HOOFSTUK V

BESLUIT

1. Vasstelling van die kernwoordeskat.

Om 'n criterium te vind vir die vasstelling van die grondwoordeskat, is besluit dat alle woorde met 'n frekwensiesyfer van 25 of hoër as basies beskou sal word. 3,130 sulke woorde is gevind. Dit is dus duidelik dat 'n heelwat omvangryker ondersoek nodig is om 'n uitgebreider en betroubaarder kernwoordeskat vas te stel. Enkele bedenklike woorde (b.v. woekeraar, mikrobe, trollie, sement) word deur veelvuldige aanwending in een of twee reekse 'n plek in die kernlys besorg. Hulle mag dan wegval ten gunste van woorde wat nou in die bylae verskyn.

2. Grootte van die woordeskat.

Telling van al die kaartjies het getoon dat die woordeskat van die vyf reekse uit 11,062 woorde bestaan. Maar hierdie syfer dui alleen woorde aan wat in die woordeboek voorkom. Daarby moet getel word: (*a*) 87 woorde (nie-samestellinge of -afleidinge), (*b*) 15,465 samestellinge[1]) en (*c*) 1,082 afleidinge[2]) wat nie in die woordeboek verskyn nie. Die totale woordeskat is dus 27,696 woorde, maar, soos vermeld, is slegs 11,062 hiervan in die woordeboek opgeteken.

Opvallend van die Afrikaanse taal — in teenstelling met Engels wat uit Romaanse sowel as Germaanse elemente bestaan en dus besonder ryk aan sinonieme is — is enersyds sy groot armoede aan grondwoorde en andersyds sy geweldige rykdom aan samestellinge en tewens ook afleidinge. Dit is hierdie rykdom van die taal wat die taak van die Afrikaans-leksikograaf bemoeilik, want hy moet gedurig kies en keur sodat sy woordeboek nie te lywig word nie. Daarby verg dit die grootste oordeelkundigheid om alleen die meesgebruikte vorme op te neem en die minder bekendes te elimineer. Dit is dan ook 'n kerngebrek van die Afrikaanse leksikografie dat die meesgebruikte en bekendste samestellinge veral nie

[1] en [2] Sien Bylae B vir voorbeelde.

almal opgeteken is nie. Uitgebreide navorsing in hierdie rigting is dus hoogs noodsaaklik.

Deurdat die opname van samestellinge en afleidinge so arbitrêr is, moes al hierdie vorme in die ontlede reekse gekontroleer word, insonderheid die volgende:

(*a*) werkwoordelike samestellinge — veral dié met aan-, af-, agter-, binne-, by-, deur-, in-, los-, mee-, mis-, na-, om-, onder-, oop-, oor-, op-, raak-, reg-, rond-, saam-, teë (teen)-, terug-, toe-, uit-, vas-, verby-, voor-, vooruit-, weg-, ens. as eerste lid en dié met -maak en -hou as tweede lid.

(*b*) saamgestelde s.nwe. Daar was 'n groot aantal met baas-, reuse-, hoof-, ens. en met stofaanduidende s.nwe. (b.v. silwer-, koper-, sy-, wol-, hout-, yster-, ens.) as eerste lid.

(*c*) alle saamgestelde bywoorde en byvoeglike naamwoorde.

(*d*) alle samekoppelinge.

(*e*) alle afleidinge — veral afgeleide b.nwe. en bwe. met on- en afgeleide werkwoorde met be-.

Die twee (of meer) lede van samestellinge (samekoppelinge) is apart opgeteken, maar slegs in enkele gevalle kon dit nie gedoen word nie, omdat een van die lede (of albei) nie in die woordeboek verskyn nie, b.v. (mee)horing, (langasem)kriek, stoom(enjin), (haarkapper)salon, (foks)terriër, (nama)bessie, (martini)-(henri). Uitsonderinge is gemaak t.o.v. *opsy* (b.v. opsyspring) en *jagluiperd*, omdat hierdie twee woorde wel in die Engels-Afrikaanse deel verskyn. Die splitsing van die samestellinge en samekoppelinge (b.v. raad-op) het dan ook meegebring dat daar aansienlik meer lopende woorde ontleed is (ongeveer 31,000) as wat daar werklik in die reekse voorkom.

