Hoofstuk 3: Rekenaargesteunde taalonderrig

3.1 Inleiding

In hierdie hoofstuk word rekenaargesteunde taalonderrig of rekenaargesteunde taalleer ten opsigte van woordeskat bespreek. Daar word gekonsentreer op woordeskatleer-programme wat van vertalings gebruik maak om woordeskat te leer. Verder word 'n program vir die opleiding van vertalers met rekenaars beskryf en word daar ook gekyk na die gebruik van rekenaars vir die opleiding van leksikograwe en hoe hierdie proses in woordeskatleerprogramme gebruik kan word.

Literatuur oor die onderwerp hanteer gewoonlik rekenaargesteunde onderrig ook ten opsigte van onderwerpe soos wiskunde en biologie. Daar is nie baie inligting oor spesifieke woordeskatleer nie; daar is heelwat meer oor die aanleer van sintaksis, oor die voordele van rekenaargesteunde onderrig, en oor die ontwerp van goeie programme.

3.2 'n Totale-uitvee woordeskatleerprogram gebaseer op vertaling

In Hardisty en Windeatt (1989) word lyste RGTO programtipes en idees met uiteensettings, en 'n lys van verskaffers van RGTO-programme en outerspakette verskaf.

Die tipe programme wat gebruik word konsentreer op die invul van gapings, die aanleer van die onreëlmatige vorme van werkwoorde, die spelling van woorde met 'n Griekse oorsprong, die aanleer van eenvoudige vraagformasietegnieke en die inoefening van taalgebruik rondom persoonlike beskrywings met die fokus op subjek-werkwoord-ooreenstemming.

'n Spesifieke groep idees fokus op woorde, of leksikale leer. Die voorgestelde programme sal onder meer die volgende leer of ondersteun: korrekte uitspraak, korrekte spelling, die vorming van getalswoorde, die gebruik van negatiewe prefikse in Engels soos *un-*, *in-*, en *im-*, die oefening van lees vir spesifieke informasie en die ontwikkeling van kennis oor hoe sosiale banderings taal kan beïnvloed, die oefening van spreekwoorde, die oefening van mondelinge vaardigheid vir intellektuele besprekings, en die aanleer van vertalingstegnieke.

Onder lg. stel die outeurs voor (Hardisty en Windeatt, 1989 : 103 & 104), dat 'n totale-uitvee (*totale deletion*) tipe program gebruik word. Sulke programme word ook teksrekonstruksieprogramme genoem, en gee aan die leerder 'n teks waarvan al die woorde uitgevee is. Die leerling kom te staan voor 'n skerm

waarop die enigste aanduidings oor die konteks van die teks die titel is, en punktuasietekens en gewoonlik 'n bepaalde merk, soos 'n kort streep wat die letters van die woorde aandui. Die student moet woorde intik wat lyk of dit in die teks kan inpas; indien die woord wel daarin voorkom word dit op al daardie plekke deur die rekenaar ingesit. Hierdie soort programme is baie gewild en bestaan in verskeie variasies.

Die voorstel oor die vertalingsprogram maak egter van 'n minder radikale weergawe hiervan gebruik. Die inhoud van die leerproses bestaan uit 'n teks wat deur die leerlinge vertaal moet word. Die vertaling word op die rekenaar gedoen. Dit word egter nie heeltemal van die begin af gedoen, soos totale-uitvee impliseer nie, maar met 'n vertaling op die rekenaar waarin foute of gapings voorkom. Die leerlinge doen die oefening en bespreek dan die herstellings wat hulle aan die teks moes doen.

3.3 Jovanovic se rekenaargebaseerde vertaalleer-program

Volgens Jovanovic (1987:86) kan vertalingsvaardighede, anders as 'n monolinguistiese taalgebruiksvaardigheid, nie verwerf word sonder instruksie nie of sonder dat 'n deurwinterde vertaler aan die leerder kan wys wat 'n goeie en aanvaarbare vertaling is en wat nie. Die outeur sê ook dat vertaalteorie nog nie (teen 1987) 'n toereikende teorie oor die proses en probleme van vertaling kon voorstel nie. Jovanovic verduidelik die ontwerp van 'n program wat vertaling onderrig, TRANS-1, wat op 'n ZX - Spectrum met 42K beskikbare RAM-geheue, beskryf is. Hoewel dit amper antieke tegnologie is, is die voorstelle wat Jovanovic maak steeds die moeite werd om in ag te neem.

Die program laat 'n leerling toe om 'n bronteks te sien, en dan 'n vertaling daarvan, waarvan geen woorde uitgevee is nie, maar waarin verskillende soorte vertaalfoute verskyn: byvoorbeeld van verkeerde woordkeuses, verkeerde konstruksies, letterlike vertalings van taaleie uitdrukkings en verkeerde register- en stylgebruike. Die leerlinge moet die foute in verskillende fases korrigeer wat trapsgewys toeneem in moeilikheidsgraad. Omdat die rekenaaromgewing van TRANS-1 so beperk was, is baie klein tekste gebruik, en word dit ook geensins deur die rekenaar self ontleed nie. In die verskillende fases kan die studente toegang kry tot kommentaar oor die vertaalfoute, waarin onder andere woordeboekinligting en voorbeelde voorkom.

Een van die ontwerpsuitgangspunte wat Jovanovic met die program het, is dat die bg. kommentaar in die program ingebou word op grond van werklike data oor die probleme wat vertaalstudente met daardie spesifieke teks het. Die teks wat in die program gebruik word, word deur studente vertaal, en op grond daarvan word besluit watter inligting in die program ingesluit kan word om die 'persoonlike' instruksie te verskaf oor al die foute wat gemaak is.

Dit is duidelik dat die rekenaar nie self enige intelligente taalmanipulering doen nie; om vir 'n verskeidenheid tekste sulke kommentare op te stel sal baie werk verg. Dit is tegnologie wat vandag toenemend deur rekenaars, veral ten opsigte van sekere aspekte, onderneem sal kan word.

Jovanovic (1986 : 90) sê oor woordeboeke (en miskien is die opmerking net van toepassing op die Jugoslawiese woordeboeke van 1986), dat, uit die oogpunt van 'n vertaler, die stand van sake in leksikografie (eentalige, tweetalige en spesifieke woordeboeke) nie goed is nie. Maar sy skryf verder dat geen woordeboek, ongeag hoe dik, al die moontlike gevalle en betekenisse van woorde wat 'n vertaler mag nodig kry nie, kan bevat nie.

Omdat die tekste wat in die program voorkom beperk is, kan sorgvuldige aanwysings gegee word oor die betekenis en vertaalekwivalente van woorde. Jovanovic stel 'n instruksie-gebaseerde, of leer-deur-kopiëringsmetode voor, in kontras met 'n metode waarin omvattende linguistiese verduidelikings van probleme gegee word - 'n metode wat deur die outeur verwerp word as uitgediend.

3.4 Rekenaargesteunde leksikografiese opleiding

Kemble sê in 'n hoofstuk oor leksikografie in *Computers as a tool in language teaching* (Kemble en Brierley, 1991:35) dat twee dimensies van woordeskatleer in taalonderrig geïdentifiseer kan word: blootstelling aan woordeskat in realistiese of werklike taalsituasies, wat die outeur taalverwerwing noem, en die aanleer van 'n woordeskat met die gebruik van woordeskatinstrumente soos verklarende en tweetalige woordeboeke en tesourusse.

In *LSP dictionaries for EFL learners*, lys Moulin (1983:40) algemene vrae wat deur studente aan hul onderwysers gestel word, soos: waarom werk die meeste tweetalige woordeboeke op die beginsel dat daar een ekwivalent (of vertaling) vir elke bronwoord gegee word? Waarom is voorbeelde in tweetalige en enkeltalige woordeboeke so skaars, so in die verbygaan en so verouderd? Waarom is volledige en betekenisvolle aanhalings die uitsondering eerder as die reël? Waarom doen leksikograwe van spesiale woordeboeke so min moeite om probleme rondom grammatika en alledaagse gebruik te verduidelik?

Moulin (loc.cit.) voorsien 'n behoefte vir twee tipes woordeboeke: die eerste 'n soort leerderswoordeboek vir 'n spesifieke veld, en die tweede 'n gevorderde, omvattende spesialiswoordeboek. Die ideale inskrywing van lg. sal die volgende bevat:

'n hoofwoord (met spelvariante):

- uitspraak (en variante)
- grammatikale informasie

- vreemdetaalekwivalente van die hoofwoorde se basiese nie-gespesialiseerde betekenisse vergesel van, waar toepaslik, 'n aanduiding van die konteks waarin hulle tipies gebruik word; kruisverwysings na sinonieme, antonieme ensovoorts wat die leerder sal aanmoedig om paradigmatiese verkennings te doen; registeretikette
- definisie van die konsep in duidelike, volledige en akkurate taal, gevolg deur 'n vertaling
- voorbeelde (met vertalings), verkieslik bestaande uit sinne, geselekteer op grond van nie net outentiekheid en frekwente gebruik nie, maar hoofsaaklik op grond van hulle bruikbaarheid vir leerders. Met ander woorde, as dit 'n weerspieëling kan wees van frekwente sintaktiese patrone, sal dit nog beter wees
- illustrasies; 'n duidelike en maklik verstaanbare tekening of kaart sal die waarde van die bg. informasie verhoog.

Ander leksikograwe, soos dié van die Cobuild woordeboekprojek (vgl. Sinclair, 1987) beskou konteks as krities. Deur te konsentreer op die formele kenmerke van 'n leksikale item en die kontekste waarin dit verskyn, het die Cobuild-span daarna gemik om by 'n redelike analise uit te kom van die verskeie betekenisonderskeidings van die leksikale item. Sinclair (1987:89) druk die beginsel uit as dat elke verskillende betekenis van 'n woord geassosieer word met 'n onderskeiding in vorm.

Hierdie voorstelle en uitgangspunte word aangewend in leksikografiese prosesse. Indien inligting soos wat Moulin voorstel op 'n sistematiese wyse in 'n rekenaarleesbare woordeboek opgeneem kan word, sou dit 'n belangrike datakorpus wees vir linguistiese ondersoek oor woordeskat en woordeskatverskille.

Hoewel 'n bespreking van die opleiding van leksikograwe nie binne die bestek van hierdie studie val nie, is dit steeds nuttig om melding te maak van die gebruik van MT-programme hiervoor. Studente kry in sulke programme 'n indruk van hoe kompleks dit is om twee tale se woorde te vergelyk, en kan hulle eie leksikale inskrywingsraamwerke vir gerekenariseerde woordeboeke ontwerp (vgl. byvoorbeeld ook Ufert, 1988 : 115).

3.5 Die LITTRE-stelsel

Volgens Farrington (1988 :119) word vertaling nie deur linguiste baie hoog aangeskryf as 'n metode om 'n tweedetaal te leer nie. Hy sê ook dat die rede hiervoor waarskynlik te vind is by die ongewildheid en die onteoretiese benadering van die grammatikavertalingsmetode. Hy skryf verder hieroor:

> 'Translation is widely discredited as a language learning activity. Looking back through the lists of published research in the field of Applied Linguistics covering the last ten or fifteen years, one is struck by the virtual absence of any mention of translation in relation to language learning.

Here and there one comes on an article making, usually indirect, reference to the use of translation in teaching, but it is usually as it were apologetically...'

Farrington (op.cit.:124 - 125) voer verder aan dat die agterweë laat van vertaaloefeninge vanuit die tweedetaal na die eerstetaal, ten minste tot drie afleidings oor die benadering van RGTO-ontwerpers aanleiding kan gee:

'... our reluctance to dabble in the processing of anything resembling meaning, our preoccupation with what the machine does best, i.e. pick out morphological and syntactic detail, and our exclusive obsession with laboursaving materials that require no preparation on the part of the teacher.'

Farrington bepleit ook 'n RGTO-programontwerp wat meer as een vertaling van 'n gegewe teks sal toelaat. Dit blyk een van die groot tekortkominge van RGTO-programme vir tweedetaalleer te wees - dat dit nie amper korrekte of variasie-vertalings kan akkomodeer nie.

Die LITTRE-stelsel kan meer as een vertaling herken en maak van vertaling gebruik vir taalleer; dit vereis ook 'n groot mate van betrokkenheid van die onderwyser. 'n Bronsin word op die rekenaarskerm aan die leerling gegee. Die leerling kan reageer deur 'n woord, 'n frase of die hele sin op een slag te vertaal. Die program gaan die voorgestelde vertaling element vir element na en, as dit aanvaarbaar is, word dit op die onderste helfte van die skerm geskryf. Verkeerde vertalings word verduidelik.

Die stelsel is ontwerp deur van die insette van taalonderwysers gebruik te maak. Daar is deur ou eksamenskrifte gesoek vir alle produksiefoute wat leerlinge met bepaalde sinne maak, sodat die program foute kan aanteken wat definitief bruikbaar sal wees uit die oogpunt van die leerder. Hierdie foute vorm die basis van kommentaar wat die program kan lewer oor foute; van die oppervlakfoute word egter rekenaarmatig hanteer.

3.6 Opsomming

Die uitgangspunt wat in RGTO gevind word, veral waar woordeskatleer met behulp van vertaling onderneem word, is moontlik 'n stap in die rigting van 'n teorie van woordeskatuiteenlopendhede. Die foute wat leerders maak wanneer hulle konsepte van hul eerstetaal probeer vertaal verteenwoordig 'n belangrike korpus data om agter te kom hoe taalleerders deur hulle eerstetaal beïnvloed word wanneer hulle 'n tweedetaal woordeskat memoriseer en om die beginsels agter te kom waarvolgens tale se woordeskat kan verskil.

In die LITTRE-stelsel wat in hierdie hoofstuk bespreek is, en ook in die WE/WT model van Widdig en Esser (1988) wat in hoofstuk 7 bespreek word, word dit as belangrik beskou dat die antwoord van 'n leerder nie net as korrek of as verkeerd beskou moet word nie. Dit is waarskynlik die gevolg van 'n meer positiewe benadering teenoor die leerderstaal en 'n poging om die polêre onbuigbaarheid wat gewoonlik met rekenaars geassosieer word te verlig.

Die diepte wat in tweedetaalleerteorie gevind word, word nie so geredelik in RGTO-teorie en programmatuur gevind nie. Daar is 'n baie groter liggaam van literatuur oor tweedetaalleerteorie as oor rekenaargesteunde taalonderrig. Een rede daarvoor is waarskynlik dat dit al 'n baie ouer studieonderwerp is as RGTO; rekenaars het eers onlangs regtig 'n lewensvatbare oorweging as onderrigmedium geword.

Deel II: Vertaalteorie en Masjienvertalingsteorie

Hoofstuk 4: Vertaalteorie

4.1 Inleiding

Christopher Candlin gee in sy voorwoord tot Roger Bell (1991:xi) se boek, *Translation and translating - theory and practice*, te kenne dat daar in die vakmanskap van vertaling nie 'n teorie van vertaling bestaan wat vergelykbaar is met die linguistiese teorieë wat in die taalwetenskap bestaan nie.

Bell stel wel in sy boek 'n vertaalmodel voor wat informasieryk is, maar dit bly 'n raamwerk. Dit kan nuttig wees vir 'n vertaler, maar om byvoorbeeld 'n masjienvertalingstelsel op grond daarvan op te rig, sal nie 'n eenvoudige enkoderingsproses wees nie. Bell beskou dit as belangrik dat 'n model na vore moet kom wat so lewensgetrou as moontlik is en wat gegrond is op linguistiese beginsels, sodat dit selfs rekenaarmatig getoets kan word (vgl. Bell, 1991:xvi).

Maar al bestaan daar nie sulke sorgvuldige modelle van vertaling soos daar in die linguistiek van onder meer sintaksis bestaan nie, is dit egter wel so dat handleidings oor die praktyk van vertaling gelaai kan wees met inligting oor vertaalkwessies en wenke oor hoe probleme opgelos kan word. As al hierdie inligting in 'n uitgebreide model ingeborduur sou word, sou daar dalk vinniger beweeg kon word in die rigting van 'n bevredigende model van taalgebruik - die *performance* van 'n taalgebruiker.

In lyn met die tema van hierdie studie, word hieronder hoofsaaklik gekyk na die rol van die woord in vertaling. Verskeie outeurs se menings oor woordbetekenis word behandel, en 'n beskrywing van die relevantste deel van Bell se vertaalmodel word verskaf. Daar word ook van die praktiese probleme beskryf waarmee vertalers te kampe het wanneer die vertaling van 'n brontaalwoord nie die eenvoudige keuse van die regte teikentaalwoord(e) en -konstruksie behels nie - die probleem van nie-ekwivalensie.

4.2 Die plek van die woord in vertaling

Bell (1991:83) maak die volgende belangrike opmerking in sy boek:

> 'The translator (and the second language learner) may begin by believing that the major problem is the word; it may be that there are words in the text which are new to the translator and whose meanings he or she does not know. However, it soon becomes clear that, although the meaning of words are problematic in themselves (there is no one-to-one correspondence between the items of one language and those of another), the greater problem is meaning which derives from the relationship of word to word rather than that which relates to the word in isolation.'

Newmark sê oor die belangrikheid van die woord weer die volgende (1988: 193):

> 'Many translators say you should never translate words, you translate sentences or ideas or messages. I think they are deceiving themselves. The SL text consists of words, that is all that is there, on the page. Finally all you have is words to translate, and you have to account for each of them somewhere in your TL text, sometimes by deliberately not translating them (e.g. sometimes words like '*schon*' en '*déjà*'), or by compensating for them, because if translated cold you inevitably over-translate them, e.g., '*Ich bin schon lange fertig*' - '*I have been ready for ages*'; '*1000 francs, 'c'est déjà mal*' - '*1000 francs, that's not bad at all*'.

Hy sê dat hy nie hiermee bedoel dat woorde in isolasie vertaal moet word nie; woorde se betekenis word wel bepaal deur linguistiese, referensiële, kulturele, en subjektiwiteitsinvloede. Die vier invloede word verderaan bespreek.

Baker (1992:53) is skepties oor die moontlikheid dat 'n woord opsigself iets kan beteken. Sy benader die kwessie in 'n bespreking van kollokasies; wanneer iemand gevra word om 'n woord se betekenis te verduidelik, sal die persoon dit kontekstualiseer met behulp van die mees tipiese kollokasies daarvan. Kollokasies word deur Baker (1992:47) gedefinieer as die geneigdheid van sekere woorde om gereeld saam voor te kom in 'n sekere taal; of tegnies gestel: 'semanties arbitrêre beperkinge wat nie logies volg uit die proposisionele betekenis van die woord nie' (op.cit.:14).

Die verskillende kollokasies waarin die woord voorkom help om die betekenis van die woord te verklaar. Kyk byvoorbeeld na die verskillende betekenisse van '*dry*' in die volgende kollokasies:

dry cow	dry sound	dry book
dry bread	dry voice	dry humour
dry wine	dry country	dry run

(Baker, 1992:53)

4.3 Benaderings tot woordbetekenis

Onder die benaderings tot die beskrywing van 'n woord se betekenis sonder Bell (1991:84) drie uit as belangrik: verwysingsteorie, komponensiële analise en betekenispostulering.

Baker (1992:13-17) steun vir haar uiteensetting van woordbetekenisse op Cruse (1986). Volgens Cruse is daar vier soorte (wat ook uitingbetekenisse insluit): proposisionele betekenis, ekspressiewe betekenis,

voorveronderstelde betekenis en ontlokte (*evoked*) betekenis. Baker beskryf hierdie kategorisering in haar eie woorde, maar sy sê terselfdertyd dat die idee van tipes betekenis teoreties onder verdenking is. Baker bespreek ook die konsep van semantiese velde, waarna later in hierdie hoofstuk verwys word.

