

troniese taaldatabasis. Hulle hoop om daardeur soektogte, kontrolering, ordening, korrekte en eenvormige definiëring te vergemaklik, taalnavrae noukeuriger en vinniger te beantwoord en veral sleurtake uit te skakel. Wat laasgenoemde betref, meld De Tollenaere (1963: 123) al dat die rekenaar "ons ... een groot gedeelte van het vervelende en weinig interessante werk uit handen (neemt)". Aitken (1978: 29) maak 'n soortgelyke opmerking.

Daar is egter bykomende redaksionele voordele verbonde aan rekenarisering, soos uit die ervaring van onder meer Aitken (1978: 29 e.v.), Jackson (1988: 236 e.v.), Van Sterkenburg (1984: 115 e.v.) en veral Landau (1984: 276 e.v.) blyk en wat vervolgens saamgevat word. Hierdie ervaring van ander toon ook duidelik dat die voordele van rekenarisering die gemelde behoeftes van die Buro in 'n ruim mate bevredig.

- (i) Groot elektroniese tekskorpusse kan op die skerm geëkserepeer word en bestaande rekords wat meer as een nuttige leksikale item bevat, kan outomaties gedupliseer word om betroubare en verteenwoordigende materiaal te bekom waarop die woordeboek gebaseer kan word.
- (ii) In die take waarvoor die rekenaar geskik is, lewer dit baie akkurater, betroubaarder en deurlopende diens as selfs die gewilligste menslike werker.
- (iii) Die snelheid van die rekenaar lei tot tyd-, personeel- en geldbesparing, wat groter toespitsing op die primêre taak moontlik maak.
- (iv) Die taaldatabasis en die voltooide artikels kan baie toegankliker wees as wat die redaksie aanvanklik vermoed het. Die snelheid en verskeidenheid van onttrekkingsmoontlikhede word vir alle praktiese doeleindes slegs bepaal deur die deeglikheid waarmee die databasisstruktuur beplan is en die mate waarin die data volgens hierdie struktuur in die databasis opgeneem word. Die rekenaar se vermoë om ook soektogte in ongestruktureerde data te kan uitvoer, moet as 'n bonus slegs vir uitsonderlike toepassing beskou word.
- (v) Die metataal in die betrokke woordeboek kan met die rekenaar nagegaan word om te verseker dat dit verklaar of ten minste verantwoord word.
- (vi) Die rekenaar kan o.a. foutiewe kruisverwysings, weglatings en onlogiese nommering kontroleer.
- (vii) Nuwe outomatiese ordeningsbeginsels is nou moontlik, bv. volgens die omgekeerde alfabet, volgens woordsoort, outeur, bron, datum, plek van optekening, getal opgawes of 'n kombinasie hiervan.
- (viii) Die rekenaar maak dit moontlik om op 'n maklike wyse inkomste te genereer deur die lewering van dataonttrekkingsdienste aan ander instansies.
- (ix) Dieselfde geld vir die samestelling van kleiner satellietprojekte van die Buro. Die WAT sou bv. maklik verkort kon word tot 'n meer

praktiese algemene handwoordeboek. Ook kan alle woordeboekinligting in vaste velde ingeskryf of op ander wyses gekodeer word, bv. deur rekenaarontleding in die geval van bestaande woordeboekteks, en dan later daarvolgens onttrek word, o.a. vir die maak van tematiese woordeboeke. Tipiese inligtingsvelde sou kon wees dié vir woordsoort, morfologie, gebruiksfrekwensie, spesialiteitsvlak (vaktaaletiket), chronolektiese afwyking van die standaardtaal (temporele etikette), streeks- of gebiedsverankertheid (geografiese etikette), formaliteits- of stylregister (stilistiese en sosiostilistiese etikette) en betekenis.

- (x) Die algemene versnelling van die werktempo en die uitskakeling van vervelende roetinerwerk kan lei tot groter werksbevreëdiging.
- (xi) Die tydperk tussen redaksionele bewerking en publikasie kan verkort word omdat alle korreksies by proefleeswerk, hetsy in gedrukte vorm of op die skerm, regstreeks op die rekenaar aangebring word.
- (xii) Direkte rekenaaroordrag van manuskrip na die drukafdeling sal moeisame en tydrawende setwerk uitskakel, asook die insluiping van foute wat daarmee gepaard gaan.
- (xiii) Opdatering en bywerking van elektroniese manuskrip is nie so 'n omvattende en duur taak soos in die geval van handsetwerk nie.
- (xiv) Die WAT-kan deels as kantoor- of tafelpublikasie gehanteer word waardeur die Buro aansienlike besparings sal bewerkstellig.
- (xv) Vakkundige en administratiewe skakeling met ander instansies per modem en selfs per satelliet is moontlik.
- (xvi) Die opleiding van nuwe personeel kan vergemaklik word deur die gebruikmaking van algemeen beskikbare of eie opleidingsprogrammatuur.
- (xvii) Manuskrip van die redakteurs sal altyd ordelik en goed leesbaar wees.
- (xviii) Nuwe teksverwerkingsmoontlikhede staan tot die personeel se beskikking, bv. soek-, sorteer-, venster-, raamwerk- en stylbladfasiliteite, asook die moontlikheid om teksgedeeltes na willekeur binne die dokument en van dokument tot dokument te verskuif. Redigeerteks kan op 'n onderskeibare wyse aangebring en na keuse weggelaat of in die bestaande teks opgeneem word. Dit is moontlik om verskillende dokumente met mekaar te integreer. Korreksies, dikwels selfs deur 'n program aanbeveel, en tipografiese verstellings kan vinnig dwarsdeur 'n dokument aangebring word, bv. om die spelling in ooreenstemming te bring met die jongste Taalkommissiebesluite. Ook kan rekenaartesourusprogramme die redaksie tydens definiëring help deur beter woordkeuses voor te stel.
- (xix) Afhangende van die programkeuse kan etlike spesialisredakteurs gelyktydig aan verskillende inskrywingsvelde in dieselfde artikel werk, terwyl elkeen bewus is van wat die ander doen.

- (xx) Akkurate statistiese berekeninge van taaldatavoorraad en werksuïtsette kan gemaak word. So kan bv. beheer uitgeoefen word oor die grootte van en die balans tussen woordeboekartikels, asook oor die totale projekomvang. Hierdie beheer maak dit makliker om 'n teikendatum vir publikasie vas te stel en die stimulering van die kopersmark presies daarop af te stem. Ook kan maklik vasgestel word of die redaksie leksikale items op 'n verteenwoordigende wyse uit bv. die streek- en vaktaal opgeneem het.

Aitken (1978: 29 e.v.), Jackson (1988: 236 e.v.), Van Sterkenburg (1984: 115 e.v.) en Landau (1984: 276 e.v.) noem ook enkele nadele verbonde aan die rekenaar wat beslis in gedagte gehou sal moet word. Die hoofbesware wentel om die neiging om die databasis deur outomatiese duplisering en intensiewer ekserperering onhanteerbaar groot te laat uitdy. Landau (1984: 287) noem verder ook die nadeel dat die spesialis-datatikker nou die werk van die gewoonlik opgevoede en wydbelese setter of vormopmaker doen en nie foute opmerk nie. Dit bring mee dat proefleeswerk op die redaksie neerkom. Hierdie nadeel is nie op die Buro van toepassing nie aangesien in elk geval slegs die redaksie vir die proefleeswerk verantwoordelik is. Regstreekse dataoordrag na die reprografiese afdeling bring mee dat nóg beroepsetters nóg onervare datatikers in die toekoms met die manuskrip sal omgaan. Die rekenaar het hier juis die voordeel dat foute net een keer reggemaak hoef te word.

4. *Voordele van rekenarisering vir die gebruiker*

Die voordele van rekenarisering vir die Buro sal bykans almal op die woordeboekgebruiker afgewentel word. Hy sal baie gouer en baie goedkoper 'n meer wetenskaplike, verteenwoordigende, moderne, eenvormige en afgeronde WAT tot sy beskikking hê. Vergelykbare terminografiese instellings sal toegang tot die Buro se uitgebreide tegniese databasis kan geniet (Swanepoel en Morris 1988: 108), wat uiteindelik ook tot voordeel van die gebruiker sal wees.

Veral die beskikbaarstelling van die woordeboek op ander wyses as in gedrukte vorm, bv. op laserskyf, hou vir die woordeboekgebruiker nuwe moontlikhede in:

- (i) Die woordeboek in masjienleesbare vorm bied aan die gebruiker die meeste van die vinnige en gevarieerde datasoek- en dataonttrekingsmoontlikhede waaroor die redaksie self beskik (Gray 1986: 48).
- (ii) 'n Groot aantal gebruikers in 'n korporatiewe instelling soos 'n universiteit kan baie maklik deur middel van 'n rekenaarnetwerk regstreekse toegang tot 'n enkele eksemplaar van 'n woordeboek op laserskyf kry (Jackson 1988: 237).

