HOOFSTUK II

PROSEDURE EN TEGNIEK

1. Begripsbepaling.

(*a*) *Wat is 'n woord?*

Aangesien dit die hoofdoel van hierdie ondersoek is om 'n lys woorde saam te stel, moet in die eerste plek rekenskap gegee word van wat as woorde beskou is.

Jespersen[1]) verwys na die definisies van Noreen, Pedersen, Wechssler, Boas, Sapir, Vendryes en Gardiner, en merk tereg op dat dit besonder moeilik is om die begrip „woord" te omskryf. Hy wys daarop dat woorde nòg suiwer fonetiese, nòg suiwer semantiese eenhede is en meen dat die eintlike definisie op sintaktiese kriteriums moet berus.

Palmer[2]) beweer: „The term ‚word' is vague and impossible of definition." „Woorde" noem hy „Lexicological Units," en hy onderskei *monologs* (woorde), *polylogs* (groepwoorde), *miologs* (dele van woorde) en *algoisms* (die uitdrukking van 'n begrip sonder 'n woord).

Die saak word inderdaad bemoeilik deur die feit dat sommige woorde semantiese waarde het (d.w.s. hulle is begripswoorde) terwyl ander slegs morfologiese waarde het (d.w.s. hulle is vormwoorde met weinig of geen betekenis nie, b.v. *is*). So het daar verskeie definisies ontstaan. Sommige taalkundiges beklemtoon die semantiese element (b.v. Le Roux),[3]) ander die morfologiese (b.v. Noreen — Pollak).[4])

Eintlik hang 'n definisie af van die standpunt wat ingeneem word. Enersyds, soos De Villiers[5]) aantoon, is daar die gesigspunt van die grammatikus of leksikoloog, andersyds kan 'n mens uitgaan van die lewende gesproke woord. Alles hang af van die doel voor oë.

Bepaal jy jou by die lewende woord, dan kan 'n woord omskryf

---

[1] The Philosophy of Grammar, 92.
[2] The Scientific Study and Teaching of Languages, 39.
[3] Oor die Afrikaanse Sintaksis, 17.
[4] Wissenschaftliche Betrachtung der Sprache, 446.
[5] Woordsoorte, werkwoorde en Tye, 6.

word as 'n klankkompleks of klankverbinding met 'n selfstandige betekenisinhoud. Hierdie betekenis, volgens De Villiers, is 'n gebruiksbegrip of Verwendungsgedanke — iets abstraks met 'n konstante waarde. Daar bestaan egter 'n verband tussen die verskillende „betekenisse" van een woord. De Villiers haal Reichling aan om te bewys dat die betekenisonderskeidinge disjunktief relevant is en disjunktief toepaslik is.[6])

Vir die doel van hierdie ondersoek, nl. die samestelling van 'n lys losstaande woorde, is die leksikologiese omskrywing, soos dié van Scholtz,[7]) Erdman[8]) en Overdiep[9]) van meer waarde. Op die oomblik dat 'n woord gebruik word, het dit 'n presiese betekenis, maar in 'n lys of woordeboek is dit nie die geval nie. Soos Wellander dit uitdruk: „Das Wort, von dem man in dem Wörterbuch oder in der Grammatik spricht, ist eine Abstraktion aus vielen Wortindividuen, die zu verschiedenen Zeiten gesprochen und gehört worden sind."[10])

In hierdie lys is dus opgeteken die konvensionele ortografiese of leksikologiese eenhede wat Palmer *monoloë* noem.

(*b*) *Watter woordkategorieë bestaan daar?*

In hierdie ondersoek word ook verwys na woordsoorte, daarom moet die klassifisering nader bepaal en verduidelik word.

Steyn[11]) verklaar dat daar tussen die verskillende woordsoorte per slot van rekening geen definitiewe grense aangewys kan word nie — m.i. 'n ietwat oordrewe uitspraak. Vir 'n onderskeiding in kategorieë beskou hy funksie as die geskikste grondslag, maar die feit dat baie woorde multifunksioneel is,[12]) maak so 'n indeling onprakties. De Villiers beweer ook: „. . . ek glo dat funksie 'n aanvullende maar geen beslissende maatstaf kan wees by ('n) indeling van die woordsoorte nie." (bl. 24).