Die faktore wat 'n woord se betekenis bepaal word deur Newmark (1988: 193 e.v.) as vier verskillende invloede beskou. Dit verteenwoordig 'n derde beskouing oor betekenis.

Hoewel die drie outeurs wat hierbo genoem word die maniere waarop na 'n woord se betekenis gekyk kan word nie op dieselfde manier aanpak nie, deel hulle nietemin heelwat opinies, en glo hulle ook almal daaraan dat 'n woord se betekenis en vertaling in 'n baie groot mate uit die konteks afgelei word. Van die drie is dit net Bell wat probeer om met voorbedagte rade 'n vertaalmodel op te stel. Baker en Newmark gee egter weer heelwat meer detail ter illustrasie van vertaalkonsepte en -probleme. Hulle staan albei heelwat plek af aan die bespreking van kollokasies. In die hoofstuk word die aard van kollokasies ondersoek en die bydrae wat die herkenning daarvan te maak het tot die vertaalproses.

4.3.1 Verwysingsteorie

Hierdie teorie probeer aandui wat die verband is tussen 'n woord en die begrip waarna die woord verwys. Daar is twee tradisionele benaderings tot hierdie kwessie: een benadering sê die verband is natuurlik en verpligtend - met ander woorde 'n '*motor*' word so genoem omdat dit die versameling van letters en klanke is wat in die geestesoog verskyn wanneer 'n mens 'n motor sien. Die ander sê die verband is arbitrêr en 'n sosiale konvensie. Lg. hipotese word deurgaans beskou as die enigste realistiese voorstel, aangesien dit duidelik is dat daar gewoonlik nie aanneemlike verbande te vind is tussen 'n woord se betekenis en die woord self nie.

Die verband is deur De Saussure (1916) beskou as een tussen die linguistiese teken en die objek waarna die teken verwys (Bell, 1991: 85). Die linguistiese teken self bestaan uit twee onlosmaaklike elemente: 'n konsep en 'n akoestiese beeld daarvan. Bell (1991:86-87) brei die model uit deur die volgende voor te stel: iemand wat meer as een taal ken het 'n skakel geskep tussen die konsep en die objek, maar die akoestiese beeld van die konsep realiseer op verskillende maniere - afhangende van die taal. Die belangrike punt is dat daar net een konsep-objek verband is, sodat dit nie nodig is dat dieselfde konsep op verskillende plekke en in elk geval nie apart in die brein gestoor hoef te word nie.

'n Moontlike oplossingstegniek vir die op-die-punt-van-my-tong verskynsel word hiermee voorgestel: dit sou behels dat 'n persoon 'n driedimensionele veelvlak visualiseer met die konsep binne in en die akoestiese beelde op die buitenste vlakke daarvan. Bell beweer dat dit makliker sal wees om 'n woord vir 'n konsep in 'n ander taal te onthou as die taalgebruiker besef dat daar net een konsep is en dat alle akoestiese realiserings aan dié een konsep verbind is.

4.3.2 Komponensiële analise

Die basiese aanname van komponensiële analise is dat die betekenis van 'n woord die som is van die elemente, of semantiese kenmerke, daarvan en dat hierdie kenmerke binêr is - óf dit het die kenmerk, óf dit is daarsonder. Dié kenmerke kan aangevul word met fonologiese, sintaktiese en morfologiese inligting, wat nie noodwendig binêr is nie. Die leksikale inskrywing vir '*verbeel*' sal byvoorbeeld soos volg lyk:

'verbeel'

a. [fərbe:l]

b. werkwoord; oorganklik, indirekte objek = subjek; direkte objek soms ingelei deur 'dat'

c. verledetydsvorm = 'verbeel'

d. subjek = letterlik: +lewendig, +menslik

figuurlik: +konkreet

indirekte objek = gewoonlik 'n voornaamwoord waarvan die antesedent die subjek is

verstandsproses: om iets te glo wat nie waar is nie

: om iets te visualiseer

Vir 'n leksikale inskrywingsmeganisme soos hierdie om bruikbaar te wees, sal dit onder andere op 'n stel semantiese primitiewe moet staat maak, om die stellings by 'verstandsproses' op 'n geformaliseerde wyse voor te stel. Die raamwerk word in moderne MT-stelsels gebruik, maar die formalisering daarvan is baie meer spesifiek.

Aanvanklik is gehoop dat universele kenmerke vir tale opgespoor sou kon word, maar volgens Bell (1991:87) was dit onsuksesvol en is komponensiële analise nou net 'n bruikbare tegniek vir die beskrywing van dele van die semantiek van 'n taal en om insigte te verkry oor die ooreenkomste en verskille tussen tale.

Een van Peter Newmark (1988:114) se menings oor komponensiële analise is dat dit nie dieselfde begrip in vertaalteorie is as in die linguistiek nie. In die linguistiek behels komponensiële analise om 'n woord te dissekteer in die betekeniskomponente daarvan, of hulle universeel is of nie. In vertaling, daarenteen, kom dit daarop neer dat 'n brontaalwoord met 'n teikentaalekwivalent vergelyk word sodat die ooreenstemmende betekeniskomponente en die komponente wat verskil uiteindelik uitspraak gee oor die mate van korrelasie.

Komponensiële analises word op verskeie maniere voorgestel, waaronder met matriks-, boom- en skalaardiagramme. Matriksdiagramme kan deur vertalers gebruik word om leksikale stelle voor te stel en

skalaardiagramme vir leksikale reekse (semantiese velde). Boomdiagramme is weer handig om aan te toon waar leksikale gapings voorkom.

Voorbeelde van 'n matriksdiagram en 'n skalaardiagram word hieronder uit Newmark (1988: 116 & 117) aangehaal.

Duits	*Engels*	*Frans*
Ton	*tone*	*ton*
Laut	*sound*	*son*
Schall		
Klang		
Geräusch	*noise*	*bruit*
Lärm (koll.)	*din* (koll.)	*vacarme* (koll.)

(Skalaardiagram uit Newmark, 1988:116)

Verskillen-de seme	om op te sit	op pote	vir een persoon	met rug	met arm-leunings	hout of me-taal
canapé	+	+	-	+	+	+
fauteuil	+	+	+	+	+	+
chaise	+	+	+	+	-	+
tabouret	+	+	+	-	-	+
siège	+	+	±	±	±	±
pouf	+	-	+	-	-	-

(Matriksdiagram uit Newmark, 1988:117)

Newmark (1988:117-123) se houding teenoor komponensiële analise is dat dit 'n hoë mate van akkuraatheid verseker; dit is meer presies en beperkend as parafrases of definisies. Die nadeel is dat dit

onekonomies kan wees in terme van tydsbesteding. Hy beskryf net soos Baker ook verskeie situasies van nie-ekwivalensie, en bespreek hulle uit die oogpunt dat komponensiële analise 'n oplossing te bied het.

Volgens Newmark (loc.cit.) kan komponensiële analise gebruik word om 'n onderskeid te maak tussen brontaal sinonieme, om die onderskeide tussen semantiese velde en die items van leksikale stelle duideliker te maak, om 'n gepaste vertaling vir afgeleide konsepte te vind, om neologismes te vertaal, en om 'mitiese' woorde te vertaal wat 'n amper onvertaalbare inhoud het of verskillende betekenisse afhangende van politieke oortuiging.

4.3.3 Betekenispostulate

'n Fundamentele probleem vir die vertaler is (Bell,1991:91) dat die verbande van ooreenkoms en verskil tussen konsepte (en die woorde wat dit uitdruk) nie noodwendig saamval in die tale waaruit en waarna vertaal word nie. Dit is egter nie moeilik om sulke verbande vir 'n bepaalde taal voor te stel in terme van 'n eenvoudige stelteorie en die konsepte van insluiting en uitsluiting nie.

Hiervolgens is drie verbande moontlik tussen twee konsepte (of tussen twee woorde): insluiting (hiponimie), uitsluiting (antonimie) en gedeeltelike insluiting en uitsluiting (sinonimie).

Volgens een definisie (wat volgens Miller, Beckwith, Fellbaum, Gross en Miller, 1993: 6 gewoonlik aan Leibniz toegeskryf word) is twee uitdrukkings sinonimies as die vervanging van een met die ander nie die waarheidswaarde van die sin waarin dit voorkom verander nie. Die skrywers (loc.cit.) meen dat ware sinonieme volgens hierdie definisie baie skaars is, en gee 'n swakker definisie waarmee sinonieme relatief tot 'n konteks gedefinieer word: twee uitdrukkings is sinonimies in 'n linguistiese konteks C as die substitusie van die een met die ander in C nie die waarheidswaarde verander nie.

Hiervolgens kan '*leer*' en '*studeer*' in die konteks van opvoeding sinonimies wees, terwyl ander kontekste van die woord '*leer*', byvoorbeeld waar sykouse ter sprake is, nie vervanging met '*studeer*' sal toelaat nie. Sinonieme behoort per definisie ook aan dieselfde grammatiese kategorie. Volgens Miller, Beckwith, Fellbaum, Gross en Miller (1993:7) moet sinonimie ook liefs as 'n gradeerbare kwaliteit beskou word; twee woorde kan in 'n meerdere of mindere mate sinonieme van mekaar wees.

Die eenvoudige definisie van 'n antoniem is dat dit die teenoorgestelde van 'n gegewe woord is. Dit is egter nie altyd waar nie, omdat iemand wat nie ryk is nie, nie noodwendig arm is nie. Antonieme kan op verskillende maniere van mekaar uitgesluit word: taksonomies (binêr, veelvoudig of hiërargies), polêr, relatief en inversief.

Hiponieme kom voor waar 'n objek (hiponiem) 'n voorbeeld van 'n ander objek (hiperniem) is. 'n Akkerboom is byvoorbeeld 'n hiponiem van die plant-klas. 'n Belangrike verband tussen hiponieme en hipernieme is dat die hiponiem die kwaliteite van die hiperniem erf. Sonder dat daar vir '*akkerboom*' gespesifiseer hoef te word dat dit groei maar nie 'n lewende wese is nie, sal hierdie kenmerke wel met '*akkerboom*' geassosieer word omdat dit die kenmerke van enige plant is (vgl. Miller, Beckwith, Fellbaum, Gross en Miller, 1993:7).

4.3.4 Proposisionele betekenis

Dit het te make met die verband tussen 'n woord en dit waarna die woord verwys. Dit is hierdie soort betekenis wat 'n mens in staat stel om te kan sê of 'n uiting waar of vals is. Wanneer 'n vertaling as onakkuraat beskryf word, is dit volgens Baker gewoonlik die proposisionele betekenis wat in twyfel getrek word.

4.3.5 Ekspressiewe betekenis

Hierdie soort betekenis kan nie as waar of vals beoordeel word nie. Dit is omdat hierdie soort betekenis te doen het met die taalgebruiker se houding teenoor iets, en nie soseer met die betekenis van die woorde self nie.

Die verskil tussen '*complain*' en '*winge*' kan dien as illustrasie. Sulke woorde word in dieselfde taal gewoonlik sinonieme genoem of byna-sinonieme. Tussen verskillende tale bestaan daar ook korrelasies ten opsigte van ekspressiewe betekenis. Daar is byvoorbeeld 'n verskil tussen '*famous*' en '*fameux*'.

'n Woord se betekenis kan beide proposisioneel en ekspressief wees, of net proposisioneel ('*boek*') of net ekspressief (byvoorbeeld vloekwoorde).

4.3.6 Voorveronderstelde betekenis

Voorveronderstelde betekenis ontstaan uit beperkings op grond van gelyktydige voorkoms of gebruik. Hierdie beperkinge val uiteen in twee groepe: seleksionele beperkinge en kollokasionele beperkinge.

Seleksionele beperkinge is funksies van die proposionele inhoud van 'n woord. 'n Mens verwag 'n menslike subjek vir die adjektief '*studious*' en 'n nie-bewegende een vir '*geometrical*'. Seleksionele

beperkinge kan verontagsaam word in die geval van figuurlike taal.

Kollokasionele beperkinge is semanties arbitrêre beperkings wat nie logies volg uit die proposisionele betekenis van 'n woord nie. 'n '*Wet*' word byvoorbeeld '*gebreek*', maar in Arabies word dit '*weerspreek*'. Net so word '*tande*' '*geborsel*', maar in Duits en Italiaans '*gepoleer*', in Pools '*gewas*' en in Russies '*skoongemaak*'.

4.3.7 Ontlokte betekenis

Ontlokte betekenis ontstaan as gevolg van dialek- en registervariasie. 'n Dialek is 'n 'variëteit' van 'n taal wat deur 'n spesifieke groep sprekers gebruik word. Dialekte kan as die volgende geklassifiseer word (Baker, 1992: 15):

- geografies, bv. 'n Skotse dialek, of Amerikaanse teenoor Britse Engels, en dat eg. '*lift*' gebruik en lg. '*elevator*'
- temporeel, bv. verskillende woorde en strukture wat deur verskillende ouderdomsgroepe in dieselfde taalgemeenskap gebruik word, of woorde wat in verskillende stadiums van die geskiedenis van 'n taal gebruik word, soos 'spreek' of 'ween' destyds en nou 'praat' en 'huil'
- sosiaal, bv. woorde en strukture wat deur sprekers van verskillende sosiale klasse gebruik word.

Register is 'n variëteit van 'n taal wat deur 'n spreker as geskik beskou word vir 'n spesifieke situasie. Registervariasie ontstaan as gevolg van veranderinge in die volgende:

- die veld van diskoers, of hoe die spreker se keuse van linguistiese items beïnvloed word deur die spreker se indruk van haar posisie ten opsigte van of betrokkenheid by die onderwerp waaroor sy praat. Bv., of sy besig is om 'n operasie uit te voer en of sy besig is met 'n gesprek oor medisyne (Baker, 1992: 16)
- die tenoor van diskoers, of die verhouding tussen die mense wat aan die gesprek deelneem. Die taal wat mense gebruik hang af daarvan of die verhouding tussen hulle kind/ma, pasiënt/dokter, werknemer/werkgewer is
- die modus van diskoers, of die rol wat die taal speel (of die diskoers 'n toespraak, essay of 'n stel instruksies is) asook die rol van die medium van oordrag (gesproke of geskrewe).

4.3.8 Semantiese velde

Baker bespreek in 'n gedeelte oor oplossings vir gevalle waar daar nie 'n teikentaalwoord vir 'n gegewe bronwoord is nie die idee van die semantiese veld.

Die woordeskat van 'n taal kan beskou word as 'n stel woorde wat verband hou met 'n stel konseptuele velde. Hierdie velde reflekteer die afdelings en onderafdelings wat deur 'n bepaalde linguistiese gemeenskap afgedruk word op die lewe soos hulle dit ervaar. Die afdelings word in linguistiek semantiese velde genoem. Baker (1992:18) lys as 'n voorbeeld die volgende velde, en sê ook dat die hierdie velde in die meeste tale bestaan: AFSTAND, GROOTTE, VORM, EMOSIE, OORTUIGINGS (*BELIEFS*), AKADEMIESE ONDERWERPE en NATUURLIKE FENOMENE. Die woorde en uitdrukkings in elke veld word soms leksikale stelle genoem.

Semantiese velde het egter ook beperkinge wanneer dit kom by die klassifisering van woordbetekenis. Dit is 'n geslaagde begrip vir woorde met 'n redelik goed gedefinieerde proposisionele betekenis; maar vir baie uitdrukkings en woorde, volgens Baker selfs die meeste woorde (1992:19), is dit onmoontlik om 'n klassifisering te doen.

Semantiese velde en leksikale stelle kan nuttige inligting aan vertalers beskikbaar maak oor die betekeniswaarde wat 'n sekere woord in 'n taalsisteem het. Wanneer 'n woord gesien word in die konteks van die leksikale stel daarvan, kan afleidings gemaak word uit die blote feit dat daardie spesifieke woord gebruik is en nie die ander moontlikhede nie.

Semantiese velde help ook om metodes te ontwikkel vir die hantering van nie-ekwivalensie. Wanneer 'n bronwoord nie 'n teikentaalekwivalent het nie, kan die hipperniem van die hiponiem byvoorbeeld gemodifiseer word om die gaping te vul.

4.3.9 Vier soorte konteks of invloede

Die linguistiese konteks kan dalk beperk wees tot kollokasies, volgens Newmark (1988:193). Hy sê dat in 90% van gevalle die konteks nie meer as dit is nie; '*un hiver blanc*' - ''*n wit winter*' is 'n voorbeeld (in Frans het '*blanc*' baie ander betekenisse). Die Afrikaanse naamwoord '*pos*' het ten minste die twee betekenisse van, enersyds, dié items wat deur 'n poskantoor hanteer word en met 'n seël van sender na ontvanger gaan, of andersyds, 'n werk of 'n vakature binne 'n kommersiële instansie. Indien die woord vertaal moes word met 'n Engelse woord, sou die vertaling vir eersgenoemde '*mail*' of '*post*' wees en vir lg., '*position*' of '*vacancy*'. Indien die woord saam met '*aansoek doen*' of '*vul*' in dieselfde frase sou verskyn, moet dit vertaal word met lg. Met woorde soos '*wegstuur*' of '*ontvang*' sou '*pos*' waarskynlik weer met '*mail*' vertaal word.

Die konteks kan aan die ander kant so groot wees soos 'n sin in die geval van 'n spreekwoord of 'n uitgebreide metafoor. Dit gebeur ook af en toe dat 'n woord linguisties gekondisioneer word deur die gebruik daarvan buite die sin, wanneer dit 'n konsep-woord is wat verskeie kere gebruik of gemodifiseer of gekontrasteer word in ander sinne of paragrawe, of waar dit as 'n stilistiese merker of *leitmotif* dwarsdeur die teks gebruik word.

Die referensiële konteks hou verband met die tema van die teks. Dikwels bepaal die tema die betekenis van baie tegniese woorde soos '*défilement*' ('*scrolling*'), '*stockage*' ('*storage*'), '*rechercher*' ('*search*'), '*fusionner*' ('*merge*'), '*appel*' ('*calling*'), wat woorde uit die elektroniese dataverwerkingsdomein is. Sulke woorde in 'n tegniese dokument wat net deur kenners verstaan sal kan word, oorskry egter gewoonlik nie meer as 5 - 10% nie.

Derdens is daar die kulturele konteks, woorde wat gekoppel word aan die bepaalde soort gedrag of denkwyse van 'n bepaalde taalgemeenskap, en woorde wat kultureel (byvoorbeeld, '*kuffiah*', 'n Arabiese kopkleed) of universeel (byvoorbeeld, '*tee*') kan wees.

Laastens is daar die idiolek van 'n skrywer, die persoonlike konteks. Mense gebruik woorde dikwels op 'n manier wat kenmerkend is van 'n spesifieke persoon.