- (iii) Dit is makliker en goedkoper om die woordeboek op laserskyf deurlopend by te werk en intekenaars op 'n inruilbasis van die jongste weergawe te voorsien as in die geval van die gedrukte weergawe (Jackson 1988: 237).
- (iv) Gebruikers soos navorsers het onmiddellike en meestal ook gelyktydige skermtoegang tot meer as een naslaanwerk in hierdie vorm. Dit stel 'n kennisbron wat potensieel gelykstaan aan 'n hele biblioteek tot hulle beskikking (Jackson 1988: 237).
- (v) Die ontwikkeling van die herskryfbare optiese skyf sal ook lei tot die era van die interaktiewe groot woordeboek (Jackson 1988: 237). Dit beteken dat die gebruiker so 'n woordeboek met nuwe of ontbrekende leksikale items kan aanvul, soos hy dit reeds kan doen met die kleiner spesialiswoordeboeke en -lyste wat deel uitmaak van meer gevorderde woordverwerkingsprogramme.

5. *Stand en resultate van fase I*

Die eerste fase van die Buro se rekenariseringsproses, wat van 1987 tot 1989 gestrek het, het hoofsaaklik om *personeelaanpassing* by bestaande kommersiële programme en die benutting van die rekenaar as *woordverwerker* gehandel. Aangesien hierdie aanwending van die persoonlike rekenaar aan die meeste rekenaargebruikers goed bekend is, word die eerste fase hier nie nader toegelig nie. Wel word 'n kort uiteensetting gegee van die rekenaarhulpmiddels waarvoor die Buro aan die einde van die fase beskik het en watter resultate daarmee behaal is.

Die volgende apparatuur en programmatuur is reeds aangekoop en in bedryf gestel:

- 15 losstaande persoonlike rekenaars, elk met 640 KG interne geheue en twee 360 KG-skyfaandrywers;
- 2 20 MG-hardeskywe vir die rekenaars in die tiksterskantoor;
- 3 40 MG-hardeskywe, twee vir die rekenaars in die taalpraktisynskantoor en een in die rekenaar van een van die redakteurs;
- 2 drukkers;
- DOS 3.2-bedryfstelselprogramme;
- Microsoft Word 4.0-woordverwerkingsprogramme (vergelyk Hartevelde 1988: 22-23 vir die motivering van die keuse van hierdie program);
- ameublement, soos rekenaar- en drukkertafels; en bedryfsmateriaal, soos data-skywe en drukkertoehore.

Die personeel, wat onmiddellik oortuig was van die voordele van rekenarisering, het gou rekenaarvaardig geraak, die nuwe woordverwerkingsmoontlik-

hede deeglik bemeester en benut, en is nog steeds entoesiasies oor die voortsetting van die proses.

Aan die einde van die eerste fase is alle manuskrip op afsonderlike persoonlike rekenaars met gebruikmaking van woordverwerkingsprogrammatuur geskep. Daarbenewens is die rekenaar ook vir alle korrespondensie, bronne-lyshantering en vir beperkte toegang tot die redaksionele besluiteregister gebruik.

6. *Beplanning en inwerkingstelling van fase II*

Die tweede fase, wat van 1990 af verder strek en waarmee reeds ver gevorder is, behels veral *programaanpassing* by die personeel se behoeftes, die uitbreiding van die rekenaarstelsel, die skepping van 'n elektroniese taaldatabasis en die optimale aanwending van die rekenaar in die leksikografiese praktyk. In hierdie fase moet die rekenaar dus al die prosesse van data-invoer tot die finale druk van die woordeboek behartig.

Intussen, op 29 Mei 1989, het die Buro van die WAT opdrag van sy Beheerraad gekry om 'n deeglike omvattende strategiese beplanning te doen en daarin ook besondere aandag te gee aan die rol wat rekenarisering kan speel in die bespoediging van die Buro se werksaamhede. Hierdeur is die ideale geleentheid geskep om die verdere rekenariseringsproses en die toekomstige redaksionele werkwysse by mekaar aan te pas. Teen die einde van die jaar het daar 'n verslag oor hierdie beplanning verskyn (Redaksie, Buro van die WAT 1989).

Tydens die strategiese beplanning is die volgende as die dringendste rekenariseringsbehoefte aangedui: 'n geselekteerde, gekontroleerde, prosesseeerbare elektroniese databasis; regstreekse rekenaaragtige toegang tot hierdie databasis vir redaksionele bewerking; regstreekse elektroniese oordrag van manuskrip na die drukafdeling; interne bladuitleg en drukwerk met optimale gebruikmaking van die rekenaar, verkieslik in die vorm van 'n kantoorpublikasie (Redaksie, Buro van die WAT 1989: 115 e.v.).

Latere behoeftes is aangedui as: die aanwending van die rekenaar by taaldiensverskaffing en vir leksikografiese en leksikologiese navorsing, rekenaar-skakeling met ander taalinstellings, 'n effektiewe gerekenariseerde bestuursinligtingstelsel en die uitgee van die WAT op laserskyf (Redaksie, Buro van die WAT 1989: 115 e.v.).

Teen die einde van dieselfde jaar is daar ook 'n uitvoerige verslag uitgebring oor hoe die Buro in sy voortsetting van die rekenariseringsproses kan inskakel in die struktuur van die RGN (LEKSENA) se LEXI-projek (Harteveld en Nieuwoudt 1989). 'n Sterk oorweging was hier die standaardisering van gerekenariseerde leksikografiese praktyke in Suid-Afrika met die oog op onderlinge versoenbaarheid en die wyer beskikbaarstelling van taaldata.

Die Administrateur van die Buro het die verslag namens die Beheerraad goedgekeur. 'n Uitvloeisel daarvan was dat die RGN mnr. B.A. Nieuwoudt, mede-ontwikkelaar van hulle *LEXI*-databasisprogram, vir 'n tydperk as programmeerder aan die Buro gesekondeer het. Sy hooftake was om hierdie program by die Buro se behoeftes aan te pas en die personeel by te staan by die aanleer en die gebruik daarvan. Ook was hy verantwoordelik vir verskeie ander programmeringstake, bv. in verband met dataoordrag van die databasis na die woordverwerkingsomgewing en die skep van 'n bibliografiese databasisprogram ter ondersteuning van die leksikografiese databasisprogram. Verder het hy hulp verleen met die keuse en inbedryfstelling van 'n rekenaarnetwerkstelsel en enkele kommersiële ondersteuningsprogramme vir die Buro.

Die meeste van hierdie take is voltooi. Die Buro beskik oor 'n eie rekenaarnetwerk (SUNWAT) in die vorm van die *Ethernet*-stelsel, 'n kragtige 386-netwerkbediener, *Novell NetWare*-programmatuur, en hardeskywe voorlopig met 'n totaal van 720 MG berguimte. Hiermee beskik die Buro oor een van die modernste en kragtigste lokale-areanetwerke op die mark en het die personeel nie net gelyktydige toegang tot dieselfde databasis en programme nie, maar kan data onderling tussen die werkstasies en desgewens met ander aangeskakelde netwerke uitgeruil word.

Aangesien die Buro van die WAT die eerste en die grootste professionele gebruiker van *LEXI* is en inderdaad meegehelp het om dit aan te pas en te verfyn, is dit wenslik om hier kortliks te motiveer waarom hy hierdie program en nie een van die ander pakkette op die mark gekies het nie:

- (i) *LEXI* is baie goedkoper as bv. Oxford University Press se *OEDIPUS*-program of sy nuwe *Open Text*-stelsel.
- (ii) Dit is die enigste plaaslike program wat dit vir 'n leksikografiese onderneming moontlik maak om altyd onmiddellike beheer oor en toegang tot sy taaldatabasis te hê vir deurlopende opdatering en onttrekking. Hierdie konsep is al in 1981 deur Oitana (1982: 48) as die mees praktiese en ekonomiese aanbeveel en intussen deur die Buro as geldig bewys. Dit is dan ook die hoofoorweging waarom nie van die Nasionale Vakterminologiesediens (NVD) se *Lexikon*-program gebruik gemaak is nie. Die NVD se lofwaardige ideaal van 'n gesentraliseerde nasionale taaldatabasis word geensins hierdeur in die wiele gery nie omdat die Buro se taaldata nog steeds op 'n beheerde wyse en met gebruikmaking van 'n oorbruggingsprogram tot die NVD se beskikking gestel kan word.
- (iii) *LEXI* is 'n basiese program wat by die spesifieke leksikografiese behoeftes van die gebruiker aangepas kan word (Alberts 1991: 48 e.v.). So 'n aanpassing is inderdaad vir die Buro gedoen en het meegebring dat die redaksie in 'n groot mate sy eie beproefde werkwyse kan

- volg. Die mededingende programme daarenteen dwing bykans elkeen sy eie konsep en metodiek aan die gebruiker op.
- (iv) *LEXI*, wat 'n *dBASE IV*-toepassingsprogram is, gebruik 'n datastruktuur en programmeertaal wat as bedryfstandaard beskou word (Alberts 1991: 48). Daarom is dit maklik aanpasbaar by kragtiger programmatuurstelsels, selfs buite die *Microsoft DOS*-omgewing. Die feit dat *dBASE* deurlopend verbeter word en ten spyte van die baie sterk mededinging op die terrein van kommersiële databasisprogramme nog steeds die grootste gewildheid geniet, vergroot egter die waarskynlikheid dat *LEXI* met die Buro se behoeftes tred sal kan hou en so 'n verandering van program nie nodig sal wees nie.
- (v) Dit is tweekantig en besonder gebruikersvriendelik. Verder loop dit op sowel 'n selfstandige persoonlike rekenaar as 'n rekenaarnetwerk en bied dit die fasiliteit dat meer as een gebruiker tegelyk in die program kan werk. Hierdie veelgebruikerfasiliteit is van die uiterste belang vir die Buro wat sy woordmateriaal op 'n groot en intensiewe skaal wil rekenariseer.