Die tradisionele indeling, op Latynse (en Griekse) lees geskoei, het uit tien rededele bestaan. Gedurende die 20e eeu kom daar egter verset, en onder die uitstaande pogings om 'n nuwe grammatika op te bou, kan ons o.a. die werke van Noreen (Duitse bewerking

---

[6] De Villiers, op. cit., 15—17.
[7] 'n Woord is „meestal die aanduiding van 'n vage kompleks van voorstellinge". (Taal en Taalverskynsels, 90.)
[8], [9], [10] Aangehaal deur De Villiers, op. cit., 6.
[11] Die Skriftelike taaluiting van Skoolkinders, 73—74.
[12] Cf. Barnes: Afrikaanse Grammatika en Taalkunde, 26.

deur Pollak),[13]) Brunot,[14]) Jespersen[15]) en Paul[16]) noem. Nuwe indelings word aangegee o.a. deur Jespersen,[17]) Kalepky[18]) en Langeveld.[19]) In Afrikaans is daar die nuwe indelings van De Villiers[20]) en Barnes.[21]) Sonder om op die meriete van die verskillende skemas in te gaan, wil ek meld dat ek my by mý eie indeling neergelê het en daardie terminologie in hierdie ondersoek gebruik het. Die indeling kan kortliks as volg ondergebring word:

1. Selfstandige naamwoorde, b.v. stoel, waarheid.
2. Beskrywende woorde by s.nwe.:
   A. Byvoeglike naamwoorde, b.v. groot.
   B. Omvangswoorde, b.v. baie (water).
   C. Hoeveelheidswoorde, b.v. ses, sesde.
   D. Lidwoorde, b.v. die.
3. Voornaamwoorde:
   A. Selfstandig, b.v. ek.
   B. Byvoeglik, b.v. my.
4. Werkwoorde, b.v. loop, word.
5. Bywoorde, b.v. more, vorentoe.
6. Tussenwerpsels, b.v. Eina![22])
7. Verhoudingswoorde:
   A. Voegwoorde, b.v. want.
   B. Voorsetsels, b.v. op, in.

(c) *Wat is woordeskat?*

Praat ons van die individu se woordeskat, dan bedoel ons die aantal woorde wat hy ken en waarmee hy hom, òf mondelings òf skriftelik in die daelikse lewe bedien. Die uitslae van ondersoeke om die omvang van so 'n woordeskat vas te stel, is so uiteenlopend dat dit onmoontlik is om gestandaardiseerde norme aan te gee.

---

[13] Wissenschaftliche Betrachtung der Sprache.
[14] La Pensee et la Langue.
[15] The Philosophy of Grammar.
[16] Prinzipien der Sprachgeschichte.
[17] op. cit., 91.
[18] Aangehaal deur De Villiers, op. cit., 21.
[19] Taal en Denken, 54.
[20] op. cit., 25—26.
[21] op. cit., 26—28.
[22] Sommige taalkundiges verwerp hierdie woordsoort, omdat dit as 'n soort oorgang tussen woorde en sinne beskou word (cf. De Villiers, op. cit., 26), maar vir die doel van hierdie ondersoek is die erkenning daarvan onontbeerlik.

Dit wissel van die gewaande 300 woorde van die (Engelse) Somersetse boer[23]) tot waarskynlik 300,000 of meer by geniale individue.

Die spreekwoordeskat van Sweedse boere word op ongeveer 26,000 woorde gestel.[24]) Charlotte Fox, daarenteë, het eksperimenteel vasgestel dat 'n aantal volwassenes (40 — 49 jaar oud) elk oor gemiddeld 144,000 woorde beskik.[25]) Shakespeare se woordeskat van 20,000[26]) is slegs dié wat in sy werke voorkom.

Eintlik word 'n individu se woordeskat deur baie faktore bepaal. „Intelligence, education and socio-economic status form a group of interrelated variables which influence the size of the vocabulary which an individual acquires,'' merk Fox[27]) tereg op. Shakow en Goldman[28]) het 'n korrelasie van ·64 tussen woordeskat en onderwys gevind, en volgens Terman en Merrill[29]) is daar 'n korrelasie van ongeveer ·81 tussen verstand en woordeskat. Sosiale en ekonomiese faktore is deur Schulman en Havighurst[30]) ondersoek, en hulle het vasgestel dat die omvang van die woordeskat tot 'n baie groot mate afhang van die maatskaplike kring waarin die individu hom beweeg.

Dit wil dus voorkom of die individu in baie opsigte deur sy woordeskat gekenskets word, want die grootte van sy woordeskat is tot 'n sekere mate die maatstaf van sy verstandelike ontwikkeling en geestelike rypheid: by benadering kan die peil van sy opvoeding en sy algemene intellektuele en sosiale status daarmee vasgestel word.

In hierdie ondersoek beteken woordeskat die lys nuwe of verskillende woorde wat in die ontlede reekse voorgekom het en opgeteken is.