4.4 Kollokasies

In linguistiek word kollokasies volgens Newmark (1988:212) gewoonlik gedefinieer as die gevestigde gesamentlike gebruik van verskillende leksikale items. Vir die vertaler is die definisie nouer: die kollokasie word beskou as die belangrikste geleentheid vir kontekstuele bepaling van woordbetekenis. Verder sê Newmark (loc.cit.) dat kollokasies bestaan uit leksikale items wat hoofsaaklik in hoëfrekwensie grammatikale strukture voorkom, nl.:

- adjektief plus naamwoord
  - '*heavy labour*', '*travail musculaire*', '*schwere Arbeit*'
  - '*runaway (galloping) inflation*', '*galoppierende Inflation*', '*l'inflation galopante*'
  - '*economic situation*', '*situation économique*', '*Konjunkturlage*'
  - '*inflationary pressure*', '*pressions (tensions) inflationnistes*', '*Inflationsdruck*'

- naamwoord plus naamwoord
  - '*nerve cell*', '*cellule nerveuse*', '*Nervenzelle*'
  - '*government securities*', '*effets publics*', '*Staatspapiere (Staatsanleihen)*'
  - '*eyeball*', '*globe oculaire*', '*Augapfel*'

- werkwoord plus objek (die werkwoorde is gewoonlik 'n aksie en die objek 'n naamwoord)
  - '*pay a visit*', '*faire une visite*', '*einen Besuch machen (abstatten)*'
  - '*score (win) a victory*', '*remporter une victoire*', '*einen Sieg erzielen*'
  - '*read a(n) (academic) paper*', '*faire une communication*', '*ein Referat halten*'
  - '*attend a lecture*', '*eine Vorlesung hören of besuchen*', '*suivre une conférence*'

Om te weet of 'n kollokasie in algemene gebruik is, of dit natuurlik is en of dit net aanvaarbaar is, is volgens Newmark (1988:213) een van die belangrikste probleme in vertaling. 'n Handige manier om kollokasies te benader is om te oorweeg wat die aanvaarbare kollokasionele grense van 'n woord is. Hierdie benadering blyk veral van toepassing te wees op adjektiewe van kwaliteit en werkwoorde wat aktiwiteite sowel beskryf as uitdruk; '*blême*' kollokeer met '*face*' en '*light*', maar gewoonlik nie met objekte nie. '*trouble*' met voorkoms, emosies, vloeistowwe, maar nie mense nie. '*grincer*' met deure ('*creak*'), tande ('*grind*' of '*gnash*'), metaalvoorwerpe ('*grate*'), maar nie met diere nie. '*keusch*' met mense en hulle uitdrukkings ('*pure*', '*chaste*'), maar nie met objekte of kos nie.

Baker (1992: 47) noem ook voorbeelde tot toeligting, en sê dat woorde wat beskou word as sinonieme, dikwels verskillende kollokate het: '*Rules*' word gebreek, maar nie '*regulations*' nie. Daar word gepraat van '*wasting time*', maar nie van '*squandering time*' nie. Insgelyks word van beide '*deliver a verdict*' en '*pronounce a verdict*' gepraat, maar terwyl '*pronounce a sentence*' aanvaarbaar is, kan '*deliver a sentence*' nie gebruik word nie.

Baker (1992:48) sê verder dat wanneer woorde kollokeer, die verband gewoonlik ook kan bestaan tussen die verskillende vorme van daardie woorde. Byvoorbeeld, '*achieving aims*', '*aims having been achieved*', '*achievable aims*', en '*the achievement of an aim*'. Maar dit kan ook wees dat die verband net tussen sekere vorme bestaan. 'n Mens '*bend rules*', maar '*rules*' kan nie beskryf word as '*unbendable*' nie; dit is eerder '*inflexible*'.

4.4.1 Kollokasionele bereik

Sommige woorde kan met 'n groter stel woorde kollokeer as ander. Die Engelse werkwoord '*shrug*', het byvoorbeeld 'n beperkte kollokasionele bestek; dit kom tipies net voor saam met '*shoulders*', en 'n menslike subjek. In teenstelling daarmee, het '*run*' 'n wye verskeidenheid van moontlike kollokate; onder meer: '*company*', '*business*', '*show*', '*car*', '*stockings*', '*tights*', '*nose*', '*wild*', '*debt*', '*bill*', '*river*', '*course*', '*water*' en '*colour*'.

Tegelykertyd is vertaling (Newmark, 1988:213) soms 'n aanhoudende worsteling om toepaslike kollokasies te vind, 'n proses om naamwoorde met toepaslike werkwoorde en werkwoorde met toepaslike naamwoorde byeen te bring, om adjektiewe met toepaslike naamwoorde en bywoorde of bywoordelike frases met toepaslike werkwoorde byeen te bring, of om toepaslike voegwoorde of voegwoordelike partikels te kollokeer.

Woorde het dikwels net een kollokaat, soos die geluide wat diere en musiekinstrumente maak. In sulke gevalle is daar nie veel van 'n vertaalprobleem nie - daar moet net vertaal word met die ooreenstemmende woord in die teikentaal. Daar is ook 'n klein aantal moontlike vertaalekwivalente vir

massanaamwoorde en versamelwoorde. Sommige werkwoorde selekteer omringende naamwoorde op grond van semantiese kenmerke soos menslikheid of abstraktheid. 'n Paar werkwoorde ('*work hard*', '*deeply regret*', '*devoutly hope*') en adjektiewe ('*profoundly unnecessary*', '*immensely disturbing*', '*totally wrong*', '*desperately unhappy*') word gekollokeer met bywoorde; baie hiervan word clichés.

Newmark bespreek ook die vertaling van akronieme, eponieme, eiename, neologismes en gewigte, mate, kwantiteite en geldeenhede, waarop nie hier ingegaan word nie.

4.5 'n Gedeelte van Bell se vertaalmodel

Terwyl moderne linguistiek 'n beskrywende houding inneem teenoor die subjek daarvan, word vertaalteorie nog soms beskou as die neerlegging van reëls vir byvoorbeeld Engels, soos 'moenie 'n sin met 'n voorsetsel eindig nie' of 'moenie die infinitief verdeel nie'. Bell sê sy benadering is soortgelyk aan dié van linguistiek: hy wil die reëls wat vertalings tot stand bring beskryf en verduidelik om die proses, die produk en die konsep te kan verklaar, en nie om vertalings in die eerste plek te beoordeel nie.

Een van die definisies van vertaalteorie sou wees om te sê dat dit moet poog om 'n lewensgetroue weergawe te wees van 'n taalgebruiker se tweetalige vermoë. Bell (1991:37) sê dit betrek 'n integrasie van die linguistiese kennis van die twee tale by spesifieke en algemene kennis van die domein en van die wêreld deur kontrastiewe linguistiese analise.

Verderaan (1991:39) sê hy dat 'n mens, gegee die hernieude belangstelling in rekenaargesteunde vertaling, kon probeer om die vakmanskap van die menslike vertaler as 'n ekspertstelsel te bestudeer - 'n gedagte wat ook deur Nirenburg (1987:20e.v.) geopper is. 'n MT-ekspertstelsel sou dan die volgende kennis en vaardighede moet bevat:

'n kennisbasis bestaande uit:

- brontaalkennis: 'n sintaktiese reëlsisteem, 'n leksikon en semantiek en teksgenererende stelsels;
- teikentaalkennis: ekwivalent aan brontaalkennis, maar vir die teikentaal
- teks-tipe kennis: of die teks byvoorbeeld poëties-fiktief of nie-fiktief en feitelik is
- domeinkennis: informasie oor die werklike-wêreldkonteks
- kontrastiewe kennis van elkeen van die bogenoemde

en 'n afleidingsmeganisme wat die volgende kan doen:

- die dekodering van tekste, dit is die lees en verstaan van brontaalteks
- die enkodering van tekste, dit is die skryf van teikentaalteks

Bell se model sluit drie meganismes in waarmee die woorde wat in 'n brontaalteks voorkom herken kan word as brontaalwoorde, sodat sintaktiese en morfologiese inligting daarop toegepas kan word en vertaalekwivalente beskikbaar word vir die vertaler. Dit is die frekwente leksis geheue (FLS), die frekwente struktuur geheue (FSS) en die leksikale soekmeganisme (LSM).

In die FLS word woorde en frases wat baie gebruik word en soms ook spesifieke vertalings gestoor. Wanneer die woorde in 'n bronteks afgebaken is en dit dan in die woordeboek opgesoek moet word (en hier word bedoel die geheue van die vertaler), word daar eers in die FLS gesoek. Die FSS pas gewoonlik meer as een woord oftewel 'n reeks woorde met 'n vaste volgorde en met dieselfde elemente en ken die ooreenkomstig gestoorde sintaktiese struktuur daaraan toe - vervolgens hoef die aftaster nie daardie stuk teks te ontleed nie. Eers wanneer daar woorde in die teks oorbly wat nie met behulp van een van hierdie twee meganismes gevind is nie, word van die normale LSM gebruik gemaak, wat woorde in die omvattende woordeboek sal opsoek, en van die ontleder of aftaster (*parser*), wat die sintaktiese struktuur sal aftas.

Woorde wat 'n hoë gebruiksvoorkoms het, soos lidwoorde en '*en*', sal byvoorbeeld in die FLS gestoor word, maar ook enige ander woorde wat dikwels gebruik word. Dit kan aangepas word by sekere soort dokumente in sekere kennisdomeine.

Bell (1991:47) gee die voorbeeld van

'*Once upon a time there was ...*'

en die Italiaanse vertaling daarvan

'*C'era una volta*'

as 'n tipiese inskrywing in die FLS ('n frase met 'n spesifieke vertaling) en die FSS ('n adjunk, 'n koppelwerkwoord en die ekspletief '*there*', dus 'n koppelwerkwoordfrase wat nog 'n subjek kortkom).

4.6 Nie-ekwivalensie vertaalprobleme

Baker (1992:17) sê in 'n gedeelte oor die probleem van nie-ekwivalensie dat dit so te sê onmoontlik is om absolute riglyne te gee vir die hantering van die verskillende tipes nie-ekwivalensie wat tussen tale bestaan. Sy (op.cit.:21 e.v.) noem en bespreek 'n aantal algemene nie-ekwivalensie probleme, oftewel vertaalprobleme waar daar nie 'n teikentaal vertaalekwivalent bestaan vir 'n brontaalwoord nie.

4.6.1 Kultureel-spesifieke konsepte

Die brontaalwoord druk 'n konsep uit wat nie in die teikentaal bestaan nie; dié vorm van nie-ekwivalensie word gewoonlik kultuur-spesifiek genoem. Die konsep wat deur die Engelse woord *'privacy'* uitgedruk word, is 'n voorbeeld hiervan; dit is volgens Baker (loc.cit.) baie moeilik om dit na ander tale te vertaal.

4.6.2 Die brontaalkonsep is nie in die teikentaal geleksikaliseer nie

'n Brontaalwoord druk 'n konsep uit wat bekend is in die teikentaal, maar die teikentaal het nie 'n woord vir die konsep nie. Die woord *'landslide'* wat ' 'n oorweldigende meerderheid' beteken, het nie 'n geredelik beskikbare vertaling in Afrikaans nie.

4.6.3 Die brontaalwoord is semanties kompleks

Woorde word dikwels deur taalgemeenskappe geskep om 'n komplekse saak op 'n bondige manier aan te dui. Indien daar nie 'n nodigheid was vir die teikentaalgemeenskap om so 'n bondige manier te vind nie, sal daar nie 'n woordekwivalent vir die bronwoord wees nie. 'n Voorbeeld hiervan is die Brasiliaanse woord *'arruação'* omslagtig beskryf in woord as 'die skoonmaak van grond onder koffiebome deur die rommel daar op te ruim en op te hoop in die middel van twee rye ten einde die versameling van bone wat gedurende oestyd afval te vergemaklik' (ITI News, 1988: 57).

4.6.4 Die bron- en teikentale maak verskillende onderskeide ten opsigte van betekenis

Die teikentaal maak meer of minder onderskeide in betekenis as die brontaal. Byvoorbeeld: in Indonesies word 'n onderskeid gemaak tussen om in die reën uit te gaan sonder die wete dat dit reën (*'kehujanan'*) en om in die reën uit te gaan met die wete dat dit reën (*'hujanhujanan'*). Nie Engels of Afrikaans maak hierdie onderskeid nie

4.6.5 Die teikentaal het nie 'n hiperniem nie

Die teikentaal kan verskeie spesifieke woorde hê (hiponieme) terwyl dit nie 'n algemene woord (hiperniem) het om die semantiese veld aan te dui nie. Russies het nie 'n ekwivalent vir *'facilities'*, menende enige toerusting, gebou, dienste ensovoorts wat voorsien word vir 'n bepaalde aktiwiteit of doel

nie. Dit het egter wel verskeie spesifieke woorde en uitdrukkings wat beskou kan word as tipes fasiliteite, byvoorbeeld '*sredstva peredvizheniya*' ('*means of transport*'), '*naem*' ('*loan*'), '*neobkhodimye pomeschcheniya*' ('*essential accomodation*'), en '*neobkhodimoe oborudovanie*' ('*essential equipment*').

4.6.6 Die teikentaal het nie 'n spesifieke term (hiponiem) nie

Dit is meer dikwels die geval dat tale neig om algemene woorde (hipernieme) te hê, maar nie spesifieke woorde nie. In Engels word onder '*jump*' byvoorbeeld meer spesifieke woorde aangetref soos '*leap*', '*vault*', '*spring*', '*bounce*', '*dive*', '*clear*', '*plunge*', en '*plummet*'; en onder '*house*' weer '*bungalow*', '*cottage*', '*croft*', '*chalet*', '*lodge*', '*hut*', '*mansion*', '*manor*', '*villa*', en '*hall*'. Daar is in baie gevalle nie direkte vertaalekwivalente vir hierdie woorde beskikbaar nie.

4.6.7 Verskille ten opsigte van fisiese en interpersoonlike perspektiewe

Fisiese perspektief is in sommige tale belangriker as in ander. Fisiese perspektief het te doen met die verhoudings tussen mense en tussen mense en plekke, soos dit uitgedruk word in woorde soos '*come/go*', '*take/bring*', en '*arrive/depart*'. Perspektief kan ook die verhouding insluit wat bestaan tussen die deelnemers aan 'n diskoers. Japanees het byvoorbeeld ses ekwivalente vir '*gee*', afhangende van wie vir wie gee: '*yaru*', '*ageru*', '*morau*', '*kureru*', '*itadaku*', en '*kudasaru*'. (McCreary, 1986).

4.6.8 Verskille ten opsigte van uitdrukkingsbetekenis

Daar kan 'n teikentaalwoord wees wat dieselfde proposisionele betekenis het as die brontaalwoord, maar dit het 'n verskillende ekspressiewe betekenis. Die verskil kan groot of klein wees, maar belangrik genoeg om 'n vertaalprobleem te veroorsaak. Dit is gewoonlik makliker om ekspressiewe betekenis by te voeg as om dit te verminder. As die teikentaalekwivalent neutraal is in vergelyking met die brontaalwoord, kan 'n vertaler 'n element, soos 'n bywoord of 'n modifiseerder, by die teikentaalekwivalent voeg sodat dit ekspressiewe betekenis van die brontaalwoord weerspieël. Die Engelse '*batter*' het byvoorbeeld die neutrale ekwivalent '*tataku*' in Japanees wat eenvoudig '*slaan*' ('*beat*') beteken; deur '*savagely*' of '*ruthlessly*' by te voeg kan die vertaling die ekspressiewe betekenis van die brontaalwoord uitdruk.

4.6.9 Verskillende vorme

Sekere prefikse en suffikse in die brontaal wat onder meer proposisionele betekenis kan oordra, het nie ekwivalente in die teikentaal nie. Arabies het byvoorbeeld nie die morfologiese potensiaal om soos

Engels '*greenish*' van '*green*' af te lei of '*retrievable*' van '*retrieve*' nie. Daarom moet parafrases in hierdie gevalle gebruik word om 'n gepaste vertaling te gee; byvoorbeeld '*can be retrieved*' vir '*retrievable*'.

Baker (1992:24) sê dat dit baie belangrik is vir vertalers om die bydrae te verstaan wat affikse maak tot die betekenisse van woorde en uitdrukkings. In Engels word hierdie affikse dikwels gebruik om aan nuwe woorde beslag te gee, sodat semantiese gapings in die taal gevul kan word.

4.6.10 Verskille ten opsigte van frekwensie en die gebruiksdoel van spesifieke vorme

Daar bestaan soms gevalle waar 'n bepaalde vorm in die brontaal 'n ekwivalente vorm in die teikentaal het, maar terselfdertyd is daar egter 'n verskil wat gebruiksfrekwensie betref. Engels gebruik die kontinuerende *-ing* byvoorbeeld baie meer as wat die ooreenstemmende Duitse vorm gebruik word. Deur so 'n vorm dus met die ooreenstemmende vorm in Duits te vertaal sal uitloop op 'n onnatuurlike vertaling.

4.6.11 Die gebruik van leenwoorde in die brontaal

Leenwoorde verleen gewoonlik 'n mate van flair en gesofistikeerdheid aan die teks waarin dit voorkom; voorbeelde hiervan is '*au contraire*' en '*dilettante*'. Wanneer die leenwoord in die brontaal bestaan en daar is nie 'n ooreenstemmende leenwoord in die teikentaal nie, word vertaalkwaliteit dikwels ingeboet.

Baker verskaf ook 'n lys van strategieë wat deur professionele vertalers gebruik word om probleme van nie-ekwivalensie op te los. Dit kom neer op die strategieë wat alreeds kortliks hierbo genoem is, maar sluit ook metodes in soos weglating en fisiese illustrasie.

4.7 Opsomming

Uit wat hierbo geskryf is lyk dit asof die kontrastering van die twee tale wat by 'n vertaling betrokke is 'n belangrike element uitmaak van die beskrywing van die vertaaltaak. Die gevaar wat by taalleer bestaan om te veel te maak van so 'n kontrastiewe vergelyking geld nie vir vertaalteorie nie. Die begrip oordrag is nie ignoreerbaar of on-nuttig in taalleer nie. 'n Mens moet egter versigtig wees om die rol daarvan in die tweedetaalleerproses nie te oorbeklemtoon nie.

In hierdie hoofstuk word verskeie universele nie-ekwivalensie gevalle verduidelik. Dit help om die vertaaltaak te vergemaklik omdat probleme gekategoriseer kan word.

Dit kan ook wees dat wanneer taalleerteorie en universele grammatika of tipologiese grammatika woordleer en leksikonvoorstelling en -gebruik beter uiteensit, daar nie veel meer gedoen kan word as om die woorde van spesifieke tale met universele semantiese konsepte in verband te bring nie, sodat die verskille tussen tale se woordeskat steeds duidelik word.

Dit is ook opmerklik dat die nie-ekwivalensies wat Baker bespreek meer verfynd is as dié van Stockwell, Bowen and Martin (1965) wat in hoofstuk 2 aangehaal is.

Dit is belangrik om die konteks in ag te neem om verantwoorde vertalings te maak, en volgens wat outeurs soos Newmark en Baker sê, veral kollokasionele konteks. Aangesien morfologiese, sintaktiese en semantiese aspekte, sowel as sosiaal-konvensionele aspekte, betrokke is by die herkenning, beperking en verduideliking van kollokasies, en kollokasies gewoonlik uit twee woorde bestaan, kan dit beskou word as 'n verdere belangrike uitgangspunt vir die inverbandbrenging van twee tale se leksikons. Die betrekking van alle belangrike aspekte van taal en die relatiewe eenvoud van konfigurasie is die onderliggende argumente vir so 'n stelling.

Hoofstuk 5: Masjienvertalingsteorie

5.1 Inleiding

Masjienvertalingsteorie (MT-teorie) het, sedert dit ontstaan het, op verskeie maniere ontwikkel sodat dit vandag 'n gevarieërde wetenskap in diens van tegnologie is. Die verskillende benaderings is die gevolg van 'n ontwikkelingsproses om vertalings te verbeter en om die uitbreiding van bestaande stelsels te vergemaklik.

MT-stelselontwerpers het kennis geneem van linguistiese teorie en natuurliketaalverwerking (NLP) literatuur wanneer dit haalbaar was om dit as 'n proses te kan implementeer op 'n rekenaar. Soos daar verskillende linguistiese teorieë bestaan, is daar ook MT-stelsels met uiteenlopende grondfilosofieë. MT moes en moet ook op 'n pragmatiese manier werk, omdat dit soms makliker is om 'n beter vertaling te genereer met 'n presies gekodeerde omseilingstegniek as om 'n ingewikkelde rekenaarmatige vertaalproses te volg wat baie tyd in beslag neem en moeilik is om te implementeer. Hierdie pragmatiese benadering het gelei tot die bevordering van kennis oor taalverwerking en dalk ook tot uitbreidings en verlengings van teorieë en modelle.

