

I. Inleiding en Doel.

Iedereen van ons wat iets op taalgebied wil presteer is begerig, en voortdurend besig, om sy woordeskat te vergroot; is steeds daarop uit om nuwe woorde sy eie te maak.

Hoewel dit natuurlik 'n prysenswaardige strewe is, ontstaan die vraag of ons daarby nie uit die oog verloor nie die onontbeerlikheid, die hoë mate van ons afhanklikheid van, die alledaagse gewone woorde—die grondslag van die taal.

'n Aansienlike aantal werkers het reeds hulle deel gedoen om, deur middel van Afrikaanse woordeboeke en ander geskrifte, wat kort-kort uit die drukperse vloei, vir ons 'n denkkeeld te gee van die uitgestrektheid van Afrikaans; het die Afrikaanse boom, met wortels en takke en al, vir ons aanskoulik voorgestel.

Maar tot dusver het nog niemand probeer om met juistheid vas te stel wat die omvang of beperktheid, en die belangrikheid, van die kern van die taal is nie. Skrywer hiervan weet in ieder geval van geen enkel uitgegewe werk in Afrikaans van dié aard nie.

Ons lewe in 'n praktiese eeu. Dit is dus die nuttigheid in die gebruik, meer as iets anders, wat vir ons die waarde van 'n woord behoort te bepaal. Hoe belangrik hierdie kern van ons taal is, uit die oogpunt van gebruiklikheid beskou, hoop ek in die navolgende bladsye duidelik aan te toon.

'n Tyd gelede was ek besig aan 'n Afrikaanse bewerking van Gregg Stenografie, dié uitmuntende en maklike stenografiestelsel wat in die laaste jare soveel opgang gemaak het. Dit was nodig om kort tekens of vorme toe te ken vir die weergawe van veelgebruikte woorde. Ek het 'n vae denkkeeld gehad en kon alleen 'n gissing maak omtrent enkele van die meesgebruikte woorde. Meer nie.

Dit het vir my, by 'n verdere versameling van 'n woordeskat vir die toepassing van die reëls vir stenografie neergelê, spoedig duidelik geword dat 'n groot gedeelte van alle lees- en skryfstof bestaan uit 'n heel beperkte aantal woorde en hulle herhalings.

Daar is maar min mense, selfs onder ontwikkelde persone wat, sonder enige voorafgaande ondersoek, voor die vuus weg kan sê wat wel die eerste meesgebruikte woord in Afrikaans is, hoewel

dié woord alleen (soos dit in lees- en skryfstof voorkom) ongeveer sewe persent uitmaak van al die woorde, soos hulle verskyn, op enige gegewe bladsy of in 'n bepaalde stuk, en dié woord, wat veelvuldigheid van gebruik betref, kop en skouers uitsteek bokant die woord selfs wat tweede op die lys staan.

As dit dan die geval is met die allereerste woord, hoeveel te moeiliker moet dit nie wees nie namate 'n mens laer op die skaal afdaal.

Hierdie vraagstuk het allengs my belangstelling so gaande gemaak dat ek besluit het om dit op 'n stelselmatige en wetenskaplike wyse te ondersoek, deur 'n aktuele telling en tabulering van woorde in leesstof te maak.

Vandaar hierdie werk.

Hoofdoel hiervan is om die 1,000 meesgebruikte woorde in gewone, alledaagse lees- en skryfstof van verskillende aard, vir volwassenes en kinders bedoel, te bepaal.

Ek het uitgegaan van die veronderstelling, dat die woorde wat in die algemeen die meeste gebruik word ook die nuttigste is en die grootste gebruikswaarde besit, en dat die meesgebruikte 1,000 vasgestel kon word deur 'n noukeurige ontleding of telling van ongeveer 150,000 woorde lees- en skryfstof van uiteenlopende aard. Hierdie hoeveelheid het in die loop van die werk geleidelik aangegroei tot dit, by voltooiing, 269,009 woorde beloop het.

II. Stof en Metode.

Na baie wik en weeg omtrent die lektuur wat vir die Onderzoek gebruik moes word het ek uiteindelik 'n program opgestel. Hierdie lys is in die loop van die werk een- en andermaal gewysig, en was by voltooiing as volg:—

Ontleding 1.	Kinderblad	10,136	woorde.
" 2.	Skoolleesboek, Std. III. .. .	19,773	"
" 3.	Dagblad—Hoofartikels .. .	6,679	"
" 4.	Kort Verhale deur Bekende Skrywers	16,558	"
" 5.	Skoolleesboek, Std. II. .. .	14,353	"
" 6.	Briewe uit Versekeringskantoor ..	21,037	"
" 7.	Toesprake—Hansard	16,851	"
" 8.	Algemene Letterkundige Werk ..	29,349	"
" 9.	Stof uit die Bybel	15,918	"
" 10.	Skoolleesboek, Std. I. (<i>Afrikaner- kind</i>)	12,186	"
" 11.	Skoolleesboek Std. I. (<i>Lees Met Lus</i>)	13,159	"
" 12.	Hofsake	15,433	"
" 13.	Tydskrifte—Persoonlike, Huislike en Boerdery-aangeleenthede ..	17,732	"
" 14.	'n Toneelstuk	18,509	"
" 15.	Dagblad—Binnelandse en Buite- landse Nuusberigte	18,018	"
" 16.	Persoonlike en Besigheidsbriewe	23,318	"
16		269,009	

Hierdie materiaal, meen ek, is taamlik verteenwoordigend van alledaagse leesstof vir kinders en grootmense. Die hele Onderzoek het dus uit 16 verskillende ontledings bestaan. 'n Breedvoerige bespreking van die bronne en afsonderlike uitslae is te vinde op bladsy 10 en volgende bladsye.

HOE DIE ONTLEDINGS GEMAAK IS.

Dit was natuurlik nodig om aantekening te hou van elke woord wat teëgekom is en dan 'n juiste opgawe te maak van hoeveel keer die woord verskyn het.

Die metode wat gebruik is om dit te doen is as volg:

As 'n stuk vir ontleding gekies is, is dit aan iemand gegee om presies af te skryf op in ruitjies gelinieerde papier („sectional paper”), d.w.s. papier wat langs en kruis in blokkies van 1 vierkante duim afgelinieer is. In elke vierkantjie moes net een woord geskryf word—nie meer nie, en nie minder nie, en in elke vierkantjie moes 'n woord staan. Syfers en voorletters moes weggelaat word.

Die gebruikte papier was in foliogrootte (13 x 8"). Daar was dus presies 104 vierkantjies op elke blad (net op een kant van die papier is geskryf). As die stuk klaar geskryf is, was dit gevolglik alleen nodig om na te gaan of al die ruitjies volgeskryf is, dan die bladsye sorgvuldig te tel en met 104 te vermenigvuldig om sekuur te weet hoeveel woorde ontleed word. As 'n woord lank was moes dit een- of meermaal afgebreek word, en die afskrywers het strenge voorskrifte gekry om nie meer as een blokkie vir 'n enkele woord te gebruik nie—hoe lank die woord ook al is. Ek het 'dan self, deur die afgeskrewene na te sien, gesorg dat hieraan voldoen is.

As alles dan behoorlik in orde is wat die afskryf betref, is die papier versigtig met 'n skêr opgesnipper langs die lyne op die papier, d.w.s. in 1-duim vierkante stukkie, sodat op elke stukkie papier 'n woord staan, en die papiertjies is in 'n kartondoos gegooi. Ek, en diegenes wat my behulpsaam was, het gedurig 'n wakende oog gehou dat daar nie van die papiertjies verlore gaan nie. (Om die belangrikheid hiervan te beklemtoon het ek aan die seuns gesê om elke papiertjie as 'n pondnoot te beskou!).

Die volgende stap was die sortering, klassifisering en telling van die woord-papiertjies. Om hierdie deel van die ontleding—'n tydrawende sleurwerk—te vergemaklik en te bespoedig, is die papiertjies eers gesorteer in 22 kleiner dosies volgens hulle beginletter (die werk is feitlik voltooi voor die allernuutste Afrikaanse spelling op die toneel verskyn het, en daar was dus geen woorde beginnende met *x* en *q* nie en slegs enkele eiename in *c* en *z*).

Toe is die kleiner dosies een vir een aangepak. Ek het spoedig uitgevind dat tyd bespaar word deur eers 'n paar van die meesvoorkomende woorde uit te soek; die ondervinding en 'n vlugge deursage het spoedig geleer watter dié woorde in iedere dosie is. By voorbeeld, in die A-dosie is die papiertjies waarop die woorde *aan*, *af*, *al*, *alle*, *alles*, *alte*, *as*, gestaan het, eerste uitgesoek, in hopies gesorteer, in pakkies afgetel en gomlastiek-bandjies om gesit. En so is aangehou tot die woorde wat slegs eenmaal voorkom, oorgeskiel het. As die aantal papiertjies, sê van die woord *aan*, 235 is, is die getal deur die teller op die

boonste papiertjie geskryf. Dit verteenwoordig dan die aantal verskynings van *aan* in dié besondere ontleding.

Ten einde juistheid te verseker is die hopies altyd nagetel, meestal deur myself persoonlik. As ek 99 tel vir 'n 100-hopie, dan tel ek dit weer oor. Kry ek weer 99, dan neem ek 99 as die juiste getal, maar kry ek, by die tweede natelling 100, dan is 100 as korrek geneem. Twee tellings moes klop—of twee van die nateller of een van die teller en een van die nateller.

En dit is miskien goed om hier daar nadruk op te lê dat alle telwerk op hierdie wyse gedoen is—daar was geen raaiwerk en geen ramings nie. Ook al die berekenings is gekontroleer, of deur myself of deur 'n ander bevoegde persoon. Slegs in een of twee onbelangrike gevalle is die tellings en berekenings nie nagegaan nie; in die gevalle word die feit gekonstateer.

Van elke woord se hopie is toe een papiertjie geneem en die totale aantal kere wat die woord in die ontleding verskyn (die woord se „verskyningskrediet” of kortweg „krediet” genoem) is daarop aangeteken en die aldus genommerde papiertjie in sy dosie teruggegooi, terwyl al die ander in 'n afsonderlike doos gaan om weggegooi te word sodra die hele ontleding klaar is.

Op al die papiertjies van woorde wat net eenmaal verskyn is die syfer 1 geskryf. Daar was, by die voltooiing van die sortering van 'n dosie, dus net een papiertjie van elke woord met die verskyningskrediet daarop aangeteken, en al wat nou nog nodig was, was om hulle almal in alfabetiese orde te rangskik en dan in 'n pakkie te maak, tydelik weg te sit tot die dosies van al die letters so deurloop is. Deur 'n *telling* van die aantal papiertjies in hierdie pakkie is die *woordeskat* van die besondere dosie verkry, en deur *optelling* van die *syfers* op die papiertjies aangeteken, kon die aantal *woorde* leesstof (in hierdie werk ook „lopende woorde” genoem) in dié letter vasgestel word.

Daar moet duidelik onderskei word tussen *woorde* en *woordeskat*. Ons kan hierdie onderskeid bes aanskoulik aantoon deur 'n ontleding op klein skaal hier te maak van die volgende aanhaling uit 'n brief deur my ontvang:—

„Ek was vir 'n week by my suster en haar familie. Dit gaan nog goed met hulle. Sussie het 'n operasie ondergaan, en nou voel sy nog maar swak. Al die seuns is nou min of meer in die werk, buite die jongste seuntjie; hy is nog jonk en op skool. Al die dogters en seuns is nog ongetroud.

„Met ons seuntjie gaan dit baie goed; hy is baie lewendig en groei mooi. Hy is al twee jaar oud. Waar gaan die tyd tog nie

heen nie, en ek word oud by die dag. My gesondheid is glad nie goed in die laaste tyd nie." (105 woorde).

Hier volg nou 'n ontleding van hierdie stukkie soos die ontledings almal deur my gemaak is:—

Woordeskat.	Verskynings-krediet.	Woordeskat.	Verskynings-krediet.
al	3	min	1
baie	2	mooi	1
buite	1	my	2
by	2	'n	2
dag	1	nie	4
die	7	nog	4
dit	2	nou	2
dogters	1	of	1
ek	2	ondergaan	1
en	6	ongetroud	1
familie	1	ons	1
gaan	3	op	1
gesondheid	1	operasie	1
glad	1	oud	2
goed	3	seuns	2
groeï	1	seuntjie	2
haar	1	skool	1
heen	1	sussie	1
het	1	suster	1
hulle	1	swak	1
hy	3	sy	1
in	2	tog	1
is	6	twee	1
jaar	1	tyd	2
jongste	1	vir	1
jonk	1	voel	1
laaste	1	waar	1
lewendig	1	was	1
maar	1	week	1
meer	1	werk	1
met	2	word	1
		62	105

Hieruit blyk dat die stukkie uit 105 woorde bestaan, maar 'n woordekat van slegs 62 oplewer.

Elkeen van die woorde is toe voor op 'n afsonderlike kaartjie, groot $2\frac{1}{2} \times 3$ duim, en agterop gelinieer, getik, en die getal wat op die ooreenstemmende papiertjie aangeteken is van die krediet

van die woord, agterop die kaartjie vir die woord, oorgeskryf, voorafgegaan van die nommer van die ontleding. (Daar is 16 ontledings altesame).

So is die krediete in 'n kolom op die kaartjie af geskryf en by voltooiing van al die ontledings opgetel om te toon wat die totale krediet vir die woord vir die ganse Onderzoek is. By elke ontleding het ek deur die kaartjies gegaan en die krediete van die papiertjies op die kaartjies oorgebring. Vandaar dat dit nodig was om die papiertjies (en natuurlik ook die kaartjies) in streng alfabetiese orde te skik. Telkens, as 'n nuwe woord teëgekom is is daar 'n kaartjie vir hom gemaak.

Vir die eerste 105,000 woorde van die Onderzoek (die eerste sewe ontledings) is daar nie afsonderlike aantekening van sulke nuwe woorde, d.w.s. woorde wat nie van tevore teëgekom is, gehou nie, maar vanaf die agste ontleding is behoorlike aantekening van die aantal gehou, en vanaf Ontleding 11 ook lyste van die woorde. Ek was oorspronklik die mening toegedaan dat, nadat omtrent honderdduisend woorde leesstof ontleed is, daar betreklik weinig nuwe woorde aangetref sou word; hierin het ek my egter, soos later sal blyk, wel deeglik vergis.

Ek het ook sorgvuldige lyste gehou, met opgawe van die krediete, van die veelgebruikte woorde in elke afsonderlike ontleding. As veelgebruikte woorde is geneem dié met 'n sekere minimumkrediet. Hierdie minima is van 3 tot 6 onderskeidelik, so min of meer na gelang van die woordetal van die besondere ontleding. Die grootte van die woordeskat in die verskillende lyste also verkry is natuurlik verskillend en loop vanaf 358 (in die kleinste ontleding) tot 665.

Deur die krediete vir hierdie lyste op te tel kon bepaal word watter persentasie van die totaal hierdie beperkte woordeskat met sy herhalings uitmaak.

Verder is daar nog afsonderlike lyste gemaak van die 10, 25, 50 en 100 meesgebruikte woorde vir elke ontleding en die persentasies wat hierdie woordeskat met sy herhalings van die totaal uitmaak, bereken. Dit dien as 'n nuttige maatstaf van vergelyking van die betreklike eenvoudigheid of moeilikheid van die behandelde stof, want aangesien die meesgebruikte woorde gewoonlik eenvoudig is, beteken 'n hoë persentasie vir so 'n beperkte woordeskat dat die ganse woordeskat betreklik klein is, en die stuk dus maklike lektuur vorm. Ook gee dit 'n denkbeeld of daar 'n mate van egaligheid of bestendigheid is in die veelvuldige gebruik van woorde, en is 'n aanduiding of so 'n Onderzoek wel al dan nie op 'n wetenskaplike grondslag berus.

III. Wat is Woordeskat?

Enkele vraagstukke het hulle voorgedoen in verband met die vraag „Wat word verstaan onder woordeskat?”

Daar bestaan vir elke gewone ontwikkelde persoon feitlik drie woordeskatte, n.l. a. die aantal verskillende woorde wat hy min of meer begryp as hy dit in leesstof teëkom, maar nie almal juis bekwaam is om self te gebruik nie; b. die woordeskat wat hy in sy skryftaal gebruik; en c. die aantal verskillende woorde wat hy as spreektaal gebruik. Die eerste klas sal baie groter wees as die tweede, maar dit is twyfelagtig of daar in Afrikaans soveel verskil bestaan tussen die tweede en die derde klasse hierbo genoem as algemeen veronderstel word.

Moet eiename van persone en plekke as deel van die woordeskat getel word? Wat word verstaan onder „verskillende woorde?” Moet b.v. *staan, staande, stand, stande, standjie, bestaan, bestaanbaar, bestaanbaarheid, bestaande, bestand, gestaan, gestand ontstaan, verstaan, verstaanbaar, verstaanbaarheid, verstand, verstandig, verstandiger, verstandigheid, verstandiglik, verstandigste*, as 'n woordeskat van 1 of 22 bereken word?

Dit sou m.i. nie reg of billik wees om eiename as deel van 'n Afrikaanse woordeskat te neem nie. Daar is so 'n menigte name in die wêreld wat in leesstof kan voorkom, baie daarvan van naturelle- of vreemde oorsprong, dat hulle nouliks as deel van die gewone woordeskat beskou kan word, en vir die doel van hierdie studie is hulle nie in aanmerking geneem nie. In sommige stukke is alle name geskrap voor hulle afgeskryf is, en waar dit nie gedoen is nie, is hulle afsonderlik gehou by die sortering van die papiertjies, en van die totaal afgetrek. Op dié manier is die „netto totaal” verkry.

Byvoeglike naamwoorde van eiename gevorm, soos b.v. *Suid-Afrikaanse*, is wel behou, asook die name van die dae van die week en die maande van die jaar, en die benamings vir nasionaliteite, tale, gelowe en instellings, soos *Roomse, Protestante, Volksraad, Senaat, Wesleyane, Engelse, Hollanders, Russies*, ens. Hulle vorm ook maar in elk geval 'n klein gedeelte van leesstof.

Wat betref die afleidinge van grondwoorde het ek besluit dat, voorlopig altans, dit die verstandigste sou wees om elke afsonderlike woordvorm as 'n woord te behandel, en afsonderlike krediete

aan toe te ken. 'n Mens kan moeilik anders. Om die enkelvoud ook die meervoud te laat insluit, en afleidinge gevorm deur die byvoeging van voorvoegsels soos *be-*, *ge-*, *her-*, *ver-*, en agtervoegsels *-ing*, *-baar*, *-heid*, *-lik* en *-loos* en ander, as almal 'n eenheid met die grondwoorde vormende te behandel, is baie onwenslik en sou die werk van die Onderzoek baie moeilik en ingewikkeld maak. Jy kan nouliks al die bostaande vorme van die woord *staan* as 'n eenheid beskou in die berekening van 'n woordeskat.

Hierdie Onderzoek is gemaak met die oog op die vergemakliking van die aanleer van die taal, van spelling, en van stenografie. 'n Aantal selfstandige naamwoorde kom net in die meervoud voor, b.v. *masels*, *reisies*, *notule*, *harsings*, en ander; terwyl van sommige selfstandige naamwoorde die meervoud heeltemal verskillend van die enkelvoud is. Die afgeleide vorm is dikwels moeiliker om te spel as die grondwoord; die stenografievorm is meestal anders, en deur die byvoeging van 'n voor- of agtervoegsel word die betekenis van 'n woord soms gans verander. Die volgende woorde dien as voorbeelde van wat hierbo gekonstateer is:—

<i>oog</i> , <i>oë</i> ;	<i>mak</i> , <i>gemak</i> ;
<i>hoog</i> , <i>hoë</i> ;	<i>gryp</i> , <i>begryp</i> ;
<i>kon</i> , <i>koning</i> ;	<i>lyk</i> , <i>gelyk</i> ;
<i>neer</i> , <i>nering</i> ;	<i>let</i> , <i>belet</i> ;
<i>wig</i> , <i>gewig</i> ;	<i>raad</i> , <i>raadgewinge</i> ;
<i>bed</i> , <i>gebed</i> ;	<i>eer</i> , <i>eerbewyse</i> ;
<i>deeg</i> , <i>deeglik</i> ;	<i>genot</i> , <i>genietinge</i> ;
<i>son</i> , <i>sonder</i> , <i>gesonder</i> ;	<i>leer</i> , <i>leerstelling</i> ;
<i>ope</i> , <i>openbaar</i> ;	<i>waar</i> , <i>verwaarloos</i> .
<i>val</i> , <i>geval</i> ;	

Daar is gevalle waar die meervoud of die afgeleide vorm van 'n woord baie meer gebruik word as die stamwoord. So b.v. word *omstandigheid* maar baie min gebruik, terwyl die meervoud, *omstandighede*, so 'n hoë krediet het dat dit onder die Duisend hoogste woorde in die taal is. *Mense* is gebruikliker as *mens*, *oë* as *oog*, *hulle* as *hul*, *beeste* as *bees*, *osse* as *os*, ens.

In hierdie Onderzoek is dus deurgaans, op enkele uitsonderinge na, wat later bespreek word, elke woordvorm as 'n afsonderlike woord behandel. Dit maak natuurlik die getalle wat gegee word vir die verskillende woordeskatte groter as hulle sou gewees het as die grondwoord, met al die vorme verkry deur die afleidinge, as 'n enkele woord sou beskou word, maar dit is m.i. die enigste gesonde, vaste grondslag vir 'n Onderzoek soos hierdie.

IV. Beskrywing en Resultate van die Verskillende Ontledings.

Soos reeds gekonstateer was die einddoel van die Onderzoek die vasstelling van die 1,000 meesgebruikte woorde. Dié sal natuurlik gekies word uit die woorde wat die hoogste totale verskyningskrediete behaal in die 16 bronne wat ontleed is, maar dit is nodig om eers 'n volledige beskrywing van hierdie bronne, en die uitslae in elkeen verkry te gee, en hulle dan met mekaar en die eindresultate te vergelyk.