Aangesien *LEXI* aanvanklik nog deels in sy ontwikkelingsfase was en dit bovendien by die Buro se veelvuldige vereistes aangepas moes word, was dit te verwagte dat daar nog leemtes en foute sou wees. Dit was inderdaad die geval, maar die probleme is intussen almal reggestel.

Die Buro gebruik die *LEXI*-program tans vir langer as 'n jaar in sy afdeling Inligtingsteundiens. Daar lewer dit steeds uitstekende diens en voorsien in al die huidige datavasleggings- en dataonttrekkingsbehoefte.

Vervolgens sal beskryf word hoe die Buro se taalmateriaal, wat uit verskillende oorde kom en in verskillende vorms gehou word, in die rekenaardatabasis opgeneem en dan verder verwerk word.

7. Oordrag van nuwe woordmateriaal in *DATABANK I*

Sedert einde Januarie 1990 word geen nuwe woordmateriaal meer in kaartvorm geliasseer nie maar direk op die rekenaar in *DATABANK I*, die databasis vir onverwerkte taaldata, ingesleutel. Hiervoor word die programweergawe *LEXI-WATI*, wat vir die Buro van die *WAT* ontwikkel is, gebruik. Dit is bedoel vir datavaslegging, -verwerking en -afvoer. Die elektroniese data in *DATABANK I* word onmiddellik gekontroleer vir korrektheid en is dan gereed vir rekenaarmatige onttrekking en leksikografiese verwerking. Tydens 'n proeflopie op 'n 386-rekenaar was *LEXI-WATI* in staat om tien soektogte binne een sekonde deur al die lemmavelde van 250 000 rekords uit te voer. Voorlopig is dit vir die Buro vinnig genoeg.

Alberts (1991: 60) gee 'n afbeelding van die eenvoudige *LEXI*-rekenarskerm wat ook deur die Buro van die WAT vir die invoer van hierdie taaldata gebruik word.

8. Oordrag van bestaande kaartmateriaal in *DATABANK I*

Die grootste en duurste rekenariseringstaak wat vir die Buro voorlê, is die beskikbaarstelling in elektroniese vorm van sy bestaande materiaalversameling op kaart. Die omvang van die vergelykbare oordragtaak in die geval van die *WNT* het die rekenarisering van daardie projek bv. laat skipbreuk ly (Van Sterkenburg 1976: 88 en 1984: 75). Die onvolledigheid en ongekontroleerdheid van die materiaal, die gebrek aan veral ouer bronne vir die kontrolering van die optekeninge, en selfs, soos in die geval van die *WNT*, die ongelukkige feit dat reeds tot manuskrip verwerkte kaartmateriaal vernietig kan wees (Van Sterkenburg 1984: 84, 117), is maar enkele van die probleme wat die omvang van so 'n taak kan bepaal.

Die Buro het egter geredeneer dat die vordering met die *WAT* regstreeks afhanklik is van die betroubaarheid en toeganklikheid van sy basismateriaal. Deur kontrolering en maksimale rekenarisering van hierdie materiaal kan die betroubaarheid verseker en die toeganklikheid en aanwendbaarheid drasties vergroot en versnel word. Om hierdie rede is die rekenarisering 'n onvermydelike taak wat met die grootste spoed afgehandel moet word.

Die haalbaarste en effektiëste metode daarvoor blyk direkte insleuteling te wees. Daarvoor moet die nodige apparatuur en personeel bekom word. Die volgende tydelike personeel word bv. benodig:

- Kaartselekteerders**, wat die materiaal vooraf selekteer en sorg dat o.a. die vorm van die lemma en woordsoortaanduiding korrek is;
- Materiaalkontroleerders**, wat die materiaal voor insleuteling aan die bronne kontroleer vir korrektheid;
- Datatikkers**, wat die insleuteling op rekenaar doen; en
- Korrektors**, wat die korrektheid van die insleuteling in die databasis kontroleer.

Datatikkers van die North American Reading Program for the OED, wat die meeste van hierdie personeelfunksies vir die Oxford University Press (OUP) verrig het, het optekeninge teen 'n tempo van gemiddeld 4 000 per persoon per maand ingesleutel (*DSNA Newsletter* 1990, 14(1): 4). Vergelyk hier ook die insiggewende statistiek van Berg et al (1988: 3 e.v.), Shenker (1989: 99) en Van Sterkenburg (1984: 122). Hieruit kan 'n indruk gevorm word van die tyd en koste wat aan die werk verbonde sal wees.

Alhoewel die Buro oor al die tegniese hulpmiddele vir kleinskaalse oordrag van die bestaande kaartmateriaal na die rekenaardatabasis beskik, ont-

breek dit hom aan die personeel, apparatuur en fondse om die 3,25 miljoen materiaalkaartjies in sy versameling op hierdie wyse en binne die gewenste tydsbestek te hanteer. Om dié rede kon daar nog nie sistematies met die taak begin word nie.

9. *Oordrag van die voltooide WAT-dele in databank II*

Die voltooide en gepubliseerde dele bevat 'n magdom geselekteerde en gekontroleerde data wat egter nog nie in elektroniese vorm toeganklik is nie.

Die koste-effektiefste wyse van oordrag na dié medium is om die dele bladsy vir bladsy met 'n opties-elektroniese teksleser af te tas (Van Sterkenburg 1984: 122). Daarom het die Buro 'n kragtige 386-rekenaar, 'n Hewlett Packard ScanJet- optiese karakteraftaster en die OmniPage-karakterherkenningsprogram aangeskaf. Deel I (A - C) van die WAT is gedeeltelik en deel VII (KOM - KOR-) volledig in die rekenaar ingelees deur effens vergrote fotokopieë van die oorspronklike gedrukte bladsye in kolomme uit te knip en dit dan af te tas. Die resultate was bo verwagting goed. Gedeeltes van hierdie teks is reeds op rekenaar gekontroleer vir korrektheid.

Die volgende stap is om hierdie data so ver moontlik outomaties met behulp van 'n rekenaarprogram te ontleed en verder met die hand op die rekenaar te struktureer volgens die *LEXI-WAT II*-veldstruktuur. Die resultaat en nut is wesenlik dieselfde as by die maak van nuwe manuskrip binne hierdie struktuur, soos in paragraaf 10 hieronder beskryf word.

10. *Redaksionele bewerking en DATABANK II*

Dit sou ideaal wees indien die program waarmee die taaldata in die databasis opgeneem en verder verwerk word, ook gebruik kon word vir die regstreekse verwerking daarvan tot woordeboekteks (Tompa en Raymond 1989: 14 e.v.). Dit was aanvanklik ook die voorneme met *LEXI*. Aangesien *dBASE IV* se woordverwerkingsfasiliteite egter nie naastenby vergelykbaar is met dié waarvan die redaksie in *MS Word 5* tydens fase I van die rekenariseringsproses gewoond geraak het nie, is die beoogde *LEXI-WAT II*-program vir die totstandbrenging van die woordeboekteks laat vaar ten gunste van 'n baie praktiese en suksesvolle oplossing: die Buro se redaksionele veldstruktuur (die struktuur waarvolgens al die inligtingstipes in 'n WAT-artikel binne ooreenstemmende inskrywingsvelde ingetik word) is in die *MS Word 5*-woordverwerkingsprogram geredupliseer. Ten einde dit van die programstruktuur van *LEXI-WAT I* te onderskei, word dit die *LEXI-WAT II*-veldstruktuur genoem.