In hierdie hoofstuk word die algemene beginsels en terminologie van MT bekyk, terwyl agtergrond ook verskaf word oor die vordering van die tegnologie sedert die ontstaan daarvan in die 1950s. Die belangrikste deel van die hoofstuk is die bespreking oor die voorstelling van leksikale kennis - die woordeboeke wat MT-stelsels gebruik.

5.2 Vertaalmodelle

Sedert MT in die 1950s 'n navorsingsonderwerp geword het in baie universiteits- en regeringsinstansies met die oog op veral kommersiële gebruik, tot ongeveer 40 jaar later, het daar drie hoofstroommodelle ontstaan. Dit is die direkte vertaal-, oordragvertaal- en interlingua vertaalmodelle. Oordragvertaling en interlinguavertaling kan saamgegroepeer word as indirekte vertaalmetodes.

5.2.1 Direkte vertaling

Veral aan die begin is daar van direkte vertalingstelsels gebruik gemaak. Hierdie soort stelsel gebruik 'n morfologiese aftaster en genereerder, leksikale vervanging en lokale, beperkte sintaktiese transformering. Daar is geen intermediêre voorstelling tussen die twee tale nie. Die benadering word beskou as naïef en primitief; dit is egter so dat direkte vertaling 'n bydrae kan maak binne 'n meer gesofistikeerde model. Die

vertaling van idiomatiese frases wat dikwels gebruik word is 'n geval waar direkte vertalingsmetodes handig te pas kan kom en dalk selfs psigologiese realiteit het.

5.2.2 Interlingua vertaling

Die bronteks word ontleed en daaruit word 'n intermediêre voorstelling geskep, onafhanklik van die aard van die teikentaal, en die teikenteks word direk vanaf die interlinguavoorstelling gegenereer. Aanvanklik is probeer om 'n interlingua te skep wat universeel sou wees, en dus enige bron- en teikentaal sou kon hanteer, maar volgens Hutchins en Somers (1992:74) is hedendaagse stelsels minder ambisieus.

Dorr (1993: 96-97) beaam dat dit oor die algemeen moeilik is om 'n stel universale te skep, selfs vir 'n afgebakende groep tale. Sy sê egter ook (loc. cit.) dat daar 'n oplewing begin het op die gebied van leksikale voorstelling en organisasie (veral met betrekking tot werkwoorde) wat 'n voortdurende poging inisieer om klasse van leksikale kennis af te baken sodat natuurliketaalverwerking vergemaklik kan word. As 'n uitvloeisel van hierdie poging word dit meer en meer haalbaar om die komponente van betekenis wat algemeen is van werkwoorde te isoleer, veral van werkwoorde in sekere klasse. Hierdie komponente van betekenis kan dan gebruik word om die leksikale voorstelling van werkwoorde oor tale in die algemeen moontlik te maak.

Die grootste voordeel van interlingua stelsels is dat die moeite vir die byvoeging van 'n nuwe taalpaar met vertalings in beide rigtings baie minder is as vir oordragstelsels. 'n Nuwe taal vereis net twee nuwe modules: 'n analiseer- en 'n genereermodule. Vir die byvoeging van twee nuwe tale, word vier nuwe modules vereis, onafhanklik van hoeveel tale alreeds deur die stelsel gehanteer word.

Die grootste nadeel is te wyte aan die interlingua voorstellingsmeganisme. Dit is baie moeilik om 'n interlingua te definieer, selfs vir tale met 'n nou verband.

Leksikale vertaling in interlinguastelsels is nie tweetalig nie. Die bronwoord word in die interlingua met 'n semantiese konsep 'vertaal' en die semantiese konsep word saam met die sintaktiese voorstelling oorweeg vir die keuse en generering van 'n gepaste teikentaalwoord.

5.2.3 Oordrag vertaling

Oordragvertalingstelsels (*transfer systems*) maak ook gebruik van 'n intermediêre voorstelling, maar hierdie voorstelling is taalafhanklik en tweetalig. Die funksie van die intermediêre module is om brontaal intermediêre voorstellings in teikentaal intermediêre voorstellings te verander.

Een van die nadele van oordragvertaling wat duidelik is, is dat MT-stelsels wat daarvan gebruik maak moeilik uitbreibaar is. Vir elke nuwe taalpaar moet daar nuwe oordragmodules, sowel as analiseer- en 'n genereermodules geskep word. Indien die stelsel uit een taalpaar bestaan, en 'n derde taal word bygevoeg wat na en vanaf albei hierdie tale moet vertaal, moet een nuwe analiseer module, twee oordragmodules vir elke taal, en hopelik net een genereermodule geskep word; dus ses modules waar 'n interlingua stelsel net twee nodig het. Vir die byvoeging van nog 'n taal sal ses oordragmodules nodig wees, terwyl 'n interlinguastelsel net die analiseer- en genereermodule sou nodig hê.

Die intermediêre voorstellings is nie in alle oordragstelsels ewe omvattend of diepgaande nie. Die eintlike oordrag word op twee vlakke gedoen: leksikale en strukturele oordrag.

Leksikale oordrag beteken dat 'n bronwoord met 'n teikenwoord vervang word. In gevalle waar daar net een teikenwoord vir 'n bronwoord is, is leksikale oordrag onproblematies, tensy daar sintaktiese komplikasies is. Ook wanneer daar baie-enkelwoorde-tot-een-ekwivalent ter sprake is, is die oordragproses nie verwikkeld nie. Die grootste probleem bestaan wanneer die situasie die omgekeerde is. Dan moet die semantiese kenmerke van woorde en die konteks van die woord in ag geneem word om die korrekte vertaling te kan kies.

Strukturele oordrag is nodig wanneer die teikentaalteks ongrammatikaal sou wees op sintaktiese vlak indien dit die struktuur van die bronteks net so oorneem. Twee gevalle waar strukturele oordrag nodig sal wees is wanneer die twee tale verskil ten opsigte van V-O (werkwoord-objek) volgordes of wanneer die vertaalekwivalent van 'n brontaalwoord 'n wysiging in die struktuur van die teikenteks afdwing.

Hoe meer diepgaande die analise van die bronteks is of hoe meer abstrak die intermediêre voorstelling van die bronteks is, hoe minder kompleks behoort strukturele oordrag te wees. Kasusrolle word byvoorbeeld as 'n algemene beskrywingsbeginsel gebruik in die saamstel van 'n strukturele interlingua om die rolle van sinskonstituente en hulle verband met mekaar te definieer. Dit is dan hierdie verbande en die keuse van vertaalekwivalente wat die struktuur bepaal wat vir die teikentaalteks gegenereer moet word. Hutchins en Somers (1992:117) beskryf die situasie soos volg:

> 'Unfortuantely, despite intensive study by linguists, there is no widely agreed set of case relations, so researchers in MT systems adopting this approach are generally obliged to devise their own set of case roles.'

5.3 Linguistiek en MT

In hierdie gedeelte word 'n kort oorsig gegee van die linguistiese teorieë en konsepte wat 'n invloed uitoefen op MT.

5.3.1 Transformasionele grammatika

Die aanvanklike transformasionele grammatika van Chomsky kon rekenaarlinguiste met nuwe oë laat kyk na konteksvrye frase-grammatikas. Die vermoë van transformasies om aan strukture te verander was egter te kragtig, en probleme is ook ondervind met die verwyderdheid van die semantiese en sintaktiese modules. Terwyl dieptestrukture nie semantiese voorstellings is nie, is dit ook nodig om sintaktiese struktuur in ag te neem wanneer 'n semantiese voorstelling gegenereer word.

Chomsky se model was egter die inisieerder van 'n hele nuwe benadering in teoretiese linguistiek, en sy eie teorie het self gedaanteverwisselings ondergaan. Een van die nuutstes daarvan is *Government and Binding* teorie, wat een transformasie postuleer, genaamd move-α. Hierdie bewegingsreëł kan soos transformasies op strukture toegepas word, maar is onderhewig aan die beginsels van Universele Grammatika en die spesifieke parameters van 'n taal.

5.3.2 Valensiegrammatikas

Hutchins en Somers (1992:37) sê oor die konsep van valensiegrammatikas dat dit op twee maniere nuttig is vir vertaling: die eerste is vir die langsmekaarstelling van werkwoorde van die twee tale. Hoewel ekwivalente werkwoorde in die twee tale maklik kan verskil ten opsigte van die sintaktiese struktuur waarin hulle gebruik word, is hulle numeriese valensie gewoonlik dieselfde. Met valensie word bedoel hoeveel argumente 'n werkwoord onder verpligting is om te neem.

Die tweede is vir die vertaling van voorsetsels, veral voorsetsels wat die partikel is in voorsetselwerkwoorde (*phrasal verbs*), soos *look at, look up, hand in, hand over.* Indien die soort valensie ooreenstem vir 'n werkwoord, met ander woorde beide die tale beskou die voorsetsel as deel van die betekenis van die werkwoord, kan die korrekte vertaling op sterkte van 'n woordeboekinskrywing gemaak word. Valensie-ooreenstemming kan ook geld ten opsigte van adjektiewe en naamwoorde.

5.3.3 Kasusgrammatika

In kasusgrammatikas word die semantiese aspekte van woorde, en veral van werkwoorde, beklemtoon. Die argumentstrukture van werkwoorde word daarvolgens gespesifiseer en dit word dan deur sinskonstituente gevul wat in 'n bepaalde posisie is met 'n bepaalde sintaktiese funksie. Bekende kasusse is AGENT, INSTRUMENT, ONTVANGER en TEMA. Kasusgrammatika maak deel uit van *Government and Binding* teorie, maar word ook in formalismes gebruik wat nie gebaseer is op hierdie

teorie nie. Een van die grootste probleme met kasusgrammatika is om 'n universeel geldige stel kasusverbande op te stel. Die totale stel verbande is gewoonlik sowat tien.

5.3.4 Leksikale Funksionele Grammatika (LFG)

In LFG is daar 'n volgorde van voorstellings wat c-strukture (*constituent structures*) en f-strukture (*feature structures*) genoem word. C-strukture word gebou met konteksvrye grammatikas (KVG), maar KVG's wat geassosieer word met aksies wat spesifiseer hoe kenmerk-toepassings gedoen moet word sodat f-strukture gebou kan word. Een van die voordele van hierdie grammatika is dat dit spesifiek gerig is op die manipulering van kenmerke (of *features*) van woorde en strukture (soos 'getal' of 'persoon' of 'lewend').

Die manipulering van kenmerke word in LFG en GPSG (hieronder) gedoen met behulp van 'n bewerking wat unifikasie genoem word. Dit bepaal watter kenmerk uit 'n groep die regte of toepaslikste is op grond van die voorkoms van dieselfde kenmerk by omringende woorde of strukture.

5.3.5 *Government and Binding* teorie (GB)

Dit word ook Beginsels en Parameters genoem. Beginsels is algemene postulate oor hoe taal beskryf kan word, met groot klem op die sintaktiese deel, en parameters is hoe uit die opsies van beginsels gekies word om 'n spesifieke taal te laat manifesteer.

Dit bestaan uit verskeie modules, waaronder *X-bar* grammatika ('n baie eenvoudige, maar hoogs geabstraheerde KVG wat gebaseer is op hoe soortgelyk die strukture van verskillende soorte sinskonstituente is) en 'n Theta-teorie (kasusgrammatika); die ander modules is Kasusteorie (wat die sintaktiese rolle van sinskonstituente spesifiseer), Bindingsteorie (wat anaforiese verhoudinge hanteer), Afbakeningsteorie (*Bounding theory*, wat beperkings plaas op die alfa-bewegingsreël, en Kontroleteorie (wat die gevolge van die toepassing van die bewegingsreël, leë argumentposisies, indekseer).

5.3.6 Generalised Phrase Structure grammatika (GPSG)

Hierdie soort grammatika probeer om sonder transformasionele of bewegingsreëls klaar te kom. Die KVG daarvan word gevorm met dominansie-reëls, wat verbande tussen die konstituente van strukture aandui, en volgorde-reëls, wat aandui wat die volgorde is waarin daardie konstituente moet verskyn in 'n oppervlaktestruktuur. Dit het verskeie ander soorte reëls of kenmerkprosedures wat byvoorbeeld gapings moet hanteer (wat deur Kontroleteorie gedoen word in GB) en die beweging van konstituente. Terwyl GB

van 'n logiese vorm gebruik maak om 'n meer semantiese voorstelling van 'n sin te kry, gebruik GPSG 'n soort Montague-grammatika.

Volgens Hutchins en Somers (1992:44) lê die aantreklikheid van Montague grammatikas daarin dat dit 'n manier voorstel om sintaktiese teorie en die gevestigde metodes van formele logika te integreer. Die basiese beginsel is dié van semantiese komposisionaliteit, waarmee bedoel word dat die betekenisse van komplekse uitdrukkings (bv. sinne) funksies is van die betekenisse van die uitdrukkings (bv. woorde) waaruit hulle bestaan.

5.4 Die leksikon in 'n paar hedendaagse MT-stelsels

Verskeie MT-stelsels is vandag in gebruik; dit is stelsels met verskillende argitekture en modelle, stelsels wat verskillende linguistiese teorieë gebruik, en stelsels wat verskillende mate van en soorte sukses behaal. In die gedeelte word die leksikons van 'n paar sulke stelsels beskryf in terme van die voorstelling van leksikale inligting. Dit is nie 'n omvattende lys nie, en dit is ook nie noodwendig 'n deursnee-lys nie; dit demonstreer wel die algemeenste soorte inligting oor leksikale items in die stelsels.

5.4.1 Systran

Systran is al sowat 20 jaar in bedryf, is al deur die EG gebruik, en is 'n direkte vertaalstelsel.

Die leksikons is 'n hoof stamwoordeboek, 'n tweetalige woordeboek met enkelwoord inskrywings en verskeie multi-woord kontekswoordeboeke. In die hoof stamwoordeboek het elke woord se stamvorm 'n volledige morfologiese, sintaktiese en semantiese beskrywing: grammatiese kategorie, valensie, subkategoriseringskenmerke en seleksiebeperkinge, en semantiese kenmerke soos 'lewendig', 'toestand' of 'abstrak' word aangedui. In die tweetalige woordeboek word 'n ekwivalente vertaalvorm verskaf.

Die kontekswoordeboeke bestaan onder meer uit 'n idioomwoordeboek, wat veral vaste uitdrukkings hanteer, 'n beperkte-semantiekwoordeboek waarin leksikale items wat uit meer as een woord bestaan wat gewoonlik saamgebruik word (kollokasies) verskyn en 'n homograafwoordeboek waarin inligting verskaf word oor die sintaktiese omgewing van homograwe sodat daar in sommige gevalle op grond van grammatiese konteks tussen homograwe onderskei kan word.

Systran het 'n oorwegend sintakties-gedrewe benadering - dit probeer soveel as moontlik probleme oplos met sintaksis voordat dit van semantiese inligting gebruik begin maak (vgl. Hutchins en Somers,1992:178 -179).

5.4.2 Météo

Die Météo-stelsel vertaal sedert 1977 tot vandag weerberigte van Engels na Frans; dit behaal 'n groot mate van sukses omdat dit tot 'n bepaalde kennisdomein beperk is met 'n eie sub-taal. Dit is 'n stelsel wat in wese 'n oordrag-ontwerp het, maar dit maak weens die aard van die tekste wat dit moet vertaal in 'n groot mate staat op direkte vertaalmetodes.

Linguistiese informasie is opgeneem in drie woordeboeke: een elk vir idiomatiese taalgebruik, plekname en algemene woordeskat. Die idiomatiese woordeboek bevat inskrywings wat vaste uitdrukkings is en leksikale items wat uit meer as een woord bestaan wat saam gebruik word; die pleknaamwoordeboek het net inskrywings vir plekname waarvan die Engels en die Frans van mekaar verskil.

Die hoofwoordeboek bevat woorde wat alreeds morfologies vervoeg is. Elke inskrywing het 'n Franse vertaalekwivalent, 'n grammatiese kategorie, semantiese kenmerke en inligting vir die morfologiese generering van die teikenwoord. Sintaktiese informasie is minimaal: inskrywings vir bywoorde dui aan of hulle met adjektiewe, werkwoorde of voorsetsels verbind kan word; werkwoorde word gespesifiseer ten opsigte van oorganklikheid. Semantiese kenmerke sluit merkers in soos '*waai*' en '*val*' wat woorde in groepe indeel volgens die weerstoestand wat daarmee geassosieer kan word; daar is byvoorbeeld verskillende soort neerslag en verskillende soorte winde.

5.4.3 Ariane (GETA)

Ariane was die produk van die Universiteit van Grenoble en het in 1987 tot stilstand gekom; dit word egter as 'n baie invloedryke stelsel beskou. Dit is 'n oordrag-stelsel, waarvan die oordragmodule uit 'n leksikale en 'n strukturele komponent bestaan.

In Ariane word woordeboekinskrywings *lexical units* genoem. Strukturele en leksikale oordrag is geskei, en lg. word dikwels beïnvloed deur die konteks en deur semantiese kenmerke. Die module wat morfologiese afleiding en generasie doen is ook in staat om nie net vervoegingsaffikse te hanteer nie, maar ook afleidingsuffikse, wat veroorsaak dat die leksikons kleiner kan wees as vir 'n stelsel wat dit nie kan doen nie.

5.4.4 Eurotra

Die Eurotra-projek is waarskynlik die grootste stelsel wat nog op die been gebring is. Dit is bekend vir die

verbintenis wat dit met die EG het om dié organisasie se vertaalprobleme die hoof te probeer bied en die verspreide en groot aantal navorsers wat aan die projek deelneem. Dit is 'n oordrag-stelsel, maar die oordraginterlingua (of *interface structure*) is gebaseer op universele beginsels sodat die verskillende taalpare in die projek eenvormigheid vertoon in verband hiermee.

Twee van die belangrikste leksikale probleme waaraan Eurotra-navorsing aandag gegee het, is tyd en modaliteit. Die tale wat in die stelsel geïnkorporeer is, het verskillende tempusstelsels en verskillende maniere en leksikale items om dit mee aan te dui; daar is 'n 'euroversele' tempusstelsel geskep om hierdie probleem die hoof te bied. Die verskillende tale het ook verskillende maniere om modaliteite uit te druk, en daar moes eweneens 'n universele voorstellingsmeganisme geskep word om die vertaling van modale konstituente goed te kan vertaal.

5.4.5 METAL

METAL is die produk van Siemens in Duitsland; dit is 'n oordrag-gebaseerde stelsel en word beskou as een van die gevorderdste stelsels wat in gebruik is.

Daar is twee enkeltalige woordeboeke wat die basiese morfologiese, sintaktiese en semantiese inligting bevat. Daar is ook 'n hiërargie van woordeskat: die basiese modules is die drie laagste vlakke: funksionele woorde, algemene woordeskat en algemene tegniese woordeskat. Enige aantal gespesialiseerde woordelyste kan hierby gevoeg word en dit kan in die hiërargie georden word volgens die behoefte van die gebruiker. Die basiese woordeboeke bevat onder meer voorkeurwaardes, waarmee tussen moontlike analises gekies kan word, informasie oor kollokasies, en 'n aanduiding van onderwerparea. Inskrywings vir werkwoorde bevat informasie oor kasusverbande (moontlike agente, instrumente, temas ensovoorts) en seleksiebeperkinge (dat die subjek byvoorbeeld 'n menslike agent moet wees), subkategoriseringsinligting (die sintaktiese strukture van die argumente - die Engelse werkwoord *'put'* subkategoriseer byvoorbeeld vir 'n naamwoordfrase en 'n voorsetselfrase) en informasie oor oorganklikheid.

'n Verdere tweetalige woordeboek bestaan waarin brontaalitems met vertaalekwivalente van die teikentaal verbind word, sonder om weer die taalafhanklike informasie te herhaal.