Dat sommige van die veelgebruikte samestellinge in die reekse nie opgeteken is nie, dui bepaald op 'n leemte in die samestellinge van die Afrikaanse woordeboek. Maar hiermee bepleit ons natuurlik nie die onvoorwaardelike opname van alle moontlike verbindinge nie, want die Afrikaanse spel- en skryfwyse, met sy sintetiese grondslag, maak die verbindingsmoontlikhede haas onbeperk. Indien alle verbindinge geboekstaaf moet word, sal die gewone woordeboek seker twaalf of meer dele beslaan i.p.v. een. Dit is hierdie groot praktiese beswaar wat volledigheid onmiddellik uitskakel en daartoe bydra dat 'n hele aantal van ons veelgebruikte samestellinge (en afleidinge) nie in die Afrikaanse woordeboek opgeteken is nie.

Slegs 'n geringe aantal woorde (nie-samestellinge en -afleidinge), meestal van uitheemse oorsprong, kon nie in die woordeboek gevind word nie, b.v. kleim, wallabie, wallaroe, teepee, wigwam, meksim, sok (= mark), enjin, pekinesie, swets, sanna, booi.

In enkele gevalle moes die woorde in die reekse verander word in ooreenstemming met die vorm in die woordeboek aangegee, b.v. soem > zoem, herstelwerk > herstellingswerk, kwikkie > kwikstertjie, leunstoel > leuningstoel, knoestig > knoesterig, wilde-sand > wilsand, asbestos > asbes, tamarind > tamaryn, piekanien > pikkenien, gelykop > gelyk-op, riempiestoel > riempiesmatstoel, harpoeneerder > harpoenier, perdoemps > pardoems, ens.

Benewens die 3,130 kernwoorde word 7,403 ander woorde (met frekwensies van 5 — 24) ook aangegee. Hulle verskyn in Bylae A.

3. Betroubaarheid van die telling.

Volgens die eerste berekening (d.w.s. vermenigvuldiging van die die velle met 104) is vasgestel dat 1,143,337 lopende woorde afgeskryf is. Telling van die frekwensiesyfers (1,066,019 vir die 3,130 kernwoorde, 75,907 vir die 7,403 woorde met frekwensies van 5 — 24 en 918 vir die 529 woorde met frekwensies van 1 — 4) dui egter 'n tekort van 493 aan. Die moontlikhede is dat daar 'n fout ingesluip het in die berekenings of dat daar van die vierkantjies verlore gegaan het.

Waar die fout lê, is moeilik om te bepaal, want die velle is opgesnipper en die vierkantjies is vernietig. Daardie syfers kan dus nie weer gekontroleer word nie. Die waarskynlikste uitleg is dat daar êrens in die berekeninge gefouteer is en dat daar ook 'n paar van die vierkantjies verlore geraak het. Hoe dit ook sy, die verskil is onbeduidend, want die tekort van 493 verteenwoordig slegs ·043% van die bruto-totaal lopende woorde volgens die eerste berekening.

4. Woordfrekwensie in Afrikaans.

Dit is interessant om daarop te let watter woorde die meeste voorkom en watter persentasie van die nettototaal lopende woorde wat ontleed is, uit die meesgebruikte woorde bestaan. Hier volg 'n paar syfers:

Die hoogste			*Lopende woorde*	*Persentasie van die netto totaal.* (1,142,844)
1	woord	verteenwoordig	85,442	7·5
5	woorde	,,	166,113	14·5
10	,,	,,	256,971	22·5
20	,,	,,	368,727	32·3
30	,,	,,	442,959	38·8
40	,,	,,	495,933	43·4
50	,,	,,	536,705	47
60	,,	,,	569,844	49·9
70	,,	,,	596,257	52·2
80	,,	,,	618,514	54·1
90	,,	,,	636,775	55·7
100	,,	,,	651,723	57
3,130	,,	,,	1,066,019	93·3