Die gerekenariseerde proses van manuskrip maak verloop dus soos volg: Onverwerkte data in *DATABANK I* word na behoefte met *LEXI-WAT I* geselekteer en daaraan onttrek. 'n Spesiale program dra dit daarna regstreeks en

met behoud van alle teksformatering na die woordverwerkingsomgewing oor. Die redakteurs verwerk dit dan binne die *LEXI-WAT II*-veldstruktuur in *MS Word 5*. Hierdie veldstruktuur word met die druk van 'n eenvoudige kode opgeroep sodat dit sigbaar op die skerm verskyn. Die gegewe volgorde van die inligtingsvelde binne die struktuur word streng gehandhaaf. Deur die veldstruktuur met die druk van 'n verdere kode in verskuilde teks om te sit terwyl die woordeboekteks in leesbare modus bly, kan die redakteur 'n indruk kry van hoe die artikel in sy finale gedrukte vorm sal vertoon. Die manuskrip maak nou deel uit van *DATABANK II*, vanwaar dit bv. na die reprografiese afdeling afgevoer kan word.

Hierdie ontwikkeling maak dit vir die redaksie moontlik om die woordeboekmanuskrip in 'n bekende omgewing en met gebruikmaking van die program se voortreflike woordverwerkingsfasiliteite tot stand te bring. Na verwagting sal dit die redaksionele produksietempo aanmerklik versnel omdat onverwerkte data nou nie meer met die hand oorgeskryf of oorgetik hoef te word nie maar elektronies en sonder die insluiping van foute vanuit *DATABANK I* in *LEXI-WAT II* ingetrek kan word.

Redaksionele manuskrip wat so tot stand kom, maak dadelik deel uit van die sogenaamde *DATABANK II*. Uit hierdie databank kan met gebruikmaking van enige versoenbare databasisprogram ook ander onttrekkings as vir reprografie gedoen word, bv. vir die lewering van taaladvies na buite of vir voorsiening van sekere materiaal en inligting aan die redaksionele komponent.

11. *Uiteensetting en toepassing van die LEXI-WAT II-struktuur*

Ten einde die vordering met die *WAT* maksimaal deur gebruikmaking van die rekenaar te versnel, moes benewens die rekenarisering van nuwe en bestaande onverwerkte taalmateriaal en van die gepubliseerde *WAT*-dele, ook sekere van die gevestigde redaksionele praktyke drasties hersien word. Hierdie hersiening het o.a. betrekking op die herorganisering van die tipiese *WAT*-artikel om dit volledig rekenaarmatig te kan bewerk en die grootste hoeveelheid informasie daaraan te kan onttrek. Aan die ander kant moes die inskrywingsvelde van *LEXI-WAT II* gestruktureer word om by die redaksionele behoeftes aan te pas.

11.1 *Uiteensetting van die voorgestelde LEXI-WAT II-veldstruktuur*

Vervolgens word 'n struktuur vir die totstandbrenging van *WAT*-manuskrip op rekenaar asook vir die effektiewe, meerdoelige, rekenaarmatige onttrekking van data daaraan voorgestel. Die struktuur is bepaal deur 'n aantal tipiese artikels van uiteenlopende aard in *WAT* deel VII (Hauptfleisch 1984) en deel VIII (Hauptfleisch 1991) te ontleed met die oog op die tipe inligting wat daarin aangebiet word asook op die gebruiklikste en vir die redakteur en gebruiker die

mees praktiese volgorde waarin dit aangebied word. Die rekenaar stel streng eise wat die logiese en konsekwente hantering van sekere inligtingstipes en die hiërgargiese strukturering van betekenisonderskeidings betref, en dit het tot heelwat noodwendige veranderinge in die redaksionele stelsel gelei. 'n Voorbeeld hiervan is die homoniemnommer wat nou voor in plaas van na die lemma gegee moet word omdat dit anders soms tussen twee uitspraakvariante, bv. *kruifery*¹, *kruierery*¹, te staan sou kom en die lemmaveld dus op ongeoorloofde wyse in twee sou verdeel.³

Ofskoon die voorgestelde struktuur voorlopig so deur die redaksie toegepas word, is daar nog nie amptelik finaal tot die presiese model besluit nie. Dit staan die Buro dus vry om die struktuur te verwerp of na behoefte aan te pas.

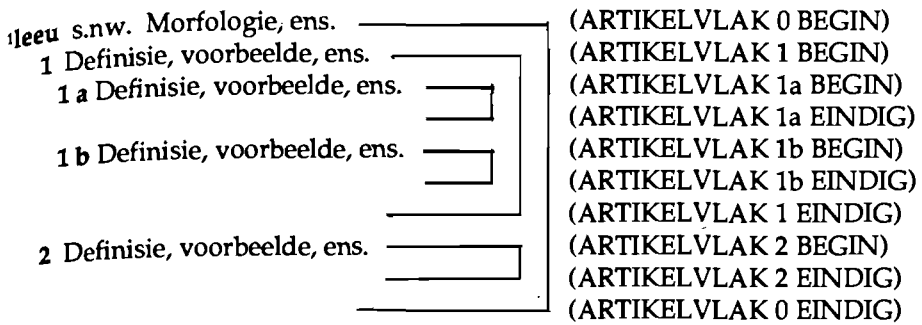
Vir doeleindes van die rekenaar word alle velde begrens deur 'n veldopening en 'n veldafsluiting tussen skerp hakies. Die veldopening en die veldafsluiting word deurgaans van mekaar onderskei deur die skuins strepie wat na 'n bepaalde kant toe hel, bv.:

- (1 α) <ARTIKELVLAK ?/>
 (2) <BETEKENISNOMMER /><BETEKENISNOMMER \>
 (3) <HOMONIEMNOMMER /><HOMONIEMNOMMER \>
 (4) <LEMMA /><LEMMA \>
 (1 Ω) <ARTIKELVLAK ?\>

Die teks word tussen die binneste twee skerp hakies ingesleutel.

Uiteengeplaaste veldgrense soos onder (1 α) en (1 Ω) hier bo dui aan dat (1 α) en (1 Ω) alle velde tussen hulle omvat, met ander woorde, dat al daardie velde (met hulle inhoud) in hierdie ARTIKELVLAK-veld (1 α)-(1 Ω) genes is. Die hele veldstruktuur is trouens op hierdie hiërgargiese of nestingsbeginsel gebaseer. Dié beginsel kom daarop neer dat 'n bepaalde veld een of meer ander kan omsluit, wat ook meebring dat die hele struktuur al dieper in homself herhaal kan word. Hierdie herhalende onderverdeling blyk die duidelikste uit die volgende vereenvoudigde skematiese voorstelling:

³ Die outeur se opregte dank aan mnr. B.A. Nieuwoudt vir sy onderrig in die toepassing van die nestingsbeginsel en sy wenke vir die rekenaarmatige strukturering van die WAT se inligtingstipes. Die volgende opmerking van Landau (1984: 276) is hier toepaslik: "The editor often begins these lengthy discussions with the idea that because the computer analysts are ignorant of dictionaries he must go through a tedious period of instructing them; but after several days of talks, he realizes with a jolt that he is the one being instructed, and not just in computer logic but in the logic of his own lexicographic ideas. It is humbling yet in the long run a beneficial exercise."



Ter wille van ruimtebesparing en oorsigtelikheid word alle velde en subvelde slegs deur die tiperende naam, dus enkel en sonder die hakies en skuins strepies, aangedui. Die Buro se afgekorte veldname word gebruik. Waar hulle nie-erkende afkortings bevat of nie deursigtig is nie, word die vol vorm tussen hakies verstrek. Die woordeboekgebruiker sien hierdie veldstruktuur nooit as sodanig in die uiteindelige woordeboek nie, maar wel die inhoud van die velde, en dan ook maar ten dele. Daarom word die velde hier vanuit die oogpunt van die leksikograaf uiteengesit, met ander woorde, daar word kursories aangedui watter soort inligting 'n redakteur van die Buro daarin moet verstrek in die proses van manuskrip maak. Daar word nie orals voorbeelde gegee of presies gewys hoe die inligting vormlik aangebied word nie.

A. VASTE VELDE (PLEKGEBONDE BINNE DIE STRUKTUUR)

(1α) ARTIKELVLAK

Die begingrens van die veld wat die logiese vlak van die hele artikel of van 'n artikelafdeling, bv. die betekenisonderskeiding, aandui. Die nommer wat in verskuilde teks in hierdie begingrens (die artikelvlak-opening) aangebring word, moet nie met die reële, gedrukte nommer in die BETEKENISNOMMER-veld verwar word nie. Die nommering van 'n artikel begin altyd by 0. Slegs wanneer daar verskillende betekenisonderskeidings voorkom, word subnommers in hierdie veld aangegee, streng volgens die nestingsbeginsel soos hier bo uiteengesit is.