5.4.6 Rosetta

Rosetta kom van die Philips Navorsingslaboratoriums in Nederland, en Hutchins en Somers (1991:279) beskryf dit as een van die innoverendste stelsels. Dit gebruik 'n interlingua as intermediêre voorstellingstruktuur en is gebaseer op Montague grammatika.

Die leksikon-items het inligting oor subkategorisering, of 'n bepaalde woord 'n lidwoord kry; dit bevat vaste uitdrukkings wat as leksikale eenhede beskou word en dit laat ook toe dat vir 'n vaste of 'n idiomatiese uitdrukking 'n sintaktiese struktuur gespesifiseer word. Uitdrukkings waarvan die struktuur, maar nie die items nie, verander na gelang van die algehele sintaktiese struktuur kan dus met een inskrywing hanteer word. Rosetta hanteer ook sekere vertalingsuiteenlopendhede - Hutchins en Somers (1991:289) dui byvoorbeeld aan hoedat die demosionele uiteenlopendheid van Dorr (vgl. hoofstuk 8) in Rosetta gehanteer word met 'n subgrammatika en met inligting oor die argumentspesifikasies van woorde soos '*like*' en '*graag*'.

5.5 Opsomming

Daar is baie soorte inligting wat woorde se betekenisse definieer of wat gebruik kan word om woorde met verskillende betekenisse van mekaar te onderskei. MT moes hierdie tipes inligting in woordeboeke inkorporeer op 'n rekenaarmanipuleerbare manier sodat tussen byvoorbeeld homograwe onderskei kan word, sodat die regte grammatikale kategorie van 'n woord vasgestel kan word, sodat die regte vertaalekwivalent gekies kan word, ensovoorts.

Dit is duidelik dat die leksikons van die MT-stelsels wat bespreek is, baie van die inligting wat deur vertalers en taalleerteoretici soos Stockwell et al. (1965) as belangrik beskou word geïmplementeer het. Komponensiële inligting en nie-ekwivalensie word geënkodeer in onderskeidelik Systran en Rosetta; in METAL word woorde gekategoriseer op grond daarvan dat hulle inhoudelik of funksioneel is, en verder verdeel in leksikons met frekwente woorde en leksikons met woorde wat min gebruik word. Baie van die stelsels hanteer ook kollokasionele inskrywings. Dit blyk ook dat van hierdie stelsels 'n magdom inligting oor woorde bevat wat nie in gewone woordeboeke altyd eksplisiet gemaak word nie. In METAL word byvoorbeeld inligting gegee oor die potensiële kasusverbande van 'n werkwoord, asook die seleksionele beperkinge en subkategorisering daarvan.

Deel III : Die simbiose tussen MT en RGTO

Hoofstuk 6: 'n Vergelyking van vertaling en taalleer

Daar kan belangrike ooreenkomste opgemerk word tussen die proses van vertaling en die proses van tweedetaalleer. Selfs die terminologieë van die twee dissiplines en hul rekenaartegnologieë deel sekere terme, hoewel nie noodwendig met dieselfde betekenis nie.

Om 'n simbiose tussen rekenaargesteunde taalonderrig en masjienvertaling met betrekking tot die leksikon te kan bewys, moet aangetoon word dat die twee wel by mekaar tegnieke en teorieë leen. Die eerste twee dele van hierdie studie dui nie aan dat dit die geval is nie. Daarenteen is dit duidelik dat sowel taalleerteorie as masjienvertaling deur die linguistiek beïnvloed is.

In hierdie laaste deel van die studie word die verbande tussen RGTO en MT bespreek, en daar word gewys hoe die dissiplines tot mekaar se voordeel aangewend kan word. Die derde deel is dus 'n uiteensetting van die potensiële simbiose tussen RGTO en MT met betrekking tot die leksikon. 'n Voorbeeld uit elke veld word gebruik om as illustrasie te dien hiervan.

6.1 Opmerklike ooreenkomste en verskille tussen vertaling en taalleer

- In MT-literatuur is die afkortings SL (*source language* of brontaal) en TL (*target language* of teikentaal) vanselfsprekend: daar word van 'n brontaal na 'n teikentaal vertaal. In tweedetaalleerteorie is verwysings na twee tale ook belangrik; die taal wat geleer word, word die L2 (tweedetaal) of die SL (*second language*) genoem, en die eerstetaal, die L1.

- Die konsepte van oordrag en die interlingua word in albei velde gebruik. In MT verwys dit na 'n soort vertaalproses of -argitektuur; in taalleer is dit egter sentrale terme van verskillende teorieë oor taalleer.

Dit is belangrik om te verstaan dat hoewel die terminologie ooreenkom, dit nie dieselfde konsepte verteenwoordig nie. In taalleer beteken oordrag dat die eerstetaal die gebruik van die tweedetaal beïnvloed. In MT beteken dit dat die oordrag-module afhanklik is van die tale wat ter sprake is, maar nie dat die vertalings minder goed gaan wees omdat die SL toegelaat sal word om 'n invloed uit te oefen nie. Daar kan egter 'n klein mate van taalbeïnvloeding ter sprake wees.

In MT verwys 'n interlingua na 'n abstrakte voorstellingstruktuur om die ontwerp van stelsels te vergemaklik en om seker te maak dat 'n neutrale voorstelling van die betekenis van 'n bepaalde teks so

ver as moontlik verkry word. In taalleer verwys dit na 'n leerderstaal of intertaal wat deur die taalleerder gevorm word op pad na gevorderde tweedetaalgebruiksvermoë.

- Die doelwitte van MT en tweedetaalleer is amper dieselfde. In MT is dit dat 'n uiting in die SL vertaal moet word na 'n uiting in die teikentaal sodat albei dieselfde betekenis oordra. En in tweedetaalleer: 'n bepaalde betekenis wat wel in die L1 geproduseer kan word, moet nou ook met 'n L2 geproduseer kan word.

MT is 'n tegnologie wat nie met psigologiese realiteit rekening hoef te hou nie. Dit het 'n doelwit, om goeie vertalings te verskaf met 'n rekenaarmatige proses, en die bereiking daarvan is wat belangrik is. 'n Mens sou kon sê dat goeie rekenaarvertalings nie moontlik sal wees sonder dat dit 'n teoretiese onderbou met psigologiese realiteit het nie; maar dan het rekenaars tot dusver ook nog nie veel van 'n neurologiese realiteit ooreenstemmend met die mens s'n nie. Dit is 'n verwikkelde saak en 'n mens kan nie bewys dat 'n teorie met selfs psigologiese realiteit tot goeie masjienvertalings sal lei nie.

Daarenteen is taalleer 'n wetenskap wat baie meer gemoeid is met 'n realistiese beskrywing van die menslike taalleerproses. Voorskrywende taalleer hoef nie altyd rekening te hou hiermee nie, maar stel dan dalk ook nie altyd die doeltreffendste leermetodes voor nie. Die bestudering van tweedetaalleer kan lei tot 'n uitbreiding van kennis oor hoe 'n mens taal prosesseer en oor die aard van taal; dit kan die teorieë van kognitiewe psigologie en linguistiek beaam en verbreed, of weerlê.

- In MT is die saamstel van 'n elektroniese woordeboek om verskeie redes 'n baie belangrike deel van die totale prentjie. 'n Woordeboek moet al die inligting bevat wat deur die ander modules van die stelsel vereis word om vertalings te kan lewer. Die vereistes wat gestel word wissel afhangende van die benadering van die stelsel, en afhangende van die kwaliteit vertalings waarna gepoog word. Maar die voorstelling van die woordeboek moet ook rekenaarwetenskaplik voordelig wees: leksikale items moet maklik gesoek en gevind kan word, en kenmerke en vertaalekwivalente wat daarmee verband hou moet maklik beskikbaar gestel kan word aan die stelsel.

In taalleer is die leer van woordeskat nie 'n leermodule wat met soveel gestruktureerdheid as byvoorbeeld die leer van sintaksis of morfologie aangebied word nie. Tog is 'n goeie woordeskat onontbeerlik vir effektiewe kommunikasie.

- 'n MT-stelsel moet dikwels elke moontlike stukkie inligting oorweeg wat oor 'n woord beskikbaar is (ensiklopediese of wêreldkennis uitgesluit) om 'n verantwoorde keuse van betekenis te kan maak. Dit sluit inligting in van die soort wat deur Bell beskryf word (sien hoofstuk 4).

Tweedetaalleerders kry dikwels tweetalige lyste van woorde wat vertaalekwivalente is wat gememoriseer moet word. En hoewel dit 'n metode is wat voldoende kan blyk, is daar gevalle waar twee woorde nie langs mekaar geplaas kan word nie - gevalle van vertaaluiteenlopendhede wat hieronder bespreek word.

- In albei is dit nodig om die leksikons op 'n rekenaarmanipuleerbare manier voor te stel. Hierdie voorstelling behels onder meer dat data op 'n ekonomiese manier gestoor moet word, en dat dit só gestoor word dat dit weer vinnig gesoek en opgeroep kan word. Werk wat in 'n MT-stelsel gedoen is, maar nie in 'n RGTO-program nie, of andersom, sou dus oorgedra word en sodoende 'n besparing aan werk meebring.

Hoofstuk 7: Bruikbare konsepte uit RGTO

Dit is redelik maklik om voorbeelde in MT te gaan soek van teoretiese toepassings wat op grond van linguistiese teorieë en pragmatiese oorwegings ontwikkel is en van toepassing kan wees op woordeskatleerprogrammatuur in RGTO. Dit is egter moeiliker om die omgekeerde te doen.

In die hoofstuk oor bruikbare konsepte uit MT word 'n MT-stelsel byvoorbeeld bespreek wat 'n voorstellingstruktuur vir leksikale materiaal ontwerp het op grond van linguistiese teoretisering oor leksikaal-semantiese kontraste tussen tale. Daar word gedemonstreer hoe hierdie voorstelling gebruik kan word om 'n woordeskatleertegniek vir RGTO-programme te skep.

Die voordeel wat RGTO vir MT inhou, volgens die literatuur wat ondersoek is, is van betrekking op bystand vir die gebruiker van die MT-stelsel met betrekking tot leksikale sake. Vertaling word beskou as 'n ingewikkelde proses wat 'n groot mate van vakmanskap vereis. Dit is 'n oogmerk van MT om 'n stelsel so gebruikersvriendelik as moontlik te maak - beide vir professionele vertalers en vir gebruikers wat nie taalteoretiese kennis het nie.

7.1 Die WE/WT - program van die *Rechenzentrum* en die *Romanisches Seminar* van die Universiteit van Keulen

Die WE (*word editor*) /WT (*word trainer*) RGTO-program van die Universiteit van Keulen dien as illustrasie. Die outeurs van die artikel oor hierdie stelsel, Widdig en Esser (1988: 265 - 266) noem van die besware op wat al teen RGTO geopper is, soos

- dat slegs een oplossing as korrek aanvaar word, en nie alternatiewe oplossings nie,
- dat 'n antwoord wat 'n sinoniem verskaf as verkeerd aangedui word, terwyl 'n verduideliking van die semantiese verwantskap 'n beter reaksie van die program sou wees, en
- dat gedeeltelike oplossings nie as deels korrek aanvaar word nie, maar as verkeerd.

Om hierdie probleme teen te werk stel hulle 'n program voor met 'n buigsame, nie-lineêre en intelligente struktuur, 'n geïnkorporeerde leerproses wat ooreenkom met die moderne psigologiese benadering tot kognitiewe leerstrukture, en 'n leerproses geskoei op moderne didaktiese metodes. Om die ontwerp

teweeg te bring, word die volgende inhoudelike vereistes gestel:

- 'n hoë mate van intelligensie bestaande uit die buigsaamheid van programfunksies en -reaksies sowel as 'n buigsame gebruikerskoppelvlak, vergesel van 'n gesofistikeerde fout-analitiese stelsel,
- leer deur die gebruik van komplekse taalstrukture, en
- die gebruik van oordrag om die geleerde woordeskat in verskillende kontekste aan te bied, sodat dit doeltreffender gebruik en onthou kan word.

Die toepassing van die drie vereistes word hieronder bespreek.

7.1.1 Intelligente vermoë

'Intelligent' moet nie te letterlik opgeneem word nie. Die intelligente aspek van die program is daarop gemik om aan die student 'n gevarieerde leeromgewing te bied - een wat by haar behoeftes pas - en om leerdersantwoorde nie net in terme van reg en verkeerd te beskou nie, maar om foute te verduidelik en gedeeltelik korrekte antwoorde te kan hanteer. Die outeurs verduidelik (1988 : 266),

> 'In developing the learning system of the WT, we anticipated the possible responses to the word asked for (as far as the storage capacity of each key allows), so that the program accepts all, or at least the most frequent solutions to the word asked for. Thus there is little danger that the program will not accept a correct, but rare response.'

Die intelligensie van die program draai ook om 'n ingeboude stelsel van fout-analise. As die leerder se invoer nie met meer as 70 % van die verwagte antwoord verskil nie, lewer die program kommentaar om die soort foute en die posisies daarvan aan te dui.

7.1.2 Leer met komplekse sintaktiese en semantiese strukture

Die program onderskryf wat, volgens die outeurs (1988: 268), 'n algemeen aanvaarde psigologiese teorie oor die leer van vreemdetale is - hoe meer georganiseerd en gestruktureerd leereenhede is, hoe meer effektief is die proses van leer en memoriseer. Daarom is die benadering van die program om materiaal aan te bied in die konteks van komplekse taalgebruik, met ander woorde, met die inkorporering van situasionele, kontekstuele, semantiese en sintaktiese faktore.

In situasionele kontekste word die leerling toegelaat om betekenisse af te lei uit tekskontekste; Widdig en Esser (loc.cit.) sê dat die kontekstuele ontdekking van semantiese kenmerke 'n standaard-tegniek is in vreemdetaalonderrig. Die leerling kan egter ook vra vir die betekenis van 'n woord. Die proses word aangehelp met 'n analitiese benadering waarin die woordeskat van 'n leksikale veld ('n groep leksikale items wat saamgegroepeer word op sterkte van 'n semantiese of kontekstuele verbintenis) voorgestel word in verband met die sintagmatiese en paradigmatiese strukture daarvan.

Die outeurs gaan van die standpunt uit (loc.cit.) dat woorde nie in isolasie beskou kan word nie, en dat daar verbande is met ander leksikale items in dieselfde taal. Hulle meen dat die semantiese verbande van teenstelling en gedeelde betekenis gebruik kan word om 'n taalleerder te help om die regte woord vir 'n bepaalde konteks te vind en ook om so 'n woord te memoriseer. Vervolgens het hulle die woordeskat van elke leksikale veld in komplekse semantiese veldstrukture georganiseer deur by elke woord waarvoor gevra word, sinonieme en antonieme daarvan by te voeg, indien daar enige is. Dié semantiese verbande word verduidelik en die gebruik daarvan word gedemonstreer. Die leerder ontvang dus ook informasie om 'n woord te relativeer met woorde van dieselfde leksikale veld, omdat sy ingelig word oor die grootte en inhoud van die veld.

Hulle skryf verder (Widdig en Esser, 1988: 268-269):

> 'In the belief that the learner of a foreign language is only able to re-use a word he has learned in a communicative situation if he knows its syntactic properties, we were anxious to formulate the question stimulus in a syntagmatic way so that the learner has to answer using a complete morphosyntactic structure. This procedure allows the user to learn not only the word asked for, but also its morphological and syntactic characteristics, a necessary precondition for the active use of a foreign language. By proceeding in this way, we avoided the use of monotonous - and ineffective - pattern drills.'

7.1.3 Versterking van leer met oordrag

Die woorde wat geleer is, word in nuwe materiaal aangebied sodat dit geïnternaliseer kan word. In WE/WT word tekste met gapings en vertalings vir hierdie doel gebruik. Dit verteenwoordig die implementering van die oordrag-konsep baie soos dit bekend staan in tweedetaalleerteorie. Hier beteken dit dat 'n vreemdetaalwoord in 'n sekere konteks geleer word, en dat die gebruik daarvan nou ook in ander kontekste gedemonstreer word.

7.2 Die voordeel vir MT

Die WE/WT stelsel het veral op twee maniere 'n bydrae te maak tot die voorstelling en gebruik van die leksikon in MT-stelsels. Die eerste is die skepping van 'n tesourusagtige rekenaarleesbare leksikon, en die tweede die aanbied van woorde binne sinskonteks en met inagneming van die sintaktiese implikasies daarvan.

7.2.1 Konkretisering van verbande tussen vertaalekwivalente

Die taak om woorde in semantiese velde te groepeer en die verbande van sinonimie en antonimie eksplisiet te maak, is ook een van die take van die ontwerper van 'n tesourus. Populêre hedendaagse woordverwerkers bevat tesourusse om opsoeke te vergemaklik terwyl die bekende Roget se tesourus beskikbaar is op CD-ROM en die Internet. Hoewel Widdig en Esser se program in 1988 waarskynlik met groot probleme te kampe gehad het ten opsigte van geheuespasie en rekenaarspoed, is dit vandag nie meer 'n probleem van dieselfde omvang nie, en is dit al makliker vir enigiemand met leksikografiese kundigheid en vaardigheid in rekenaarprogrammering om 'n tesourus op te stel.

Die WE/WT program bevat egter selfs in die lig van die moontlikheid van 'goedkoop'-tesourusse 'n element wat voordelig is vir implementering in MT. Hierdie voordelige element lê daarin dat die verbande tussen leksikale items in dieselfde semantiese veld eksplisiet gemaak is. Die programmatuur moet dié eksplisiete verbande in ag neem om die korrekte antwoorde te kan verskaf of om antwoorde te kan toets. Indien die leerder dus '*house*' in die gaping invul by die sin '*We went ...*' in plaas van '*home*', sal die inligting in die leksikon 'n sinonimiese verband aandui tussen die twee woorde, sodat dit ten minste nie die leerder se antwoord as totaal verkeerd sal beskou nie. Die leerder self kan ook eksplisiet in kennis gestel word van die verband wat tussen twee woorde ter sprake is.

Terwyl sinonimie en antonimie nie die enigste twee moontlike verbande tussen woorde is nie, is dit terselfdertyd belangrik uit die oogpunt van WE/WT dat selfs net hierdie inligting eksplisiet voorgestel word. Dit is 'n stap verder as om die inligting net saam te groepeer.

Inligting soos hierdie kan in MT-stelsels gebruik word om die vertaalresultate te verbeter en om die gebruiker van die stelsel te help om geproduseerde vertalings te verbeter, of om die bronteks te korrigeer voordat dit vertaal word.

7.2.2 Inverbandbrenging van woordbetekenisse met sintaktiese kontekste

Terwyl die ontwerpers van 'n MT-stelsel seker moet maak dat hulle voldoende en noukeurige inligting oor woorde in die leksikon van die stelsel inskryf, en hulle van taalteoretiese inligting gebruik maak om dit

doeltreffend te kan doen, is die gebruiker van die stelsel nie noodwendig bekend met die inligting nie. Die handboeke van MT-stelsels bevat gewoonlik 'n gedeelte oor die grammatikas van die tale wat ter sprake is. Dit verduidelik grammatikaterminologie, soos die definisies van naamwoorde of werkwoorde, morfologie en eenvoudige sintaksis, soos dat 'n werkwoord onoorganklik kan wees of deur 'n sekere voorsetsel gevolg word.

Die tegnieke wat in WE/WT gebruik word om aan leerlinge die toepaslike betekenis van 'n woord te leer met behulp van die semantiese en sintaktiese kenmerke daarvan, kan gebruik word om die gebruiker van 'n MT-stelsel te begelei wanneer woorde in die stelsel se leksikons ingeskryf moet word. Die vermoë om nuwe woorde te kan insleutel is funksionaliteit wat alle MT-stelsels aan gebruikers moet voorsien. Woordeboeke is selde so volledig, en sal miskien nooit wees nie, dat gebruikers nie nodig sal hê om woorde in te sleutel nie.