1. TYDSKRIF VIR KINDERS.

By die oorweging van geskikte stof vir die bereiking van die doel van die Onderzoek, n.l. die verkryging van die meesgebruikte woorde, het dit my voorgekom dat dit nodig sou wees om heelwat kinderlektuur in die Onderzoek op te neem, in die mening nie juis dat al die woorde in sodanige leesstof die meesgebruikte, alledaagse woorde sou wees nie, maar dan tog dat 'n groot aantal van wat die alledaagse grondwoorde is, woorde vir die beduiding van baie van die grondbegrippe en alledaagse dinge van die lewe, hierin aangetref sou word. Van die 16 ontledings bestaan vyf uit kinderlektuur en die eerste van hulle, en van die hele Onderzoek is twee uitgawes van *Die Kleinspan*, 'n blaadje vir kinders bedoel, wat as byvoegsel tot *Die Huisgenoot* deur die Nasionale Pers te Kaapstad, weekliks uitgegee word. Nos. 34 en 37, dato 31 Januarie 1930 en 21 Februarie 1930 onderskeidelik, is geneem vir die ontleding. Sover nagegaan kan word is dit die werk van tien skrywers.

Alle opskrifte, hofies en onderskrifte is weggelaat, asook voorletters en syfers, en alleen die aaneenlopende leesstof geneem. Alle woorde is afgeskryf en getabuleer net soos hulle verskyn het, met inbegrip van verkortings, en alle eiename van persone en plekke is, soos in al die ontledings gedaan is, van die totaal afgetrek by die sortering van die papiertjies. Die totaal woorde voor die sortering begin is was 10,266. Die totaal van die getalle vir die verskillende dosies (beginletters van woorde) na die sortering, na aftrek van 'n totaal van 278 vir name, d.w.s. die netto totaal, was 10,136. Hierdie getal plus die 278 behoort, as die sortering, telling en berekenings reg was, 10,266 te maak, maar in werklikheid beloop dit 10,414, sodat daar 'n verskil (fout) van

148 te veel is. Aangesien die berekening van die aantal woorde (vierkantjies) voor die papier opgesnipper word 'n eenvoudige saak is, soos reeds verduidelik op bladsy 4, kan die getal 10,266 met gerustheid as juis beskou word, en die feit dat die uitslag nie met hierdie oorspronklike getal klop nie, kan alleen verklaar word deur 'n mistelling of misrekening êrens, iets wat maklik kan gebeur met die omslagtige werk wat nodig is om tot die resultaat te geraak.

Om die hele boel weer oor te doen om hierdie verskil te vind sou ondenkbaar wees en nouliks die moeite werd, aangesien die fout maar slegs 1.4% van die oorspronklike totaal is. Daar is in al die ontledings 'n klein foutjie of verskil, in sommige 'n tekort, in ander 'n teveel, sodat die een waarskynlik opmaak wat die ander te kort skiet. In sommige van die ontledings is die fout so klein dat dit nouliks noemenswaardig is. In elk geval, aangesien die sorterings so 'n omslagtige en tydrowende boel is, en dit menslik is om te dwaal, moet hierdie foutjies as onvermydelik beskou word, en is buitendien so gering dat hulle in 'n ontleding op so 'n skaal maar heel weinig afdoen van die juistheid van die uitslae. 'n Volledige opsomming van die foute word op bladsy 65 gegee.

Die netto totaal van hierdie ontleding, dan, is 10,136 woorde leesstof. Op die grondslag op bladsy 9 uiteengesit—die grondslag wat deurgaans vir al die ontledings gebruik is—is die woordeskat van die twee uitgawes van *Die Kleinspan* 1,745. Daar is 'n afsonderlike lys gehou van alle woorde met 'n krediet van 3 en hoër. Dit wil sê, daar is natuurlik enkele woorde wat 'n verskyningskrediet van honderde het en 'n aantal ander ook met krediete baie hoër as 3, maar woorde met 'n verskyning van driemaal is die laagste wat in hierdie lys opgeneem is.

Daar is 508 verskillende woorde in hierdie lys (uit 'n woordeskat van 1,745). Daar is reeds op gewys dat die minimumkrediete vir hierdie lyste van veelgebruikte woorde hoër of laer is na gelang van die grootte van die ontleding (die laagste 3 en die hoogste 6) en dat die aantal woorde alsoo verkry tussen 358 en 665 is. Dit sal te veel ruimte in beslag neem om hierdie lys vir elke ontleding te gee, hoe graag ek dit ook sou wil doen, maar dit kan nuttig wees om een sodanige lys te gee. In Ontleding No. 1 het die volgende woorde 'n krediet van minstens 3. Die lys is in alfabetiese orde gerangskik en van die hoogste krediete af in elke beginletter.

aan	47	belangstelling	3	daaraan	4
al	37	behandel	3	daarvoor	3
af	35	bene	3	dieselfde	3
ander	21	besonder	3	dop	3
as	15	beste	3	dorp	3
altyd	14	bladsy	3	draggie	3
aarbei	13	blompot	3	duim	3
almal	12	bokant	3	druk	3
a	11	boos	3	doodgemaak	3
ag	7	broer	3	dinge	3
arm	7	benoud	3	daarheen	3
alles	6	die	747	en	390
anders	5	dit	112	ek	105
agterlyf	5	daar	77	een	27
asseblief	5	dat	65	eers	15
agterna	5	dan	41	elke	8
artikels	5	deur	35	eie	6
Afrikaans	4	dis	19	eet	5
agter	3	daardie	18	eerste	5
awonture	3	doen	17	elkeen	3
au	3	dam	12	erdewerk	3
agteruit	3	draai	12	fluit	6
aangetrek	3	dadelik	11	fatsoen	4
armoedig	3	drie	10	fyn	3
baie	48	dr.	9	fluister	3
by	47	daarin	8	gaan	50
begin	21	dosie	8	gemaak	26
bye	19	dans	6	groot	21
bietjie	9	dag	6	gee	21
boom	9	daarom	6	gat	18
besig	8	dink	6	gou	17
bly	7	dra	6	geld	16
bek	6	daarop	5	groter	11
blink	5	daarna	5	gehad	10
bring	5	donker	5	gang	10
binne	5	darem	5	goed	8
borde	5	dae	4	gewoon	8
byt	5	dele	4	geantwoord	7
beveel	4	diertjie	4	gedink	7
byna	4	daarvan	4	genoeg	7
betaal	4	ding	4	goud	7
boomstam	4	dood	4	gooi	7
bosse	3	dorpie	4	geneem	7
bestaan	3	duur	4	gekom	7
bang	3	dog	4	gegee	7
blaas	3	dikwels	4	geweet	7
buite	3	dak	4	geword	6

gesien	6	hoor	10	kamer	3
gebruik	6	hê	9	kant	3
grond	5	hande	7	kap	3
gekry	5	heuning	7	kleinspan	3
geen	5	haal	6	laat	20
geraak	5	heeltemal	5	lê	14
gepraat	4	hele	5	langs	12
gesê	4	huistoe	5	lag	12
gesit	4	huise	4	loop	10
gevang	4	hol	4	lank	10
glad	4	helfte	3	lekker	9
gryp	4	homself	3	lyf	7
goeie	4	huisie	3	lig	7
gelede	4	huil	3	lyn	5
gehou	4	in	156	ly	5
gegaan	4	is	120	lang	5
gehoor	4	iets	7	land	4
graag	3	inhalig	4	lug	4
goue	3	jou	55	lyk	3
gevra	3	jy	47	lus	3
gehaat	3	julle	25	lui	3
gewillig	3	jul	10	lief	3
geweldig	3	jong	4	lewê	3
gevoelig	3	ja	4	liewe	3
gestuur	3	kom	49	lande	3
gesoek	3	koning	40	met	81
gesig	3	kop	34	maar	79
gelukkig	3	kan	32	my	57
gedoen	3	kry	24	maak	39
gebly	3	kyk	23	man	25
gedraai	3	kon	16	mannetjie	24
gebeur	3	kinders	19	moet	22
gedagte	3	klein	14	mense	13
het	311	klaar	13	mammie	12
hy	239	klei	13	meer	12
hulle	84	krap	9	meneer	11
hom	56	koek	9	moes	11
hul	54	koppies	8	mos	11
haar	28	kos	8	mooi	10
hoe	21	koppie	7	mens	8
hierdie	20	klas	6	mond	7
huis	17	klere	6	muis	7
hout	17	klim	6	ma	7
hier	16	kort	6	meteens	5
hardloop	12	kwaad	6	mol	5
hard	11	kat	4	maag	4
hou	10	klomp	3	manier	4

moeg	4	oopmaak	3	speel	7
moeder	4	opeet	3	soldate	7
mak	3	opgemerk	3	son	6
mekaar	3	pappa	19	soek	6
merk	3	paar	13	slaap	6
middel	3	pottebakker	12	stoep	6
min	3	porselein	12	stukke	6
maklik	3	plan	12	stukkies	6
mamma	3	pragtig	7	smeltoond	6
moenie	3	pierings	6	stuk	5
mensoom	3	paleis	4	steek	5
'n	216	pad	4	spring	5
nie	181	praat	4	slim	5
na	62	pappie	4	seker	5
nou	42	plek	4	soveel	5
net	29	prys	3	skree	4
nog	25	poeier	3	saam	4
neem	18	probeer	3	sagte	4
nee	13	pap	3	stukkend	4
nuwe	11	prinsipaal	3	stam	3
nooit	9	reier	11	stof	3
niks	8	rond	8	sterk	3
nes	6	rukkie	7	stem	3
naderhand	5	roep	6	sout	3
nonnie	4	ry	6	skyn	3
neer	3	reguit	5	skool	3
nadat	3	rug	5	skielik	3
nek	3	resitasie	4	self	3
om	111	regop	3	saggies	3
op	77	reg	3	stadig	3
ons	50	ruk	3	strate	3
of	35	rus	3	sulke	3
ou	25	sy	150	stuifmeel	3
ook	23	so	76	skobbetjies	3
onder	20	sal	48	te	154
onderwyl	17	sien	45	toe	96
oor	17	se	43	tog	17
oondjies	7	sê	40	tot	13
oop	7	slang	35	twee	12
oubaas	6	sit	17	trek	11
oë	5	soos	16	tussen	8
oud	5	staan	10	terug	7
oupa	4	sodat	8	tuin	4
O	4	somaar	8	tafel	4
opgetel	4	stukkie	8	toegelaat	3
ongelukkig	3	snaakse	8	treurig	3
oondjie	3	swerm	7	tros	3

totdat	 3	veilig	 4	waar	 18
uit	 48	vandag	 4	want	 17
u	 5	vuur	 4	wil	 16
uitgeroep	 4	vertel	 4	wees	 15
uur	 4	voel	 4	wou	 11
uitasem	 3	ver	 4	wag	 10
uitdraai	 3	vormlose	 4	weg	 9
uitstootwiel	 3	veel	 4	wonderlike	 8
van	 158	vlak	 4	wiel	 8
vir	 102	vloeibare	 4	wys	 7
voor	 22	vliegtuig	 4	warm	 6
vang	 12	verander	 3	wyfie	 6
vinnig	 11	volk	 3	wit	 5
vorm	 11	vleis	 3	water	 5
vra	 10	verskrik	 3	wêreld	 5
vuurvliegies	 6	vlerke	 3	waarheid	 5
vloer	 6	vloeistof	 3	waterdraer	 4
visse	 6	vrolike	 3	wakker	 4
voordat	 6	verglaas	 3	wasem	 3
vuurvlieg	 5	verder	 3	weet	 3
vreeslike	 5	vanoggend	 3	watter	 3
verby	 5	verbaas	 3	wanneer	 3
vas	 5	verhoog	 3	waarin	 3
vat	 5	verkeerde	 3	werkers	 3
vorentoe	 5	was	 123	waarom	 3
voete	 5	wat	 92	ydel	 3
vet	 5	word	 67	ystervark	 3
vier	 4	weer	 35		
vervolg	 4	werk	 18		
7,981		8,399		508	8,585

Uit hierdie 508 dikwels-voorkomende woorde is toe gekies die 100 wat die hoogste krediete het, en hulle is, soos aangetoon word, volgens hulle waarde (of krediete) gegradeer van die hoogste af na onder. Op dié manier kon die gesamentlike krediete vir die eerste 10, 25 en 50 woorde ook verkry word. Die lys is as volg:—

ONTLEDING 1: 100 MEESGEBRUIKTE WOORDE.

1. die	 747	8. in	 156
2. en	 390	9. te	 154
3. het	 311	10. sy	 150
4. hy	 239		
5. 'n	 216		2,702
6. nie	 181	11. was	 123
7. van	 158	12. is	 120

13.	dit	112	57.	een	27
14.	om	111	58.	gemaak	26
15.	ek	105	59.	nog	25
16.	vir	102		man	25
17.	toe	96		julle	25
18.	wat	92		ou	25
19.	hulle	84	63.	mannetjie	24
20.	met	81		kry	24
21.	maar	79	65.	kyk	23
22.	daar	77		ook	23
	op	77	67.	voor	22
24.	so	76		moet	22
25.	word	67	69.	gee	21
		4,104		hoe	21
26.	dat	65		groot	21
27.	na	62		begin	21
28.	my	57		ander	21
29.	hom	56	74.	hierdie	20
30.	jou	55		laat	20
31.	hul	54		onder	20
32.	gaan	50	77.	kinders	19
	ons	50		dis	19
34.	kom	49		pappa	19
35.	baie	48	80.	daardie	18
	sal	48		gat	18
	uit	48		neem	18
38.	aan	47		werk	18
	by	47		waar	18
	jy	47	85.	doen	17
41.	sien	45		hout	17
42.	se	43		huis	17
43.	nou	42		gou	17
44.	dan	41		tog	17
45.	koning	40		want	17
	sê	40		sit	17
47.	maak	39		oor	17
48.	al	37		onderwyl	17
49.	af	35	94.	kon	16
	deur	35		hier	16
		5,284		geld	16
	of	35		soos	16
	slang	35		wil	16
53.	kop	32	99.	wees	15
	kan	32		as	15
55.	net	29	100.		6,341
56.	haar	28			

Hierdie uitslae is veelseggend uit die oogpunt van die waarde (gebruiklikheid) van woorde beskou. Uit die ganse woordeskat van 1,745 kom slegs 508 verskillende woorde in die lys van veelgebruikte woorde (krediet van minstens 3). Hierdie 508 woorde (met hulle veelvuldige herhalings) maak 8,585 lopende woorde, d.w.s. soos hulle in leesstof voorkom, van die 10,136 netto totaal uit. Dit kom te staan op 84.6%. Die res van die woordeskat, n.l. 1,237 (1,745 min 508) kom so selde voor dat hulle, met herhalings, maar altesame 1,551 lopende woorde verteenwoordig, of minder as 16% van die hele ontleding. As ons die verhouding neem tussen woordeskat en lopende woorde in die lys van veelgebruikte woorde dan sien ons dat dit 1:16 is, d.w.s. elkeen van hierdie 508 woorde is gemiddeld 16 keer gebruik in hierdie stuk, terwyl die mingebruikte woorde elkeen maar 1.2 maal gebruik is.

Dit is dus duidelik dat wie 'n studie van Afrikaans maak vir die menigvuldige gebruike waartoe 'n taal hom leen, hy hom nie soveel hoef te kwel oor die 1,237 mingebruikte woorde as hy hom behoort toe te lê, veral by die aanvangstadium van die studie, op die 508 veelgebruikte woorde wat so 'n groot gedeelte van die hele aantal vorm. Dit is die verstandige weg, die weg van maksimumresultate met 'n minimum van tyd en inspanning.

In die algemeen gesproke besef ons maar weinig, meen ek, hoeveel ons aan hierdie grondwoorde van die taal verskuldig is; intendeel, ons sien dikwels met minagting op hulle neer.

Hierdie syfers vir die 508 woorde is verbasingwekkend. Maar hoe hoër ons klim op die lys, des te merkwaardiger word die uitslae. Laat ons 'n slaggie stilstaan by die volgende groep—die 100 meesgebruikte woorde.

Hierdie 100 woorde verteenwoordig 6,341 uit die totaal van 10,136 lopende woorde. Dit is 62.5%. Die 100 woorde alleen (uit 'n woordeskat van 1,745) vorm dus oor die drie-vyfdes van die leesstof in die twee eksemplare van *Die Kleinspan*.

Die 50 hoogste woorde beslaan 52.1% van die netto totaal, 25 woorde neem 40.4% op, en die 10 gebruiklikste woorde maak 26.6% van die geheel uit.

Dit is inderdaad merkwaardig, en hierdie woorde is van die allergrootste belangrikheid en nut in die gebruik van die taal.

Ek het, by die voltooiing van die eerste ontleding, hierdie uitslae baie belangwekkend gevind, dog het gemeen dat hulle maar bloot toevallig is en sulke hoë persentasies van die geheel uitmaak omdat dit kinderlektuur en dus betreklik maklike leesstof is. Ek het gemeen dat die persentasies vir die verskillende ontledings baie van mekaar sou verskil, en dat daar g'n standhoudendheid sou wees nie.

Ek het dus met spanning uitgesien na die resultate vir die volgende ontleding, om dan die twee met mekaar te vergelyk. Maar aangesien elke ontleding van vier tot ses weke van my beskikbare tyd in beslag geneem het, moes ek 'n bietjie geduld beoefen. Ook op 'n ander punt sou vergelyking belangwekkend wees. Selfs al sou die persentasies van die eerste 10, 25, 50 en 100 woorde nie besonder uiteenlopend wees nie, was die vraag nog of hierdie lyste min of meer uit dieselfde woorde sou bestaan. Hierdie punt word later breedvoerig bespreek.

2. SKOOLLEESBOEK, STANDERD III *value 5*

Ontleding No. 2 is 'n skoolleesboek, en die boek hiervoor gebruik is *Jong Suid-Afrika, Std. III—Bloemlesing vir die Laerskool*, deur C. J. Alers, uitgewers J. L. van Schaik, 1927. Alle met sinne aaneengeskakelde stof, met die uitsondering van poësie, tot op bladsy 101 is gebruik, gewende 'n gros totaal van 20,387 lopende woorde. Die boek beslaan 145 bladsye en sal dus ongeveer 29,000 woorde bevat.

Die aantal verskillende skrywers wat hierdie materiaal bygedra het is, sover nagegaan kan word, tien.

Die orde waarin hierdie ontledings gedaan is is van geen betekenis nie. Ek was hiervoor afhanklik vir die afskryf van verskillende buitestaande persone, en die ontledings is eenvoudig gedaan soos ek die afskrywings ingekry het.

Hierdie was die hoogste standerd skoolleesboek wat ontleed is. Die woordeskat hierin gevind, n.l. 2,588, is dan ook die grootste van alle skoolleesboeke wat ontleed is. Dit is oor die 700 woorde meer as die woordeskat in die twee uitgawes van *Die Kleinspan*, die woordeskat waarvan 1,745 is, sodat 'n kind wat hierdie leesboek deurloop het vermoedelik *Die Kleinspan* met gemak sal kan lees en verstaan, veral aangesien maar omtrent driekwart van die leesstof in die boek behandel is. Dit wil egter nie noodwendig sê dat die verkreeë woordeskat maar driekwart van die ganse woordeskat sal wees nie, aangesien die woordeskat nie na verhouding tred hou met die hoeveelheid leesstof nie.

Daar is 520 veelgebruikte woorde, d.w.s. woorde in hierdie geval met 'n minimumkrediet van 5; hierdie 520 verskillende woorde verteenwoordig 16,598 uit die netto totaal van 19,773 lopende woorde, of 83.9%.

Die 100 meesgebruikte woorde maak 12,194 (61.6%) van die netto totaal uit; die 50 hoogste woorde 10,017 (50.6%); die 25 hoogste 7,632 (38.5%), en die eerste 10 maak 4,828 of 24.4%.

Die verhouding tussen woordeskat en lopende woorde (soms die „herhalingsverhouding” genoem) vir hierdie stuk in sy geheel geneem is 1:7; vir die veelgebruikte woorde alleen 1:31; vir die 100 hoogste woorde ongeveer 1:122; vir die 50 woorde 1:200; vir die hoogste 25 alleen 1:305 en vir die eerste tien woorde omtrent 1:483.

Die volgende tabel gee 'n vergelyking van die persentasies wat die hoogste groepe woorde in Ontledings 1 en 2 van die netto totale uitmaak:—

Ontleding No.	Veelgebruikte Woorde.	Die hoogste			
		100 woorde.	50 woorde.	25 woorde.	10 woorde.
1 (minimum- krediet 3; 508 woorde)	84.6%	62.5%	52.1%	40.4%	26.6%
2 (minimum- krediet 5; 520 woorde)	83.9%	61.6%	50.6%	38.5%	24.4%

Ofskoon die totaal vir Stuk 2 ontleed byna tweemaal so groot is as dié van Stuk 1, en 'n mens dus redelikerwys kon verwag dat die herhalings groter sou wees, is hulle inderdaad deur 'n bank 'n bietjie laer—'n verdere aanduiding van die moeilikheidsgehalte van die stof. Hoewel hulle laer is, is die verskille egter maar baie klein, en dit was vir my besonder verrassend om deur Ontleding No. 2 feitlik 'n bevestiging van die persentasies in Ontleding 1 te kry, te meer omdat daar so 'n groot verskil is in die hoeveelheid ontlede stof.

Ek het dus rede gehad om te verwag dat hierdie groepe, wat ek as 'n vergelykingsgrondslag gebruik het, in al die ontledings ongeveer as volg sou wees:—

Veelgebruikte woorde	84%
100 meesgebruikte woorde	60%
50 " "	51%
25 " "	40%
10 " "	25%,

en dat hulle hoër of laer sou wees na gelang van die maklikheid of moeilikheid en moontlik ook die hoeveelheid van die behandelde stof.

Die woordeskat van mingebruikte woorde in Ontleding No. 2 (d.w.s. die totale woordeskat min die veelgebruikte woorde) is 2,068, verteenwoordigende 3,175 lopende woorde. Dit beteken dat 79% van die *woordeskat* slegs 16% van die *lopende woorde* lewer; die herhalingsverhouding is 1:1.5.