(2) BETEKENISNOMMER

Die reële betekenisnommer, soos dit gedruk moet word. Dit korreleer logies maar nie tipografies nie met die nommer by ARTIKELVLAK. Die logiese artikelvlak aanduiding 0 kry nooit die reële teenhanger 0 in die BETEKENISNOMMER-veld nie.

Betekenisonderskeidings word alfanumeries aangedui, bv. I en 1 a i. Betekenisstoepassings (of gebruikstoepassings) word konsekwent van

- eg. onderskei deurdat dit met hakies om elke eenheid van die nommer gegee word, bv. (1)(a)(i).
- (3) **HOMONIEMNOMMER**
Homoniemaanduiding in die vorm van 'n verhewe syfertjie voor die lemma.
- (4) **LEMMA**
Definiendum (leksikale, subleksikale of multileksikale item) met toepaslike klemaanduiding.
- (5) **SORTEERLEMMA**
Herhaling van die lemma in sy basiese vorm, bv. sonder enige klem-aanduiding. Dit word vir rekenaarsoeke toegte gebruik en word nie in die woordeboek gedruk nie.
- (6) **UITSPRAAK**
Fonetiese transkripsie, maar nie klem nie, word hier gegee.
- (7) **WOORDSOORT**
Woordsoortaanduiding, wat verkieslik by alle soorte lemmas gegee behoort te word. Die Buro verstrek dit egter nie by subleksikale en multileksikale lemmas nie.
- (8) **MORFOLOGIE**
Morfologiese inligting (verbuigings-, vervoegings- en iteratiewe vorme), met of sonder spesifieke gebruikskwalifikasies.
- (9) **SINTAKSIS**
Sintaktiese inligting, bv. oor predikatiewe of attributiewe gebruik, vaste sintaktiese volgorde of geykte sintaktiese verbindings.
- (10) **REGISTER**
Stilistiese en sosiostilistiese etikette, bv. *eufemisme, klanknabootsing, kollektief*; en *geselstaal, skertsend, verhewe, neerhalend, seksisties, rassisties, skeltaal, vloek, plat, vulgêr*.
- (11) **SPESIALITEITSVELD**
Vak-, dissipline-, beroeps- en groepsaanduidende etikette, bv. *plantkunde, skeepvaart, hekelwerk, krieket* en *kerk-, vissers-, skool-, studente- en kindertaal*.
- (12) **TYDSAANDUIDING**
Temporele etikette, bv. *histories, verouderd, verouderend*.
- (13) **ANDER ETIKETTE**
Streektaalaanduiding en gebruiklikheidsetikette, bv. *streektaal, onge woon*. Eersgenoemde maak die invul van die veld **GEOGRAFIESE VOORKOMS** verpligtend.

- (14) **WISSELVORME**
 Vermelding van of verwysing na een of meer gelykwaardige of ongelikwaardige leksikale items wat ten opsigte van uitspraak en skryfvorm variante van die betrokke lemma is maar dieselfde betekenis het. Wisselvormvermeldings, bv. "Ook *tasaterwater*", word in leesbare teks gegee en sal dus in die gedrukte WAT op hierdie plek in die artikel verskyn. Wisselvormverwysings, bv. "Sien **KASATERWATER**", word hier in verskuilde teks gegee en sal dus nie in die gedrukte woordeboek op hierdie plek in die artikel sigbaar wees nie. Dit word in verskuilde teks gegee uitsluitlik met die oog op rekenaarmatige *onttrekking*, bv. vir die maak van 'n wisselvormwoordeboek. Vir die WAT in gedrukte vorm word die wisselvormverwysings wel in leesbare teks in die DEFINISIE-veld herhaal.
- (15) **LOS-VAS**
 Vermelding van 'n moontlike losgeskrewe vorm van die lemma, bv. van *krom buig* onder die lemma **krombuig**.
- (16) **ETIMOLOGIE**
 Etimologiese inligting m.b.t. onveranderde ontleding (bv. by *establishment*), volksetimologie (by *katjiepiering*), afleidings van eiename (by *klistron*, *krabbel*) en Afrikaanse letterwoorde (by *vigs*).
- (17) **DEFINISIE**
 Definiens of verklaring.
- (18) **GRAFIKA**
 'n Verwysing na 'n rekenaarlêer wat 'n illustrasie of 'n tabel bevat. Dié grafiese materiaal word tydens die drukproses outomaties op die aangeduide plek in die teks ingetrek.
- (19) **SINONIEME**
 Volwaardige sowel as gedeeltelike sinonieme van die lemma. 'n Definiens wat in die vorm van 'n sinoniemverwysing in die DEFINISIE-veld staan, word in verskuilde teks in hierdie veld herhaal met die oog op rekenaarmatige *onttrekking*, bv. vir die maak van 'n sinoniemwoordeboek.
- (20) **SEMANTIESE OPPOSISIE**
 Antonieme en ander semanties teenstellende, maklik met die lemma verwarbare leksikale items.
- (21) **KOLLOKASIES**
 Geykte of tipiese verbindings in frasevorm, bv. m.b.t. werkwoord- en voorsetselgebruik. Sintakties vormvaste verbindings word egter onder **SINTAKSIS** gegee.

- (22α) **VOORBEELDE**
Poëme (wat uit gebruiksvoorbeelde saamgestel word) en gedokumenteerde sitate. Binne hierdie veld kan enigeen of almal van die onderstaande subvelde genes word, in die volgorde soos aangegee. Elke formele sitaat word verder gestruktureer binne die spesifieke Sitaat-, Outeur- en Bronsubvelde met die oog op bv. rekenaarmatige statistiese ontleding van aangehaalde materiaal.
- (23) **Nabepaling**
Subveld, ingelui deur 'n aanlopie, bv. "As nabepaling", vir voorbeelde waarin die lemma, bv. die b.nw. goed, as 'n nabepaling na bv. *alles, iets, min, niks, veel* voorkom.
- (24) **Byw.**
Subveld, ingelui deur 'n aanlopie, bv. "Ook soms byw.", vir voorbeelde waarin die lemma in 'n bywoordelike toepassing gebruik word, terwyl die res van die betekenisvlak 'n ander woordsoortlike funksie beskryf.
- (25) **Selfst.**
Subveld, ingelui deur 'n aanlopie, bv. "Ook soms selfst. met meervoudige betekenis", vir voorbeelde waarin die lemma in 'n selfstandige toepassing gebruik word, terwyl die res van die betekenisvlak 'n ander woordsoortlike funksie beskryf.
- (26) **Werkw. (Werkwoordelik)**
Subveld, ingelui deur 'n aanlopie, bv. "Ook soms werkw.", vir voorbeelde waarin die lemma in 'n werkwoordelike toepassing gebruik word, terwyl die res van die betekenisvlak 'n ander woordsoortlike funksie beskryf.
- (27) **Fig.**
Subveld, ingelui deur 'n aanlopie, bv. "In fig. verband", vir voorbeelde waarin die lemma in 'n figuurlike verband gebruik word.
- (28) **Ander aanlope**
Subveld vir ander aanloopgevalle as die voorafgaande wat nie van 'n stilistiese aard is nie, bv. "I.d. negatief", "In 'n ou bron", "Ook as stofnaam", "Ook soms byv.", "Soms i.d. verkleiningsvorm met bykomende modale waarde", "Van nie-tasbare dinge".
- (29) **Styl**
Subveld waarin aanlope van stilistiese aard, maar nie Fig.- en Selfst.-gevalle nie, gegee word, bv. "Metonimies gebruik", "Van vermenslikte diere", "Met persoonifikasie", "Ook soms met omsetting v.d. voorwerp as onderwerp".

- (22Ω) **VOORBEELDE**
Dit sluit die veld vir voorbeelde af.
- (30) **GEOGRAFIESE VOORKOMS**
Verpligte aanduiding van die plek, omgewing of streek van voorkoms in die geval van streektaal — sien **ANDER ETIKETTE**. Die geografiese voorkoms kan egter ook blyk uit die subveld "Streektaalvb." — s.d.
- (31) **VGL.**
Vgl.-verwysings. Sulke verwysings is deur die nestingsbeginsel aan 'n artikelvlak gekoppel. Hulle kan dus betrekking hê op een of meer volledige artikelvlakke, bv. betekenisonderskeidings, maar nie op spesifieke inskrywings binne so 'n artikelvlak nie. Vir laasgenoemde word die subveld "Spes. vgl." gebruik.
- (32) **OPMERKINGS**
Enige nuttige inligting oor onreëlmatige of eiesoortige taalverskynsels waarvoor nie reeds d.m.v. 'n inskrywingsveld voorsiening gemaak is nie, of wat bykomend tot of verhelderend van reeds vermelde inligting is, word hier in die vorm van 'n opmerking verstrekkend met die aanloop "OPM." gemerk. Dit word net soos die Vgl.-gevalle deur die nestingsbeginsel aan 'n artikelvlak gekoppel en kan dus op een of meer artikelvlakke slaan. Deur pertinente verwysing kan dit op enige ander veld se inhoud kommentaar lewer. Dit kan egter nie fisies binne 'n ander veld genes word nie. Vir so 'n spesifieke opmerking binne 'n ander veld moet die subveld "Spes. opm." gebruik word.
- (33α) **UITDRUKKINGS**
Veld waarin gespesialiseerde en idiomatiese uitdrukkings binne die onderskeie subvelde opgeneem word.
- (34α) **Gespesialiseerd**
Beginmerker van die subveld vir alle gespesialiseerde uitdrukkings.
- (35α) **Gespes. vlak (Gespesialiseerde vlak)**
Subveld waarin die logiesevlaknommer van die gespesialiseerde uitdrukking in verskuilde teks aangebring word.
- (36) **Gespes. soekwoord**
Die soekwoord van die gespesialiseerde uitdrukking word hier aangebring.
- (37) **Gespes. uitdr.**
Die gespesialiseerde uitdrukking in sy basiese vorm word hier gegee.