Die WE/WT-program se reaksie op die '*home/house*' verwarring kan op die volgende manier benader word deur van die sintaktiese kontekskenmerke van die woorde te beskryf. Van '*house*' in die enkelvoud kan gesê word dat dit gedetermineer moet word met byvoorbeeld 'n lidwoord of 'n besitlike voornaamwoord, terwyl '*home*' van dié twee net met 'n besitlike voornaamwoord gedefinieer kan word binne die konteks van '*home/house*'; dit mag ook sonder 'n determineerder gebruik word. In die gapingsin wat hierbo gegee is, is dit verder onmoontlik om '*house*' in te vul aangesien dít en die determineerder daarvan voorafgegaan moet word deur 'n voorsetsel, terwyl dit nie die geval is met '*home*' nie. 'n Mens kan dus nie sê '*We went the house*' of '*We went to house nie*'

Behalwe daarvoor dat die enkodering van hierdie inligting op 'n konsistente en rekenaarmanipuleerbare manier van onskatbare waarde kan wees vir die vertaalproses self, kan dit ook gebruik word om die gebruiker van die MT-stelsel te help om die woordeboeke daarvan self uit te brei. Deur gapingsinne te verskaf soos die bogenoemde kan die gebruiker haar kennis van die taal gebruik om die MT-stelsel 'n woord reg te laat gebruik. In hierdie geval is die MT-stelsel dus die leerder en die gebruiker die onderwyser.

Die proses kan soos volg verloop: die gebruiker voer die woord '*house*' in as die woord wat sy wil hê die MT-stelsel moet kan gebruik. Die stelsel sal dan gebruik wat as 'n ingeboude taalvermoë beskryf kan word, en aan die gebruiker vra of die sin '*We went house*' 'n grammatikale sin is. Die gebruiker kan die sin korrigeer na '*We went to the house*', wat deur die stelsel se ontleder afgetas kan word, sodat dit sal herken dat 'house' gedetermineer is en dat die naamwoordfrase waarvan dit die hoofwoord is, voorafgegaan word deur die voorsetsel '*to*'.

Om die gebruik van die woord verder te verifieer, kan dit die naamwoordfrase-moontlikhede van Engels gebruik om verdere konfigurasies met behulp van die gebruiker te toets; dit kan ook ander voorsetsels gebruik en ander determineerders. Dit sou die volgende sinne aan die gebruiker kon wys vir verifiëring:

'*We went inside the house*', '*We went around the house*', '*We went to our house*', '*We went to a house*', '*We went to a colourful house*'. Al hierdie sinne sou deur 'n gebruiker as korrek aangedui word.

Terselfdertyd sou dit 'n sin kon genereer soos: '**We went up the house*' wat deur die voorsetsel ongrammatikaal gemaak word. Indien die subkategoriseringsinligting van '*went*' volledig is, en daar ook genoegsame beperkinge is op hoe telwoorde saam met naamwoorde gebruik word, sal sekere ander sinne weer nie gegenereer word nie: '**We went on the house*' en '**We went to two house*'.

Indien die gebruiker '*huis*' sou invoer as 'n nuwe woord, sou sy ook die twee vertaalekwivalente daarvan moet verskaf. Die MT-stelsel sal dan met die proses wat hier beskryf is verskillende sintaktiese kenmerke aan die twee verskillende woordpare toeken, sodat dit korrek deur die stelsel afgetas en gebruik sal kan word.

'n Proses soos hierdie behoort baie meer kompleks te wees as wat hier voorgegee word omdat daar soveel grammatiese inligting is wat oorweeg moet word. Indien hierdie inligting egter wel beskikbaar is, kan dit gebruik word om baie nuttige inligting oor woorde by die gebruiker te kry sonder dat die gebruiker self ingewikkelde terminologie hoef te gebruik, en sonder dat die gebruiker noodwendig weet watter inligting die MT-stelsel onttrek uit of 'n gegenereerde sin korrek is of nie. Uit die reaksie van 'n gebruiker dat '*The house eats*' nie sin maak nie, sou die stelsel kon aflei dat '*house*' nie lewend en bewegend is nie.

7.2.3 Fout-voorspelling en intelligente hulpfasiliteite

In die WE/WT-program is gebruik gemaak van werklike data oor die moontlike foute wat leerders maak; hierdie inligting word gebruik om aan die RGTO-program 'n illusie van intelligensie te gee en om die verantwoordbaarheid daarvan te verhoog. Hierdie tegniek kan ook aan die woordeboekfasiliteit van 'n MT-stelsel 'n mate van intelligensie verskaf. Indien dit foute kan antisipeer wat deur 'n tipiese gebruiker gemaak kan word met die inskrywing van 'n woord, sal dit kan voorkom dat verkeerde inskrywings in die woordeboeke van die stelsel beland, en dit sal ook die gebruiker op 'n betekenisvolle manier kan help om daardie inskrywing te korrigeer.

Hoofstuk 8: Bruikbare konsepte uit MT

'n Praktiese voorbeeld in masjienvertaling waar die benadering tot leksikale inligting moontlik van nut kan wees vir die leer van 'n woordeskat van 'n tweedetaal word in hierdie hoofstuk bespreek. Daar word gekyk na Bonnie Dorr se navorsing oor leksikaal-semantiese vertalingsuiteenlopendhede en die voorstelling van die leksikale items wat dit veroorsaak. Die navorsing is toegepas binne die UNITRAN-vertaalstelsel.

8.1 Dorr se formalisering van masjienvertalingsuiteenlopendhede

In Dorr (1993:233) se boek *Machine translation: a view from the lexicon* word onder meer gepoog om twee soorte informasie te formaliseer:

- die linguisties-gebaseerde klasse waarop leksikaal-semantiese uiteenlopendhede gebaseer word

- die tegnieke waarmee leksikaal-semantiese uiteenlopendhede opgelos word.

Volgens Dorr (loc.cit.) verskaf dit die basis vir die bewys dat die leksikaal-semantiese uiteenlopendheidklassifikasie, soos voorgestel in die raamwerk, alle moontlike vertaal-nie-ekwivalensies dek wat met leksikaal-semantiese eienskappe te doen het. Met leksikaal-semantiese eienskappe word bedoel die dele van leksikoninskrywings wat nie te doen het met sintaktiese informasie, idiomatiese gebruik, aspektuele kennis, diskoerskennis, domeinkennis of wêreldkennis nie.

Die konsep van leksikaal-semantiese vertalingsuiteenlopendhede staan sentraal in Dorr se boek. Met die term word bedoel dat 'n woord in die brontaal nie eenvoudig met 'n enkele ooreenstemmende woord in die teikentaal vertaal kan word nie, maar dat die regte vertaalekwivalent kan lei tot 'n struktuurverandering en/of 'n veranderde uitdrukkingswyse.

Die klasse wat hierbo vernoem word, is LCS (*lexical conceptual structure*) klasse wat 'n indeling van woorde is op grond van basiese tipes en semantiese primitiewe. 'n LCS is 'n raamwerk waarbinne 'n abstrakte voorstelling van die inhoud en struktuur van 'n taalteks gedoen kan word; dit is komposisioneel van aard sodat 'n paar LCS'e saamgevoeg kan word om 'n groter LCS te vorm. LCS'e konstitueer die interlinguavoorstellings wat in Dorr se model gebruik word, en is ook die raamwerke vir leksikale inskrywings. Die naam LCS dui op die belangrike rol wat aan leksikale items toegeken word; konseptuele struktuur word beskryf as 'die taal van verstandelike voorstelling' (Dorr,1993:32) en as hiperstel van

semantiese voorstellings. 'n Konseptuele struktuur is nie net 'n weergawe van die semantiese inhoud van 'n taaluiting nie, maar dit neem ook die sintaktiese struktuur van die uiting in ag.

Die LCS-klasse is die vlak waarop leksikaal-semantiese verskille tussen twee tale aangedui word. Die UNITRAN-model sluit bepaalde meganismes in om te verseker dat vertalingsuiteenlopendhede korrek gehanteer word wanneer dit nodig is. Dit sluit onder meer in skakelroetines tussen natuurlike taalteks en interlinguavoorstellings en semantiese parametrisering en die prosessering daarvan. Dit verteenwoordig 'n proses wat nie binne die omvang van die studie bespreek word nie.

Sewe soorte leksikaal-semantiese uiteenlopendhede word onderskei en Dorr neem die standpunt in dat dit die volledige lys van moontlike uiteenlopendhede is; haar argument hiervoor is uitgewerk op grond van die linguistiese teorie van *Government and Binding*. 'n Volledige uiteensetting van hierdie argument kan gevind word in Dorr (1993:265 - 281).

Dorr onderskei ook 'n stel sintaktiese vertalingsuiteenlopendhede. Dit is sintaktiese asimmetrieë tussen twee tale wat die gevolg is van die sintaktiese idiosinkrasieë van daardie tale. Anders as die leksikaal-semantiese soort oefen leksikale items nie 'n invloed daarop uit nie.

Sy (Dorr,1993:19) gee onder meer konstituent-volgorde aan as 'n voorbeeld van strukturele vertalingsuiteenlopendheid:

Afrikaans : Nadat ek hom besoek het ...,
Engels : After I have visited him ...,
('Nadat ek het besoek hom')

Wat dus bedoel word daarmee dat leksikale items nie 'n invloed het op die uiteenlopendheid nie, is dat '*hom*' net sowel '*my vriend*' of '*ons kliënt*' kon wees - dit sou steeds in die Engelse sin ná die werkwoord geplaas moes word.

Oor uiteenlopenheidsklassifisering sê Dorr (1993:21) dat hoewel dit nie 'n nuwe konsep is nie, dit eers onlangs die onderwerp geword het van sistematiese navorsing. 'n Omvattende oorsig van uiteenlopendheidsvoorbeelde word aangebied deur Barnet et al. (1991a,b), Dorr (1990a,b), Kameyama et al. (1991), Kinoshita et al. (1992), Lindop en Tsujii (1991), Tsujii en Fujita (1991), Beaven (1992a,b) en Whitelock (1992) en besprekings in byvoorbeeld Melby (1986) en Nirenburg en Nirenburg (1988). Die term wat gewoonlik gebruik word vir uiteenlopendheid is 'komplekse leksikale oordrag', maar Dorr (1993:21) beskou dit as 'n ongewenste naam aangesien dit 'n probleem is wat inherent is aan MT en nie net aan die oordrag-benadering nie. Hoewel die vertaling van sulke uiteenlopendhede vir menslike vertalers meesal 'n onbewustelike proses moet wees, is die probleem uiteindelik een wat inherent is aan vertaling.

8.2 Leksikaal-semantiese vertalingsuiteenlopendhede

Die tipes uiteenlopendhede word genoem, verduidelik en geïllustreer met voorbeelde. Van die tegniese aspekte wat onderliggend is aan die konsepte, word uiteengesit sodat die uiteenlopendhede makliker verstaan kan word.

8.2.1 Leksikale inskrywings

'n Leksikale inskrywing word in UNITRAN 'n RLCS (*root word lexical conceptual structure*) genoem en dit word op twee vlakke gedoen: 'n taalonafhanklike vlak en 'n taalafhanklike vlak wat die sintaktiese realisering van die woord en die argumente daarvan spesifiseer. (Die raamwerk vir 'n RLCS is dieselfde as vir die interlingua-voorstellings, wat CLCS'e genoem word (*composite lexical conceptual structure*).) Dié struktuur lyk só:

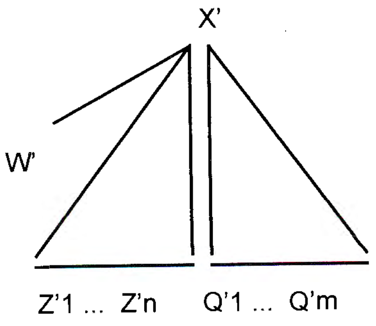
$[_{T(X')}\ X'$

$\quad([_{T(W')}\ W'\ ],$

$\quad[_{T(Z'1)}\ Z'_1\ ],\ \ldots\ [_{T(Z'n)}\ Z'_n\ ],$

$\quad[_{T(Q'1)}\ Q'_1\ ],\ \ldots[_{T(Q'm)}\ Q'_m\ ])]$

'n Boomvoorstelling hiervan lyk só:



Die inskrywing van 'n werkwoord kom neer op die voorstelling van 'n enkelvoudige sin waarin die betrokke werkwoord die hoofmoment is. In so 'n sin sou X' hierbo die logiese hoof wees, W' die logiese subjek, Z'1 ... Z'n die logiese argumente en Q'1 ... Q'm die logiese modifiseerders en T(Φ) die LCS tipe wat korrespondeer met die primitief Φ.

In die sin '*Die brugwerker vang die non met sy vry arm.*' is '*vang*' = X, '*die brugwerker*' = W, '*die non*' = Z_1 en '*met sy vry arm*' = Q_1. Die LCS tipe sal moontlik die UNITRAN primitief EVENT-$STAY_{INSTR}$ wees om aan te dui dat 'n aksiewerkwoord ter sprake is en dat dit veroorsaak dat die argument (in hierdie geval die direkte voorwerp) na 'n toestand van beperking gebring word met behulp van een of ander instrument (in hierdie geval 'n arm).

Die bg. struktuur sal soos volg ingevul kan word vir 'vang':

$[_{Event}$ $STAY_{INSTR}$
    $([_{Thing}$ * W],
    $[_{Thing}$ * Z])]

(Die modifiseerder-gedeelte kan weggelaat word omdat dit 'n opsionele konstituent is.)

Die '*' hierbo is 'n merker wat in UNITRAN gebruik word om ruimte te laat vir taalspesifieke veranderlikes en om aan te dui waar twee tale van mekaar sou verskil ten opsigte van die invul van die veranderlike. Dorr gaan byvoorbeeld in op die volgende voorbeeld:

Engels: I stabbed John.
Spaans: Yo le di puñaladas a Juan
    'I gave knife-wounds to John'

Die leksikale inskrywing vir die Engelse '*stab*' moet voorsiening maak daarvoor dat die Spaanse vertaling nie net 'n woord-vir-woord weergawe sal wees daarvan nie. Die inskrywing lyk só:

$[_{Event}$ CAUSE
    $([_{Thing}$ * W],
    $[_{Event}$ GO_{Poss}
        $([_{Thing}$ KNIFE-WOUND],
        $[_{Path}$ $TOWARD_{Poss}$
            $([_{Position}$ AT_{Poss} $([_{Thing}$ KNIFE-WOUND], $[_{Thing}$ * Z])])])],
    [$WITH_{Instr}$ * $([_{Event}$ *HEAD*], $[_{Thing}$ * U SHARP-OBJECT])])] (Dorr,1993:132)

Dit kom rofweg ooreen met '*Thing* W veroorsaak *Thing* Z om 'n meswond te besit d.m.v. 'n skerp voorwerp U'.

Hoewel die moontlike uiteenlopendhede dus beperk is, sal dit steeds 'n uitgebreide proses wees om tussen enige twee tale uiteenlopendhede soos hierdie in die RLCS'e te laat realiseer. Dit is egter nie 'n probleem waarby 'n mens in elk geval sou kan verbykom, sonder om die verskil volledig te spesifiseer nie.

Die belangrike punt is dat Dorr 'n beperkte raamwerk voorstel waarbinne 'n mens enige uiteenlopendheid sou kon klassifiseer.

8.2.2 Vertalingsuiteenlopendhede

1. *Tematiese uiteenlopendheid*

Dit is die herposisionering van twee satelliete van die hoof (byvoorbeeld die subjek en die objek); met ander woorde W → Z en Z → W. Een van die voorbeelde wat Dorr (deur Lindop en Tsujii) verskaf (1993:236) is:

Engels : I like the car
Duits : Mir gefällt der Wagen

Die subjek van die Engelse sin word die komplement van die werkwoord in die Duitse sin, terwyl, die objek van die Engelse sin die objek word van die Duitse sin.

2. *Promoverende uiteenlopendheid*

Dit word gekarakteriseer deur die promovering van 'n logiese modifiseerder. In so 'n situasie word die logiese modifiseerder geassosieer met die sintaktiese hoofposisie, en die logiese hoof word dan geassosieer met 'n interne argument posisie; met ander woorde X → Z en Q → X.

'n Voorbeeld van Dorr (1993:236) is:

Nederlands: Hij is toevallig ziek
Engels: He happens to be ill.

Die hoof '*ziek*' word die argument '*to be ill*', en die modifiseerder '*toevallig*' word die hoof '*happen*'.

3. *Demoverende uiteenlopendheid*

Dit word gekarakteriseer deur die demovering van 'n logiese hoof. In so 'n situasie word die logiese hoof geassosieer met die sintaktiese modifiseerderposisie en die logiese argument word dan geassosieer met die sintaktiese hoofposisie; met ander woorde X → Q en Z → X.

‘n Dorr-voorbeeld (1993:236) is:

Duits: Er liest gern
Engels: He likes reading

Die Duitse werkwoord *'liest'* word dus in die Engelse sin die modifiseerder van *'like'*. Terselfdertyd word die Duitse *'gern'* die hoofwerkwoord *'like'* in die Engelse sin.

4. Strukturele uiteenlopendheid

Strukturele uiteenlopendheid verskil van die laaste drie soorte daarin dat dit nie die posisies van die argumente verander nie, maar die verhoudinge tussen daardie posisies. Voorbeelde van Dorr (1993:167) is:

Engels: John entered the house
Spaans: Juan entró en la casa
('John entered (into) the house')

Duits: Der Student beantwortet die Frage
Frans: L'étudiant répond à la question
('The student answered the question')

5. Samesmeltende uiteenlopendheid

Dorr (1993:258) beskryf samesmeltende uiteenlopendheid as die onderdrukking van ‘n CLCS konstituent; met ander woorde dit verteenwoordig ‘n geval waar ‘n sekere argument nie net so in die oppervlakvoorstelling (die uiteindelike gegenereerde teks) hoef te verskyn nie. Dit is egter nie noodwendig net die weglating van ‘n konstituent nie, maar eerder die inkorporering van ander noodsaaklike argument(e) van ‘n gegewe handeling. Hieronder gebruik Engels byvoorbeeld die enkele woord *'stab'* vir die twee Spaanse woorde *'dar'* (gee) en *'puñaladas'* (meswonde); dit is omdat die *'puñaladas'* gedeelte in die Engelse werkwoord geïnkorporeer is.

Engels: I stabbed John.
Spaans: Yo le di puñaladas a Juan
'I gave knife-wounds to John'

'n Ander voorbeeld (Dorr:1993:260) is:

Engels: John swam across the river
Frans: John a traversé la rivière à la nage
('John crossed the river by swimming')

6. Kategoriale uiteenlopendheid

Kategoriale uiteenlopendheid word gekarakteriseer deur die assosiasie van 'n logiese konstituent met 'n sintaktiese kategorie wat verskil van die verstek (*default*) kategoriale waarde. Byvoorbeeld:

Engels: I am hungry
Duits: Ich habe Hunger
('I have hunger') (Dorr:1993:174)

Terwyl '*am*' en '*habe*' nie leksikale ekwivalente is nie, verskil die predikate van die woorde ook ten opsigte van sintaktiese kategorie; '*hungry*' is 'n adverbiale frase terwyl '*Hunger*' 'n naamwoordelike frase is.

7. Leksikale uiteenlopendheid

Leksikale uiteenlopendheid word opgelos tydens die proses van leksikale seleksie. Dit duik net op in die teenwoordigheid van ander uiteenlopendhede.

Engels: John broke into the room
Spaans: Juan forzó la entrada al cuarto
('John forced entry to the room') (Dorr,1993:187)

Dié tipe is leksikaal omdat die keuse van 'n teikentaalwoord nie 'n letterlike vertaling is van die brontaalwoord nie.