Die volgende is 'n lys van die 100 meesgebruikte woorde in hierdie ontleding verkry (volgens waarde gerangskik, van die hoogste krediete af na onder):—

ONTLEDING 2: 100 MEESGEBRUIKTE WOORDE.

1. die	1395	33. hom	103
2. en	690	34. na	99
3. 'n	459	al	99
4. nie	409	36. as	98
5. sy	398	37. dat	93
6. het	347	38. net	90
7. hy	302	word	90
8. is	290	40. kan	85
9. in	280	uit	85
10. dit	258	42. se	79
		weer	79
	4,828	44. jy	78
11. hulle	237	45. ook	74
12. te	234	46. my	72
13. van	230	47. moet	71
14. maar	201	48. aan	70
15. was	200	49. baie	68
16. om	194	nog	68
17. vir	188		
18. op	176		10,017
19. so	174	51. sal	67
toe	174	52. soos	66
21. ek	169	53. een	65
22. wat	166	ou	65
23. haar	161	55. groot	62
24. sê	157	56. by	58
25. daar	143	maak	58
		58. sit	56
	7,632	hoor	56
26. met	142	60. af	53
27. kom	134	sien	53
28. gaan	128	62. kyk	51
29. ons	127	63. dis	50
30. dan	119	64. of	49
ma	119	staan	49
32. nou	115	tog	49

67.	jou	48	85.	bly	35
68.	hoe	47		wil	35
69.	mooi	46	87.	dink	34
70.	ander	45		begin	34
	ja	45		hele	34
	vra	45	90.	voor	33
73.	deur	44		klein	33
74.	oor	43	92.	more	32
	oom	43		onder	32
76.	julle	41		twee	32
77.	hier	40	95.	altyd	31
78.	lê	39		meer	31
79.	pa	38		roep	31
	waar	38		teen	31
	water	38		tot	31
82.	gee	37	100.	goed	30
	O!	37			
	weet	37	100.		12,194

3. DAGBLAD: HOOFARTIKELS.

Die stof vir hierdie ontleding het bestaan uit elf hoofartikels in *Die Burger* n.l. dié in genoemde blad se uitgawes van 19 en 22 Februarie 1930, en 5 en 7 Maart 1930, asook die stof in die rubriek *Buitelandse Oorsig* in die uitgawe van 19 Februarie 1930.

Laasgenoemde is ingesluit omdat dit byna op dieselfde voet staan as hoofartikels en, na ek meen, gewoonlik geskryf word deur 'n vooraanstaande lid van die redaksiepersoneel.

Dit is nie moontlik om te sê hoeveel skrywers hierdie stof gelewer het nie aangesien die identiteit van die skrywers van hoofartikels gewoonlik 'n redaksiegeheim bly behalwe vir artikels van 'n politieke strekking in verkiesingstye. Dit kan egter as byna seker aangeneem word dat hulle die werk is van minstens twee of drie verskillende persone.

'n Groot deel van die artikels is twisgeskryf soos grotendeels die manier is van hoofartikels, en die meeste handel oor inheemse aangeleenthede soos die koste van lewensonderhoud, die regte en handhawing van Afrikaans, spoorwegsake, lewe in die stede en op die platteland, die naturellekwessie, geldsake, loterye, vrouestemreg, ens., terwyl die buitelandse oorsig handel oor die ondergang van die Russiese landbou.

Dit is die kleinste van al die ontledings en beloop slegs 6,732 woorde. Vir so 'n klein hoeveelheid lopende woorde moet die verkreeë woordeskate—1,747—as goed beskou word, hoewel nie buitengewoon groot nie. Na aftrek van 42 verskynings van name is die netto totaal, na die verskil (tekort) van 11 in aanmerking geneem is, 6,679.

Vir die ontleding in sy geheel geneem is die herhalingsyfer een van die laagste wat verkry is in al die sestien ontledings waaruit die Onderzoek bestaan. Dit is n.l. 'n ietsie onder die 1:4. Dat dit so laag is moet verklaar word uit die verskeidenheid van die woordekeus in hierdie soort lektuur, wat gewoonlik, wat taal en styl betref, as die beste in die blad beskou word; 'n bykomende faktor sal waarskynlik wees die kleinheid van die ontleding, hoewel dit glad nie met sekerheid gekonstateer kan word nie aangesien daar geen ontleding op 'n groter skaal van dieselfde genre leesstof gemaak is tot staving van so 'n bewering nie, en ook aangesien die persentasie vir die hoogste tien woorde (30.2%) en dié vir die 25 woorde (42.8%), hoër is as dié deur Ontledings Nos. 1 en 2 getoon.

Die hoë persentasies vir hierdie twee groeopies is moeilik om te verklaar; 'n mens sou werklik nie verwag dat die redaksie van *Die Burger* hom vir 30% en 42% op 'n woordeskat van 10 en 25 woordjies onderskeidelik verlaat nie; dit toon egter hoe 'n groot gebruik goeie skrywers maak van hierdie hoekstene van die taal.

Nog onverklaarbaarder word hierdie twee uitslae as 'n mens hulle vergelyk met 74.8% vir die 358 veelgebruikte woorde (krediet van 3 en hoër), 59.2% vir die 100 hoogste woorde en 51.2% vir die eerste 50 woorde, elkeen waarvan laer is as dié vir die ooreenstemmende lysie van die meeste van die ontledings.

Die woordeskat van mingebruikte woorde is 1,389, gewende 'n gesamentlike krediet van 1,679. Hieruit blyk dus dat 72% van die woordeskat slegs 25% van die lopende woorde gee.

Die 100 meesgebruikte woorde in hierdie ontleding (volgens waarde gerangskik) is:—

ONTLEDING 3: 100 MEESGEBRUIKTE WOORDE.

1. die	723	15. dit	64
2. van	265	16. sal	58
3. en	159	17. op	56
4. nie	153	18. ons	54
5. 'n	143	19. vir	50
6. in	135	word	50
7. het	124	21. hy	48
8. is	116	22. aan	46
9. te	106	23. met	42
10. dat	95	24. hulle	39
		25. meer	33
	2,019		
11. om	83		2,860
12. as	75	26. maar	30
13. wat	73	eie	30
14. sy	70	28. party	29

29.	daar	28	65.	stad	11
30.	was	27		koelaki	11
31.	Suid - Afrikaanse	26		kon	11
32.	hul	25		hoe	11
	minister	25		staat	11
34.	kan	24		onder	11
	hierdie	24		vandag	11
36.	ook	23		oor	11
37.	deur	22		platteland	11
	een	22		land	11
39.	wees	21	75.	mense	10
40.	sou	20		paar	10
	al	20		net	10
42.	hom	19		teen	10
	na	19		landbou	10
	by	19		beleid	10
45.	geen	18		laat	10
	tot	18		kom	10
	uit	18	83.	hoop	9
	dan	18		lewe	9
49.	ander	17		doen	9
	alleen	17		blad	9
		—		groter	9
		3,419	88.	sels	8
	nou	17		wetsontwerp	8
	moet	17		eerste	8
53.	of	16		leier	8
	sowjet	16		beskaafde	8
55.	taal	15		daarvan	8
	sonder	15		groot	8
57.	nog	14		uitgawe	8
	so	14	96.	kry	7
	soos	14		gaan	7
60.	baie	13		gee	7
	boer	13		werk	7
62.	elke	12		voor	7
	maak	12			—
	dus	12	100		3,953

4. KORT VERHALE.

Hierdie klas leesstof speel 'n belangrike rol in die hedendaagse gewone lektuur, en dit het my dus gerade gelyk om 'n afsonderlike ontleding daarvan te maak. Hiervoor is gebruik een verhaal elk van vier verskillende Afrikaanse skrywers—twee daarvan beroemd in die Afrikaanse letterkunde—Langenhoven en Leipoldt; en die ander twee—J. Wapenaar en A. H. Jonker—

bedrywige jong skrywers, maar glad nie so goed bekend as die twee eersgenoemdes nie.

Hier is 'n volledige opgawe van die gebruikte stof:—

1. „Die Verdwyning van Flip Loots” (C. J. Langenhoven), *Die Huisgenoot*, 7 Februarie 1930.
2. „Die Galgalsmaller” (C. L. Leipoldt), *Die Huisgenoot*, 13 Desember 1929.
3. „Tom se Erfenis” (J. Wapenaar), *Die Huisgenoot*, 21 Maart 1930.
4. „Die Student” (A. H. Jonker), *Die Huisgenoot*, 21 Maart 1930.

Hierdie ontleding, groot 16,558 woorde netto lewer, soos 'n mens kan verwag, 'n groot woordeskat, naamlik 2,922 verskillende woorde. Die herhalingsyfer is 1:5 vir die ontleding as 'n geheel. Daar is 492 woorde met 'n krediet van minstens 4. Hierdie 492 woorde met herhalings maak 13,232 (79·9%) van die netto totaal van 16,558 lopende woorde, en die herhalingsverhouding vir hierdie woorde alleen geneem is 1:26. Die woordeskat van mingebruikte woorde is 2,430, gewende 3,326 lopende woorde, d.w.s. 83% van die woordeskat gee slegs 20% van die lopende woorde. Die herhalingsyfer vir hierdie deel van die leesstof is 1:1·3.

Vir die aantal lopende woorde is die woordeskat van veelgebruikte woorde (492) nie groot nie, en die persentasie wat hierdie woorde van die totaal uitmaak (79·9%) heelwat onder die verwagte norm, maar die persentasies wat die 10, 25, 50 en 100 hoogste woorde van die totaal uitmaak—25·8%, 40·3%, 51·4% en 61·4% onderskeidelik—toon 'n merkwaardige ooreenstemming met die voorgaande ontledings.

Uit al die gegewens blyk dus dat die woordeskat, in sy geheel geneem, groot is; dat die aantal veelgebruikte woorde beperk is en hulle persentasie laag in vergelyking met ander ontledings, maar wat betref die binneste sirkel van die woordeskat, die skrywers hulle min of meer in dieselfde mate van die onontbeerlike kern van die taal bedien het as in ander geskifte gedaan is.

Hierdie ontleding is die eerste van wat losweg “sosiale” lektuur genoem kan word, d.i. waarin woorde, wat maatskaplike en persoonlike benamings, denkbeelde en verhoudings weergee, in groot mate gevind word.

Vir 61·4% het hierdie vier skrywers (gesamentlik) op die volgende 100 woorde met herhalings staat gemaak:—

ONTLEDING 4: 100 MEESGEBRUIKTE WOORDE.

1.	die	903	42.	gaan	63
2.	nie	554	43.	hulle	62
3.	het	461	44.	kan	61
4.	'n	443	45.	dan	60
5.	en	441	46.	weer	55
6.	hy	387	47.	weet	53
7.	sy	314	48.	tannie	52
8.	van	275	49.	sou	50
9.	te	258		al	50
10.	ek	243			
		4,279			8,521
11.	in	238	51.	dis	49
12.	is	227	52.	oor	46
13.	om	225		ander	46
14.	dit	209	54.	julle	45
15.	was	200		moet	45
16.	wat	184		deur	45
17.	maar	169		soos	45
18.	met	153	58.	tog	43
19.	op	123		wees	43
20.	hom	118	60.	uit	42
21.	vir	115	61.	twee	41
	dat	115	62.	jou	40
23.	nou	110	63.	laat	39
24.	as	108	64.	meer	38
25.	so	106	65.	een	37
				man	37
		6,679	67.	net	35
26.	sal	105		voor	35
27.	my	99		ou	35
28.	of	98	70.	sien	34
	daar	98	71.	waar	33
30.	jy	97	72.	stap	32
31.	aan	93		hier	32
32.	haar	84	74.	kon	31
33.	ons	82		iets	31
34.	na	77		hul	31
	by	77		meester	31
36.	se	74		maak	31
	sê	74	79.	hierdie	30
38.	nog	73		oë	30
39.	toe	70		af	30
40.	kom	68	82.	bly	29
41.	ook	67		hoe	29
			84.	dink	28

	wil	28	95.	weduwee	24
86.	geen	27		word	24
	sit	27		vra	24
	ja	27		hou	24
	huis	27	99.	goed	23
	niks	27		nooit	23
	mos	27			
92.	bietjie	26	100.		10,179
	kry	26			
	alles	26			

5. SKOOLLEESBOEK, STANDERD II.

Vir hierdie ontleding is gebruik 'n Standerd II. skoolleesboek van die serie *Lees Met Lus*, deur Hoogenhout, Botha en van der Lingen. (Uitgewer, Maskew Miller, Bpk., Kaapstad, 5de Druk). Al die stof in die boek, met die uitsondering van poësie, is ontleed, en die aantal lopende woorde is 14,353 netto. Dit is, sover nagegaan kan word, die werk van minstens vier skrywers.

Die woordeskat is groot 2,067, d.w.s. omtrent 300 meer as die woordeskat van twee uitgawes van *Die Kleinspan*. Daar is 631 woorde met 'n krediet van 3 en meer.

Dié 631 veelgebruikte woorde vorm 87.1% van die netto totaal; die 100 meesgebruikte woorde 63.3%; die eerste 50 maak 52.4%; die 25 en 10 meesgebruikte woorde maak 39.2% en 24.6% onderskeidelik uit van die netto totaal.

Daar val nie veel aan te merk op die syfers van hierdie ontleding nie; hulle is min of meer op die norm—die persentasies van die eerste 10 en 25 woorde 'n ietsie onder, en dié vir die 100 meesgebruikte woorde en vir die 631 veelgebruikte woorde heelwat bokant. Die volledige syfers is te vinde in die Tabel van die uitslae vir die verskillende ontledings, teenoor bladsy 66.

ONTLEDING 5: 100 MEESGEBRUIKTE WOORDE.

1.	die	826	11.	is	178
2.	en	517	12.	van	174
3.	het	422	13.	te	167
4.	'n	349	14.	haar	161
5.	nie	326	15.	maar	146
6.	sy	248	16.	dat	141
7.	hy	238	17.	ek	140
8.	hulle	224	18.	nou	139
9.	was	203	19.	wat	133
10.	in	191	20.	om	128
			21.	as	124
			22.	vir	117
		3,544			

	ons	117	61.	waar	38
24.	toe	115		sou	38
25.	so	112	63.	of	37
		5,636	64.	gou	36
26.	dan	102		laat	36
27.	daar	101	66.	dag	34
28.	op	100	67.	se	33
	hom	100	68.	ma	31
30.	met	96		kinders	31
31.	sal	95		groot	31
32.	dat	94	72.	wil	31
	julle	94		kon	30
34.	aan	84		weet	30
35.	sê	82		hierdie	30
36.	my	78	75.	kyk	28
37.	gaan	73		begin	28
38.	kom	70	77.	almal	27
39.	baie	66		vertel	27
	nog	66		hier	27
41.	ou	64	80.	mens	26
	ook	64		klein	26
43.	word	62		maak	26
	weer	62		want	26
45.	jy	61	84.	ander	25
46.	na	60		gehad	25
	kan	60		goed	25
48.	een	55		moes	25
	net	55		loop	25
50.	by	54	89.	deur	24
		7,534		staan	24
	al	54		vra	24
52.	hoe	53		altyd	24
53.	uit	48		onder	24
54.	tog	44		moeder	24
55.	sien	43	96.	kry	24
	jou	43		daardie	23
57.	moet	42		dis	23
58.	mooi	40	98.	oom	22
	pa	40		hoor	22
60.	ja	39	100.	soos	21
			100.		9,091

6. BRIEWE: LEWENSVERSEKERING.

Dit was oorspronklik my plan om een ontleding te maak van besigheidsbriewe van verskillende aard, en nie soveel van 'n enkele

soort besigheid nie. Met die oog hierop het ek verskillende sake-manne wat ek ken genader, want ek was daarop gesteld dat die briewe vir ontleding briewe moes wees wat werklik deur besighede uitgestuur en ontvang is, met die oog op die praktiese doel van die Onderzoek.

Ek het egter spoedig uitgevind dat dit amper makliker is vir 'n kameel om deur die oog van 'n naald te gaan as om besigheids-manne te beweeg om 'n klompie afskrifte van alledaagse briewe, wat deur hulle ontvang of uitgestuur is, tot my beskikking te stel. 'n Onderneming in likwidasië, op wie se personeel ek vroeër diens gedoen het, laat liewers sy ou briewe by karvragte wegry en verbrand as om 'n vyftigtal tot my beskikking te stel vir 'n wetenskaplike ondersoek. 'n Ander, aan wie ek ook nie 'n onbekende is nie, laat liewers, by die verskuiwing van sy kantoor, ou briewe by hope vernietig as om aan my smekende versoek om, al is dit maar 'n dosyn of so, vir hierdie Onderzoek beskikbaar te stel, te voldoen.

Daar is, dit weet ek, soiets in sake as vertrouë en geheimhouding, maar dit kan oordryf word. In elk geval, hierdie onnodige en kinderagtige argwaan en geheimsinnigheid is nie bereken om die saak van wetenskaplike ondersoek te bevorder nie.

Gelukkig vir die werk was daar 'n skitterende oase in hierdie dorsland van agterdog en onverskilligheid. Ek het my nood bekla by mnr. M. S. Louw, bestuurder van die S.A. Nasionale Lewensassuransie Maatskappy, Beperk (S.A.N.L.A.M.), en hierdie keer het ek nie tevergeefs aangeklop nie. Mnr. Louw het nie alleen deur sy warme belangstelling my gedurig aangemoedig in die werk nie; hy het ook daadwerklike hulp verleen. Hy het bereidwilliglik 'n aantal briewe uit die saak waarvan hy bestuurder is, tot my beskikking gestel. Meer, deur sy bemiddeling het die tiksterspersoneel die briewe op die ruitjiespapier getik, en het hulle en ander lede van die staf met die opsnippering en die sortering gehelp.

Soos reeds gekonstateer, was dit nie my plan om 'n enkele ontleding aan die briewe van een soort besigheid te wy nie; die ontleding is eintlik gemaak op versoek van mnr. Louw. Terwyl dit tog gemaak is het ek besluit om dit in die ganse Onderzoek op te neem en die resultate daarby in te sluit.

Die ontleding bestaan uit 110 briewe altesame. Hiervan is 77 uitgestuur en 33 deur die kantoor van buite ontvang. Die uitgestuurde briewe is geskryf (of dikteer) deur twaalf persone, verantwoordelike lede van die personeel, en die ontvangte briewe is dié van ongeveer 33 persone, sodat die stof vir hierdie

ontleding die werk is van omtrent 45 persone. Die inkomende briewe strek oor 'n tydperk van min of meer nege maande; die uitgestuurde briewe oor omtrent 14 dae.

Aangesien die briewe op ruitjespapier oorgetik is deur die Maatskappy se tiksters, en die papier onder my toesig op hulle kantoor opgesnipper is, het ek persoonlik nie insage van die briewe as *briewe* gehad nie; wel as 'n woordemassa sonder sinsverband. My is egter meegedeel dat die briewe handel oor die volgende afdelings van 'n versekeringsaak: Nuwe besigheid, lenings en afkopings, vernuwings, groepversekering, eise, pensioene, reklamewerk, briewe aan agente, algemeen. Op 'n vraag het mnr. Louw die mening uitgespreek dat hierdie materiaal goed verteenwoordigend is van die assuransiebesigheid. Al die woorde in die briewe, met inbegrip van die adres, die aanhef en die slot, maar met die uitsondering van die name van plekke en persone, is geneem.

Die netto aantal lopende woorde (na die aftrek van 139 name en 23 papiertjies wat deur die opsnippering onleesbaar gemaak is) is 21,037. Die gemiddelde aantal woorde per brief vir die gros totaal is dus 193.

Die ganze woordeskat in hierdie stof aangetref is 2,324. As die aantal lopende woorde in aanmerking geneem word (21,037 netto—die grootste ontleding tot dusver) is die woordeskat bepaald nie groot nie; dit is, inderdaad, kleiner as dié aangetref in die eerste 101 bladsye van 'n Standerd III. skoolleesboek en slegs 257 groter as dié van 'n Standerd II. skoolleesboek. Waar dit, soos in hierdie geval, briewe van slegs 'n enkele soort besigheid betref, kan 'n mens so 'n resultaat, ofskoon dit merkwaardig is, verwag. Waar die omvang en verskeidenheid van die taalskat so betreklik beperk is, stel dit nie besonder sware eise aan die tiksterspersoneel nie. By so 'n afwesigheid van omvang behoort 'n tikster, wat in so 'n saak diens neem, binne korte tyd 'n hoë mate van bedrewenheid te bereik. Dit is geen wonder dat (soos baie tiksters in verskillende takke van die sakewêreld werksaam my al self verseker het) met 'n *algemene* stenografiese snelheid van slegs circa 60 woorde per minuut, hulle maklik na enige tyd al die diktee van hulle besondere besigheid kan afneem.

Soos gemerk sal word uit die Staat teenoor bladsy 66 is die persentasiesyfers en herhalingsverhouding baie hoog vir hierdie ontleding—met 'n enkele uitsondering, dié vir die 10 meesgebruikte woorde—die hoogste tot sover aangetref. Die 579 veelgebruikte woorde (krediet van 5 en hoër) maak 86.1% uit van die netto totaal van 21,037 lopende woorde ontleed; die 100 hoogste woorde vorm 63.8% van die geheel; die 50 meesgebruikte woorde 55.0%; die eerste 25 woorde met herhalings 44.8% en die 10

hoogste woorde 28.7%. Daar is 'n woordeskat van 1,745 mingebruikte woorde (75% van die ganse woordeskat), en hierdie woordeskat verteenwoordig slegs ongeveer 14% van die netto totaal lopende woorde. Vir hierdie ontleding is, in plaas van die lys van die 100 meesgebruikte woorde, 'n lys van die 180 meesgebruikte woorde gemaak; hulle vorm 71.2% van die netto aantal lopende woorde.

ONTLEDING 6: 180 MEESGEBRUIKTE WOORDE.