Die meesgebruikte woord, volgens hierdie ontleding, is *die*; daarna volg *en*, *'n*, *hê* (*het*), *is*, *hy*, *nie*, *van*, *sy* (he) en *hul*(*le*). Tesaam verteenwoordig hulle 22·5% van die netto totaal lopende woorde. Hierdie lysie verskil heelwat van Aucamp s'n, want sy eerste tien woorde is *die*, *en*, *nie*, *het*, *'n*, *van*, *is*, *in*, *ek*, en *dit* (wat 25·4% van die netto totaal lewer). Die gebrek aan sameval is waarskynlik te wyte aan die verskillende aard van die stof wat in die twee ondersoeke ontleed is.

Hierdie studie toon verder aan dat die hoogste 60 woorde helfte van die lopende woorde beslaan—by Aucamp is dit die eerste 50 woorde — en dat die kernwoordeskat van 3,130 woorde (d.i. 28·2% van die leksikologiese woordeskat of 11·3% van die totale woordeskat) vir 93·3% van die lopende woorde (m.a.w. van die geskrewe taal in die ontlede reeks) verantwoordelik is. Die gemiddelde frekwensiesyfer vir die kernwoordeskat is 340·6 teenoor 9·7 vir die orige woorde. Die bevindinge dui tewens aan met hoe min basiese woorde die skrywers van leesseries (en vermoedelik ook die meeste skrywers, om te oordeel na Aucamp se ondersoek) hulle inderdaad bedien. Dit sou belangwekkend wees om, indien daar soortgelyke syfers beskikbaar word, 'n vergelyking met die gesproke taal te tref.

Op grond van sy eie bevindinge beweer Aucamp[3]) dat dit met sekerheid aangeneem kan word dat „in elke gewone geskrif van 'n soort" die eerste tien woorde gemiddeld nooit onder die 25% sal daal nie, die gebruiklikste 50 woorde nie onder die 51% nie en die 100 gebruiklikste woorde nie onder die 60% nie. Die dogmatiese karakter van so 'n stelling maak dit aanvegbaar. Trouens, dit mag geldig wees vir drukwerk in die algemeen, maar vir hierdie ondersoek is dit beslis nie. Wat nodig is om vergelykings te tref, is uitgebreide frekwensietellings van verskillende soorte stof, d.w.s. van algemene drukwerk, sake- en persoonlike briewe, skoolhandboeke, tydskrifte, ens.

Wat die aard van die kernwoordeskat betref, kan aangetoon word dat die woorde uit alle soorte rededele bestaan. Dit is egter opvallend dat daar onder die eerste 100 besonder min s.nwe. en wwe. voorkom: die oorgrote meerderheid is vnwe., b.nwe. of wat Jespersen partikels noem (d.w.s. bywoorde, voorsetsels, voegwoorde en tussenwerpsels). Ontleding van die woordsoorte (kernwoordeskat) gee naastenby die volgende persentasies: s.nwe.: 41%; wwe.: 28%; b.nwe. (telwoorde en lidwoorde): 11%; ander woordsoorte: 20%.

Die kernlys verteenwoordig ook allerlei terreine, b.v. liggaamsdele, kledingstukke, groente, vrugte, diere, voëls, insekte, reptiele, plaas- en stadslewe, huislike lewe, gereedskap, minerale, die jag, sport, weersgesteldheid, die gevoelslewe, e.d.m.

