die horen bij deze tekstkenmerken, namelijk respectievelijk 2, 26, 5 en 44.

In het bestand zijn de cellen D14, D15 en D16 leeg. Daar moeten formules komen die horen bij de omschrijvingen van C14, C15 en C16.

- 4p **18** Schrijf op welke Excel-formules in D14, D15 en D16 in de spreadsheet ingevuld moeten worden zo, dat met de uitkomsten ervan de reading-ages van FOG, FK en SMOG in de cellen D20, D21 en D22 berekend kunnen worden.
Schrijf op welke uitkomsten je krijgt voor de reading-age in de cellen D20, D21 en D22 bij de gegeven beginstand van D7 tot en met D10.

Leesindices

De *FreS*-formule uit het filmpje, aan het begin van deze opgave, is voor Engelse teksten ontworpen. Brouwer ontwikkelde een variant van deze formule - *Leesindex Brouwer* - die beter bij de Nederlandse taal zou passen. Ook deze index wordt berekend met behulp van W en Z . Opgemerkt kan worden dat naarmate de waarden voor W en Z groter zijn, de waarden van de leesindex kleiner zijn.

De uitkomsten kunnen op dezelfde wijze geïnterpreteerd worden als in figuur 5.

Uit onderzoek van onder andere kinderboeken en universitaire studieboeken stelde Brouwer vast dat bij Nederlandse teksten Z ligt tussen 5 en 30 en dat W ligt tussen 1,27 en 2,02.



Open het bestand *LEESINDEX.XLS* blad 1.

Met bestand *LEESINDEX.XLS* blad 1 kun je in de cellen D15 en D17 getallen invullen voor W en Z om de leesindex van Brouwer te berekenen.

Bij het openen van het bestand zie je dat in cel D15 voor W de waarde 1,29 is ingevuld en in cel D17 voor Z de waarde 10. In cel D19 wordt dan de Leesindex Brouwer (88,96) bij deze waarden van W en Z uitgerekend.

Daarnaast zie je de grafiek van de leesindex als Z verandert, maar W constant is. Deze geeft bij de gekozen waarde van W (in D15) het verband weer tussen de leesindex van Brouwer en Z . Je ziet dat bij het toenemen van de waarden van Z de leesindex afneemt.

Voor de leesindex van Brouwer geldt de volgende formule:

$$\text{Leesindex Brouwer} = 195 - 66,7 \cdot W - 2 \cdot Z$$

In een bibliotheek krijgen boeken een groene sticker als ze voldoen aan de volgende drie criteria:

- de leesindex van Brouwer: tussen 85 en 95;
- de gemiddelde woordlengte W : tussen 1,29 en 1,35 lettergrepen;
- de gemiddelde zinslengte Z : van 7 tot 9 woorden.

- 5p **19** Onderzoek of de criteria b) en c) samen er al voor zorgen dat automatisch criterium a) geldt, zodat criterium a) feitelijk overbodig is.

Voordat Brouwer zijn leesindex ontwikkelde, had Douma begin jaren zestig al een variant van de *FreS*-formule proberen te ontwikkelen die beter bij de Nederlandse taal zou passen:

$$\text{Leesindex Douma} = 206,8 - 77 \cdot W - 0,93 \cdot Z$$

De formule van Douma en die van Brouwer verschillen.

We kijken nu naar die verschillen.

Dit doen we door naar de leesbaarheid van een stukje uit de krant te kijken. Het is een zin uit een artikel over de miljoenennota:

tekstfragment Volgens het kabinet blijft vooral de export achter bij de wereldhandel en daardoor is er nog geen einde gekomen aan de neergaande fase van de conjunctuur.

Voor degenen die dat handig vinden is deze zin in LEESINDEX.XLS blad 2 te vinden.

De leesindex van Douma voor deze zin is: 52,3.

De leesindex van Brouwer voor deze zin is: 30,1.

- 4p **20** Leg aan de hand van de formules uit waarom bij deze zin de uitkomst van Douma groter is dan die van Brouwer (52,3 versus 30,1).

De zin van het tekstfragment kan enigszins aangepast worden, zó dat de betekenis niet verandert terwijl de leesbaarheid groter wordt. Je zou meer zinnen kunnen gebruiken of sommige woorden vervangen door synoniemen die minder lettergrepen bevatten.

Bij de woorden in de tekst die moeten meetellen bij de leesindex-berekening staat een '1' in kolom C. Als je in deze kolom een woord niet wilt laten meetellen, moet je de 1 in een 0 veranderen. De leesindex zal veranderen als men in kolom C enen in nullen verandert. Je kunt ook de woorden veranderen en woorden vervangen. Je moet dan wel de getallen in de kolommen C, E en I passend maken.

- 4p **21** Kies een manier waarbij de leesbaarheid van deze zin volgens beide formules groter wordt. Je kunt hierbij de spreadsheet LEESINDEX.XLS blad 2 gebruiken.



Sluit de Excelbestanden af. De bestanden niet opslaan.