1.	die	1,548	31.	by	90
2.	van	722		uwe	90
3.	u	720	36.	moet	89
4.	ons	578		daar	89
5.	'n	506	38.	deur	87
6.	en	430	39.	maatskappy	83
7.	is	413	40.	tot	81
8.	in	404	41.	na	75
9.	te	393	42.	of	73
10.	dat	331	43.	bestuurder	72
			44.	sy	68
		6,045	45.	nog	67
11.	sal	318	46.	hy	61
12.	dit	277	47.	maar	60
13.	om	269	48.	was	58
14.	nie	264		betaal	58
15.	wat	253		dan	58
16.	het	250			
	vir	250			11,574
18.	op	241	51.	bedrag	56
19.	word	221	52.	onder	55
20.	ek	210		ook	55
21.	aan	186		so	55
22.	as	177	55.	dus	50
23.	met	165	56.	premies	49
24.	my	156	57.	nou	47
25.	polis	152		brief	47
			59.	deser	46
		9,434	60.	maak	43
26.	kan	149	61.	doen	42
27.	wees	124		kom	42
	hierdie	124		bly	42
29.	jaar	107		enige	42
30.	waarde	99	65.	baie	41
31.	heer	95	66.	voor	40
32.	premie	92	67.	graag	38
33.	wil	91		tyd	38

69.	assistent	37	114.	bogenoemde	23
70.	besigheid	36		weer	23
	stuur	36		sluit	23
	verder	36	117.	waar	22
	al	36		besit	22
74.	polisse	35		daarop	22
75.	ontvang	34		eerste	22
	ander	34		daardie	22
	neem	34	122.	daarvan	21
	soos	34		winste	21
	rente	34		aansoek	21
80.	bonusse	33		berig	21
	moontlik	33		gemaak	21
	se	33		lewe	21
83.	assuransie	32		mag	21
	gaan	32		nodige	21
	hom	32		volgende	21
	saak	32		sou	21
87.	geval	31		voltooi	21
	betalbaar	31	133.	egter	20
89.	skrywe	30		erken	20
	ten	30		voorbaat	20
	maand	30		uit	20
	asseblief	30		uitstaande	20
	een	30	138.	hiermee	19
	laat	30		net	19
	hulle	30		ouderdom	19
	agent	30		stel	19
	weet	30		verskaf	19
98.	gee	29		teen	19
99.	nuwe	28		sê	19
100.	antwoord	27	145.	wys	18
		13,431		bewys	18
	deel	27		goed	18
	sake	27		boeke	18
103.	meneer	26		hou	18
	aktuaris	26		predikant	18
	end	26		verneem	18
	verband	26	152.	dankie	17
107.	elke	25		want	17
	geen	25		informatie	17
	vind	25		agente	17
	jaarlikse	25		eers	17
111.	lening	24		hoop	17
	skema	24		krag	17
	twee	24		meer	17
				ontvangs	17

	raad	17		groot	15
	nommer	17		dank	15
	volle	17		oor	15
	versoek	17		vra	15
165.	alle	16		volgens	15
	geleentheid	16		sien	15
	reeds	16		sonder	15
	trek	16		verkry	15
	vereffen	16	180.	stoporder	15
170.	betref	15			
	kennis	15	180.		14,996

7. TOESPRAKE: HANSARD.

Aangesien daar 'n woordelike stenografiese opname van die toesprake in die Volksraad gemaak word het ek gemeen dat dit nuttig sou wees as hiervan ook 'n stuk ontleed word, en die begrotingsdebatte is hiervoor gekies omdat daar in die loop van hierdie besondere debatte gewoonlik 'n groot verskeidenheid alledaagse sake onder bespreking kom. Slegs 'n gedeelte van die ontlede stof kan egter as die gesproke woord (as onderskeie van die geskrewe woord) beskou word, aangesien die begrotingsredes van die Ministers van Finansies en Spoorweë gewoonlik sorgvuldig voorberei en afgelees word.

Die twee toesprake beslaan dan ook die grootste gedeelte van die ontleding, en daar is slegs enkele toesprake van ander lede (en handelende oor ander sake) in die ontleding opgeneem. Hulle is, weliswaar, vermoedelik ook voorberei, en mis dus enigszins die ongedwongenheid van gewone onvoorbereide toesprake uit die vuur uit gelewer.

Die stof bestaan n.l. uit die begrotingsredes van die Ministers van Finansies en Spoorweë gedurende die Parlementsitting van 1930 en toesprake deur wyle dr. Lamprecht en dr. Stals, en die here Strydom (gedeelte) en Wolfaard. Die begrotingsredes is in Engels gehou (die Hansard-vertaling is gebruik) en die ander is almal oorspronklik in Afrikaans gehou. Die stof is dus die uiting van ses persone. Dit is alles ontleen aan die weeklikse Afrikaanse uitgawe van Hansard No. 10 (1930), bladsye 2619 tot 2640; 2673; 2674; 2679 tot 2683; 2691 tot 2693; 2698 tot 2701; 2708 tot 2716.

Alle syfers is geskrap en hier en daar selfs 'n hele staat met woorde en al, omdat hulle nie in sinne aaneengeskakel is nie.

Die ander toesprake handel oor oproerige byeenkomste en het grotendeels betrekking op die verhouding tussen blankes en nie-blankes in Suid-Afrika.

Die ontleding beloop netto (na aftrek van 73 verskynings vir name) 16,851 woorde. Die hele woordeskat is 2,314—nie groot nie as die verskeidenheid van die onderwerpe onder bespreking in aanmerking geneem word—selfs 10 minder as die woordeskat in 110 assuransiëbriewe aangetref, ofskoon natuurlik die aantal lopende woorde van hierdie ontleding heelwat kleiner is as dié van Ontleding No. 6.

Daar is 607 veelgebruikte woorde (krediet van minstens 4) makende 85·1% van die netto totaal uit.

Die persentasies wat die 100, 50, 25 en 10 hoogste woorde met hul veelvuldige herhalings van die totaal uitmaak (61·9%, 53·7%, 45·6% en 31·1% onderskeidelik) is almal bokant die norm—vir die twee laasgenoemde die hoogste wat dusver aangetref is. 'n Mens sou meen dat woorde van besondere toepassing (vakterme) vir hierdie uitslag verantwoordelik is. 'n Deursage van die 100 hoogste woorde staaf nouliks hierdie mening. In die eerste 25 is daar geen enkele woord van besondere toepassing op dié soort stof nie. Die eerste sodanige woord—*inkomste*—staan nommer 31 op die lys, en laer af in dieselfde lys van 100 is daar slegs 'n verder 21 wat as besondere woorde beskryf kan word, t.w. *waarde*, *agbare*, *raming*, *uitgawe*, *vermindering*, *land*, *vermeerdering*, *verdrag*, *persent*, *invoer*, *naturelle*, *begroting*, *lede*, *uitvoer*, *diamante*, *lid*, *minister*, *wetsontwerp*, *miljoen*, *prys*, *hoeveelheid*.

Hiervan val slegs 7 in die eerste 50, en enkele, soos *waarde* en *land*, het ook groot krediete in ander ontledings opgedoen.

ONTLEDING 7: 100 MEESGEBRUIKTE WOORDE.

1. die	1,745	16. as	173
2. van	818	word	173
3. en	460	18. sal	166
4. in	423	19. om	165
5. 'n	386	20. ek	162
6. is	371	met	162
7. nie	276	22. aan	154
8. het	275	23. jaar	109
9. dat	257	24. tot	98
10. wat	232	25. daar	93
	5,243		7,696
11. te	229	hierdie	93
12. op	212	27. was	79
13. vir	193	28. maar	72
14. dit	186	29. uit	69
15. ons	178	30. hulle	68

31.	inkomste	66	66.	invoer	30
	moet	66		oor	30
33.	kan	63	68.	groot	29
34.	wees	57	69.	naturelle	28
35.	of	52		hier	28
36.	deur	51		mense	28
37.	daardie	50		toe	28
38.	waarde	49	73.	kom	27
39.	meer	48		begroting	27
	na	48		verwag	27
41.	agbare	47		sê	27
	hy	47	78.	ter	27
43.	ook	46		terwyl	26
44.	raming	45		lede	26
45.	uitgawe	44		gevolg	26
46.	vermindering ..	43	81.	uitvoer	25
47.	sy	42		diamante	25
48.	land	41		lid	25
49.	onder	39	84.	minister	24
50.	my	38		geen	24
		9,059		koring	24
51.	gaan	36		wetsontwerp ..	24
	per	36	88.	teenoor	23
	vermeerdering ..	36		ten	23
	ander	36		waar	23
55.	by	35		meel	23
	dan	35		miljoen	23
57.	gedurende	34		maande	23
	minder	34		nou	23
	wil	34		daarvan	23
60.	teen	33	96.	prys	22
	ton	33		hoeveelheid ..	22
62.	toestand	32	98.	soos	21
63.	bedrag	31		een	21
	persent	31		eerste	21
	so	31	100.		10,442

8. 'N ALGEMENE LETTERKUNDIGE WERK.

Hiervoor is gebruik dié welbekende Afrikaanse verhale—*Sewe Duiwels* deur Oom Jan, m.i. een van die beste boeke van kort verhale in Afrikaans, en 'n werk wat nog nie genoeg waardeur word nie. Ek weet van g'n enkele boek in Afrikaans wat die suiwere atmosfeer van die plaas- en die algemene geselligheidslewe van ons boermense beter as hierdie verhale weergee nie; wat 'n getrouer weerspieëling is van die egte, ongekonstelde Afrikaanse

volkstaal van die platteland, in teenstelling met die Afrikaans met 'n sterk Engelse en/of Nederlandse geur van baie van ons hedendaagse skrywers.

Die jongste uitgawe van hierdie werk (dertiende druk, deur die Nasionale Pers, Beperk, Bloemfontein, uitgegee) is hiervoor gebruik. Al die stof behalwe opskrifte en Nederlandse gedeeltes (en natuurlik name van plekke en persone) vanaf bladsy 10 tot op bladsy 98 is geneem. Dit behels die verhale *Geldduiwel*, *Liegduiwel* en die grootste gedeelte van *Praatduiwel*, en beloop gros 30,672 woorde leesstof. Die hele boek beslaan 280 bladsye, sodat dit by benadering in sy geheel 95,000 woorde bevat.

Dit is die grootste van al die verskillende ontledings, en is so groot gemaak om te dien as 'n aanwysing of die vergelykingsyfers baie deur 'n groter ontleding beïnvloed word.

Die netto totaal lopende woorde verkry is 29,349. Die woordeskat in die behandelde gedeelte van die boek aangetref is 2,639, d.w.s. 51 groter as wat van 'n Standaard III. leerling verwag word, maar kleiner as gevind is in die 16,558 lopende woorde waaruit die ontleding van vier kort verhale bestaan.

Hoe groot die woordeskat van die hele boek sal wees is moeilik om uit die verkreeë uitslag te skat, want dit moet nie vergeet word nie dat die verhouding van die aantal nuwe woorde skerp daal hoe meer lopende woorde van een stuk ontleed word. Daar is, na raming, ietwat onder 'n derde van die ganse werk ontleed, maar dit kan met volslae sekerheid gekonstateer word dat die woordeskat in die hele boek nie driemaal so groot sal wees as dié van die ontlede gedeelte nie. Inderdaad, ek betwyfel sterk of die woordeskat vir die hele *Sewe Duiwels* veel groter as 3,500 woorde sal wees—natuurlik op die grondslag wat deurgaans in die Ondersoek toegepas is—dat elke woordvorm as 'n selfstandige en afsonderlike woord tel.

Soos verwag kan word is die herhalingsverhouding in hierdie ontleding hoog. *Verwag kan word*, sê ek, en wel veral om twee redes. Eerstens omdat die aantal lopende woorde so groot is, en tweedens omdat die stof bestaan uit eenvoudige lektuur, en die verhoudings heel waarskynlik, sonder die bykomende faktor van die groot aantal lopende woorde, tog hoog sou gewees het.

Die herhalingsyfers vir die ontleding as 'n geheel, vir die veelgebruikte woorde, en vir die mingebruikte woorde, is 1:11, 1:47 en 1:1·7 onderskeidelik—in ieder geval heelwat hoër as tot dusver aangetref is, en hoër as deur enige van die hele sestien ontledings aangetoon is—met een uitsondering, dié vir die ontleding as 'n geheel, wat alleen vir Ontleding No. 12 hoër is.

Vir hierdie ontleding is 'n lys gehou van alle woorde met 'n krediet van minstens 6. Daar is 544 sulke woorde, en hulle lewer altesame (met herhalings) 25,747 (87.7%) van die netto totaal van 29,349 lopende woorde.

Die persentasies vir die 100, 50, 25 en 10 meesgebruikte woorde (n.l. 66.5%, 54.6%, 41.3% en 25.7% onderskeidelik) is vry hoog, maar nie in alle gevalle die hoogste tot dusver verkry nie, en in geen enkel geval die hoogste nie van al die ontledings waaruit die Ondersoek bestaan. Uit die voorgaande skyn die voor die hand liggende gevolgtrekking te wees dat, ofskoon die grootte van die ontleding wel die herhalingsverhouding beïnvloed, dit nie die persentasies van die mees- en veelgebruikte woorde besonder veel aantass nie.

Die 2,095 mingebruikte woorde lewer slegs 3,602 van die netto totaal van 29,349 lopende woorde. Dit wil sê, 79% van die *woordeskat* gee slegs 12% van die gebruikte *woordetal*. Dit vorm 'n sterke teenstelling met die veelgebruikte woorde, waar 'n *woordeskat* van slegs 544 of 20% van die hele *woordeskat*, 87.7% van die gebruikte *woordetal* vorm.

Vanaf hierdie ontleding is ook aantekening gehou van die *aantal* woorde wat vir die eerste keer teëgekom is. Vir elke sodanige woord moes natuurlik 'n afsonderlike kaartjie gemaak word. Hierdie aantalle het verrassende uitslae gelever, en het duidelik aangetoon dat my oorspronklike mening dat, na die eerste honderd-duisend woorde leesstof ontleed is, daar betreklik weinig nuwe woorde sou voorkom, 'n lelike vergissing was. In hierdie ontleding is daar b.v. 767 sulke nuwe woorde aangetref. Hierdie aantal is groot as die hoeveelheid en aard van die reeds behandelde leesstof en die betreklike maklikheid van die stof vir hierdie ontleding, in aanmerking geneem word.

ONTLEDING 8: 100 MEESGEBRUIKTE WOORDE.

1. die	1,607	11. te	432
2. nie	936	12. ek	416
3. en	822	13. is	388
4. 'n	757	14. oom	349
5. hy	693	15. dat	328
6. het	629	16. wat	314
7. was	621	sê	314
8. dit	517	18. van	298
9. maar	506	19. om	293
10. sy	478	20. in	253
		21. toe	252
		22. op	238
	7,566		

23.	vir	235	61.	uit	84
24.	so	229	62.	laat	81
25.	met	218	63.	weet	79
			64.	jou	78
		12,123	65.	een	76
26.	jy	214	66.	kry	75
27.	as	213	67.	vra	73
28.	haar	208	68.	meneer	71
29.	daar	199	69.	oor	70
30.	ook	193		sit	70
31.	kom	192		antwoord	70
32.	gaan	184		kyk	70
33.	hom	179	73.	ou	69
34.	sal	178	74.	doen	68
35.	dan	158		deur	68
36.	al	155	76.	wees	67
37.	ons	152		bietjie	67
38.	na	151		mense	67
	hulle	151	79.	hoe	66
40.	my	146	80.	goed	64
41.	nou	141		paar	64
42.	kan	140		se	64
43.	word	132	83.	hou	63
44.	aan	129	84.	dink	61
	by	129		hier	61
46.	sou	116		maak	61
	moet	116		vertel	61
48.	nog	113	88.	alles	57
	wel	113		bly	57
50.	of	101		eers	57
			91.	moes	56
		16,026	92.	wou	55
	wil	101		waar	55
52.	tante	99	94.	mooi	54
	net	99	95.	niks	51
	kon	99	96.	help	49
55.	tog	97		stil	49
56.	had	92		geen	49
57.	praat	91	99.	hoor	48
58.	gee	90		geld	48
	weer	90			
60.	sien	88	100.		19,525

9. STOF UIT DIE BYBEL.

Die Bybel is ongetwyfeld een van die alledaagse, meesgelese boeke in ons land, 'n soort lektuur waarin woorde in verband met die grondbegrippe van die menslike bestaan, sy siele-, sedelike en

godsdienstige lewe, baie gebruik word, en ek het dit dus raadsaam en nodig geag om hiervan ook 'n monster te ontleed. Ongelukkig was die hele Afrikaanse vertaling van die Bybel nog nie beskikbaar toe die ontleding gemaak is nie; alleen die vier Evangelies en die Psalms. Maar aangesien hierdie tog die meesgelese gedeeltes van die Bybel is, het ek besluit om 'n aantal hoofstukke hieruit vir die werk te neem, en hiervoor is die uitgawe van 1929—sover ek weet die jongste hersiening destyds—gebruik.

Ek wou hierdie ontleding 15,000 woorde groot maak en het, sover hulle my bekend is, dié hoofstukke gekies wat die meeste gelees word en, wat die Evangelies betref, het ek ook probeer om sover moontlik hoofstukke te kies wat nie in die verskillende Evangelies oormekaar slaan nie, want al is die bewoording van dié gedeeltes nie presies gelykluidend nie, moes dit noodwendig die herhaling onbehoorlik beïnvloed. Opskrifte, en opsommings bo-aan en in die verskillende hoofstukke, is nie gebruik nie.

Die stof vir hierdie ontleding bestaan uit Psalms 10, 11, 12, 17, 18, 19, 23, 40, 42, 49, 51, 62, 73, 90, 91, 100, 103, 119, 120 en 121 (circa 7,900 woorde), en die volgende hoofstukke uit die Evangelies is ingesluit: Matteus 5, 6 en 7; Markus 10 en 16; Lukas 2 en 15 en Johannes 3 en 4 (omtrent 8,042 woorde); altesame 29 hoofstukke.

Ek beskou hierdie stof as die werk van sewe skrywers—die vier Evangeliste, Dawid, Asaf (Ps. 73) en Moses (Ps. 90).

Die aantal lopende woorde van die ontleding is 15,918 netto, en die ganse woordeskat vir hierdie 29 hoofstukke groot 1,820; nie minder as 600 hiervan is vir die eerste keer aangetref, d.w.s., omtrent 'n derde deel van die woordeskat is nie wat genoem kan word alledaagse woorde nie.

Hoe eenvoudig die grootste deel van die leesstof van die Bybel egter is blyk nie alleen uit die betreklike kleinheid van die woordeskat nie; dit tree des te skerper aan die dag by 'n vergelyking met die woordeskat van ander ontledings wat by benadering uit dieselfde aantal lopende woorde bestaan.

Ontleding No. 4 (16,558 lopende woorde netto) lewer 'n woordeskat van 2,922 op; Ontleding No. 5, 'n kleiner stuk (14,353 lopende woorde)—'n Standaard II. skoolleesboek, gee 'n woordeskat van 2,067; Ontleding No. 7 (toesprake) met 16,851 woorde, verskaf 'n woordeskat van 2,314, terwyl 'n Standaard I. skoolleesboek (*Lees Met Lus*), met slegs 13,159 lopende woorde (Ontleding No. 11) 1,883 verskillende woorde, 63 meer as in hierdie Bybelse stof aangetref, oplewer.

'n Merkwaardige kommentaar voorwaar op die eenvoud van die taal van hierdie groot godsdienstige en letterkundige werk—

die Bybel, 'n boek wat die grootste en diepste vraagstukke van die lewe behandel; 'n vingerwysing vir dié Afrikaanse skrywers wat meen dat dit „geleerd en fraai” is om ellelange saamgestelde woorde te gebruik, soms vir die eenvoudigste geskrifte. Laat hulle die Bybel lees!—so nie vir die saligheid van hulle siele nie dan wel vir die hervorming van hulle woordekeus.

As ons die herhalingsyfers en die persentasies vir die veelgebruikte woorde nagaan, tree hierdie eenvoud nog duideliker aan die lig.

Elke woord in hierdie stof as 'n geheel geneem is gemiddeld 8 keer herhaal. Daar is 631 woorde met 'n minimumkrediet van 3, en die herhalingsverhouding vir hierdie woorde is 1:22. Dit wil sê, vir hierdie woorde alleen geneem het elkeen van die 631 gemiddeld 22-maal voorgekom. Dit laat 1,189 mingebruikte woorde, wat gemiddeld 1.2 maal herhaal is.

Vir 90.7% het die skrywers van genoemde hoofstukke hulle slegs op dié 631 veelgebruikte woorde, en herhalings daarvan, verlaat; vir 69.3% is van slegs 100 verskillende woorde gebruik gemaak, terwyl die 50, 25 en 10 meesgebruikte woorde (met herhalings) onderskeidelik 59.7%, 47.9%, en 31.5% van die netto totaal gebruikte woorde uitmaak.

Hierdie persentasiesyfers is sonder uitsondering die hoogste tot hiertoe in die Onderzoek aangetref, en toon duidelik hoe 'n groot gedeelte van die leesstof uit die Bybel uit 'n baie beperkte woordeskat bestaan.

ONTLEDING 9: 100 MEESGEBRUIKTE WOORDE.