Die vraag, of daar belangrike woorde uitgelaat is, is moeilik om te beantwoord. Funksieverdeling het meegebring dat 'n geringe aantal bekende woorde nie in die basiese lys verskyn nie, b.v. baan, deug, eis, flits, gaap, haak, hef, e.a.; maar behoudens hierdie groepie is dit onmoontlik om na ander te verwys. Die optimum woordeskat is nog onbekend, en ontbrekende woorde kan alleen deur verdere navorsing vasgestel word.

5. Aanwending van die lys.

Die kind op die laerskool wat Afrikaans as tweede (of vreemde) taal aanleer, moet kennis maak sowel met dié woorde wat die meeste in die taal voorkom as met die onontbeerlike woordeskat van sy onmiddellike omgewing (soos i.v.m. die liggaam, kledingstukke, diere en voëls, vertrekke in die huis en hul inhoud, die skool, tuin-

[3]) Woordeskat en woordherhaling, 120.

en ander gereedskap, spele, ens.). (Vergelyk Palmer se uitspraak in „The Scientific Study and Teaching of Languages," bl. 122: „The Vocabulary.... should include none but the commonest and most characteristic units.... This microcosm should be formed in accordance with and as a compromise between the principles of Frequency, Ergonic combination, Concreteness, Proportion and General Expediency."). M.a.w., in die keuse van 'n optimum-woordeskat moet beide die frekwensie- en bruikbaarheidsbeginsel toegepas word.

Vir die samestelling van gegradeerde leesseries, spellyste, taalboekies, ens. in suiwer Engelssprekende centra sou die aangegewe kernwoordeskat met die byvoeging van 'n lys utiliteitswoorde naastenby voldoende wees. Waar die omgewing tweetalig of oorwegend Afrikaans is, kan 'n heelwat uitgebreider woordeskat gebruik word.

Vir moedertaalonderwys op die laerskool is die kernlys natuurlik onvoldoende en verder as st. II of III kan dit ons nie voer nie. Tot op daardie stadium bied dit (t.o.v. die woordeskat) 'n objektiewe grondslag aan vir die samestelling van toetse, oefeninge, leerplanne, leesboeke, spellyste, e.d.m. en vir die waardebepaling van reeds bestaande of nuwe handboeke vir taalonderrig. Dit beteken natuurlik nie dat die woordeskat van handboeke, oefeninge, ens. uitsluitlik uit so 'n lys woorde moet bestaan nie. Dit sou veels te kunsmatig en onoordeelkundig wees. Die lys moet nietemin as basies beskou word en verder uitgebrei word in ooreenstemming met die eise van die leerplan en die behoeftes van die leerlinge.

Vir die samestellinge van handboeke, e.d.m. tot en met st. VI sal minstens 10,000 woorde nodig wees. In hierdie ondersoek word wel 10,533 woorde aangegee, maar vir die betroubaarheid van die 7,403 woorde in die bylae veral kan die skrywer nie instaan nie. Daarvoor is, soos reeds vermeld, uitgebreider navorsing nodig. Hierdie studie is dus net aanvoorwerk waarop ander moet voortbou. Met 'n omvangryker lys sou dit nie alleen moontlik wees om betroubaarder handboeke (tot en met st. VI) saam te stel nie, maar ook om die geskiktheid van bestaande of nuwe handboeke te meet. (Sien verder pp. 24 — 28.)

6. Aanverwante ondersoeke.

Terloops is daar alreeds na allerlei aanverwante vraagstukke verwys. Talle probleme (wat direk uit hierdie studie voortvloei)

bly in Afrikaans nog onopgelos, en daarom is wetenskaplike navorsing t.o.v. die volgende dringend noodsaaklik:

(*a*) Uitbreiding van die kernwoordeskat tot minstens 10,000 woorde. Vir hierdie doel sou die ontleding van minstens vyfmiljoen lopende woorde nodig wees.

(*b*) Die samestelling van kernwoordeskatte ten opsigte van al die vakke op die laerskool.