(*d*) *Frekwensie*(*syfer*) *of* (*verskynings*)*krediet.*

Hierdie terme dui aan: die aantal kere wat 'n bepaalde woord in 'n bepaalde stuk, boek of ontleding voorkom. In „Die man eet die appel en die peer,'' b.v., het *die* 'n frekwensie van 3, *man* het 'n frekwensie van 1, ens.

---

[23] Jespersen: Language, its Nat., Development and Origin, 126.
[24] Jespersen, op. cit., 126.
[25] Fox: Vocab. Ability in later Maturity. (J.E.P., Dl. 38, 486.)
[26] Jespersen, op. cit., 126.
[27] op. cit., 488.
[28] Aangehaal deur Fox, op. cit., 489.
[29] Aangehaal deur Fox, op. cit., 489.
[30] Relations between Ability and Sosial Status n a Midwestern Community. (J.E.P., Deel 38, 442.)

2. Prosedure en tegniek.

(*a*) *Stof*:

Aangesien hierdie lys bedoel is om as hulpmiddel te dien by die taalonderwys op die laerskool, is die volgende leesreekse ontleed:

(I) Die Dagbreek-serie (8 boeke).
(Pienaar, Dippenaar, Laubscher en Joubert — J. H. de Bussy, Pretoria, 1948.)

(II) Die Silwerboom-reeks (9 boeke).
(Heese, Pohl en De Vos — Nasionale Pers, Bpk., Kaapstad, 1949.)

(III) Jong Suid-Afrika (10 boeke).
(Alers — J. L. van Schaik, Pretoria, 1949.)

(IV) Die Springbok-serie (9 boeke).
(Du Toit — Nasionale Pers, Bpk., Bloemfontein, 1948.)

(V) Lentedae Skoolleesboekies (10 boeke).
(Theron, Theron en Groeneweg — Unie-Boekhandel, Bpk., 1946.)

(*b*) *Vasstelling van nuwe woorde* (*die woordeskat*):

Die hoofbeginsel wat hier toegepas is, is dat elke woord wat vir die eerste maal ingevoer is, as 'n nuwe of aparte woord beskou is. Alle soorte rededele is ingesluit, en waar daar die geringste onsekerheid bestaan het, is die woord getoets aan die lys woorde opgeteken in die jongste uitgawe van die Tweetalige Woordeboek van D. B. Bosman en I. W. van der Merwe. Die kriterium was dus of die woord in hierdie boek verskyn of nie.

Die samestelling van die frekwensielys was egter onderhewig aan sekere beperkinge. Die volgende is nie as nuwe (aparte) woorde beskou nie:

(I) meervoude.

(II) verkleinwoorde.

(III) attributiewe vorme van die adjektief met die e-verbuiging.

(IV) reëlmatige vergrotende en oortreffende trappe van vergelyking. Die stellende trap is afgeskryf.

(V) swak verlede deelwoorde (alle funksies). Net die stamvorm is opgeteken, behalwe as die deelwoordvorm in die woordeboek verskyn.

(VI) teenwoordige deelwoorde (alle funksies). (Sien V.)

(VII) gesubstantiveerde b.nwe. (Sien V).

(VIII) telwoorde met *en*, b.v. een-en-twintig. Drie aparte woorde is opgeteken.

(IX) reduplikasies, b.v. spring-spring. Slegs enkeles wat in die woordeboek verskyn, is opgeteken, b.v. gou-gou.

(X) die volgende buigingsvorme van die pers. vnw.: myne, jou(n)e, syne, hare. Die onverboë vorme is opgeteken.

(XI) name van persone, plekke, lande, tale, gelowe, diere (b.v. Wagter), hul afleidinge en samestellinge.

(XII) wisselvorme. Net een vorm is behou.

(XIII) afkortings en voorletters. Afkortings is voluit geskryf.

(XIV) syfers.

(XV) sametrekkinge, b.v. dis, moenie. Die aparte woorde is opgeteken, b.v. dit is, moet nie.

(XVI) vreemde woorde uit ander tale, insluitende Nederlands.

(XVII) toutologiese e.a. samekoppelinge, b.v. kinderlik-naïef, laag-op-die-grondse. Die aparte woorde is opgeteken.

(XVIII) nuutskeppinge (veral samestellinge, samekoppelinge, tussenwerpsels, afleidinge, ens. wat nie in die woordeboek verskyn nie).

(XIX) adjektiewe wat deur middel van funksieverandering as bywoorde gebruik is. Sulke bywoorde, b.v. mooi, oud, groot, ens., is beskou as adjektiewe.

(XX) partikels (b.v. as, toe, al, ens.) met 'n verskeidenheid van betekenisse. Sulke woorde verskyn net een maal.

(XXI) bepaalde rangtelwoorde. Die stamvorm is behou, behalwe *eerste*.