1. die	1,026	16. hom	177
2. en	775	17. 'n	172
3. het	478	18. sal	171
my	478	19. julle	151
5. u	456	20. Here	138
6. nie	446	21. want	126
7. van	423	22. jou	116
8. is	333	op	116
9. ek	308	na	116
10. hulle	292	25. dit	115
	5,015		7,638
11. wat	290	26. sê	110
12. vir	261	27. maar	102
13. in	258	28. dat	100
14. hy	222	God	100
15. sy	194	30. te	92

31.	word	91	66.	wet	35
	was	91	67.	deur	34
	met	91	68.	moet	33
34.	om	89		baie	33
	aan	89		hierdie	33
36.	ons	84	71.	so	32
37.	as	83	72.	siel	31
38.	toe	81		geen	31
39.	laat	72	74.	man	30
40.	uit	69	75.	lewe	29
41.	jy	64		sodat	29
42.	kan	61	77.	of	28
43.	oor	58		mense	28
44.	daar	53	79.	omdat	27
	wees	53		elkeen	27
46.	by	51	81.	wil	26
	al	51		seun	26
48.	Jesus	50		hul	26
	soos	50	84.	vrou	25
50.	kom	45		hemel	25
		9,518		aarde	25
	gaan	45	87.	weg	24
52.	vader	42		haar	24
	O!	42		af	24
	gesê	42		gekom	24
55.	woord	41	91.	insettinge	23
56.	doen	40		bevele	23
	dan	40		gebooie	23
58.	maak	39	94.	geregtigheid	22
	voor	39		groot	22
60.	hart	38		mag	22
	gee	38		kyk	22
62.	tot	37		oë	22
	sien	37		meer	22
	ook	37	100.	bewaar	21
65.	een	36	100.		11,042

10. SKOOLLEESBOEK, STANDERD I.

Vir hierdie ontleding is 'n Standaard I. boek van die reeks *Die Afrikanerkind*, 15de Druk (deur C. J. Alers, Nasionale Pers, Bloemfontein), gebruik. Al die stof (behalwe opskrifte en poësie) in die boek—12,186 woorde netto—is geneem.

Die boek is, sover nagegaan kan word, die werk van een skrywer.

Daar dit vir onderwysdoeleindes nuttig kan wees, gee ek vir hierdie ontleding 'n volledige lys van die ganse woordeskat, om dit dan te vergelyk met dié van Ontleding No. 11, wat ook van 'n Skoolleesboek, Standaard I. is, maar van 'n ander serie (*Lees Met Lus*, sien bladsy 43), en ook met die woordeskat van 'n Standaard II. leesboek. Hierdie lyste vorm Bylae I. agter in die boek.

Hierdie boekie is die eenvoudigste, wat woordekeus betref, van al die ontledings skoolboeke en, op een na openbaar dit die kleinste woordeskat in enige van die sestien ontledings aangetref, n.l. slegs 1,296. As al die woordvorme van 'n woord as een woord behandel is (wat nie gedoen is nie) sou die woordeskat nog aanmerklik kleiner wees.

Hierdie woordeskat is byna presies net die helfte van dié van *Jong Suid-Afrika, Standaard III.* (Ontleding No. 2) en omtrent 800 kleiner as Standaard II, van *Lees Met Lus* oplewer, terwyl dit ongeveer 600 kleiner is selfs as die woordeskat van 'n Standaard I. boek van die *Lees Met Lus* serie. Dit toon die beperkte woordeskat en dus die eenvoudigheid van hierdie leesboekie, wat populêr moet wees aangesien dit 'n eksemplaar van die 15de druk is.

Die netto aantal woorde is 12,186. Daar is 509 veelgebruikte woorde (krediet van 3 en meer).

Die persentasie wat die 509 veelgebruikte woorde van die geheel uitmaak (91.7%) is die hoogste tot hiertoe aangetref, en dié van die 100 meesgebruikte woorde (69.3%) is presies dieselfde syfer as dié van Ontleding No. 9, wat tot sover ook die hoogste was. Vir die 50, 25 en 10 hoogste woorde is die persentasies 56.7%, 42.6% en 27.6% onderskeidelik—ook almal goed bokant die verwagting.

Vir 'n opgawe van die herhalingsverhoudings, ens. word die leser verwys na die Tabel teenoor bladsy 66.

In hierdie ontleding is 194 woorde aangetref wat nie tevore in die Onderzoek teëgekom is nie. Met die oog op die feit dat reeds meer as 160,000 woorde leesstof ontleed is, en daaronder inbegrepe is skoolleesboeke Standaards II. en III., is hierdie aantal nuwe woorde opvallend, en 'n aanduiding van die uitgebreidheid van die woordeskat van 'n taal.

ONTLEDING 10: 100 MEESGEBRUIKTE WOORDE.

1. die	993	8. daar	202
2. hy	367	9. dit	198
3. nie	321	10. het	197
4. sy	302		
5. en	270		3,364
6. 'n	263	11. in	189
7. is	251	12. ek	152

13.	op	144	56.	maak	41
14.	kom	136		oom	41
15.	maar	134		onder	41
16.	nou	121	59.	kat	40
	jy	121	60.	se	38
18.	dan	119	61.	ja	36
19.	vir	116		huis	36
20.	so	105		oor	36
21.	ma	104	64.	baie	34
22.	toe	99		deur	34
	kyk	99		na	34
	sê	99	67.	laat	33
25.	dis	98	68.	moet	32
				wil	32
		5,200		hou	32
26.	van	92		lê	32
27.	loop	85	72.	voor	31
28.	hom	84	73.	goed	28
29.	hoe	81		daardie	28
	met	81		kop	28
31.	hulle	79		eers	28
32.	kan	77	77.	nee	27
	my	77		vat	27
34.	te	75		water	27
35.	sal	74		boom	27
36.	om	68		as	27
37.	net	67	82.	kombuis	26
38.	jou	66		ons	26
	staan	66		val	26
	gaan	66	85.	soos	25
41.	wat	65		waar	25
	by	65		stil	25
43.	sit	63		word	25
44.	sien	61		ander	25
45.	ook	59		bietjie	25
46.	aan	55		lekker	25
47.	vra	54	92.	lyk	24
48.	uit	53		roep	24
49.	hier	51	94.	mooi	23
50.	weer	50		weg	23
				klein	23
		6,914	97.	dag	22
51.	haar	49		trek	22
	al	49		af	22
53.	een	43		groot	22
	hoor	43			
55.	nog	42	100.		8,448

11. SKOOLLEESBOEK, STANDERD I.

Na die vorige ontleding het dit my voorgekom dat dit baie nuttig sou wees om 'n vergelyking te tref tussen die woordeskatte van leesboeke van verskillende reekse wat vir dieselfde standerd bedoel is. Vandaar hierdie ontleding, waarvoor geneem is 'n Standerd I. van die reeks *Lees Met Lus* deur P. I. Hoogenhout, B.A., J. P. Botha, B.A., en P. M. van der Lingen, B.A. (Uitgewers, Maskew Miller, Bpk., Kaapstad, 6de druk).

Die stof in hierdie boek is blykbaar die werk van drie skrywers, maar aangesien Ontleding No. 5 die werk van dieselfde skrywers is, kan hulle nie vir hierdie Onderzoek as afsonderlike skrywers bygetel word nie.

Die netto aantal lopende woorde is 13,159, en die woordeskat van hierdie boek is groot 1,833.

Enkele vergelykings hiervan met ander ontledings is belangwekkend. Die woordeskat is byna 600 groter as 'n Standerd I. boek van die reeks *Die Afrikanerkind* (Ontleding No. 10), maar slegs 184 kleiner as dié van 'n Standerd II. boek van hierdie reeks (*Lees Met Lus*) soos in Ontleding No. 5 aangetoon, sodat die verskil in die uitgebreidheid van die woordeskat van 'n Standerd I. van verskillende reekse skoolleesboeke groter is as dié tussen Standerd I. en Standerd II. van dieselfde reeks.

Dit toon hoe 'n groot verskil daar is tussen die gehalte van moeilikheid in boeke van verskillende reekse vir dieselfde standerd gemerk, en ook hoe nodig dit is vir skrywers van skoolboeke om op meer wetenskaplike wyse te werk te gaan in die samestelling van skoolleesboeke.

Daar is 585 veelgebruikte woorde (krediet van 3 en meer).

Wat die persentasies wat die veel- en meesgebruikte woorde van die totaal uitmaak betref, is dit opvallend dat terwyl dié vir die 585 veelgebruikte woorde (87.7%), en vir die 100 en 50 meesgebruikte woorde (64.6% en 52.9%) beide bokant die norm is, dié vir die 25 hoogste woorde (39.7%) op twee ander ontledings na, die laagste in die hele Onderzoek is, en dié vir die 10 hoogste woorde (24.9%) ook op twee ander ontledings na, die laagste is. Dit beteken dat, hoewel die skrywers nie in so 'n groot mate op die herhaalde gebruik van die eerste 10 en 25 woorde teruggeval het nie, hulle wel 'n groter gebruik gemaak het van die enigsin wyer woordeskatkring.

Daar is in hierdie ontleding 341 woorde wat vir die eerste keer teëgekome is. Vir al die verdere ontledings van hier af is nie alleen aantekening gehou van die *aantal* nuwe woorde (d.w.s. woorde wat nie tevore aangetref is nie), maar daar is ook afsonderlike *lyste* van dié woorde gemaak.

Ofskoon dit buite die bestek van hierdie Onderzoek val om 'n studie van mingebruikte woorde te maak sou dit tog interessant (en miskien nuttig) wees as ek 'n lysie gee van alle eenlettergrepige woorde of woorde met nie meer as ses letters uit hierdie lysie van nuwe woorde. Dit toon aan dat nie juis alle kort woorde veelgebruikte woorde is nie, net soos nie alle veelgebruikte woorde kort woorde is nie. Benewens hierdie lysies vir die verskillende ontledings vanaf 11 tot 16 word nog in Bylae V. 'n lys gegee van sulke mingebruikte woorde, d.w.s. wat in die 269,009 woorde wat ontleed is, 'n krediet van slegs 1 het of glad nie in die hele Onderzoek teëgekóm is nie.

Die volgende 34 eenlettergrepige woorde of woorde met nie meer as ses letters nie is vir die eerste keer in hierdie ontleding teëgekóm:—

beur	girts	knol	plaat	strand
bol	gorts	kors	plaps	suie
bom	hane	kliks	sein	ta
buie	hase	lap	sie	tata
daag	halms	leie	skaar	vore
duif	jam	miet	sports	vreug
gars	kese	peer	steil	

Hierdie woorde het meestal ook 'n baie lae verskyningskrediet en kan dus nie as veelgebruikte woorde aangemerkt word nie.

'n Lys van woorde in hierdie Ontleding (*Lees Met Lus*, Standerd I.) aangetref, maar nie die *Afrikanerkind*, Standerd I. nie, is te vinde in Bylae I (b), en van woorde wat deur Standerd II. *Lees Met Lus* (Ontleding No. 5) opgelewer word, maar nie in een van die Standerds I. is nie, word gegee in Bylae I (c).

Die volgende is 'n lys van die 100 meesgebruikte woorde in die ontleding tans onder behandeling:—

ONTLEDING 11: 100 MEESGEBRUIKTE WOORDE.

1. die	785	sy	220
2. het	410	9. ek	191
3. en	404	10. is	188
4. nie	336		
5. 'n	297		3,280
6. hy	229	11. was	184
7. hulle	220	12. dit	175

13.	te	155	word	45
14.	van	151	57. ma	43
15.	maar	147	mooi	43
16.	in	137	59. twee	38
17.	om	128	een	38
18.	vir	122	61. ja	37
19.	toe	120	62. of	34
20.	so	114	moeder	34
21.	wat	113	64. weet	33
22.	ons	107	wil	33
23.	op	104	66. soos	32
24.	daar	99	67. lekker	31
	dat	99	voor	31
			69. hoor	30
		5,235	laat	30
26.	dan	97	hier	30
27.	as	93	72. bly	29
28.	nou	88	boetie	29
29.	haar	85	tog	29
	met	85	kon	29
31.	hom	83	76. ander	28
32.	sê	82	mammie	28
33.	nog	80	wees	28
34.	net	76	kyk	28
35.	al	70	80. moes	27
36.	kom	68	staan	27
	my	68	groot	27
38.	kan	67	klein	27
39.	aan	66	84. loop	26
40.	jy	63	gou	26
	julle	63	86. sit	25
	ook	63	niks	25
43.	weer	62	hoe	25
	gaan	62	sou	25
45.	baie	60	seker	25
46.	oom	56	vra	25
47.	sal	54	92. onder	24
48.	uit	50	mens	24
49.	na	49	waar	24
	ou	49	95. af	23
			begin	23
		6,974	97. slaap	22
51.	moet	47	doen	22
	se	47	want	22
53.	sien	46	100. meer	21
	by	46		
55.	jou	45	100.	8,510

12. SIVIELE SAKE IN DIE HOOGGEREGSHOF.

In die Hooggeregshof word 'n woordelike stenografiese verslag gemaak van alles wat gesê word deur getuies en andere, en dit was met die doel om ook van enkele monsters van die *gesproke* woord ontledings te maak wat hierdie stof in die Onderzoek opgeneem is en ten einde eweneens 'n idee te vorm van die aard van die woordeskat van hierdie soort materiaal vir stenografiese werk.

Mnr. Buys, stenograaf in die Kaapse Hooggeregshof, was so vriendelik om afskrifte van die verslae van twee siviele sake, wat hiervoor gebruik is, beskikbaar te stel. Hulle is dié van 1. Breed vs. Die Holv Cross Klooster ('n saak oor brandsiek skape) wat in Maart 1930 deur regter van Zijl verhoor is, en 2. De Beer vs. van der Merwe, 'n waterdispuut (appèlsaak) wat in April 1922 voor regter Gardiner gedien het.

'n Deel van die getuienis in hierdie sake was in Engels; alleen Afrikaans (in elk geval maar 'n deel van die getuienis) is gebruik.

Soos begryplik is waar getuies, wat soms heeltemal of byna ongeletterd is ondervra word, is die gehalte van die taal nie hoog nie, en gevolglik het daar heelwat anglisismes en verbasterde woorde soos *plots*, *plotte*, *Sister* (Gertrude), *convent*, *culvert*, voorgekom. Hierdie woorde is almal geskrap en nie in die ontleding inbegrepe nie, omdat hulle g'n Afrikaanse woorde is nie. Die stof is min of meer in gelyke dele deur die twee sake bygedra, en is die uiting van elf persone altesame met inbegrip van die twee regters en die twee advokate, want daar is vrae van almal van hulle te boek gestel. Die grootste deel van die getuienis is egter gelewer deur twee persone.

Die netto aantal lopende woorde waaruit hierdie ontleding bestaan is 15,433. Die woordeskat is 1,246—die kleinste woordeskat van al die ontledings, selfs kleiner as dié van 'n eenvoudige Standaard I. skoolleesboek (*Die Afrikanerkind*—woordeskat 1,296), asook dié van Ontleding No. 3, wat uit slegs 6,679 lopende woorde, of minder as die helfte van hierdie ontleding, bestaan.

Vir die ontleding as 'n geheel geneem is elke woord gemiddeld twaalfmaal herhaal—die hoogste herhalingsverhouding van al die 16 ontledings, terwyl die herhaling vir die 492 veelgebruikte woorde (krediet van 3 en meer) 1:29 is—die vierde hoogste van al die ontledings.

Die persentasies wat die veelgebruikte woorde van die totaal gebruikte woorde uitmaak (93.7%), en dié vir die 100, 50, 25

en 10 hoogste woorde, n.l. 73·7%, 61·5%, 49·5% en 34·4% onderskeidelik, is sonder uitsondering die hoogste vir enige van die 16 ontledings verkry.

Uit al die voorgaande blyk dus dat hierdie soort stof die maklikste is wat teëgekom is. As 61% van die woordetal die gebruikmaking van slegs 50 verskillende stenografievorme beteken en oor die 49% slegs 25 vorme, spreek dit vanself dat rapporteerwerk in die hof, uit die suiwer stenografiese oogpunt beskou, nie besonder sware eise aan die snelskrywer stel nie. Gelukkig ook maar, want hy het in ander opsigte, soos hoë snelheid, slegte en onduidelike sprekers, ens., soms met moeilike toestande te kampe.

Aangesien 'n groot deel van die getuienis uit kruisverhoor bestaan, waar die advokate elke puntjie van alle kante haarfyn uitpluis, en die taal van die getuies in baie gevalle van die eenvoudigste is, kan 'n mens hierdie uitslae verwag. Ek moet desnietemin erken dat hulle vir my 'n verrassing was.

Daar is selfs in hierdie ontleding 249 nuwe woorde aangetref, w.o. die volgende 22 kort woorde:—

alla	kope	ooi	soas	vors
dip	male	ooie	spons	vrot.
grou	morg	pale	stoet	
jeuk	naaf	skets	suip	
kamp	noord	skut	swel	

Vir al hierdie woorde is die verspreiding, en in die meeste gevalle die krediete ook, baie laag. Inderdaad, met die uitsondering van twee woorde van hierdie lysie—*dip* en *ooie*—het nie een van hulle meer as 5 keer voorgekom nie in die hele 269,009 lopende woorde, en ofskoon laasgenoemde twee woorde krediete van 25 en 28 onderskeidelik verkry het, is hierdie stuk die enigste waarin hulle voorgekom het, sodat hulle as blote vakterme in verband met die boerdery beskou kan word.

ONTLEDING 12: 100 MEESGEBRUIKTE WOORDE.

1. die	1,472	8. 'n	357
2. het	755	9. is	342
3. ek	465	10. van	318
4. en	443		
5. nie	407		5,319
6. was	401	11. daar	288
7. dit	359	12. water	241

13.	in	193
14.	toe	181
15.	hulle	174
16.	u	165
17.	op	160
18.	wat	142
19.	my	138
20.	na	118
21.	ja	113
22.	hy	112
23.	by	110
24.	gemaak	97
25.	dat	96

7,647

	huis	96
	wal	96
28.	te	92
	vir	92
30.	om	88
31.	as	87
32.	se	84
33.	skape	83
34.	deur	80
35.	aan	78
36.	met	76
37.	maar	75
38.	of	71
39.	loop	70
40.	kom	68
	sy	68
42.	voor	67
43.	kan	63
	grond	63
45.	nooit	62
46.	tussen	61
47.	gesien	60
	jaar	60
49.	gaan	59
	gekom	59

9,505

	ramme	59
52.	mnr.	58
	omtrent	58
54.	brandsiekte	56
55.	sê	54

56.	net	51
	ram	51
58.	weer	48
59.	geloop	46
60.	hom	45
61.	nog	44
62.	so	43
63.	tot	41
	winkel	41
65.	sien	39
	oor	39
67.	want	38
	plaas	38
	kloof	38
70.	verkoop	36
71.	ook	35
	twee	35
	tevore	35
	pad	35
75.	ons	34
	dan	34
	weg	34
78.	uit	33
	moes	33
	maak	33
	laat	33
	af	33
	een	33
84.	kon	32
	jy	32
	dag	32
	tyd	32
	sand	32
89.	moet	31
	hamels	31
91.	kraal	30
	gewees	30
	waar	30
94.	ander	29
	huise	29
	gesê	29
97.	gekry	28
	ou	28
	mense	28
	word	28

100.

11,379

13. TYDSKRIFTE: NUUS EN PRAATJIES OOR DIE BOERDERY, DIE HUISLIKE EN GESELLIGE LEWE, ENS.

'n Aansienlike gedeelte van die alledaagse gesprekke van 'n groot deel van die bevolking, en van die stof in dag-, week- en maandblaaie word aan hierdie onderwerpe gewy, en dié soort leesstof moet dus noodwendig 'n groot deel van die alledaagse woorde uitmaak. Dit is om dié rede wat ook hiervan 'n ontleding gemaak is.

Die volgende is 'n volledige opgawe van die stukke wat gebruik is: „Kyk in Jou Spieël” (*Huisgenoot*, 10 April 1931, bladsy 69); die rubriek „Ons Geselskolom,” uit *Die Huisgenoot* van 10 April en 8 Mei 1931, paginas 71 en 69 onderskeidelik; „Blomme Wat Bloei in die Winter,” (*Huisgenoot*, 17 April 1931, pagina 67); „Gemak in die Winter” en „Hoe om Trekke te Verban,” uit die *Huisgenoot* van 8 Mei 1931, pagina 71; die rubriek „Uit die Vrouewêreld,” uit *Die Afrikaner* van 17 April 1931; 'n artikel uit *Die Huisgesin*, pagina 355; die rubriek „Boere-aangeleenthede,” in die uitgawes van 3, 17 en 21 April 1931 van *Die Afrikaner*; uit *Die Landbouweekblad* artikels uit die volgende uitgawes gedurende 1930 gepubliseer: 15 Januarie (beginnende op pagina 2255); 12 Februarie (beginnende op pagina 2533); 5 Februarie (pagina 2450); 26 Februarie (pagina 2643).

Die nuus en praatjies oor die huis het gegaan oor 'n menigte dinge in verband met die huislike lewe en die huisgesin: oor gesondheid, kinders, klere, modes, gemak, die kookkuns, rate, skoonheidsontwikkeling, verloofskappe, die verhouding tussen die geslagte, huwelike, ens., en dra hoofsaaklik die stempel van vroulikheid. Hierdie gedeelte het omtrent die helfte van die woordetal gelewer.

Die stof handelende oor boeresake (wat die ander helfte van die materiaal vorm), het ook op talryke afdelings van die boerdery betrekking. Daar is o.m. besprekings van deskundiges oor bemesting, landbou-onderwys, die mielieboerdery, marke, reëns, plaasieindomme, landbouwetgewing, grondverspoeling, besproeiing; oor wol, voorbehoedmiddels teen siektes, die hoenderboerdery, vee-siektes; daar is verslae van die Afdeling Ekonomie en Marke; 'n uittreksel uit die verslag van die bestuurder van die Landbank; 'n aankondiging van die Minister van Landbou oor bek- en klouseer.

Die taal is oor die algemeen taamlik goed, maar die lang, gekoppelde woorde, en ander aanduidings, gee aanleiding tot die mening dat 'n groot deel van die stof vertalings is. Alle opskrifte, Engelse woorde, name en syfers is geskrap of afgetrek.

Hierdie leesstof is die uiting van 'n vry groot aantal persone wat ongelukkig nie met juistheid vasgestel kan word nie. In sover dit nagegaan kan word het minstens neëntien verskillende skrywers daar 'n hand aan gehad.