(*c*) In hierdie ondersoek is die aard en inhoud van die stof in die reekse nie ontleed nie. Dit is 'n studie op sigself en sluit o.a. die volgende in: (I) vasstelling van die voorkeure van leerlinge op die laerskool, m.a.w., in watter soort stof die kind op verskillende stadia die meeste belangstel; (II) bepaling van dié eienskappe van die stof wat belangstelling opwek en die leesles gewild maak, b.v. spanning, humor, verrassing, dialoogvorm, ens.; (III) onderlinge verhouding van die lestipes; (IV) bepaling van die waarde van die dialoogvorm teenoor die gewone verhaaltrant op verskillende stadia; (V) vasstelling van 'n geskikte woordeskat om die optimum lestipes te dek, ens.

(*d*) Vasstelling van die omvang van kernwoordeskatte op verskillende stadia, m.a.w. aanduiding van die strale vir die verskillende standerds.

(*e*) Vasstelling van die optimum gemiddelde frekwensie in leesreekse, m.a.w. hoe dikwels elke nuwe woord herhaal moet word om doeltreffende retensie te bewerkstellig.

(*f*) Die samestelling van 'n utilitaristiese woordeskat t.o.v. tweede-taalonderwys, aanduiding van die aard en omvang daarvan en bepaling van die verhouding tussen en die plek en nut van die twee tipes (d.w.s. 'n frekwensie- en bruikbaarheidswoordeskat).

(*g*) Frekwensietellings van onopgetekende samestellinge (en afleidinge) om leemtes in die gangbare woordeboeke te vul.

(*h*) Frekwensietellings van verskillende soorte stof, b.v. koerante, tydskrifte, persoonlike en sakebriewe, algemene drukwerk, ens.

Met hierdie lysie kan ons volstaan. Dit is natuurlik nie volledig nie, maar genoeg is aangebied om te bewys dat daar nog veel op die vorser in Afrikaans wag en dat daar nog talle pedagogiesbelangrike vrae onbeantwoord bly.

7. Varia.

(*a*) By sommige woorde (b.v. enkel, grond, trekker) is die betekenisse of funksies per abuis op sommige vierkantjies uitgelaat. Sulke woorde verskyn net een maal, omdat die verdelings nie aangetoon kon word nie.

(*b*) Om lang omskrywinge in die lyste by homonieme te vermy, word enkele Engelse ekwivalente gebruik.

(*c*) Die laaste boek van die Jong Suid-Afrika-reeks is vir sts. VI en VII bedoel. Om dié rede is net helfte van die boek ontleed, d.i. tot aan die einde van die les op bl. 127.

(*d*) Waar nodig, is anglisismes aangesuiwer en die korrekte vorm opgeteken, b.v. gretig i.p.v. angstig (Ek was angstig om hom weer te sien).

(*e*) Die feit dat daar slegs 529 woorde met 'n krediet van 4 of laer voorkom, getuig dat daar in al die reekse voorsiening gemaak is vir 'n redelike mate van woordherhaling. Die gemiddelde frekwensiesyfer vir hierdie ondersoek is 41 wat goed vergelyk met die aanbevole 25 — 60 (d.i. een nuwe woord vir elke 25 — 60 lopende woorde.)

(*f*) Wat intensiewe vorme van b.nwe. betref, is net een vorm opgeteken, b.v. *stokoud* maar nie *horingoud* nie, ens.

(*g*) Dit is onmoontlik om die bevindinge van hierdie ondersoek met dié van Aucamp te korreleer, omdat die criteria, wat nuwe woorde betref, so hemelsbreed verskil. In sy eerste 1,000 woorde, b.v., verskyn (I) 'n hele aantal verlede deelwoordsvorme en stamvorme van dieselfde woord (b.v. doen, gedoen; hou, gehou; bly, gebly, ens.); (II) enkelvouds- en meervoudsvorme van dieselfde woord (b.v. arm, arms; been, bene; boer, boere, ens.) Homonieme en funksies word ook nie onderskei nie. Hierdie vorser het alle woorde en vorme voor die voet as nuwe woorde beskou.