- (38) Gespes. register
Sosiostilistiese etikette by die gespesialiseerde uitdrukking, bv. *gesels- taal, skertsend, verhewe, neerhalend, seksisties, rassisties, plat, vulgêr*, kom in hierdie subveld.
- (39) Gespes. spes.-veld (Gespesialiseerde spesialiteitsveld)
Spesialiteitsveld vir vak-, dissipline-, beroeps- en groepsaanduidende etikette wat by die gespesialiseerde uitdrukking aangegee word. Die etikette is van dieselfde tipe as by SPESIALITEITSVELD — s.d.
- (40) Gespes. tydsaand. (Gespesialiseerde tydsaanduiding)
Tydsaanduiding in die vorm van temporele etikette, bv. *histories, ver- ouderd, verouderend*, by die gespesialiseerde uitdrukking.
- (41) Gespes. ander etik. (Gespesialiseerde ander etikette)
Streektaalaanduiding en gebruiklikheidsetiketete, bv. *streektaal, onge- woon*, by die gespesialiseerde uitdrukking.
- (42) Gespes. los-vas
Vermelding van 'n moontlike vasgeskrewe vorm van die gespesiali- seerde uitdrukking.
- (43) Gespes. etim.
Etimologiese inligting by die gespesialiseerde uitdrukking.
- (44) Gespes. def.
Definiens of verklaring van die gespesialiseerde uitdrukking.
- (45) Gespes. sin.
Sinonieme van die gespesialiseerde uitdrukking word hier gegee.
- (46) Gespes. opp. (Gespesialiseerde opposisie)
Semantiese opposisies van die gespesialiseerde uitdrukking word hier gegee.
- (47) Gespes. vbe. (Gespesialiseerde voorbeelde)
Poëme en gedokumenteerte sitate met die gespesialiseerde uitdruk- king. Elke formele sitaat word verder gestruktureer binne die spesi- fieke Sitaat-, Outeur- en Bronsubvelde.
- (48) Gespes. aanloopvbe.
Subveld vir voorbeelde wat ingelui word deur aanlopies, bv. "In 'n ou bron". Verskillende aanloopgevalle by 'n gespesialiseerde uit- drukking word saamgegroepeer. Elke formele sitaat word verder ge- struktureer binne die spesifieke Sitaat-, Outeur- en Bronsubvelde.
- (49) Gespes. geogr. voorkoms
Verpligte aanduiding by 'n gespesialiseerde uitdrukking van die plek, omgewing of streek van voorkoms in die geval van streektaal.

Die geografiese voorkoms kan egter ook blyk uit die subveld "Streektaalvb." — s.d.

- (50) Gespes. vgl.
"Vgl."-verwysings by 'n gespesialiseerde uitdrukking. Dit is deur die nestingsbeginsel aan die logiese artikelvlak van so 'n uitdrukking gekoppel. Nie "Vgl." nie maar die subveld "Spes. vgl." moet dus gebruik word waar 'n verwysing betrekking het op spesifieke inskrywings binne die artikelvlak.
- (35Ω) Gespes. vlak
Afsluiting van die subveld waarin die logiesevlaknommer van die gespesialiseerde uitdrukking in verskuilde teks aangebring word.
- (34Ω) Gespesialiseerd
Eindmerker van die subveld vir alle gespesialiseerde uitdrukkings.
- (51α) Idiomaties
Beginmerker van die subveld vir alle idiomatiese uitdrukkings.
- (52α) Idiom. vlak
Subveld waarin die logiesevlaknommer van die idiomatiese uitdrukking in verskuilde teks aangebring word.
- (53) Idiom. soekwoord
Die soekwoord van die idiomatiese uitdrukking word hier gegee.
- (54) Idiom. uitdr.
Die idiomatiese uitdrukking in sy basiese vorm word hier gegee.
- (55) Idiom. register
Sosiostilistiese etikette, bv. *geselstaal*, *skertsend*, *verhewe*, *neerhalend*, *seksisties*, *rassisties*, *plat*, *vulgêr*, by die idiomatiese uitdrukking kom in hierdie subveld.
- (56) Idiom. tydsaand. (Idiomatiese tydsaanduiding)
Tydsaanduiding in die vorm van temporele etikette, bv. *histories*, *verouderd*, *verouderend*, by die idiomatiese uitdrukking.
- (57) Idiom. ander etik. (Idiomatiese ander etikette)
Streektaalaanduiding en gebruiklikheidsetiketette, bv. *streektaal*, *ongewoon*, by die idiomatiese uitdrukking.
- (58) Idiom. los-vas
Vermelding van 'n moontlike vasgeskrewe vorm van die idiomatiese uitdrukking.
- (59) Idiom. def.
Definiens of verklaring van die idiomatiese uitdrukking.

- (60) Idiom. vbe. (Idiomatiese voorbeelde)
Poëme en gedokumenteerde sitate met die idiomatiese uitdrukking. Elke formele sitaat word verder gestruktureer binne die spesifieke Sitaat-, Outeur- en Bronsubvelde.
- (61) Idiom. aanloopvbe.
Subveld vir voorbeelde wat ingelui word deur aanlopies, bv. "In 'n ou bron". Verskillende aanloopgevalle by 'n idiomatiese uitdrukking word saamgegroepeer. Elke formele sitaat word verder gestruktureer binne die spesifieke Sitaat-, Outeur- en Bronsubvelde.
- (62) Idiom. geogr. voorkoms
Verpligte aanduiding by 'n idiomatiese uitdrukking van die plek, omgewing of streek van voorkoms in die geval van streektaal. Die geografiese voorkoms kan egter ook blyk uit die subveld "Streektaal-vb." — s.d.
- (52Ω) Idiom. vlak
Afsluiting van die subveld waarin die logiesevlaknommer van die idiomatiese uitdrukking in verskuilde teks aangebring word.
- (51Ω) Idiomaties
Eindmerker van die subveld vir alle idiomatiese uitdrukkings.
- (33Ω) UITDRUKKINGS
Dit sluit die uitdrukkings se veld af.
- (63) SAMESTELLINGS
Veld vir onverklaarde samestellings en onverklaarde afleidings. Klem en woordsoort word aangedui. Ook die volgende tipe afleidings kan vir hierdie hantering oorweeg word: *hidrotakties* (onder HI-DROTAKSIE), *kerkhofwaarts* (onder KERKHOF), *kieserskap* (onder KIESER), *kramming* (onder KRAM), *krateragtig*, *-vorm*, *-vormig* en *-vorming* (onder KRATER), *kropvattery* (onder KROPVAT). Verder verdien morfologiese inligting en SPESIALITEITSVELD-etiket hier oorweging.
- (64) AFKORTINGS
Afkortings van die lemma vir sover dit onmiddellik en sonder moeite tot die redaksie se beskikking is. Hierdie veld word bloot ingevul (in verskuilde teks) vir rekenaarmatige onttrekking. Indien 'n afkorting egter gelemmatiseer word, moet dit in die veld LEMMA ingevul word.
- (65) ENGELS
Die Engelse vertaalekwivalent vir sover dit onmiddellik en sonder moeite tot die redaksie se beskikking is. Dit word in verskuilde teks gegee en is bedoel vir rekenaarmatige onttrekking. Dit word dus nie in die WAT opgeneem nie.