Die netto aantal lopende woorde is 17,732, en die woordeskat uit hierdie ontleding verkry is 3,133—met een uitsondering, dié van Ontleding No. 15—die grootste in die hele Ondersoek vir 'n enkele ontleding. Waar dit die uiting van soveel persone is, en daar soveel kante van die boerdery en die gesellige lewe bespreek word, ten dele deur vakmanne, kon 'n mens natuurlik 'n groot woordeskat verwag, te meer aangesien die meeste persone deur wie se hande hierdie woordemassa gegaan het goed ontwikkelde mense is—landboudeskundiges en beroepskrywers (joernaliste) en vertalers. Die woordeskat is selfs groter as dié van vier verhale deur bekende Afrikaanse skrywers (Ontleding No. 4), wat uit naasteby dieselfde aantal woorde leesstof bestaan. Ook is dit groter as dié van Ontleding No. 8 (*Sewe Duiwels*) wat by die dertigduisend woorde leesstof beloop.

Daar is 615 veelgebruikte woorde (krediet van 4 en hoër) en hierdie woorde, met herhalings, maak 80·4% van die netto totaal uit—heelwat onder die norm, soos dan ook die persentasies vir die 100 en 50 meesgebruikte woorde is, n.l. 58·5% en 50·3% van die totaal onderskeidelik, terwyl die persentasies wat die 25 en 10 meesgebruikte woorde met hulle herhalings van die totaal uitmaak—40·6% en 26·9%—'n ietsie bokant die verwagte norm is.

Hierdie ontleding het die groot aantal van 1,104 nuwe woorde opgelewer, ondanks die feit dat daar reeds die woordeskat van 12 ontledings, belopende oor die 200,000 woorde leesstof, byeenversamel was. Onder hierdie 1,104 nuwe woorde is daar baie lang, aaneengekoppelde woorde, maar ook die volgende 60 kort woorde word in hierdie lys aangetref:—

bale	hals	o.m.	skaf	tef
bleik	herfs	plak	skans	tenks
boud	heup	pels	slaai	trag
drag	hooi	pere	skeer	vaas
dm.	jaart	pint	snit	vase
doods	jig	poef	soda	velt
dorp	klae	reik	soep	verg
drang	knip	repe	soom	vloot
eg	kuit	reuk	spoeg	weeg
eien	leed	rog	staal	wes
fles	lomp	rûe	streef	w.o.
gom	nate	rye	teel	wyne

Die volgende is die 100 meesgebruikte woorde wat in hierdie leesstof gebruik is (58·5% van die totaal lopende woorde):—

ONTLEDING 13: 100 MEESGEBRUIKTE WOORDE.

1. die	1,264	ek	57
2. en	587	43. nou	56
3. van	482	44. jou	54
4. is	452	45. al	52
5. 'n	418	46. wees	51
6. nie	406	47. by	50
7. in	339	48. net	46
8. te	288	49. gebruik	45
9. dit	287	was	45
10. ons	262		
	4,785		8,936
11. word	240	na	45
12. wat	239	52. groot	42
13. om	196	53. boere	40
14. vir	183	laat	40
15. as	167	55. oor	39
16. op	163	haar	39
17. het	162	57. kom	38
18. dat	153	jaar	38
19. kan	145	59. soos	37
20. of	140	mooi	37
21. met	139	61. uit	36
22. maar	138	62. hy	35
23. daar	128	goeie	35
24. hulle	124	64. vel	34
25. baie	114	meer	34
		66. goed	33
	7,216	67. onder	31
26. sal	111	my	31
27. moet	104	69. hou	29
28. sy	100	70. elke	28
29. so	97	doen	28
30. ook	86	72. mens	27
31. dan	84	tog	27
32. hierdie	83	enige	27
33. jy	81	wil	27
34. aan	80	76. dieselfde	26
35. deur	77	gaan	26
36. maak	66	kry	26
37. ander	62	79. tyd	25
38. nog	60	waar	25
39. want	58	81. iets	24
een	58	weer	24
41. tot	57	boer	24
		eiers	24

	geval	24	94.	sê	22
86.	teen	23		dus	22
	beter	23		sien	22
	geen	23		altyd	22
	af	23	98.	mense	21
	lewe	23		se	21
	alle	23		klein	21
	grond	23			
	water	23	100.		10,376

14. 'N TONEELSTUK.

Die taal van 'n toneelstuk is naasteby die gewone daaglikse, persoonlike omgangstaal, of behoort dit altans te wees. Dit is om dié rede wat 'n toneelstuk in die Onderzoek opgeneem is.

Vir hierdie ontleding is gekies die bekende *Lenie* deur H. A. Fagan (Uitgewers, Nasionale Pers, Beperk, derde duisend). Al die woorde in die boek wat op die spel betrekking het is gebruik met inbegrip van die toeligtings van die skrywer. Alle name is geskrap, en ook is as name behandel woorde soos *tante*, ens., waar hierdie woorde in plaas van name gebruik word, b.v. „*Tante: Oppas tog vir die ruite, kind.*” „*Tante lag.*” Ook *mnr.* is geskrap waar dit voorkom om die naam van die spreker aan te dui en tog nie deel van die gesprekke self vorm nie, b.v. „*Mnr. Brand: So, jy het darem 'n opinie. Wat is die opinie nogal?*” In sulke gevalle is beide die *mnr.* en sy naam geskrap.

Hierdie stuk word geneem as die werk van een skrywer.

Die boek bevat 18,509 woorde leesstof (na die aftrek van bogemelde skrappings) en daar is 'n mistelling (teveel) van slegs 1 vir hierdie ontleding gewees. Die woordeskat is groot 2,214.

Enkele vergelykings met ander stukke op hierdie punt is nie van belang ontbloot nie. Ontleding No. 15 (sien pagina 54) bestaan uit omtrent 500 minder woorde leesstof, maar sy woordeskat is meer as 'n duisend groter. Ontleding No. 13 (praatjies oor die huis en die boerdery) bestaan uit omtrent 800 minder lopende woorde as hierdie een en toon 'n woordeskat van 919 groter as dié van hierdie stuk, terwyl die versekeringsbriewe (Ontleding No. 6, 21,037 lopende woorde) 'n woordeskat van 2,324 aantoon.

Alle woorde in hierdie ontleding met 'n krediet van minstens 4 word as veelgebruikte woorde beskou. Die resultaat toon 514 sodanige woorde, en hulle maak, met herhalings, 16,123, of 87.1%, van die totaal gebruikte woorde uit, terwyl die 100 meesgebruikte woorde 66.6% van die totaal vorm, die 50 woorde 24.4%, en die 25 hoogste woorde 41.0%. Hierdie persentasies

is almal bokant die verwagte norm; die drie eersgenoemde heelwat, en die laasgenoemde een nie veel nie. Opvallend is dit egter dat die persentasie vir die 10 hoogste woorde in hierdie geval slegs 24.9% is. Daar is slegs twee ander ontledings in die hele Onderzoek wat 'n ietwat laer persentasie vir hierdie groepie woorde aantoon.

In hierdie ontleding is 565 nuwe woorde aangetref. Baie hiervan is ongewone en moeilike woorde soos b.v.—om maar enkele te noem—*perorasie*, *pierrette*, *pierretteklere*, *pierrettekostuum*, *plaasvervangster*, *plesierjag*, *mouklossie*, *redenering*, *ryswepie*, *toekomsprobleem*, *toilettafel*, *teësparteling*, *saambaalbroek*, *syleuning*, *verstandsvermoë*, ens., ens., maar in hierdie lys van 565 nuwe woorde verskyn ook die volgende 32 kort woordjies:—

afsê	klag	opry	rak	stik
bowe	klug	orig	ramp	stomp
deus	lote	pag	skrams	tok
dui	moois	prag	sloof	tug
gaat	naai	raam	slyt	vink
gul	naand	raat	sport	wae
kiet	noot			

Die 100 meesgebruikte woorde (wat 66.6% van die totaal uitmaak) is:—

ONTLEDING 14: 100 MEESGEBRUIKTE WOORDE.

1. die	847	20. dat	170
2. ek	623	21. haar	169
3. nie	583	22. nou	164
4. en	452	23. so	160
5. het	391	24. hy	152
6. 'n	383	op	152
7. jy	375		
8. dit	354		7,607
9. my	318	26. dan	142
10. maar	296	27. kan	140
		28. kom	136
	4,622	29. sal	128
11. is	289	30. aan	124
12. te	276	31. as	123
13. om	233	32. nog	114
14. sy	221	33. dis	112
15. jou	217	34. ook	108
16. in	212	met	108
17. vir	204	36. by	94
18. wat	190	37. al	93
19. van	176	38. moet	91

39.	gaan	90	70.	vader	48
	uit	90		oor	48
41.	ons	87	72.	kind	47
42.	ja	86		toe	47
43.	dare	85		na	47
44.	wil	84	75.	hoe	46
45.	mos	82	76.	goed	43
46.	tog	75		bly	43
47.	net	74	78.	praat	42
48.	sit	70	79.	nee	41
49.	sê	69	80.	neem	40
50.	laat	67		voor	40
				sien	40
		10,079	83.	moeder	39
51.	self	65	84.	hou	38
	deur	65	85.	venster	37
53.	gang	64		hom	37
54.	weer	62		baie	37
55.	wees	61		tyd	37
56.	maak	60	89.	daardie	36
57.	daar	59	90.	soos	35
58.	was	57		dink	35
59.	sou	56		man	35
	weet	56	93.	af	34
61.	staan	55		julle	34
	O!	55	95.	hier	33
63.	of	54	96.	iets	32
64.	ag	53		meer	32
65.	se	50	98.	somaar	30
	ou	50		stap	30
	word	50	100.	waar	29
	bietjie	50			
	kyk	50	100.		12,343

Soos 'n mens kan verwag in hierdie soort lektuur kom daar etlike voornaamwoorde voor in die lys. Sewe daarvan—*ek, jy, my, jou, sy, haar, hy*—val binne die hoogste 25 woorde. Die enigste ander woorde hierin wat nie van algemene toepassing is nie, en wat 'n mens kan verwag om 'n goeie herhaling van in 'n werk van hierdie aard te vind is *gang, vader, ag, kind, moeder* en *venster*, hoewel die meeste van hulle ook goeie krediete in ander ontledings het.

15. DAGBLAD: BINNELANDSE EN BUITELANDSE NUUSBERIGTE.

Daar is deesdae seker min soorte lektuur wat, in die algemeen gesproke, meer gelees word as die koerant. Sonder die nuusblad

voel die gewone mens hom afgesny van die res van die grote wêreld—'n soort Robinson Crusoe.

Die dagblad behoort 'n getroue weerspieëling te gee van die daaglikse doen en late en dinke van die woelende en strewende mensheid in ons land en deur die hele wêreld heen; dit handel oor alles en nog wat. Dit was dus vanselfsprekend dat van hierdie lektuur ook 'n goeie monster ontleed moes word bowe en behalwe die hoofartikels (Ontleding No. 3), en leesstof van huislike en gesellige aangeleenthede en die boerdery (Ontleding No. 13), wat reeds in die Onderzoek inbegrepe was. En om te verseker dat die ontleding goed verteenwoordigend sal wees van die algemene nuusstof van die daaglikse koerant, is buitelandse sowel as binnelandse nuusberigte geneem en dit uit enkele koerante strekkende oor 'n tydperk van omtrent 16 maande. Die buitelandse en binnelandse nuusberigte het elkeen omtrent die helfte van die ontlede stof bygedra.

Alleen deur sinne aaneengeskakelde stof is gebruik; alle hofies bo-aan en in 'die kolomme, asook alle losstaande woorde, is geskrap en, natuurlik, ook name en syfers.

Die binnelandse nuus is ontleen aan *Die Burger* van 9 Februarie 1930 (pagina 8) en dié van 16 Mei 1930 (pagina 11) en sluit o.m. in die volgende: stukke, berigte en verslae in die rubriek „Van Alle Kante”; verslae van die saak teen die konstabels te Wellington, van Kandalie in die hof en 'n gesellige aand deur die Nasionale Party op Woodstock, waarop enkele voor-aanstaande Nasionaliste toesprake gehou het; berigte omtrent 'n geveg met 'n tier, ontslag van onderwysers in die Vrystaat, 'n krieketwedstryd; 'n bespreking van die behuisingsvraagstuk, die draadlose uitsaaidiens, verkiesingsberigte, en enkele ander kort beriggies oor allerhande sake.

Die buitelandse nuus is uit die volgende uitgawes van *Die Burger*: 7 Maart, 5 Junie, 18 Julie en 13 Augustus (almal van 1930); 26 en 30 Januarie, 6, 13 en 18 Mei, en 3 en 4 Junie, 1931. Die meeste items is Reuter-berigte wat uit die Engels vertaal en die meeste waarvan terselfdertyd baie verkort is, vandaar dat die stof uit soveel uitgawes van die blad geput moes word. Hulle bestaan uit 'n groot aantal verskillende beriggies, te veel om almal te klassifiseer; daar is 'n hele klomp in verband met Gandhi en die moeilikhede in Indië, 'n taamlke lang verslag in verband met die saak teen lord Kysant; ander berigte sluit o.m. die volgende in: 'n gedenkteken in Pruise onthul, Amy Johnson in Australië onthaal, siekte van prinses Beatrice, mynwerkers verongeluk, ontploffing op 'n boot, vlotebeperking, geweldplegings deur kommuniste, onluste in Egipte, 'n oud-minister beboet, S.A. boerereis

in Engeland, snelheidsrekord op die water, vliegtuigongelukke, 'n swaaibrug beswyk, Britse handel in Suid-Amerika, nuwe grondwet vir China, die omwenteling in Spanje, die wêreld se ekonomiese moeilikhede, gelukwense aan die koning.

Hierdie is maar enkele om 'n aanduiding van die leesstof te gee, en is tipies van die nuusberigte van enige dagblad.

Die netto aantal lopende woorde is 18,018, en die woordeskat in hierdie ontleding, wat slegs die vyfde grootste van al die ontledings in die Onderzoek is, is die grootste van almal n.l. 3,271. Die taal van ons nuusblaaie, oor die algemeen geneem, is so goed soos die meeste Afrikaanse lektuur; dit is dan ook meestal die werk van beroepskrywers (joernaliste); maar ofskoon daar lang woorde in die lys van moeilike woorde aangetref word, soos b.v. *akkompagnement*, *arbeidsfederasie*, *admiraliteit*, *aktuarisse*, *bereidwilligheid*, *buitensporighede*, *handelsmededinging*, *Jesuietklooster*, *konseratorium*, *gewetensvryheid*, *naturellereserwes*, *passasiers-vliegtuig*—om maar enkele te noem—is die woordeskat m.i. nie groot nie soveel weens die moeilikheid van die stof as weens die feit, dat dit bestaan uit 'n menigte kort berigte oor allerhande sake oor die ganse wêreld.

Ons het reeds gesien dat waar die stof alles oor een saak handel of in een rigting loop, die woordeskat betreklik klein is. Voorbeelde hiervan is Ontleding No. 6 (Versekeringsbriewe) en No. 8 (*Sewe Duiwels*). Laasgenoemde bestaan uit slegs 2½ “Duiwels”—die werk van 'n enkele skrywer, en ofskoon dit verreweg die grootste van die ontledings is, lewer dit 'n woordeskat van slegs 2,639 d.w.s. die vierde grootste. Die versekeringsbriewe is die derde grootste ontleding en lewer 'n woordeskat van 2,324, of die sesde grootste van die woordeskatte, 'n woordeskat wat selfs kleiner is as dié van 'n vir Standaard III. bedoelde leesboek.

Hieruit blyk dus dat dit 'n groot woordeskat vereis om al die berigte in 'n dagblad met verstand te volg, en dat hierdie stof uitstekend is vir die verruiming van 'n mens se woordeskat.

Dit is jammer dat dit onmoontlik is om die aantal skrywers wat tot hierdie materiaal bygedra het met juistheid te bepaal. Die buitelandse nuus is almal Reuter-berigte, en is deur Reuter-korrespondente uit 'n menigte plekke in alle dele van die wêreld na Londen deurgestuur (alles natuurlik in Engels). Daar is dit waarskynlik geredigeer en klaargemaak vir deursending na Suid-Afrika, waar dit nogeens bewerk is (altyd nog in Engels) vir aflewering aan die koerante, waar dit weer vertaal en in Afrikaans bewerk is deur verskillende lede van die redaksiepersoneel.

Van die binnelandse nuus is daar blykbaar nie soveel vertaalwerk nie. Die meeste verslae is gemaak deur die blad se verslaggewers, en die berigte van die blad se eie (Afrikaansskrywende) beriggewers, terwyl die rubriek „Van Alle Kante,” as ek dit nie mis het nie, onder die redaksie staan van 'n vooraanstaande lid van die personeel, hoewel soms ook 'n skrywe van lede van die algemene publiek daarin 'n plek vind. In hierdie deel van die stof het Reuter dus blykbaar g'n handjie gehad nie. Die aantal medewerkers vir die woordemateriaal van hierdie ontleding kan dus nie vasgestel word nie. Seker is dit egter dat hulle nie minder as, sê, vyftien was nie.

In hierdie stuk is 665 woorde met 'n verskyningskrediet van minstens 4. Hierdie 665 lewer 79.5% van die totaal van 18,018 lopende woorde. Dié persentasie is heelwat onder die norm, wat geen verbasing hoef te wek nie met die oog op die groot woordeskat vir 'n nie alte groot aantal lopende woorde nie—op een na, dié vir hoofartikels—die laagste in die Ondersoek verkry, terwyl die persentasie vir die 100 meesgebruikte woorde—58.1%—die laagste is vir enige ontleding in hierdie groep woorde verkry.

Met die oog op die voorgaande is dit wel merkwaardig dat die persentasies wat die 25 en 10 hoogste woorde van die totaal vorm, hoër is as in die meeste ontledings aangetref n.l. 44.4% en 32.4% onderskeidelik—'n ietsie hoër as dié vir hoofartikels (Ontleding No. 3) verkry. Hierdie benaderende ooreenstemming toon in hoe 'n hoë mate beroepskrywers gebruik maak van die binneste kern van die taal. Eersgenoemde persentasie is die vyfde hoogste in enige van die 16 ontledings teëgekom en laasgenoemde die tweede hoogste. Dit is dus van besondere belang vir koerantmanne en lettersetters om te weet watter hierdie 10 en 25 meesgebruikte woorde is. Hier is hulle—die eerste 10 en 25 van die lys van die 100 meesgebruikte woorde:—

ONTLEDING 15: 100 MEESGEBRUIKTE WOORDE.

1. die	1,872	11. wat	236
2. het	668	12. op	230
3. van	727	13. om	225
4. in	473	14. hy	194
5. is	442	15. met	164
6. 'n	441	16. word	156
7. en	418	17. sal	150
8. dat	298	18. aan	134
9. nie	258	19. sy	108
10. te	247	20. vir	105
		21. was	100
		na	100
	5,844		

23.	daar	93	nou	27
24.	dit	91	62. weer	25
25.	deur	84	waar	25
			Dinsdag	25
		8,014	65. getuie	24
26.	as	82	alle	24
27.	hulle	79	verband	24
28.	maar	76	68. kom	22
29.	ons	70	eerste	22
30.	my	68	ten	22
31.	ek	67	voor	22
32.	tot	66	72. daardie	21
33.	by	63	geld	21
34.	uit	62	al	21
35.	toe	60	gister	21
36.	nog	53	of	21
	ook	53	spr.	21
38.	mnr.	49	78. drie	20
	oor	49	gesluit	20
40.	hierdie	48	ander	20
41.	verklaar	47	egter	20
42.	kan	43	tussen	20
43.	se	42	onder	20
44.	groot	41	84. nuwe	19
45.	regering	40	gaan	19
46.	hom	39	waarin	19
47.	wees	38	minister	19
48.	moet	37	88. wil	18
	geen	37	uur	18
50.	baie	36	verslag	18
			haar	18
		9,359	92. persone	17
51.	gesê	35	saak	17
52.	twee	34	soos	17
53.	teen	33	toestand	17
	een	33	gehou	17
55.	so	30	gegee	17
	vandag	30	motor	17
57.	jaar	29	verdere	17
58.	geneem	28	gekom	17
	meer	28		
60.	hul	27	100.	10,475

Soos 'n mens uit die persentasies kan aflei is die hoogste 25 almal woorde van algemene toepassing, terwyl daar in die hele 100 alleen die volgende tien as vakterme vir koerantelektuur beskryf kan word, t.w. *mnr.*, *verklaar*, *regering*, *getuie*, *verband*,

spr., *minister*, *verslag*, *motor*, *persone*; hoewel enkele van die woorde soos b.v. *mnr.*, *verband* en *persone* ook dikwels in ander ontledings verskyn, veral in briewe.

Daar is 1,073 nuwe woorde in hierdie ontleding aangetref—30 minder as die aantal nuwe woorde deur Ontleding 13 opgelewer. Hierdie getal sou seker oneindig groter gewees het as die ontleding vroeër in die Onderzoek geval het, en nie deur Ontleding No. 13 voorafgegaan is nie.

Hierdie 1,073 nuwe woorde sluit die volgende 50 kort woorde in:—

aas	etse	loods	prof.	Skot
art.	get.	lord	prok.	skroei
a.s.	gewy	lyke	rype	slaags
besk.	hits	mnre.	sale	spil
bar	jaarts	Nas.	S.A.P.	spr.
b.v.	jute	n.m.	sel	stig
demp	kapt.	noue	sen.	veto
duld	klub	Pole	sersj.	v.m.
eerw.	konst.	pous	sinds	wrak
e.k.	laai	pres.	skors	wurg

'n Aantal hiervan is verkorting, maar hulle word gegee omdat hulle afsonderlik gehou is in die vorm waarin hulle aangetref is.

16. PERSOONLIKE EN BESIGHEIDSBRIEWE.

Die materiaal van hierdie ontleding—briewe—beskou ek van die allergrootste belang vir die Onderzoek. Daar is seker geen beskaafde mens wat nie briewe lees en briewe skryf nie; hoewel daar 'n menigte mense is wat nie artikels, opstelle en ander stukke skryf nie, is briewe 'n skryfvorm wat algemeen beoefen word deur elke gewone persoon. In die handel- en sakewêreld is briewe natuurlik 'n noodsaaklikheid. Briewe raak die alledaagse lewe op elke punt, en private briewe behoort 'n meer persoonlike en intieme woordeskat te lewer as byna enige ander soort lektuur.