(XXII) In ooreenstemming met die jongste gebruik is samestellinge soos *totsiens*, *so'n*, *byvoorbeeld*, *terwille*, *tweemaal*, *eenmaal*, ens. elk as twee woorde beskou. *Eenmaal* (*eenkeer*) gelyk aan *eendag* is as een woord getel.

(c) *Vasstelling van die frekwensielys:*

Sjablone (wasvelle) is langs en kruis gelinieer om ruite van een duim vierkant te verkry. Daarna is 'n Gestetner-kopieermasjien en afdruk- of kopieerpapier gebruik om die verlangde aantal kopieë af te rol. Op elke afgerolde vel (folio-formaat, 8″ × 13″) was daar presies 104 vierkantjies.

Elke serie is genommer (1, 2, 3, ens.) en apart ontleed. Al die lesse (opskrifte inbegrepe, maar nie begripstoetse nie) in elke boekie van die reeks is woord vir woord op die geruite velle afgeskryf — een woord in elke blokkie — met weglating of wysiging van sekere woorde soos onder 2(*b*) verduidelik. Lang woorde is een of meer maal afgebreek, en net een kant van die vel is gebruik.

Die velle is genommer om by voltooiing van die reeks (d.w.s. die afskryf daarvan) die aantal lopende woorde vas te stel. Dit is gedaan deur die getal velle (bladsye) met 104 te vermenigvuldig. Honderd velle, b.v. gee presies 10,400 lopende woorde. Deur optelling van die frekwensiewaardes van al die nuwe woorde aan die einde van die ontleding (van die reeks) is die getal lopende woorde weer uitgewerk. As die twee getalle klop, dan kon daar vasgestel word dat daar geen papiertjies verlore gegaan het nie.

By die afskryf van die woorde uit 'n reeks is die betekenisse van homonieme en die funksies van sekere woorde (b.v. *boer*: s.nw. of ww.; *skelm*: s.nw. of b.nw.) opgeteken. Voor die opsnippering is die velle weer 'n keer deurgelees om hierdie betekenisse en funksies te kontroleer.

Nadat hierdie werk (die afskryf en kontrolering van die woorde) voltooi is, is elke vel al langs die gesjabloneerde lyne met 'n skêr opgesnipper om 104 los papiertjies van 1 duim vierkant te kry.

22 skoendosies is gemerk met die letters van die alfabet. Op 21 dosies het net een letter verksyn. Op die 22e is die letters C, Q, X, Y en Z geskryf. Die 104 los papiertjies van elke vel is volgens beginletter gesorteer en in die korrekte dosies gegooi. Alle woorde, b.v., wat met *a* begin, kom in die A-dosie, ens.

Daarna is die inhoud van die dosies een vir een ontleed. Deur gebruik te maak van Aucamp se lys kon 'n hele paar woorde met waarskynlik die hoogste frekwensies eerste uitgesoek word. Eindelik het daar net woorde met 'n frekwensie van 1 oorgebly.

As daar in die A-dosie, b.v., 200 papiertjies met die woord *aan* gevind word — hulle is nagetel totdat twee tellings klop — dan word

die syfer 200 op een van die papiertjies geskryf om die frekwensie van die woord in daardie telling aan te dui. Hierdie papiertjie word in 'n aparte dosie bewaar, terwyl die ander weggegooi word. Al die woorde in die A-dosie is so vasgestel.

Op hierdie wyse is die woordeskat van 'n reeks vasgestel, want na die ontleding was daar net een papiertjie vir elke nuwe woord met die verskyningskrediet of frekwensiesyfer daarop.

Elke (nuwe) woord is op 'n kaartjie (5″ × 3″ groot) getik of geskryf. Onder die woord is ook die nommers van die reekse getik (geskryf), en na elke ontleding is die frekwensie van die woord naas die betrokke of ooreenkomstige nommer geskryf. Op dié manier kon die frekwensie van die woord in al die reekse saam bepaal word. Die verspreidingsfaktor is by die vasstelling van die eindfrekwensie buite rekening gelaat. Die homogene aard van die stof het dit nie nodig gemaak nie.

Na die eerste ontleding is die kaartjies alfabeties gerangskik. Nuwe woorde wat later (in ander reekse) verskyn het, is ook op kaartjies getik, en hierdie kaartjies is by die ander gevoeg.

By voltooiing van die ondersoek (d.w.s. na ontleding van al die reekse) kon die grootte van die woordeskat en die frekwensies van al die (nuwe) woorde in al die reekse gesamentlik bereken word. Telling van al die kaartjies het die woordeskat aangegee en optelling van die getalle op elke kaartjie het die frekwensie van daardie woord aangedui.