- (66) **ANDER TALE**
Veld vir *besondere, onreëlmatige of tiperende* vertaalekwivalente in Duits, Latyn, Nederlands en inheemse Suid-Afrikaanse tale. Dit word in verskuilde teks gegee, met die erkende afkorting vir die betrokke taal tussen hakies. Dit is bedoel vir rekenaarmatige onttrekking en word dus nie in die WAT opgeneem nie.
- (67) **KOMMENTAAR**
Kommentaar vir huishoudelike gebruik, bv. onthounotas, motiveerings, moontlike alternatiewe formuleringe. Dit kom in verskuilde teks en word dus nie in die WAT opgeneem nie.
- (68) **BRONNE**
Geraadpleegde bronne, bv. vir die etimologie of by definiëring, word hier vermeld. Dit is vir huishoudelike gebruik. Die besonderhede kom in verskuilde teks en word dus nie in die WAT opgeneem nie.
- (69) **SUBVLAKAANLOOP**
Veld vir enige aanloop wat 'n subvlak soos dié vir betekenistoepassing insluit.
- (1Ω) **ARTIKELVLAK**
Dit sluit die betrokke artikelvlak af.

B. VRYE VELDE (NIEPLEKGEBONDE BINNE DIE STRUKTUUR)

- (70) **Taksonomie**
Wetenskaplike name. By soortonderskeiding word slegs die familie-, geslags- en soortnaam gegee. Dié subveld met sy inskrywing word binne die DEFINISIE- of Gespes. def.-veld genes.
- (71α) **Sitaat**
Die begingrens van die sitaatsubveld. Elke formele sitaat word in so 'n veld gegee, bv. met die oog op rekenaarmatige statistiese ontleding van aangehaalde outeurs en bronne.
- (72) **Outeur**
Die outeur van die sitaat word hier gegee.
- (73) **Bron**
Die afgekorte bron van die sitaat word hier gegee.
- (71Ω) **Sitaat**
Die eindgrens van die sitaatsubveld.
- (74) **Streektaalvb.**
Die subveld waarin die streek aangedui word waarin die spesifieke gebruiksvoorbeeld opgeteken is.

- (75) Spes. vgl.
Vrye veld wat in enigeen van die vaste velde ingevoeg kan word en waardeur 'n inskrywing in so 'n vaste veld met enige ander inskrywing vergelyk kan word. "Vgl." lui so 'n vergelyking in.
- (76) Spes. opm.
Vrye veld wat in enigeen van die vaste velde ingevoeg kan word en waarin 'n opmerking spesifiek oor 'n inskrywing in die betrokke vaste veld gegee kan word. So 'n spesifieke opmerking word altyd deur "Opm." ingelui.

11.2 Redaksionele toepassing van die *LEXI-WAT II*-veldstruktuur

Vervolgens word 'n verkorte voorbeeld van manuskrip wat binne hierdie veldstruktuur geskep is, gegee om aan te toon wat die redakteur voor hom op die rekenaarskerm sien en hoe hy enkele van die velde moet invul. Waar van een artikelvlak na 'n volgende beweeg moet word, hoef hy bloot die sjabloon van al die velde deur die druk van 'n kode binne die bestaande artikelvlak te herhaal. Velde wat nie ingevul word nie, word uitgegee. Teksformatering vind op die gewone manier ooreenkomstig die redaksionele stelsel plaas. Teks wat nie in die gedrukte weergawe van die woordeboek opgeneem moet word nie, bv. al die veldgrense, die redakteur se huishoudelike kommentaar en bronverwysings, verskyn in verskuilde teks en word gevolglik nie gedruk nie. Dit sou wel deel kon uitmaak van 'n eventuele elektroniese woordeboek, bv. op laser-skyf.

```
<ARTIKELVLAK 0/>
  <LEMMA />leeu'bos <LEMMA \>
  <SORTEERLEMMA />leeubos<SORTEERLEMMA \>
  <WOORDSOORT />s.nw. <WOORDSOORT \>
  <ARTIKELVLAK 1/>
  <ARTIKELVLAK 1a/>
    <BETEKENISNOMMER />1 a <BETEKENISNOMMER \>
  <MORFOLOGIE />(ook i.d. vorm leeubossie) <MORFOLOGIE \>
  <DEFINISIE />Inheemse struik,
  <Taksonomie /> Zygophyllum morgsana (fam. Zygophyllaceae)<Taksonomie \>,
  tot 1 m hoog met vlesige liggroen blare, liggeel blomme en viervlerkige doos-
  vrugte; <DEFINISIE \>
  <SINONIEME />sin. skilpadbos; vetbos (minder gebruiklik).<SINONIEME \>
  <ARTIKELVLAK 1a \>
  <ARTIKELVLAK 1b/>
  <BETEKENISNOMMER />b <BETEKENISNOMMER \>
```

<DEFINISIE />Inheemse wydvertakkende bos of goed gevormde boom,
 <Taksonomie /> *Salvadora angustifolia* (fam. *Salvadoraceae*)<Taksonomie \>,
 met 'n growwe vaal bas en smal blougroen of vaalgroen effens vlesige blare
 en ronde vruggies. <DEFINISIE \>
 <ARTIKELVLAK 1b \>
 <ARTIKELVLAK 1 \>
 <ARTIKELVLAK 2 />
 <BETEKENISNOMMER />2 <BETEKENISNOMMER \>
 <DEFINISIE />Enige bos waarin leeus graag hou: <DEFINISIE \>
 <VOORBEELDE />
 <Sitaat />Hulle (die leeus) (soek) skuiling in die doringlose ruie leebos
 <Outeur />(O. Pirow: <Outeur \>
 <Bron />Sjangani, 1953, 31). <Bron \>
 <Sitaat \>
 <Sitaat />Hoe goed het ek die haak-en-steekplate ... later leer ken! Dit was die leeu-
 bos van Ngésomit
 <Outeur />(Sangiro: <Outeur \>
 <Bron />Simba, 1944, 29).<Bron \>
 <Sitaat \>
 <ARTIKELVLAK 0 \>

Die gedrukte weergawe van bostaande rekenaarmanuskrip lyk dan soos volg:

leeu'bos s.nw.

1 a (ook i.d. vorm leebossie) Inheemse struik, *Zygophyllum morgsana* (fam. *Zygophyllaceae*), tot 1 m hoog met vlesige liggroen blare, liggeel blomme en viervlerkige doosvrugte; sin. *skilpadbos*; *vetbos* (minder gebruiklik).

b Inheemse wydvertakkende bos of goed gevormde boom, *Salvadora angustifolia* (fam. *Salvadoraceae*), met 'n growwe vaal bas en smal blougroen of vaalgroen effens vlesige blare en ronde vruggies.

2 Enige bos waarin leeus graag hou: *Hulle* (die leeus) (soek) skuiling in die doringlose ruie leebos (O. Pirow: Sjangani, 1953, 31). *Hoe goed het ek die haak-en-steekplate ... later leer ken!* Dit was die leebos van Ngésomit (Sangiro: Simba, 1944, 29).

11.3 Voordele van die LEXI-WAT II-veldstruktuur

Die doelmatige gestruktureerdheid van die inskrywingsvelde van LEXI-WAT II blyk een van die belangrikste winste van die hele rekenariseringsprojek te wees. Manuskripskepping binne dié struktuur in MS Word 5 lei vanself tot 'n sistematiese werkmethode en tot 'n streng georganiseerde, afgebakende en eenvormige aanbiedingswyse wat sowel die redakteur as die woordeboekgebruiker dien. Die redakteur hoef nie kosbare redaksionele tyd te bestee aan die be-

paling van die volgorde van inligtingstipes nie en die gebruiker leer gou waar hy sy inligting konsekwent in die artikel sal vind.

Ook vergemaklik die doelmatige gestruktureerdheid personeelopleiding en spesialisering binne spanverband in dié sin dat 'n bepaalde redakteur 'n spesifieke veld in verskillende artikels kan invul om die taak te leer bemeester of juis omdat hy 'n spesialis in dié werk is. Verder is 'n gestruktureerde skerm 'n groter aansporing tot werk as 'n blanko vel papier.

Data kan ook met gebruikmaking van *LEXI* of enige ander kragtige databasisprogram veldgewys aan *DATABANK II* onttrek word, bv. vir die lewering van taaladvies of die studie van taalverskynsels.

Ten slotte outomatiseer die gestruktureerdheid van *DATABANK II* grotendeels die samestelling van kleiner satellietprojekte. So 'n samestellingsproses sou uit 'n paar eenvoudige stappe bestaan: Die rekenaar word die instruksie gegee om bv. al die lemmas te soek en uit te druk waarby daar in die *REGISTER*-veld die etiket "*(seksisme)*" vermeld word, om sodoende outomaties 'n lys van seksistiese terme te verskaf. Of die rekenaar kry die instruksie om al die lemmas te soek waarby daar onder *SPECIALITEITSVELD* die etiket "*(rugby)*" voorkom; daarby moet dan die betrokke definisie en die Engelse vertaalekwivalent uit hulle onderskeie velde gehaal en alles gedruk word vir die outomatiese produksie van 'n tweetalige woordeboek van rugbyterme met Afrikaanse definisies. Op dieselfde koste-effektiewe wyse kan sinoniem-, streektaal-, vaktaal-, kollokasie- en uitdrukkingswordeboeke saamgestel word.