Dit was oorspronklik my voorneme om aparte ontledings te maak van besigheidsbriewe en van persoonlike briewe, maar in die loop van die Onderzoek het dit reeds baie groter geword as ek in eerste instansie bedoel het, deurdat ek dit wenslik gevind het om van tyd tot tyd nuwe stukke in te sluit, sodat ek dit raadsaam geag het om hierdie twee saam te smelt. Aangesien die hoofdoel van die Onderzoek is om die Duisend woorde van die mees algemene gebruik in die alledaagse lewe vas te stel, het ek gemeen dat daar geen besondere beswaar hierteen kon wees nie, ofskoon ek erken dat dit wensliker sou gewees het om hierdie twee afsonderlik te doen. By 'n onderneming van so 'n omvang vind 'n mens egter uit dat daar perke aan jou tyd, kragte en geduld is.

Hierdie ontleding is uitgestel tot die laaste om my meer tyd te gee vir die versamel van briewe, want ek wou g'n kunsmatig opgestelde briewe hê nie, maar werklike briewe wat deur persone, sake en inrigtings uitgestuur en ontvang is. Wat meer is, ek wou graag hiervoor briewe gebruik wat in Afrikaans *geskryf is*—nie vertalings nie. Die magtige Engelse taal dring reeds van alle kante op ons aan, en veral in die sakewêreld voer dit nog die bowetoon, nie alleen in ons land nie, maar ook in die internasionale sake-verkeer. Buitendien wou ek graag briewe van soveel verskillende klasse besigheid moontlik in die ontleding opneem.

Soos ek reeds vroeër in hierdie werk gemeld het, word alle pogings om besigheidsbriewe te bekom deur 'n netwerk van onbereidwilligheid, wantroue en geheimsinnigheid versper, en wat dié klas briewe betref is dit my nie geluk om bogenoemde ideaal heeltetal te verwesenlik nie. Wel was ek instaat om self enkele briewe van besighede bymekaar te maak; dan het ek 'n klompie gekry van 'n aptekersaak in Kaapstad wat deur 'n vriend bestuur word, maar dié was almal in Engels en moes deur my vertaal word; verder was ek gelukkig genoeg om, deur die vriendelike bemiddeling van mnr. M. S. Louw van S.A.N.L.A.M., insage te kry van briewe van 'n algemene handelaar se besigheid in Noord-westelike Kaapland. Dit alles was egter lank nie genoeg nie, want hierdie ontleding tel altesame 23,318 lopende woorde netto. Die enigste uitweg dus was om gebruik te maak van 'n aantal briewe uit 'n brieweboek. Dit was baie teen my sin, maar ek moes dit noodgedwonge doen by gebrek aan beter.

Die volgende is 'n nadere uiteensetting van die stof vir hierdie ontleding gebruik: Die geheel bestaan uit 146 briewe wat deur omtrent 82 skrywers opgestel is oor die tydperk 1927–1931, en beslaan 23,349 woorde gros, sodat die gemiddelde aantal woorde per brief 159 is.

Van die 146 is 40 private briewe, die werk van 32 skrywers (omtrent 15,300 woorde) en 106 besigheidsbriewe, opgestel deur omtrent 50 persone, en belopende circa 8,000 woorde leesstof.

Die materiaal van die private briewe is ontleen aan briewe wat ek self van familiebetrekkinge en vriende ontvang het, aan afskrifte van enkele van my eie briewe aan andere, en aan die briewe wat goedgunstiglik vir my deur 'n paar vriende en vriendinne van hulle eie ontvangte briewe op ruitjespapier afgeskryf en versnipper is. Dit is die gewone briewe tussen familie en vriende, en sommige daarvan is, volgens my verseker is, van 'n hoogs persoonlike en private aard. Die meeste skrywers van hierdie briewe is ontwikkelde mense, party van hulle persone met universiteitsgrade; 22 van die opstellers van die briewe is lede van die skoon geslag en 10 is mans. Omtrent 5,000 woorde van

hierdie deel van die ontleding is uit die briewe van een man. Die skrywers en skryfsters is persone van verskillende ouderdomme, en die briewe afkomstig uit verskillende dele van Suid-Afrika. Hierdie deel van die stof beskou ek as taamlik bevredigend.

Baie onbevredigend egter is die stof vir besigheidsbriewe, wat slegs ongeveer 8,000 woorde bygedra het. Dit word opgemaak deur 40 briewe ontvang deur die Kommissiewinkels van die Arbeidskolonie te Kakamas, waarvan ek, deur die welwillende toestemming van mnr. W. F. van Riet, die algemene bestuurder, insage gehad het en op die ruitjespapier kon afskryf, benewens 29 briewe deur die bestuurder geskryf. Verder, 'n brief van die Klerk van die Volksraad ontvang, 'n brief (vertaling) deur 'n landbougenootskap, 'n ander van 'n kleremaker te Kaapstad (vertaling) en 'n vierde van die Boeresaaamwerk, Bpk., te Port Elizabeth aan 'n boer. Buitendien is daarin inbegrepe sewe briewe deur 'n Kaapstadse apteker ontvang en uitgestuur.

Dit is al die praktiese briewe—briewe wat werklik deur sake uitgestuur en ontvang is. Dit was egter nog nie genoeg nie, en daar is dus, om 'n groter aantal lopende woorde te kry, die volgende briewe geneem uit *Juta se Briefsteller* (uitgawe van 1922): Nos. 29, 35, 36, 37, 38, 39, 40, 42, 43, 44, 45, 46, 47, 50, 51, 52, 54, 55, 59, 60, 62, 64, 65, 81, 95 en 96.

Hierdie briewe sluit in bestellings, briewe van handelsreisigers, briewe in verband met rekenings, verdragings, wissels, krediethriewe; betreffende verskepings, betrekkings, eise, aksynsopgawes, ens. Die briewe verskyn in beide offisiële tale in die boek, en is vermoedelik vertalings uit die Engels.

Die materiaal van besigheidsbriewe is oor die algemeen baie onbevredigend. Daar is te veel vertalings, en die "boekbriewe" is ongewens. Die briewe van die Kommissiewinkels is nog die beste, maar daar is te veel van hulle van dieselfde trant en nie genoeg verskeidenheid nie. Dit is egter die beste wat ek bymekaar kon sukkel, en dit is my plan om nog 'n beter ontleding of ontledings van briewe te maak.

Al die stof in die briewe is gebruik, insluitende die aanhef en die slot, maar nie die name van plekke en persone en reekse van prysopgawes nie wat in besigheidsbriewe nogal dikwels voorgekom het.

Die woordeskat uit hierdie ontleding verkry is 2,711. Gesien dit die tweede grootste van al die ontledings is, en dit besigheids- en persoonlike briewe insluit, is hierdie woordeskat, ofskoon hoog, nie buitensporig groot nie. Dit is 72 groter as die woordeskat van die grootste ontleding (No. 8—29,349 lopende woorde) en 387 groter as dié van Ontleding No. 6 (versekeringsbriewe—21,037

lopende woorde), maar heelwat minder as dié verkry uit die kleiner ontledings Nos. 4 (kort verhale), 13 of 15 (dagblaai en tydskrifte).

Die 534 veelgebruikte woorde (krediet van 5 en meer) is verantwoordelik vir 19,925 (of 85.4%) van die 23,318 lopende woorde waaruit die ontleding bestaan—net bokant die verwagte norm, maar nie die hoogste wat verkry is nie. Die persentasies wat die drie groepe—die 100, 50 en 25 meesgebruikte woorde van dit totaal uitmaak—64.9%, 54.7% en 42.9% onderskeidelik, is goed bokant die norm, terwyl die 10-woorde groep, met 25.9%, net omtrent op die norm is.

Daar is 'n woordeskat van 2,177 mingebruikte woorde wat, met die lae herhalingsverhouding van 1:1.5 slegs 3,393 van die totaal van 23,318 lopende woorde lewer.

Tot aan die end van Stuk 15 is omtrent 250,000 gros woorde leesstof van allerlei soorte ontleed, en 'n mens kon redelikerwyse verwag dat jy geen of baie weinig nuwe woorde sou teëkom, veral vir betreklik maklike stof soos hierdie. Dit was dus vir my 'n openbaring om nog 'n lys van 591 nuwe woorde hierin te verkry, en daarby nogal sowat 'n vyftigtal heel kort woorde. Hier is 'n paar van die langer woorde van die lys: *Aandeelhoudersvergadering, akseptasiewissel, allerinteressantste, beliggaming, bewonderingswaardig, daardieselfde, eetaartappels, formaliteite, gespesifiseerde, gewaarwording, grootwildjagters, hanepootrosyne, hulpvaardigheid, karaktertrekke, kommissiewinkels, oorbôgem, openhartigheid, polisiekonstabels, reikhalzend, rokkie, seergemaak, sinikus, sielsgeheime, slimmigheidjies, suinig, spoorwegadministrasie, swaargekry, vergenoegdheid, voetgangersprinkane, waarlik-waar, wrintigwaar*. Maar genoeg hiervan. Ook die volgende klompie kort woorde en verkortings is vir die eerste keer in hierdie ontleding aangetref:—

a.u.b.	Feb.	keus	part	tere
Aug.	films	ko.	P.S.	toer
baal	gal	lêer	room	tolk
bpk.	grof	mat	sig	tronk
bone	gros	mere	slap	uie
boot	hore	Nov.	std.	vele
dnr.	Jan.	N.S.	stg.	vies
dons	Janu.	Okt.	straks	vlek
dos.	j.l.	olik	string	vlyt
ds.		oulik	stuks	v.o.s.
dw.				wê

Die volgende is 'n lys van die 100 meesgebruikte woorde van hierdie ontleding.

ONTLEDING 16: 100 MEESGEBRUIKTE WOORDE.

1. die	1,098	42. sy	93
2. ek	819	43. hy	91
3. en	636	na	91
4. nie	608	45. wil	87
5. het	563	uwe	87
6. is	490	47. net	86
7. van	469	48. by	85
8. ons	455	49. heer	84
9. 'n	453	50. dan	78
10. u	451		
			12,771
	6,042	51. weer	76
11. my	406	52. goed	75
12. dit	378	53. waarde	70
13. te	364	54. ou	68
14. vir	329	55. hier	66
15. in	277	weet	66
om	277	57. sê	65
17. sal	248	58. want	64
18. jy	243	59. laat	63
jou	243	60. kom	60
20. maar	228	61. hulle	58
21. dat	222	62. bly	57
22. so	200	toe	57
23. op	192	64. sien	55
wat	192	65. mnr.	54
25. met	177	66. een	53
		67. soos	51
	10,018	kry	51
26. baie	167	haar	51
27. nou	164	70. word	50
28. as	148	71. meel	48
29. was	143	per	48
30. kan	134	73. tog	47
31. ook	122	74. maak	46
32. aan	119	75. oor	44
33. of	118	gee	44
gaan	118	77. beste	43
35. nog	115	ontvang	43
36. wees	112	79. hierdie	42
37. al	105	doen	42
brief	105	81. sak	41
39. daar	104	tot	41
40. moet	100	83. hoop	40
41. stuur	97	84. bestuurder	39

	twee	39		dink	33
	pond	39	95.	hom	32
87.	skryf	37		seker	32
88.	se	36	97.	laaste	31
	tyd	36		paar	31
	hoe	36		groot	31
91.	darem	34		sou	31
92.	dank	33	—		
	deur	33	100.		15,133

Hierdie is die laaste ontleding van die Onderzoek en bring die totaal op 273,191 woorde gros, en 269,009 lopende woorde netto, d.w.s. na die aftrek van name, ens.

V. Opsomming.

Ek beskou hierdie ontleding van oor 'n kwartmiljoen woorde leesstof—uit sestien bronne geput, van allerlei aard, die werk van minstens 220 skrywers van beide sekse van verskillende ouderdom en uit verskillende dele van die land, persone van verskillende grade van geleerdheid, aanleg en beroep, en lopende oor 'n lang tydperk—voldoende vir die vasstelling van die Duisend meesgebruikte woorde, as die vasstelling van so 'n aantal woorde van algemene gebruik wel moontlik is.

Soos reeds vroeër gekonstateer, is daar 'n klein foutjie of verskil in elke ontleding as gevolg van mistellings of misrekenings, of beide. Die volgende staat toon die verskille aan en die benaderende persentasie van die fout in elke ontleding.

FOUTE IN DIE WERK.

Ontleding No.	Totaal ruitjies volgens berekening voor opsnippering.	Gros totaal papiertjies van tellings en optellings verkry.	FOUT.		Persentasie van Fout.
			Teveel.	Tekort.	
1.	10,266	10,414	148	—	1·4%
2.	20,387	20,292	—	95	·46%
3.	6,732	6,721	—	11	·16%
4.	17,040	17,106	66	—	·38%
5.	14,576	14,616	40	—	·27%
6.	21,198	21,176	—	22	·10%
7.	16,938	16,924	—	14	·08%
8.	30,672	30,786	114	—	·36%
9.	15,942	16,041	99	—	·62%
10.	12,214	12,202	—	12	·09%
11.	13,399	13,391	—	8	·06%
12.	15,427	15,440	13	—	·08%
13.	17,752	17,736	—	16	·09%
14.	18,513	18,514	1	—	·005%
15.	18,488	18,477	—	11	·06%
16.	23,349	23,355	6	—	·02%
	272,893	273,191	487	189	

Die grootste fout—1·4%—is in die eerste ontleding begaan, maar ondervinding in die werk het blykbaar 'n heilsame uitwerking gehad, want in die meeste verdere ontledings is die fout baie kleiner, in almal—met die uitsondering van Ontleding 9, waar dit 62% is—onder die 5%. Op al die ontledings is daar 'n gesamentlike te veel van 487 en 'n tekort van 189. As ons sou aanneem dat die te veel deels deur die tekort uitgewis word, bring dit die fout op 'n netto te veel van 298 vir die 273,191 lopende woorde gros. Dit is 10% of 'n mistelling van 1 in elke duisend woorde. Dit is so onbeduidend klein dat dit feitlik niks afdoen van die wetenskaplike waarde van die werk nie.

KLASSE-INDELING VAN DIE STOF.

Hoewel die stof uit sestien ontledings bestaan, is sommige van hulle materiaal van dieselfde aard, en dit kan min of meer in die volgende ses klasse ingedeel word:—

Klas.	Ontleding Nos.	Netto Totaal lopende Woorde.	Persentasie van Netto Totaal.
1. Kinderlektuur ..	1, 2, 5, 10, 11	69,607	circa 26%
2. Maatskaplik ..	4, 8, 14	64,416	„ 24%
3. Briewe	6, 16	44,355	„ 16%
4. Dagblaai en Tydskrifte ..	3, 13, 15	42,429	„ 16%
5. Algemeen saaklik	7, 12	32,284	„ 12%
6. Godsdienstig ..	9	15,918	„ 6%
6 klasse.	16 ontledings.	269,009	100%

Hierdie is maar 'n taamlike los indeling; die stof van sommige van die ontledings slaan gedeeltelik met dié van ander oormekaar. So kon b.v. 'n deel van klas 4 (dagblaai en tydskrifte), veral dié van Ontledings Nos. 3 en 15, onder klas 5—algemeen saaklik—inbegrepe word; ook die stof van besigheidsbriewe kon onder hierdie klas gekom het, terwyl persoonlike briewe onder die klas „maatskaplik” ingedeel kon word, en selfs kinderlektuur of onder „maatskaplik” of „algemeen saaklik” kon val. Maar hierdie lyk my die beste indeling om die woorde-families bymekaar te hou.

Ook wil ek nie beweer dat die hoeveelheid van elke klas presies die belangrikheid van dié klas in die gewone alledaagse

lewe weerspieël nie; die hoeveelheid is in sekere mate willekeurig, hoewel ek meen dat, met die uitsondering van kinderlektuur en briewe, die hoeveelhede by benadering in die regte verhouding is. Wat betref kinderleesstof, die hoeveelheid lopende woorde daarvoor is so groot omdat twee ontledings van 'n standerd I. boek gemaak is, om vergelykingsdoeleindes; maar dit, nòg enige ander klas, kan die keuse van die 1,000 meesgebruikte woorde onbehoorlik beïnvloed, aangesien die krediet van 'n enkel ontleding of 'n enkel klas tog beperk word by die keuse, soos verduidelik word op bladsy 75. Van briewe sou ek graag 'n baie groter hoeveelheid, met 'n baie groter verskeidenheid, wou ingesluit het, maar kon hierdie begeerte nie verwesenlik sien nie weens die moeilikheid om dié soort lektuur te bekom.

STAAT VAN DIE UITSLAE.

Om vergelyking van die uitslae te vergemaklik laat ek hier 'n volledige staat volg van die resultate van die 16 verskillende ontledings waaruit hierdie Ondersoek bestaan.

TOELIGTING VAN DIE STAAT.

'n Korte opheldering in verband met hierdie opsomming skyn my nodig te wees.

Die letters heel aan die linkerkant is vir die aanduiding van die reël of item wat bespreek word, en word gebruik om verwysings te vergemaklik.

Daar is reeds vroeër met nadruk gewys op die verskil tussen woordeskat en woorde, laasgenoemde dikwels „lopende woorde” of „woorde leesstof” genoem. Die eerste reël toon die woordeskat wat elke ontleding opgelewer het, terwyl die tweede reël die aantal netto lopende woorde (d.i. na aftrek van name) aantoon.

Vir elke ontleding is 'n lys gemaak van die meesgebruikte woorde. Omdat die ontledings van verskillende grootte is, varieer die minimumkrediet vir, en die aantal van, die meesgebruikte woorde natuurlik ook. Die aantalle van hierdie meesgebruikte woorde word in reël (d) gegee, en net daaronder, in reël (e), wát die kleinste krediet is wat 'n woord vir die verskillende ontledings moes hê om as 'n veelgebruikte woord vir dié ontleding gereken te word.

Die veelgebruikte woorde met hulle herhalings maak in ieder geval die grootste deel van die lopende woorde van die ontledings uit. Die getalle wat hulle van die netto totaal maak word in reël (f) gegee, terwyl die persentasies hiervoor in reël (h) verskyn, en dié wat die 100, 50, 25 en 10 meesgebruikte woorde van die

netto totaal vorm, in reëls (i), (j), (k) en (l) onderskeidelik te sien is. Alleen die *persentasies* vir hierdie vier groepe is gegee, nie die werklike getalle nie; dié is te vinde in die verskillende lysste van die 100 meesgebruikte woorde.

Die herhalingsverhoudings, d.w.s. die verhouding tussen die woordeskat en die lopende woorde, of, met ander woorde, hoeveel maal elke afsonderlike woord gemiddeld herhaal is, is ook bereken, eerstens vir die ontleding in sy geheel geneem, d.w.s. die veel- en mingebruikte woorde deurmekaar; dan vir die veelgebruikte woorde alleen (wat natuurlik die hoogste 100, 50, 25 en 10 woorde insluit); en vervolgens vir die mingebruikte woorde alleen geneem. Die getalle is natuurlik verkry deur die getal van die lopende woorde in ieder geval met die woordeskat te verdeel, en die syfer na die dubbelpunt in reëls (c), (g) en (o) toon aan hoeveel maal elke woord in iedere groep gemiddeld gebruik of herhaal is.

Die woordeskat van mingebruikte woorde word verkry deur die woordeskat van veelgebruikte woorde van die ganse woordeskat af te trek. Mingebruikte woorde beteken dié wat minder kere voorkom in elke stuk as die minima in reël (e) gegee vir die verskillende ontledings onderskeidelik. Die aantal lopende woorde vir die mingebruikte woorde word insgelyks verkry deur die aantal lopende woorde van die veelgebruikte woorde van die netto totaal vir die ontledings af te trek.

Die totaal in een op die laaste kolom syfers na word natuurlik verkry deur die aantalle vir die 16 ontledings langs die reël af op te tel, en die gemiddelde syfer deur hierdie totaal met die aantal ontledings—16—te verdeel.

BESTENDIGHEID VAN PERSENTASIES.

Wat in hierdie opsomming veral opmerklik is, is die egaligheid en bestendigheid van die persentasies wat die groepe woorde, as die veelgebruikte, die 100, 50, 25 en 10 meesgebruikte, of hoogste, woorde beskryf, van die netto totaal lopende woorde uitmaak. Die laagste persentasie vir die veelgebruikte woorde (n.l. 74·8%) is dié in Ontleding No. 3 (hoofartikels), maar dit moet in aanmerking geneem word dat genoemde ontleding baie klein is (slegs 6,679 woorde leesstof), en dat die minimumkrediet 3 is—net so hoog soos vir ontledings wat oor die 15,000 woorde beloop. Begryplik dus dat dit slegs die klein getal van 358 veelgebruikte woorde oplewer, en die persentasie laag is. Behalwe dié lae persentasie vir hierdie groep is daar slegs twee ontledings wat 'n persentasie van onder die 80 persent toon (Ontledings 4 en 15), wat beide net oor die 79% oplewer, en ontledings is met baie

groot woordeskatte. Behalwe nog drie ander, wat so hoog loop soos tot in die 90%, staan die persentasies vir al die ander tien hoër of laer in die taggentig. Die hoogste vir hierdie groep is 93·7%, en die laagste soos hierbo gemeld, maar aangesien die aantal also gekieste woorde in aantal varieer tussen 358 en 665, is hierdie benaderende eenvormigheid merkwaardig. Die gemiddelde persentasie vir hierdie groep vir die 16 ontledings is 85·6%.

As ons die 100-woorde groep nagaan sien ons dat hierdie standhoudendheid van die persentasies net so opvallend is. Hier beloop hulle almal—met die uitsondering van dié van vier uit die 16 ontledings—'n persentasie van in die sestig, vanaf 61·4% tot 69·3%. Van die vier uitsonderings daal twee met 'n breuk onder die 59 persent en vir een—dié vir Ontleding 12 (hofsake), waarin die herhaling uit die aard van die saak hoog moet wees, en dit ook is—klim die persentasie tot oor die sewentig persent, n.l. 73·7%. Vir die res is die persentasies almal nader aan die 60 persent as aan die 70 persent, en is gemiddeld 64·2% vir al sestien.