8.2.3 Verdere voorbeelde van leksikaal-semantiese vertalingsuiteenlopendhede

In die volgende tabel word voorbeelde gegee van vertalingsuiteenlopendhede tussen Afrikaans en Engels.

oplet -> pay attention	samesmeltend en leksikaal
verlief -> in love	struktureel of kategoriaal
enable -> in staat stel	samesmeltend
face -> in die gesig staar	samesmeltend
aansoek doen -> apply	samesmeltend
afstand doen -> renounce	samesmeltend
beswaar maak -> protest	samesmeltend
inkopies doen -> shop	samesmeltend
agtermekaar -> in order	samesmeltend
beskikbaar -> on hand	struktureel of kategoriaal
pluis -> in order	struktureel of kategoriaal
toevallig -> happens (to be)	promoverend
graag -> like (to)	demoverend

8.3 Die voordeel vir taalleer

Al sewe leksikaal-semantiese uiteenlopendhede (LSU's) wat deur Dorr geklassifiseer word het in 'n mindere of meerdere mate voorspelbare implikasies vir sintaktiese konteks. Wat daaruit afgelei kan word is dat die betekenis van 'n woord, en dus die vertaling wat dit moet kry, afhang van die sintaktiese omgewing waarin dit voorkom. Dit verteenwoordig die eerste rede waarom die teorie van Dorr toepasbaar kan wees op taalonderrig - dat 'n woord se betekenis nie soveel beklemtoon sal word dat die sintaktiese beperkinge daarvan vergeet word nie.

'n Tweede en derde rede vloei hieruit voort. Omdat die verskillende leksikaal-semantiese uiteenlopendhede onderskei word van strukturele uiteenlopendhede, kan verhoed word dat 'n leksikaal-semantiese uiteenlopendheid aan 'n leerder onderrig word, terwyl die leerder nog nie genoegsame kennis het van die tweedetaal se sintaksis om die nuwe woordeskat te kan gebruik nie. Min of meer dieselfde geld daarvoor dat die verskillende leksikaal-semantiese uiteenlopendhede geassosieer word met verskillende struktuur- of sintaktiese konteksveranderinge.

8.3.1 Klassifisering van woordeskat volgens invloed op sintaktiese omgewing

In hoofstuk 2 word Gleitman (1993) se *syntactic bootstrapping* teorie uiteengesit waarvolgens beweer word dat die sintaktiese beperkinge van 'n woord, of die sintaktiese omgewing wat 'n woord afdwing, bydra tot die leer van die woordeskat van 'n eerstetaal. In die voorafgaande hoofstuk word ook aangedui dat Widdig en Esser (1988) die standpunt inneem dat die leerder van 'n vreemdetaal ook die sintaktiese karakteristieke van 'n woord moet leer as sy dit effektief wil gebruik.

Dorr se LSU's maak dit moontlik om tussen woorde te onderskei wat ingrypende sintaktiese verwringing van die teikentaalsin veroorsaak en woorde wat dit nie doen nie. Die moeilikheidsgraadhiërargie van Stockwell, Bowen en Martin (1965) wat in hoofstuk 2 aangehaal is, is gemik op die laaste groep woorde en is, in teenstelling met Dorr se klassifisering van uiteenlopendhede, 'n redelik eenvoudige klassifikasie. Baker (1992) se klassifisering van nie-ekwivalensie vertaalprobleme lê klem op die afwesigheid van een-tot-een korrelasies tussen woorde, maar gaan nie in op die rekonstruksie van die TL-teks op dieselfde stapsgewyse manier as Dorr nie.

Die onderskeid tussen Dorr se klassifisering aan die een kant, en Baker (1992) en Stockwell et al. (1965) s'n aan die ander kant kan met Afrikaans-Engelse voorbeelde geïllustreer word. Stockwell et al. beskou die moeilikste kategorie van nie-ooreenstemming as een waar die eerstetaal een woord het met twee moontlike tweedetaalekwivalente. 'n Voorbeeld hiervan is '*huis*' wat twee spesifieke vertalings kan hê - '*house*' en '*home*' - onder die hiperniem '*dwelling*'.

Die verskil tussen die twee Engelse woorde is semanties, maar hulle word ook verskillend gebruik. Die uiteenlopendheid van

Engels: The man goes home
Afrikaans: Die man gaan huis toe

kan gekategoriseer word as 'n strukturele uiteenlopendheid, terwyl daar geen strukturele uiteenlopendheid is by

Engels: They want a small house with a view
Afrikaans: Hulle soek 'n klein huis met 'n uitsig

nie. In lg. geval is '*house*' en '*huis*' sinonieme, terwyl '*home*' en '*huis*' in eg. geval sinonieme is.

'n Ander voorbeeld is '*dwing*' wat met '*force*' of '*compel*' vertaal kan word. Nie een van die twee het strukturele uiteenlopendheid tot gevolg nie; substitusie van die vertaalekwivalent is genoeg om die

korrekte vertaling te produseer. Die keuse wat tussen die twee moontlikhede gemaak moet word, is 'n keuse wat nie opgelos kan word deur leksikaal-semantiese uiteenlopendhede te oorweeg nie, omdat daar ook nie 'n uiteenlopendheid van hierdie aard is nie. Die keuse sal eers gemaak kan word na verdere semantiese en kontekstuele analise.

Die belangrikheid van die verskil tussen Dorr en Baker en Stockwell et al. se klassifisering van leksikale uiteenlopendheid, lê in die eerste plek daarin dat 'n onderskeid tussen sinonieme in verskillende tale, of vertaalekwivalente, gemaak kan word wat enersyds sintaktiese veranderinge tot gevolg het, en andersyds vervang kan word sonder dat die sintaktiese konteks verander. Dit sal nie noodwendig so eenvoudig wees as om te sê dat die woordpare wat nie strukturele verandering tot gevolg het, eerste geleer moet word en daarna die ander soorte woordpare nie. Die woord '*home*' is byvoorbeeld 'n baie eenvoudige Engelse woord, en 'n mens kan nie 'n Afrikaanse leerder van die woord af weghou totdat sy strukturele uiteenlopendheid kan hanteer nie. 'n Meer haalbare voorstel is om die klassifiserings te doen, en om dan die woorde in daardie groepe aan te bied.

8.3.2 Aanbieding van woordeskatgroepe met inagneming van die vlak van die taalleerder

Volgens die ZISA Multidimensionele Model wat in hoofstuk 2 bespreek is, word werkwoordskeiding eers op vlak X+2 aangeleer. Dit is moontlik dat hierdie voorstel van sintaktiese struktuurverwerwingsvolgorde 'n invloed kan hê daarop of 'n woord op 'n gegewe stadium aangeleer kan word of nie. Veronderstel 'n tweedetaalleerder moet die vertaalekwivalent '*oplet-pay attention*' leer, terwyl sy op vlak X is. Aangesien geteoretiseer word dat 'n leerder op vlak X+*n* slegs kan aanleer wat op vlak X+(*n*+1) is, sal hierdie taalleerder net 'n sintaktiese struktuur kan begin leer wat op vlak X+1 is.

'*Oplet-pay attention*' word dus nou aan die taalleerder onderrig. Die sin

Engels:	The girl did not pay any attention
Afrikaans:	Die meisie het glad nie opgelet nie

is waarskynlik 'n struktuur wat net op vlak X+2 gebruik sal kan word, waar werkwoordskeiding aangeleer word. Alhoewel die skeiding van 'n hulpwerkwoord en 'n werkwoord, soos in '*The girl must always pay attention*' met 'n bywoord nie ooreenkom met die skeiding van 'n werkwoord en die objek daarvan met 'n modifiserende adjektief nie, is dit 'n moontlikheid vir ondersoek. By

Afrikaans:	Daar moet opgelet word
Engels:	Attention must be paid

word die naamwoord egter wel geskei van die werkwoordkollokaat met 'n hulpwerkwoord. Dit is egter ook so dat die Engelse sin eerder geskryf kan word as '*You must pay attention*'.

Dieselfde argument geld vir die verskillende sintaktiese transformasies wat deur die verskillende leksikaal-semantiese uiteenlopendhede self veroorsaak word. Die woord '*toevallig*' word hierbo aangedui as 'n woord wat promoverende uiteenlopendheid veroorsaak; in die sin 'Hy is toevallig siek' word dit as 'n bywoord gebruik. Net 'n leerder wat alreeds op vlak X is sal kan leer om die bywoord voorop te plaas, vir die sin '*Toevallig is hy siek*'.

'n Bywoord is egter een van die woorde waarsonder niemand 'n taal effektief sal kan gebruik nie. Beide '*toevallig*' en '*pay attention*' kan op sekere vlakke van ontwikkeling gebruik word en op ander nie. Hierdie situasie kan gebruik word om te argumenteer dat woorde nie noodwendig uit 'n woordeskatleersessie uitgesluit kan word omdat hulle nie in sekere strukture gebruik kan word nie. Maar 'n standpunt wat wel ingeneem kan word is dat sekere woorde makliker sal wees om te leer, of onder sekere omstandighede makliker sal wees om te gebruik as ander. Dat bywoord-vooropplasing alreeds op vlak X+1 geleer word, maar werkwoordskeiding eers op vlak X+2, is moontlik 'n refleksie van die algemene gebruik van bywoorde, en die meer komplekse gebruik van konsepte soos '*pay attention*'.

Die oorweging van leksikaal-semantiese uiteenlopendhede vir woordeskatleerprogramme kan eerder gesien word as 'n metode vir die programontwerper om woorde aan 'n tweedetaalleerder aan te bied wat vir haar moontlik sal wees om te gebruik, as 'n metode om aan 'n tweedetaalleerder te verduidelik wat die tegniese linguistiese detail is van dit wat geleer moet word.

Hoofstuk 9: Slot en oorweging van hipotese

Die grootste deel van hierdie studie word daaraan gewy om 'n agtergrond te verskaf waarteen die hipotese ondersoek kan word. Die ander moontlikheid was om voorstelle oor simbiotiese raakpunte te maak, en om dan grotendeels te konsentreer op die empiriese toetsing en verfyning van hierdie stellings.

Die nadeel verbonde aan die eerste opsie is dat die voorstelle wat uiteindelik gemaak moet word om die hipotese te ondersteun of nie te ondersteun nie, opsigself hipotetiese voorstelle sal wees. Die nadeel verbonde aan die tweede opsie is dat die hipoteses oor simbiotiese raakpunte gemaak sou moes word sonder voldoende ondersoek oor die navorsingsonderwerp. Sodoende sou voorstelle moontlik nie aktueel kon wees nie, en sou dit ook nie in ooreenstemming wees met die teoretiese agtergrond van die onderwerp nie.

Die beste oplossing is waarskynlik om albei hierdie keuses in 'n enkele navorsingstuk te kombineer, maar die aard van hierdie studie skakel so 'n moontlikheid uit. 'n Verdere beperkende faktor is dat die navorsing grootliks interdissiplinêr is, sodat baie navorsingsliteratuur gedek moet word. Daar is dus besluit om 'n teoretiese studie te doen, wat as basis kan dien vir verdere navorsing oor die betrokke onderwerp.

9.1 Opsomming

Die navorsing in hierdie studie maak dit duidelik dat woordeskatleer as ondersoekveld binne taalleer nog nie baie aandag geniet het nie. Dit is verder ooglopend dat die veld van tweedetaalleer geskakeerd is - 'n verskeidenheid van uiteenlopende teorieë is gevestig. Die veld het ook 'n groot mate van volwassenheid bereik, wat weerspieël word in die feit dat dit in sekere opsigte al die volle sirkel voltooi het. 'n Voorbeeld hiervan is die aanvanklike gewildheid van die oordrag-begrip, die latere beklemtoning van ander konsepte en toe die inkorporering van die begrip in 'n multidimensionele teorie.

Rekenaargesteunde taalonderrig ten opsigte van woordeskat bied nie 'n groot verskeidenheid van programme aan nie, maar wel heelwat innoverende idees. Van die programme waarop hierdie studie ingaan maak ook gebruik van vertaling as 'n moontlike woordeskatleer metode. Sommige van hierdie programme is ook al meer as vyf jaar oud, wat beteken dat dit nog moes funksioneer met beperkende rekenaarhardeware, as dit vergelyk word met die kragtige masjiene wat vandag beskikbaar is.

Groot hoeveelhede inligting kan opgediep word om woorde se betekenisse te verklaar en van ander betekenisse te onderskei. Werke oor vertaling bevat 'n rykdom van hierdie soort inligting. Dit is interessant dat die vertaalhandleidings wat ondersoek is, nie te veel konsentreer op abstrakte semantiese kwessies nie, maar eerder poog om praktiese metodes te verskaf vir die oplos van probleme rakende betekenis en nie-ekwivalensies in vertaling.

Die rekenaarmatige voorstelling van natuurlike taal leksikons is in min rekenaartoepassings so belangrik as in MT. Baie van die inligting wat deur vertalers as belangrik beskou word, is al in hierdie stelsels geënkodeer; verskeie soorte vertaal-nie-ekwivalensies word hanteer, sowel as kollokasies, subkategorisering, seleksiebeperking, ensovoorts.

Die informasie wat in deel een en twee uiteengesit is, word gebruik om in deel drie aan te dui hoe interdissiplinêre samewerking tussen RGTO en MT voordele vir beide kan inhou.

9.2 Samevatting van belangrike temas

Verskeie temas het herhaaldelik in literatuur voorgekom wat vir hierdie studie ondersoek is. Hieronder word van die belangrikstes daarvan kortliks saamgevat.

9.2.1 Gebalanseerde teoretiese standpunte

Tweedetaalleer kan, in elk geval by die aanvang van die proses, voorkom asof dit 'n proses van vertaling is vanaf die eerstetaal na die tweedetaal. Hierdie waarneming sluit aan by die stellings van taalleerteoretici met 'n behavioristiese uitgangspunt, soos Lado (1957). Dit het gelei tot die gebruik van konsepte soos taaloordrag en navorsing soos kontrastiewe analise. Toe behaviorisme in onguns begin kom het in die vroeë tweede helfte van die eeu, het navorsers van tweedetaalleer begin soek na plaasvervangers vir hierdie konsepte. Die fel kritiek van vroeër het egter later plek gemaak vir 'n meer realistiese benadering waarin erken word dat oordrag wel ook plaasvind.

Vertaling as tweedetaalleer metode het nie noemenswaardige navorsingsdekking nie (Farrington, 1988). Dit is moontlik 'n reaksie op die eertydse oorbeklemtoning van taaloordrag, maar dalk ook 'n resultaat van die gebrek aan 'n vertaalteorie wat vergelyk kan word met standaarde van taalleer- en linguistiese teorieë.

Nadat die transformasionele beweging in die laat 1950s en vroeë 1960s in die linguistiek posgevat het, en die behavioristiese uitgangspunt in onguns laat kom het, is nuwe idees oor taalleer geopper. 'n Vername voorbeeld hiervan is die voorstel van 'n aangebore taalvermoë en 'n universele grammatika. Daar is egter

steeds nie uitsluitsel daaroor of hierdie vermoë wel bestaan nie en ook nie of die vermoë ook gebruik word vir die leer van 'n tweedetaal nie.

Kontemporêre taalleerteorie is gevarieerd en ook gebalanseerd in dié sin dat daar vandag erken word dat oordrag wel plaasvind tydens tweedetaalleer. Indien die veldtog gepolariseer was, sou twee hoofgroepe woordeskatleer metodes waarskynlik kon bestaan; hulle word kortliks hieronder uiteengesit. 'n Benadering tot woordeskat sal nou egter 'n kombinasie wees van hierdie twee.

9.2.1.1 Die oordragmetode

Die woordeskat van die tweedetaal word aangebied met inagneming van die ooreenstemmende woordeskat van die eerstetaal. Dit kan prakties gedoen word deur tweedetaalwoordeskat aan te gee as vertalings van eerstetaalwoordeskat, of deur taalleerders eerstetaalsinne te laat vertaal wat die woorde bevat wat geleer moet word. Hierdie benadering skep die indruk dat dit behavioristies is en dat hier van konsepte soos oordrag en metodes soos kontrastiewe analise gebruik gemaak moet word.

Deur middel van kontrastiewe analise kan uitgekom word by die identifisering van 'n verskeidenheid leksikale uiteenlopendhede, soos dié van Stockwell et al. (1965), en moontlik selfs nog ingewikkelder uiteenlopendhede, soos dié van Dorr (1993). Deur die woordeskatte van die eerste- en tweedetaal te kategoriseer ten opsigte van ekwivalensie of nie-ekwivalensie, soos byvoorbeeld deur Baker (1992) gespesifiseer word, kan die gevare van negatiewe oordrag vermy word en die voordele van positiewe oordrag benut word.

9.2.1.2 Die universele grammatika metode

Die woordeskat van die tweedetaal kan egter ook in isolasie aangebied word. Die betekenis van woorde kan dan verklaar word op die wyse waarop 'n enkeltalige verklarende woordeboek dit sou doen.

Die betekenis van die woorde wat geleer moet word, word verduidelik met die tweedetaal self. Die woordvorm word dus direk met die betekenis daarvan in verband gebring; die eerstetaalekwivalent tree dus nie meer op as 'n moontlike skakel tussen die betekenis en die tweedetaalwoord nie. Dit is moontlik dat hierdie soort woordeskatonderrig binne 'n benadering sal plaasvind wat nie behavioristies georiënteerd is nie, die gevare van oordrag sal probeer vermy, en 'n sterker of swakker standpunt sal inneem ten gunste van 'n nativistiese teorie met die bestaan van universele grammatika as belangrike bestanddeel.

9.2.2 Ontwikkelingsekwense

Ontwikkelingsekwense en sekwensiële ontwikkelingsvlakke word gepostuleer deur Krashen (1985) en die ZISA-model (Clahsen, 1980; Pieneman, 1980, 1987; Meisel, Clahsen en Pieneman, 1981; Pieneman en Johnston, 1987; Hyltenstam en Pieneman, 1985). Veral lg. gee gedetailleerde beskrywings van die konsepte wat op elke vlak deur die taalleerder bemeester word. Hierdie ontwikkelingsvlakke maak dit moontlik vir die onderwyser om taalleermateriaal op sinvolle maniere aan te bied - net materiaal wat toeganklik is vir 'n bepaalde leerder word aan daardie leerder beskikbaar gestel.

9.2.3 Sintaktiese konteks en kollokasies

Vertaalteorie en MT-teorie dui aan dat die sintaktiese konteks van 'n woord gebruik kan word, nie net om te bepaal wat die woordsoort van 'n woord is nie, maar ook om te bepaal watter een van 'n verskeidenheid betekenisse die betrokke woord wel het. Wanneer die betekenis van 'n woord bepaal is, is dit makliker om die vertaalekwivalent daarvan te kry.

Dit is ook moontlik om die korrekte vertalings van woorde te vind, sonder om te veel op sintaktiese of semantiese inligting staat te maak. Indien twee of meer brontaalwoorde binne 'n redelike afstand van mekaar af voorkom, en die woorde aangeteken is as kollokate, kan die dienooreenkomstige aangetekende vertalings van daardie woorde gebruik word as vertaalekwivalente.

Albei hierdie soorte konteksinligting kan eksplisiet aan woordeskatleerders beskikbaar gestel word, of dit kan gebruik word om leermateriaal sinvol te groepeer.

9.2.4 Verskeidenheid van nie-ekwivalensie

In die veld van vertaling word ook baie noukeurige inligting aangetref oor die verskillende gevalle van nie-ekwivalensie wat tydens die vertaalproses raakgeloop kan word. Hierdie soort inligting sal veral nuttig wees wanneer 'n leermetode gevolg word wat gebaseer is op vertaalmetodes.