Dit sal vir die Buro 'n nuttige bykomende inkomstebron kan vorm, veral in die lig van die ondervinding van die OUP met die *OED*: "Dit is onmoontlik om die geld wat aan die *OED* bestee is, ooit uit sy verkoopsprys te verhaal. So 'n projek kan nooit 'n lonende projek wees nie. Dit moet deur ander projekte gedra word" (Van Schalkwyk 1991: 8). Die OUP verhaal dan ook inderdaad die geld wat hy op die nuwingsgewende *OED*-projek verloor, dubbel en dwars uit sy verkope van kleiner woordeboeke, veral die *Oxford Advanced Learner's Dictionary of Current English*, waarvan reeds meer as vyftien miljoen eksemplare van die hand gesit is (Shenker 1989: 95).

12. Oordrag van manuskrip na die reprografiese afdeling

Een van die groot frustrasies van die Buro is die feit dat geen *WAT*-manuskrip wat met gebruikmaking van persoonlike rekenaars geskep is, tans in elektroniese vorm, bv. op rekenaarskyf, na die drukkerij gestuur kan word vir regstreekse oordrag in 'n rekenaarsetmasjien nie. Die situasie is eenvoudig dat die Buro se huidige eksterne drukker nog nie vir hierdie stap gereed is nie en verkies om alles met die hand oor te set. Dit lei vanselfsprekend tot menslike fouttering en het 'n vertragende en duur proses van herhaalde proeflees en korrigering tot gevolg.

Hierdie breuk in die produksieprosesse maak dus een van die belangrikste rekenariseringsvoordele vir die Buro ongedaan en kan nie langer bekostig word nie.

Die Buro is gevolglik van voorneme en tegnies reeds in staat om die WAT vanaf deel IX self uit te gee. Die voltooide manuskrip sal dan regstreeks uit LEXI-WAT II langs die rekenaarnetwerkstelsel na sy eie reprografiese afdeling oorgedra en in die vorm van 'n kantoorpublikasie in bladsye uitgelê en vir proefleesdoeleindes gedruk word. Vir die uitleg word gebruik gemaak van MS Word 5. Die reduplisering van die Buro se veldstruktuur vir woordeboek-inligtingstipes in MS Word 5 het dus meegebring dat die manuskrip van meet af aan in dié formaat tot stand kom en gevolglik net so in die vorm van 'n kantoorpublikasie uitgegee kan word. Dit het die verdere voordeel dat alle laat teksveranderinge onmiddellik deel uitmaak van die finale teks van DATA-BANK II. Die teksveranderinge hoef dus nie in sowel die woordverwerkingsteks as die databasisteks aangebring te word nie.

Drukwerk word op 'n laserdrukker in die PostScript-standaard gedoen. Dié standaard verseker dat die redaksioneel vasgestelde bladuitleg presies in daardie vorm op enige PostScript-drukker of -setmasjien weergegee sal word. Die resultate wat met hierdie prosesse behaal is, het alle kwaliteitsverwagtinge oortref.

Nadat die Buro die bladsye finaal opgemaak het, vind daar dus geen ingryping in die teks meer plaas nie en word alle verdere insluiping van foute so uitgeskakel.

Die enigste prosesse wat dan nog buite die Buro sal plaasvind, is moontlik die maak van hoëkwaliteitdrukstukke en die fotografering daarvan, en daarna die grootskaalse druk- en bindwerk. Dit is egter al tegnies moontlik om die maak van drukstukke uit te skakel omdat teks wat in PostScript-formaat uitgelê is, tans regstreeks van rekenaarskyf af op 'n drukkersfilm vasgelê kan word, d.w.s. sonder gebruikmaking van 'n kamera.

Volgens alle berekening sal die Buro se eie bladuitleg baie tyd en geld bespaar.

13. Ten slotte

Die sukses van enige projek kan bepaal word deur die mate waarin dit aan die gestelde verwagtinge of doelwitte voldoen, gemeet teen koste en tyd. Volgens hierdie maatstaf kan beweer word dat die rekenariseringsprojek van die Buro van die WAT tot dusver baie suksesvol verloop het. In die kort bestek van minder as vier jaar is die meeste van die verwagte voordele van rekenariseringsprojek van die Buro (vergelyk paragraaf 3) behaal of lê hulle goed binne bereik. Die voordele vir die gebruiker (vergelyk paragraaf 4) sal eers realiseer as die eerste rekenaarmatig geproduseerde dele van die WAT op die mark kom.

Verder word ver wag dat die rekenaar soos geen ander hulpmiddel nie in toenemende mate sal meehelp om die voltooiing van die WAT te bespoedig en die projek op 'n lonende grondslag te plaas.

Bibliografie

- Aitken, A.J. 1978. Historical Dictionaries, Word Frequency, Distribution and the Computer. *Cahiers de Lexicologie* 32: 28-47.
- Alberts, Mariëtta. 1991. Die LEXI-program. *LVSA: Byblad* 10(1): 38-68.
- Berg, Donna Lee, Gaston H. Gonnet en Frank Wm. Tompa. 1988. *The New Oxford English Dictionary Project at the University of Waterloo*. Waterloo, Ontario: UW Centre for the New Oxford English Dictionary.
- De Tollenaere, F. 1963. *Nieuwe wegen in de lexicologie*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.
- DSNA Newsletter 1990, 14(1): 4.
- Gray, J.C. 1986. Creating the Electronic New Oxford English Dictionary. *Computers and the Humanities* 20: 45-49.
- Harteveld, P. 1988. Die rekenariser ing van die werksaamhede van die Buro van die WAT. *Die Taalpraktisyn* 2: 10-27.
- Harteveld, P. en B.A. Nieuwoudt. 1989. *Die rekenariser ing van die werksaamhede van die Buro van die WAT. Verslag van samesprekings gevoer en aanbevelings gedoen oor die voortsetting van die rekenariseringsproses binne die struktuur van die RGN se LEXI-projek*. Ongepubliseerde verslag. Stellenbosch: Buro van die WAT.
- Hauptfleisch, D.C. (Red.). 1984. *Woordeboek van die Afrikaanse Taal. Sewende deel. KOM - KOR*. Pretoria: Die Staatsdrukker.
- Hauptfleisch, D.C. (Red.). 1991. *Woordeboek van die Afrikaanse Taal. Agtste deel. KOS - KYW*. Stellenbosch: Buro van die WAT. (Op die pers.)
- Jackson, Howard. 1988. *Words and Their Meaning*. Londen: Longman.
- Landau, S.I. 1984. *Dictionaries: The Art and Craft of Lexicography*. New York: The Scribner Press
- Oitana, C. 1982. Introduction to Session II: New Technologies. Goetschalckx, J. en L. Rolling (Reds.). 1982. *Lexicography in the Electronic Age: Proceedings of a Symposium held in Luxembourg, 7-9 July 1981*: 47-50. Amsterdam: North-Holland Publishing Company.
- Redaksie, Buro van die WAT. 1989. *Verslag oor die strategiese beplanning vir die Buro van die Woordeboek van die Afrikaanse Taal*. Ongepubliseerde verslag. Stellenbosch: Buro van die WAT.
- Shenker, Israel. 1989. Annals of Lexicography: The Dictionary Factory. *The New Yorker* 3 April 1989: 86-100.
- Swanepoel, P.H. en R. Morris (Reds.). s.j. *Die rekenariser ing van leksikografiese prosesse (WORDNET-projek)*. LEXINET en die rekenariser ing van taal. Verslag 1. Ongepubliseerde tussentydse verslag. Pretoria: Raad vir Geesteswetenskaplike Navorsing (INTAK).
- Swanepoel, P.H. en R. Morris (Reds.). 1988. *Die rekenariser ing van leksikografiese prosesse. WORDNET-projek van die LEXINET-program*. Verslag LEXI-4. Pretoria: Raad vir Geesteswetenskaplike Navorsing.

- Tompa, Frank Wm. en Darrell R. Raymond.** 1989. *Database Design for a Dynamic Dictionary*. Waterloo, Ontario: UW Centre for the New Oxford English Dictionary.
- Van Schalkwyk, D.J.** 1991. *Verslag oor die oorsese inligtingsreis van dr. D.J. van Schalkwyk 3-27 Mei 1991*. Ongepubliseerde verslag. Stellenbosch: Buro van die WAT.
- Van Sterkenburg, P.G.J. (Red.)**. 1976. *De Nederlandse lexicologie tussen handwerk en machine*. Groningen: H.D. Tjeenk Willink.
- Van Sterkenburg, P.G.J.** 1984. *Van woordenlijst tot woordenboek: Inleiding tot de geschiedenis van woordenboeken van het Nederlands*. Leiden: E.J. Brill.