Wat betref die 50-groep is die persentasies almal behalwe een (alweer Ontleding 12, waarvoor dit 61·5% is) in die vyftig, en hier ook meestal taamlik laag in die vyftig, en loop vanaf 50·3% tot 59·7%. Dié van slegs drie ontledings gaan oor die 55%. Gemiddeld vir die 16 ontledings, 54·0%.

Vir die 25-groep hoogste woorde is daar slegs drie wat onder die 40% daal, n.l. 38·5% (die laagste), 39·2% en 39·7%. Al die ander dertien ontledings se persentasies vir die groep staan hoër of laer in die veertig, die meeste van hulle alweer nader aan die 40% as aan die 50%, met slegs drie oor die 45%. Die hoogste is natuurlik weer vir dié van Ontleding 12. Waarom die persentasie so laag is vir hierdie groep vir 'n Standaard III. leesboek (Ontleding 5) kan ek nie verklaar nie. Gemiddeld, 42·6%.

Daar is heelwat variëring in die 10-woorde groep. Die persentasies van elf van die ontledings vir hierdie groep is hoër of laer in die twintig, meestal in die omtrek van 25%—die laagste van hierdie elf 24·4% en die hoogste 28·7%. Daar is vier met net 'n breuk onder die 25 persent (24·4% is die laagste). Die ander vyf is almal in die dertig; Ontleding 12 natuurlik weer die hoogste met 34·4%. Die ander vier in die dertig is 30·2%, 31·1%, 31·5% en 32·4%.

Wat my in hierdie groep verras het is die hoë persentasies vir Ontledings 3 en 15—altwee koerantestof met groot woordeskatte in verhouding tot die aantal lopende woorde ontleed—een daarvan nogal hoofartikels!

Die totaal van die woordeskat vir die 16 ontledings is 35,920, maar aangesien 'n menigte woorde oormekaar slaan, spreek dit vanself dat die werklike woordeskat (aantal kaartjies) verkry, baie kleiner sal wees. Dieselfde geld vir die woordeskat van die veelgebruikte woorde, die totaal waarvan, deur die getalle van die 16 ontledings op te tel, 8,784 is. Hierin sal die sameval van die krediete op dieselfde woorde wel groter wees, egter lank nie in so 'n mate dat dit die aantal verskillende woorde van hierdie groep tot op 549, soos dit gemiddeld uitwerk, sal afbring nie, soos later sal blyk.

Hierdie staat toon dat elkeen van die 16 ontledings gemiddeld uit 16,813 lopende woorde bestaan, met 'n woordeskat van gemiddeld 2,245, en dat in sy geheel geneem elke afsonderlike woord omtrent $7\frac{1}{2}$ maal herhaal of gebruik is, terwyl die gemiddelde herhalingsverhouding vir die veelgebruikte woorde alleen 1:26.2, en dié vir die mingebruikte woorde alleen geneem 1:1.4 is. Uit die gemiddelde woordeskat van 2,245 is gemiddeld 549 veelgebruikte woorde, latende 'n woordeskat van gemiddeld 1,696 mingebruikte woorde gewende, met herhalings, slegs gemiddeld 2,411 lopende woorde. Hieruit blyk weer duidelik hoe onbelangrik, uit die oogpunt van gebruiklikheid beskou, hierdie gedeelte van die leesstof is.

KREDIETE OORGEDRA.

Soos reeds vroeër gemeld, is alle woorde in die loop van die ontledings behou in die vorm waarin hulle aangetref is, met inbegrip van spelfoute (behalwe in die briewe van enkele baie ongeletterde persone, in watter gevalle die spelfoute deur my in die oorskrif sommer verbeter is), verkortings, dubbelvorme van een woord, ens., en vir elkeen is 'n afsonderlike kaartjie voorsien.

Met die voltooiing van al die ontledings het dit my egter voorgekom dat dit raadsaam is om die krediete van sommige van hulle bymekaar te tel, en dit is dan ook gedoen. Verkorte woorde se krediete is by die woord voluit geskryf, gevoeg, maar dit alleen waar die woorde beide voluit geskryf en in verkorte vorm teëgekrom is. So is *mnr.* se krediete by *meneer* gevoeg; *adv.* se krediete by *advokaat*; *genl.* se krediete by *generaal*, ens. Waar 'n woord alleen in die verkorte vorm verskyn, is dit in dié vorm behou. Waar dit egter 'n verkorting van 'n sinsnede is soos *o.m.* (*onder meer*) *d.v.s.* (*dit wil sê*) *d.i.* (*dit is*) ensovoorts, is dit onveranderd behou.

Op dieselfde manier is gehandel met woorde waarvan daar twee vorme, of wisselvorme in die spelling, bestaan, soos *skryf*, *skrywe*; *kafee*, *kafé*; *regiment*, *resjiment*; *toorry*, *towery*, *toordery*;

vergeef, vergewe; uitgedryf, uitgedrywe; vrind, vriend, ens., dit wil sê, waar die verskillende vorme ôf toegelaat word deur die jongste Woordelys van die Suid-Afrikaanse Akademie vir Taal, Lettere en Kuns (1931), ôf dit bloot afhang van die skrywer se willekeur, maar nie waar die wisselvorme verskillende betekenisse het of verskillend gebruik word nie, soos *jong, jonk; lang, lank; jul, julle; hul, hulle; leef, lewe; liever, lievers; gou, gou-gou*—hierdie woorde is almal afsonderlik gehou, asook 'n paar wisselwoorde waar die verskillende vorme aansienlik van mekaar verskil n.l. die woorde *skadu, skaduwee; miete en miedens (badde, baddens, bedde, beddens,* is glad nie in die hele Onderzoek teëgekom nie).

In die geval van verkeerd gespelde woorde is die krediete van sulke woorde oorgedra na die regte spelling volgens die jongste Woordelys van die Akademie. Waar daar twee of meer deur die Akademie erkende wisselvorme bestaan, is die krediete oorgebring na die woord ôf in sy eenvoudigste, onverbuigde vorm ôf dié een wat die gebruiklikste is. Hierin het ek maar na eie oordeel te werk gegaan. By voorbeeld, die krediete vir *skrywe* is by *skryf* gevoeg; dié vir *resjiment* by *regiment*; dié vir *reisies* by *resies*; dié vir *vrind, vrinde, vrindelik, vrindin, vrindinne, ens.* is by *vriend, vriende, vriendelik, vriendin, vriendinne, ens.* onderskeidelik, getel.

Die woorde *bevoorbeeld* en *vereers* is orals in die ontlede stof aangetref en voorlopig onveranderd behou. Volgens die nuutste spellingdekrete word hulle vermoedelik nou by *voorbeeld* en *vir eers*, en die behaalde krediete deur eersgenoemde twee is dus op laasgenoemde vier oorgedra.

By die oordra van die krediete is ook die verspreiding, d.i. in hoeveel ontledings die woord voorkom, in aanmerking geneem. Die ontledings waarin beide die verkorting en die volle woord voorkom, word net eenmaal getel, terwyl dié waarin die woord en sy verkorting nie gemeen is nie, afsonderlik getel word.

Die volledige lys van woorde waarvan die krediete oorgedra is, word in Bylae II. gegee.

Al die kaartjies vir die woorde waarvan die krediete op dié manier oorgebring is, is toe verwyder uit die ander kaartjies wat die woordeskat vir die Onderzoek verteenwoordig. Toe is die krediete vir die verskillende ontledings, wat in 'n kolom op die kaartjies aangeteken is soos hulle in die loop van die Onderzoek voorgekom het, opgetel en was alles gereed om die kaartjies te tel en dusdoende uit te vind hoe groot die woordeskat is van die 269,009 woorde leesstof wat ontleed is.

TOTALE WOORDESKAT VERKRY.

Dit het 'n totale woordeskat (met die uitsondering natuurlik van dié woordvorme waarvan die krediete oorgedra is) van 14,222 verskillende woorde opgelewer op die grondslag reeds vroeër uiteengesit n.l. dat elke verskillende woordvorm as 'n afsonderlike, selfstandige woord geneem is.

Hierdie 14,222 woordkaartjies is toe in sewe groepe verdeel volgens hulle krediete, en die uitslag toon dat daar 959 woorde is met 'n krediet van 25 en hoër; 229 woorde met 'n krediet van tussen die 20 en 24 (beide getalle in ieder geval inbegrepe); 307 woorde met 'n verskyningskrediet van 15 tot 19; 576 woorde met 10 tot 14; 1,592 woorde met 'n krediet van 5 tot 9; 3,990 woorde met 'n verskyning van 2 tot 4; en 6,569 verskillende woorde wat net 1 maal in die hele Onderzoek teëgekóm is. (Die getalle vir laasgenoemde twee groepe is nie nagetel nie—dit is nie die moeite werd nie; hulle is egter werklike tellings en nie ramings nie, en daar hulle met sorg getel is, sal hulle naasteby of heeltemal reg wees).

VI. Die Keuse van die Duisend Meesgebruikte Woorde.

Uit hierdie woordeskat van 14,222 moes die 1,000 meesgebruikte woorde toe gekies word. Nogal heel maklik, sou 'n mens dink, veral aangesien hulle reeds in groepe verdeel en gegradeer is; al wat jy doen is om die 1,000 woorde met die hoogste krediete van bo af te tel.

Maar so eenvoudig is die saak glad nie. Daar is allerhande faktore wat in aanmerking geneem moet word. As alle woorde met 'n hoë krediet ook 'n hoë verspreiding gehad het, sou die keuse maklik genoeg gewees het. Maar, hoewel dit by baie woorde aangetref word, is dit nie altyd die geval nie. Daar is heelwat woorde met 'n hoë krediet maar 'n lae verspreiding—meestal vakterme. Aan die ander kant is daar woorde met 'n hoë verspreiding maar deurgaans lae krediete.

As 'n woord sê 150 maal voorkom, maar net in een ontleding aangetref word, en 'n ander woord kom, sê, 30 maal voor, maar word in 10 uit die 16 ontledings teëgekom, watter een is nou die algemeenste en die meesgebruikte woord? Of neem die geval van 'n woord wat 'n totale verskyningskrediet van 170 het en in 3 ontledings voorkom, en 'n ander wat 80 maal voorkom en in 16 bronne aangetref is, of 'n woord wat 'n krediet van 100 het en in 9 bronne voorkom, en 'n ander met 'n krediet van 90 en 'n verspreiding van 10.

Hoe gaan jy beslis watter van hierdie woorde die gebruiklikste is? Watter maatstaf gaan jy gebruik? Dit het my spoedig duidelik geword dat dit nodig sal wees om 'n aantal reëls neer te lê waaraan woorde getoets moet word om in die lys van die 1,000 meesgebruikte woorde opgeneem te word.

Ek het baie gevalle van hierdie aard teëgekom, en dit het my laat sien dat daar twee faktore is wat veral in aanmerking geneem moet word in die vasstelling van die algemeenheid en gebruiklikheid van 'n woord, naamlik die verskyningskrediet en die verspreiding oor die ontlede bronne, dit wil sê hoeveel maal die woord werklik aangetref is en in hoeveel van die verskillende

ontledings dit verskyn het. Die vraagstuk was hoe om die twee faktore te verbind en tot 'n omvattende eenheid saam te voeg, en also 'n maatstaf te kry om die waarde van 'n woord vas te stel.

Een of twee metodes is hiervoor probeer, soos die vermenigvuldiging van die krediet met die verspreiding, maar dit het geblyk dat dié metode 'n te groot gewig aan die verspreiding verleen. Na raadpleging met 'n aktuaris en 'n statistikus het ek besluit om die volgende metode vir die lading, of vasstelling van die waarde van woorde, te gebruik.

HOE 'N OMVATTENDE MAATSTAF VERKRY IS.

In plaas dat die krediet vermenigvuldig is met die verspreidingstal is dit vermenigvuldig met die vierkante wortel daarvan. Die hoogste verspreiding moontlik is 16, en volgens hierdie metode moes die totale krediet van 'n woord vermenigvuldig word met die vierkante wortel van die aantal ontledings waarin die woord voorkom. As die juiste vierkante wortel elke keer moes gebruik word sou die arbeid van die berekening vir al die woorde baie groot gewees het, en gevolglik is dit alleen geneem tot die naaste half, soos in die volgende tabel aangetoon.

<i>Verspreiding.</i>	<i>Getal gebruik vir Lading.</i>
1	1
2-3	$1\frac{1}{2}$
4-5	2
6-7	$2\frac{1}{2}$
8-10	3
11-14	$3\frac{1}{2}$
15-16	4

Om met 'n paar voorbeelde op te helder: As 'n woord in 7 ontledings voorkom en 'n krediet van 24 het, is sy belade krediet (of waarde) $2\frac{1}{2} \times 24 = 60$; as 'n woord in 15 bronne verskyn en 'n krediet van, sê, 20 het, is sy waarde 80; 'n woord met verspreiding van 11 en 'n krediet van 25 het 'n waarde van 87 (halwes word nie gereken nie). Hierdie metode verseker dat die meeste waarde geheg word aan die verskyningskrediet, dog met behoorlike inagneming van die verspreiding.

Dit was natuurlik onnodig om vir die keuse van die hoogste 1,000 woorde, waardes vir die hele 14,222 woorde te bereken. Dit is alleen gedaan vir dié wat 'n krediet van 10 en hoër opgehoop het, waarvan daar 2,071 is. Hierdie getal, het ek gemeen, is heeltemal voldoende vir die keuse van 'n duisend woorde.

Na die berekening van die waardes van hierdie 2,071 woorde was dit 'n eenvoudige saak om hulle te gradeer van die hoogste tot die laagste om dan die eerste 1,000 af te tel. Dit is gedaan en 'n merkie hier ingesit om hulle van die res af te skei.

By die deursien van hierdie 1,000 woorde het ek egter gemerk dat ek nog nie alle moeilikhede onder die knie het nie. Ek het by voorbeeld onder hulle woorde gevind soos *polis*, met 'n verspreiding van slegs 1, maar 'n krediet van 152; die woord *uwe* (uit briewe) met 1:177 (die eerste syfer gee die verspreiding en die tweede, net na die dubbelpunt, die krediet); *bestuurder* 3:115; *heer* 3:180; *premie* 2:94, en daar ek meen dat woorde met so 'n lae verspreiding nie die aanspraak, om as 'n algemene woord beskou te word, regverdig nie, het ek besluit dat 'n verdere beperking nodig is.

VERDERE BEPERKINGS.

Ek het dus vasgestel dat 'n woord, om in die lys van Duisend opgeneem te word, moes verskyn in minstens drie van die ses klasse lektuur uiteengesit op bladsy 66 en in nie minder as ses ontledings nie. Hierdie beperking, meen ek, is niks te streng nie, veral wat betref die aantal ontledings. Immers, die ontlede stukke was die meeste vry lank, en as 'n woord nie verskyn in meer as 6 uit die 16 nie, dit kom daar nie op aan hoe 'n groot krediet hy het nie, sal niemand tog wil beweer dat hy 'n algemeen veelgebruikte woord is nie.

Vervolgens het ek die kaartjies tot by die 1,000-merk sorgvuldig nagegaan en al die woorde wat nie aan hierdie eise voldoen nie, uitgehaal. Daar is 52 sulke woorde in hierdie eerste eliminerings. Hulle is met 'n sterretjie gemerk in die lys van 1,071 woorde (sien bladsy 134) wat na die eerste Duisend kom.

By die deurgaang van die kaartjies by hierdie eerste eliminerings het ek gesien dat vir sommige woorde verreweg die grootste deel van die krediet deur net een of twee ontledings of klasse lektuur bygedra is, terwyl die krediet deur 'n ander ontleding of klas, of ander ontledings of klasse lektuur bygedra, baie klein is. Dit het my voorgekom dat dit noodsaaklik is om die maksimumkrediet van 'n enkele bron of 'n enkele klas te beperk ten einde te verhoed dat die krediet totaal deur net een bron of een klas geheel oorskadu word, met 'n onbeduidende krediet uit ander bronne of klasse verkry.

Ten einde dus die ewewig van die krediet vir 'n woord te bewaar het ek, na enige oorweging, neergelê dat nie meer as 40% ($\frac{2}{5}$ de) van die krediet totaal van 'n woord uit een bron, en nie

meer as 60% (of 3/5de) uit een klas lektuur toegelaat sou word nie. Gevolglik het ek nog 'n keer dié kaartjies, die krediete waarvan hierdeur getref kon word, deurgegaan en waar die krediete hierdie persentasies oorskry het, is hulle van die totaal afgetrek en 'n nuwe waardesyfer aan die woord toegeken. Op dié manier is 212 woorde in waarde verminder, en hierdie kaartjies is toe uitgehaal en op hulle regte plekke, volgens hulle nuwe waarde, weer ingesit. Dit wil sê, hulle is almal laer af verskuif, sommige meer en ander minder, maar hulle het almal nog, met die uitsondering van 16, bokant die oorspronklike 1,000-merk gebly. 'n Lys van hierdie 212 woorde van die tweede eliminerings, met 'n opgawe van hulle oorspronklike en verminderde waardes, en hoeveel plekke hulle in die 1,000-woorde Lys gedaal het, verskyn in Bylae III. agter in die boek, en dié wat nie uitgesak het nie, is kursief gedruk in die lys van Duisend.

Deur hierdie twee eliminerings is 68 woorde uit die lys van die Duisend woorde van hoogste waarde verwyder (52 in die eerste en 16 in die tweede eliminerings). Hierdie 68 woorde is toe gekies uit dié onderkant die 1000-woorde merk, uit woorde wat die naaste aan die neergelegde vereistes voldoen, en die Duisend algemeenste en meesgebruikte woorde was toe voltallig, en behoorlik gegradeer.

HOE DIE WOORDE GEGRADEER IS.

Die waardes is alleen gebruik vir die keuse en gradering van die lys van 1,000 woorde; die syfers wat gegee word toon die werklike aantal bronne of ontledings waarin die woorde aangetref is (die syfer voor die dubbelpunt), en die werklike aantal kere wat die woord voorgekom het (na die dubbelpunt). Die belade krediete (berekende waardes) word nie gegee nie, omdat ek meen dat so 'n boel syfers net die leser verwar. Om egter aan te toon hoe die gradering uitgevoer is, word net 'n kort lysie woorde, met hulle werklike verspreiding en krediete, en hulle berekende waardes, as voorbeeld hier gegee. Die hoogste waardes kom bo-aan en woorde neem laer plekke in na gelang van kleiner waardes.

Hierdie woorde is sommer hier en daar uit die lys geneem en dien alleen as 'n voorbeeld van hoe die gradering geskied het. Hulle is gegradeer volgens die waardegetalle wat in die laaste kolom verskyn.

VOORBEELD VAN GRADERING.

<i>Stand.</i>	<i>Woord.</i>	<i>Verpreiding.</i>	<i>Krediet.</i>	<i>Waarde.</i>
1.	nuwe	16	134	536
2.	klaar	15	113	452
3.	oomblik	16	100	400
4.	vas	13	88	308
5.	uur	15	69	276
6.	swaar	15	52	208
7.	oog	13	56	196
8.	Junie	8	62	186
9.	skoon	11	53	185
10.	middel	15	44	176
11.	blomme	8	51	153
12.	verneem	10	46	138
13.	taamlik	10	31	93
14.	besonder	10	30	90
	betrekking	6	36	90
16.	bou	12	25	87
	dik	8	29	87
18.	breek	12	20	70
19.	snaaks	7	22	55

Die reëls vir die keuse van die Duisend woorde kan as volg opgesom word: Die duisend woorde met die hoogste waardes is gekies, maar die woorde moet tewens aan die volgende vereistes voldoen t.w. 1. 'n woord moet 'n verspreiding hê van minstens 6 uit die 16 ontledings en van minstens 3 uit die 6 klasse leesstof; 2. nie meer as 40% van die krediet totaal deur 'n enkele ontleding, of 60% deur 'n enkele klas lektuur gelewer, is vir 'n woord toegelaat nie—wat hy meer het is afgetrek vir die vasstelling van die waarde van die woord en dus sy stand, maar die *werklike krediete* word in die Lys gegee.

Ek moet erken dat ek die reëls graag strenger sou gesien het. By voorbeeld, dat woorde in minstens een meer as die helfte van die ontledings en van die klasse lektuur moes verskyn; dat nie meer as 30% of 25% van die krediet van 'n enkele ontleding, en nie meer as 50% of 40% van 'n enkele klas leesstof, in aanmerking kom nie. Ook sou dit miskien goed gewees het as 'n vereiste

minimum vir die krediet wat 'n enkele ontleding of klas bydra, vasgestel kon word, want ek het gemerk dat die krediete van sommige ontledings of klasse onbeduidend klein is.

As hierdie reëls of beperkings egter toegepas is sou hulle die gevolg hê om 'n aantal woorde verder uit die lys te verwyder en 'n groot aantal ander 'n laer plek in die lys laat inneem, en as dit gebeur het sou daar nie uit die totaal van 14,222 verskillende woorde, wat die Onderzoek opgelewer het, 'n duisendtal woorde beskikbaar wees nie, soos nader aangetoon word op bladsy 114.

Die lys van die 1,000 meesgebruikte woorde word hieronder gegee met hulle verspreidings en krediete. Hierdie is nie berekende waardes of belade krediete en in geen geval skattings nie, dog die werklike verskynings soos getel en aangeteken in die verskillende ontledings waaruit die Onderzoek bestaan, en opgetel. Die belade krediete, of waardes, word nie in die volle lys gegee nie, want hulle is alleen gebruik as 'n maatstaf vir die keuse en gradering van die eerste Duisend woorde.

Ek beskou hierdie Duisend woorde van so 'n oorwegende belangrikheid as die grondslag van die gebruiklike Afrikaanse woordeskat, dat ek dit wenslik ag om hulle eers te gee volgens hulle waarde gerangskik, van die hoogste af tot die laagste gegra-deer, en daarna dieselfde Lys in alfabetiese orde, met hulle Engelse betekenisse, en met behoud van die nommers wat hulle stand aanwys.