9.2.5 Semantiese verbande tussen woorde

Verbande soos sinonieme en antonieme is redelik bekende verbande wat aan die gebruikers van 'n taal bekend behoort te wees. Ander, soos hiponimiese, hipernimiese, polisemiese en taksonomiese verbande, en ook vormlike verbande, soos homografie en homofonie, is minder bekend, maar kan ook handig wees om verbande tussen die woorde van 'n nuwe taal te skep, en om sodoende hul betekenisse te kan onderskei.

9.3 Slot

Die voorstelle wat gemaak word oor 'n moontlike simbiose tussen masjienvertaling en rekenaargesteunde taalonderrig met betrekking tot die leksikon ondersteun die hipotese wat in die inleiding tot hierdie studie gemaak is. Daar is probeer om vas te stel of die relevante dissiplines enige finale antwoorde het op kwessies soos wat die regte of enigste manier is om 'n tweedetaal se woordeskat te leer of wat die omvattendste model van die vertaalproses is.

Soos verwag kon word, bestaan daar nie sulke antwoorde nie. Die navorsing was egter noodsaaklik omdat dit aandui wat die rigting van ontwikkeling is en wat die slaggate van die verlede was. Die voorstelle in hoofstuk 7 en hoofstuk 8 word gemaak op grond van hierdie riglyne. Dit dui daarop dat 'n wedersyds voordelige samewerking tussen masjienvertaling en rekenaargesteunde taalonderrig 'n aantreklike moontlikheid is.

BIBLIOGRAFIE

ADJEMIAN, C. 1976. On the nature of interlanguage systems. *Language learning* 26: 297-320.

ALATIS, J.E. (*ed.*). 1968. Contrastive linguistics and its pedagogigal implications. Report of the 19th Round Table on Linguistics and Language Studies. Washington D.C.: Georgetown University Press.

ANDERSEN, R. 1984. The one to one principle of interlanguage construction. *Language learning* 34: 77-95.

ANDERSON, R.C. en AUSUBEL, D.A. (*eds.*). 1965. Readings in the psychology of cognition. New York: Holt, Rinehart & Winston.

AUSUBEL, D.P. 1967. Learning theory and classroom practice. Bulletin, no. 1. Toronto, Ontario: Institute for Studies in Education. p. 307 - 309.

ARABSKI, J. 1979. Errors as indications of the development of interlanguage. Katowice: Uniwersytet Slaski.

ASHER, J. 1977. Learning another language through actions: the complete teacher's guidebook. Los Gatos, CA: Sky Oaks Productions.

BAHNS, J. en WODE, H. 1980. Form and function in L2 acquisition: the case of do-support in negation. (*In* Felix, S. (*ed.*). Second language development. Tubingen: Gunter Narr. p. 81-92).

BAKER, M. 1992. In other words - a coursebook on translation. Londen en New York: Routledge.

BARNETT, J., INDERJEET, M., MARTIN, P. en RICH, E. 1991a. Reversible machine translation: what to do when the languages don't line up, Proceedings of the Workshop on Reversible Grammars in Natural Language Processing, ACL-91, University of California, Berkeley, CA. p. 61 - 70.

BARNETT, J., INDERJEET, M., RICH, E., AONE, C., KNIGHT, K. en MARTINEZ, J.C. 1991b. Capturing language-specific semantic distinctions in interlingua-based MT, Proceedings of Machine Translation Summit, Washington, DC. p. 25 - 32.

BEAVEN, J. 1992a. Lexicalist unification-based machine translation, Ph.D. thesis, University of

Edinburgh.

BEAVEN, J. 1992b. Shake and bake machine translation, Proceedings of Fourteenth International Conference on Computational Linguistics, Nantes, France. p. 603 - 609.

BELL, R.T. 1991. Translation and translating - theory and practice. London en New York: Longman.

BEVER, T. 1970. The cognitive basis for linguistic structures. (*In* Hayes, J. (*ed.*). Cognition and the growth of language. New York: John Wiley. p. 279-352).

BLAAS, L. 1982. Fossilization in the advanced learner's lexicon. Ongepubliseerde manuskrip, Universiteit van Utrecht, Department of English.

BLEY-VROMAN, R. 1983. The comparative fallacy in interlanguage studies: the case of systematicity. *Language learning* 33: 1-17.

BLEY-VROMAN, R. 1989. What is the logical problem of foreign language learning? (*In* Gass, S. en Schachter, J. (*eds.*). Linguistic perspectives on second language acquisition. Cambridge: Cambridge University Press. p. 41 - 68).

BLOOMFIELD, L. 1933. Language. New York: Holt. Londen: Allen and Unwin, 1935.

BLOOMFIELD, L. 1942. Outline guide for the practical study of foreign languages. Special Publications of the Linguistic Society of America. Baltimore: Linguistic Soceity of America.

BOWERMAN, M. 1987. Commentary: mechanisms of language acquisition. (*In* McWhinney, B. (*ed.*). Mechanisms of language acquisition. Hillsdale, N.J.: Lawrence Erlbaum Associates).

BROWN, H.D. 1987 . Principles of language learning and teaching. Englewood Cliffs, New Jersey: Prentice-Hall.

BRUNER, J.S. 1960/1977. The process of education. (Heruitgegee in 1977 met 'n nuwe voorwoord). Cambridge, Mass.: Harvard University Press.

BRUNER, J.S. 1966. Toward a theory of instruction. Cambridge, Mass.: Harvard University Press.

BURT, M.K., DULAY, H.C. en FINOCCHIARO, M. (*eds.*). 1977. Viewpoints of English as a second language. New York: Regents.

BUTEAU, M. 1970. Students' errors and the learning of French as a second language. *International review of applied linguistics* 8:133-145.

BUTTERWORTH, G. 1972. A Spanish-speaking adolescent's acquisition of English syntax. MA TESL tesis, University of California at Los Angeles.

CAMERON, K. 1989. Computer assisted language learning: program structure and principles. Oxford: Intellect Books; Blackwell Scientific Publications.

CARROLL, J.B. 1971. Current issues in psycholinguistics and second language teaching. *TESOL Quarterly* 5: 101-114.

CHAMOT, A.U. 1978. Grammatical problems in learning English as 'n third language. (*In* Hatch, E. (*ed.*). Second language acquisition: a book of readings. Rowley, Mass.: Newbury House. p. 176-189.)

CHASTAIN, K. 1971. The development of modern language skills: theory to practice. Philadelphia: Center for Curriculum Development.

CHASTAIN, K. 1976. Developing second-language skills: theory to practice. Chicago: Rand McNally. (2de uitgawe van Chastain, 1971).

CHOMSKY, N. 1965. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press.

CHOMSKY, N. 1980. Rules and representations. Oxford: Blackwell.

CHOMSKY, N. 1981. Lectures on government and binding. Dordrecht: Foris.

CLAHSEN, H. 1980. Psycholinguistic aspects of L2 acquisition. (*In* Felix, S. (*ed.*). Second language development. Tubingen: Gunter Narr. p. 57-79).

CLAHSEN, H. 1984. The acquisition of German word order: a test case for cognitive approaches to L2 development. (*In* Andersen, R. (*ed.*). Second language: a crosslinguistic perspective. Rowley, Mass.: Newbury House. p. 283-332).

CORDER, S. 1975. 'Simple codes' and the source of the second language learner's initial heuristic hypothesis. Lesing voorgedra by die Colloque 'Theoretical linguistics models in applied linguistics' IV, Université de Neuchâtel. (In *Studies in second language acquisition*. 1977. 1(1): 1-10).

CORDER, S. 1983. A role for the mother tongue. (*In* Gass, S. en Selinker, L. (*eds.*). Language transfer in language learning. Rowley, Mass.: Newbury House. p. 85-97).

CORDER, S.P. 1967. The significance of learners' errors. *International Review of Applied Linguistics*, 5, 161-170.

CRUSE, D.A. 1986. Lexical Semantics. Cambridge: Cambridge University Press.

DATO, D. 1970. American children's acquisition of Spanish syntax in the Madrid environment: a preliminary edition. Finale verslag, Projek no. 3036, Kontrak No. O.E.C.2-7-002 637,USHEW.

DATO, D.P. (*ed.*). 1975. Developmental psycholinguistics: Theory and applications. Georgetown University Round Table on Languages and Linguistics. Washington D.C.: Georgetown University Press.

DE SAUSSURE, F. 1916. Cours de linguistique générale. Payot, Parys.

DI PIETRO, R.J. 1968. Contrastive analysis and the notions of deep and surface grammar. (*In* Alatis, J.E. (*ed.*). Contrastive linguistics and its pedagogigal implications. Report of the 19th Round Table on Linguistics and Language Studies. Washington D.C.: Georgetown University Press. p. 65 - 80)

DI PIETRO, R.J. 1971. Language structures in contrast. Rowley, Mass.: Newbury House.

DORR, B.J. 1990a. Solving thematic differences in machine translation, Proceedings of the 28th Annual Conference of the Cognitive Science Society, Ann Arbor, Michigan, p. 74 - 81.

DORR, B.J. 1990b. Lexical Conceptual Structure and machine translation. Ph.D. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

DORR, B.J. 1993: Machine translation: a view from the lexicon. Cambridge, Massachusetts: MIT Press.

DULAY, H. en BURT, M. 1973. Should we teach children syntax? *Language learning* 23: 245-258.

DULAY, H.C. en BURT, M.K. 1974. You can't learn without goofing: an analysis of children's second language learning strategies. (*In* Richards, J.C. (*ed.*). Error analysis: perspectives on second language acquisition. Londen: Longman. p. 95 -123).

DULAY, H.C. en BURT, M.K. 1975. A new approach to discovering universals of child language acquisition. (*In* Dato, D.P. (*ed.*). Developmental psycholinguistics: theory and applications. Georgetown University Round Table on Languages and Linguistics, Washington D.C.: Georgetown University Press. p. 209-233).

DULAY, H.C. en BURT, M.K. 1977. Remarks on creativity in second language learning. (*In* Burt, M.K., Dulay, H.C. en Finocchiaro, M. (*eds.*). Viewpoints of English as a second language. New York: Regents. p. 95-126).

DUSKOVA, L. 1969. On sources of errors in foreign language learning. *International review of applied linguistics* 7:11-36.

ECKMAN, F. MORAVSCIK, E., en WIRTH, J. 1989. Implicational universals and interrogative structures in the interlanguage of ESL learners. *Language Learning*, 39, 173 - 205.

ELLIS, R. 1985. Understanding second language acquisition. Oxford en Kaapstad: Oxford University Press.

ELSHOUT-MOHR, M. en VAN DAALEN-KAPTEIJNS, M.M. 1987. Cognitive processes in learning word meanings. (*In* McKeown, M.G. en Curtis, M.E. (*eds.*). The nature of vocabulary acquisition. Londen en Hillsdale, N.J.: Lawrence Erlbaum Associates. p. 53-71).

ERVIN-TRIPP, S. 1974. Is second language learning like the first? *TESOL Quarterly* 8: 111-127.

FARRINGTON 1988. Translation and the microcomputer. (*In* Jung, U.O.H. (*ed.*). Computers in applied linguistics en language teaching. New York, Parys, Bern: Verlag Peter Lang).

FELIX, S. 1985. More evidence on competing cognitive systems. *Second language research* 1(1): 47-72.

FRIES, C. 1945. Teaching and learning English as a foreign language. Ann Arbor: University of Michigan Press.

FRIES, C. 1949. The Chicago investigation. *Language Learning* 2:89-99.

GAGNÉ, R.M. 1977. The conditions of learning. New York: Holt, Rinehart and Winston.

GASS, G. & SELINKER, L. 1994. Second language acquisition: an introductory course. Lawrence Erlbaum Associates: Hillsdale, New Jersey.

GATTEGNO, C. 1972. Teaching foreign languages in schools: The Silent Way. Second Edition. New York: Educational Solutions.

GATTEGNO, C. 1976. The common sense of teaching foreign languages. New York: Educational Solutions.

GEORGE, H. 1972. Common errors in language learning. Rowley, Mass.: Newbury House.

GINGRAS, R.C. (*ed.*). 1978. Second language acquisition and foreign language teaching. Arlington, Va.: Center for Applied Linguistics.

GIVON, T. 1981. Typology and functional domains. *Studies in Language* 5: 163-193.

GIVON, T. 1985. Function, structure and language acquisition. (*In* Slobin, D. (*ed.*). The cross-linguistic study of language acquisition. Hillsdale, N.J.: Lawrence Erlbaum. p. 1005-1028).

GLEITMAN, L. 1993. The structural sources of verb meanings. (*In* Bloom, P. (*ed.*). Language acquisition. New York en Londen: Harvester Wheatsheaf. p. 174-221).

GOUIN, F. 1880. L'art d'enseigner et d'étudier les langues. Parys. Vertaal deur Swan, H. en Betis, V. 1892. The art of teaching and studying languages. Londen: George Philip.

GRIMSHAW, J. 1981. Form, function and the language acquisition device. (*In* Baker, C.L. en McCarthy, J.J. (*eds.*). The logical problem of language acquisition. Cambridge, Mass.: MIT Press. p. 183-210).

HAKUTA, K. 1974. A preliminary report on the development of grammatical morphemes in a Japanese girl learning English as a second language. *Working papers on bilingualism* 3: 18-43.

HALLIDAY, M.A.K., MCINTOSH, A. en STREVENS, P. 1964. The linguistic sciences and language teaching. Londen: Longman.

HARDISTY, D. en WINDEATT, S. 1989. Computer assisted language learning. Oxford: Oxford University Press.

HATCH, E. 1978a. Discourse analysis, speech acts and second language acquisition. (*In* Ritchie, W. (*ed.*). Second language acquisition research. New York: Academic Press. p. 137-155).

HATCH, E. (*ed.*). 1978b. Second language acquisition: a book of readings. Rowley, Mass.: Newbury House.

HATCH, E. 1983. Psycholinguistics: a second language perspective. Rowley, Mass.: Newbury House.

HATCH, E., FLASHNER, V. en HUNT, L. 1986. The experience model and language teaching. (*In* Day, R. (*ed.*). Talking to learn: conversation in second language acquisition. Rowley, Mass.: Newsbury House. p. 5 - 22.

HOLLAND, J.G. en SKINNER, B.F. 1961. The analysis of behavior. New York: McGraw-Hill.

HUANG, J. 1970. A Chinese child's acquisition of English syntax. Ongepubliseerde MA TESL tesis, University of California at Los Angeles.

HUEBNER, T. 1980. Creative construction and the case of the misguided pattern. (*In* Fisher, J. Clarke, M. en Schachter, J. (*eds.*). On TESOL '80. TESOL, Washington, D.C. p. 101-110).

HUTCHINS, W.J. en SOMERS, H.L. 1992. An introduction to machine translation. Londen, San Diego, New York, Boston, Sydney, Tokio, Toronto: Academic Press.

HYLTENSTAM, K. 1977. Implicational patterns in interlanguage syntax variation. *Language learning* 27(2): 383-411.

HYLTENSTAM, K. 1987. Markedness, language universals, language typology and second language acquisition. (*In* Pfaff, C. (*ed.*). First and second language acquisition processes. Cambridge, Mass.: Newbury House. p. 55-78).

HYLTENSTAM, K. en PIENEMAN, M. (*eds.*). 1985. Modelling and assessing second language acquisition. (nie verstrek): Multilingual Matters.

ITI News. Desember 1988. 3(3). (ITI is die Institute of Translation and Interpreting in die Verenigde Koninkryk.)

JOHNSTON, M. 1985. Syntactic and morphological progressions in learner English. Research report, Department of Immigration and Ehtnic Affairs, Commonwealth of Australia.

JOVANOVIC, M. 1987. Translation and computers - an encounter of the third kind. (*In* Legenhausen, L. en Wolff, D. (*eds.*). Computer assisted language learning and innovative EFL methodology. Augsburg: Augsburg Universtät.)

JUNG, U.O.H. (*ed.*). 1988. Computers in applied linguistics en language teaching. New York, Parys, Bern: Verlag Peter Lang.

KAMEYAMA, M. OCHITANI, R., PETERS, S. en SIRAI, H. 1991. Resolving machine translation mismatches with information flow, Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics, University of California, Berkeley, CA, p. 193 - 200.

KEMBLE, I.R. en BRIERLEY,W. (*eds.*). 1991. Computers as a tool in language learning. New York, Londen, Toronto, Sydney, Tokio, Singapoer: Ellis Horwood.

KINOSHITA, S., PHILLIPS, J. en TSUJII, J. 1992. Interaction between structural changes in machine translation, Proceedings of Fourteenth International Conference on Computational Linguistics, Nantes, France, p. 679 - 685.

KRASHEN, S. 1976. Formal and informal linguistic environments in language acquisition and language learning. *TESOL Quarterly* 10: 157-168.

KRASHEN, S. 1978. The monitor model for second language acquisition. (*In* Gingras, R.C. (*ed.*). Second language acquisition and foreign language teaching. Arlington, Va.: Center for Applied Linguistics. p. 1 - 26.)

KRASHEN, S. 1985. The input hypothesis: issues and implications. New York: Longman.

KRASHEN, S. en SCARCELLA, R. 1978. On routines and patterns in language acquisition and performance. *Language learning* 28: 283-300.

KRASHEN, S. en TERREL, T.D. 1983. The Natural Approach: language acquisition in the classroom. Oxford: Pergamon Press.

LADO, R. 1957. Linguistics Across Cultures. Ann Arbor: University of Michigan Press.

LANDAU, B. en GLEITMAN, L. 1985. Language and experience: evidence from the blind child. Cambridge, Mass.: Harvard University Press.

LARSEN-FREEMAN, D. en LONG, M.H. 1991. An introduction to second language acquisition research. Londen en New York: Longman.

LEVENSTON, E. 1979. Second language acquisition: issues and problems. *Interlanguage Studies Bulletin*, 4, 147-160.

LINDOP, J. en TSUJII, J. 1991. Complex transfer in MT: a survey of examples, Center for Computational Linguistics, UMIST, Manchester, CCL/UMIST Report 91\5.

LOZANOV, G. 1979. Suggestology and outlines of suggestopedy. New York: Gordon and Breach Science Publishers.

MACKEY, W.F. 1965. Language teaching analysis. Londen: Longman.

McCREARY, D.R. 1986. Improving bilingual Japanese-English and English-Japanese dictionaries. *Papers in linguistics* 19(1): 55-66.

MEARA, P. 1984. The study of lexis in interlanguage. (*In* Davies, A., Criper, C. en Howatt, A. (*eds.*). Interlanguage. Edinburgh: University of Edinburgh Press. p. 225 - 235).

MEISEL, J. , CLAHSEN, H. en PIENEMAN, M. 1981. On determining developmental stages in natural second language acquisition. *Studies in second language acquisition* 3(1): 109-135.

MEISEL, J. 1983. Strategies for second language acquisition: more than one kind of simplification. (*In* Andersen, R. (*ed.*). Pidginization and creolization as language acquisition. Rowley, Mass.: Newbury House).

MELBY, A.K. 1986. Lexical transfer: missing element in linguistics theories, Proceedings of Eleventh International Conference on Computational Linguistics, Bonn, Germany.

MILLER A., BECKWITH, R., FELLBAUM, C., GROSS, D. en MILLER, K. 1993. Introduction to WordNet: an on-line lexical database. Internet: http://www.cogsci.princeton.edu/~wn/.

MILON, G. 1974. The development of negation in English by a second language learner. *TESOL Quarterly* 8: 137-143.

MOULIN, A. 1983. LSP dictionaries for EFL learners. (*In* Hartmann, R.R.K. Lexicography: principles and practice.)

NEMSER, W. 1971. Approximative systems of foreing language learners. IRAL 9:115-123. (*Ook in* Richards, J.C. (*ed.*). 1974. Error analysis: Perspectives on second language acquisition. Londen: Longman. p. 55-63).

NEMSER, W. 1971. Approximative systems of foreing language learners. *International review of applied linguistics* 9: 115-124.

NEWMARK, P. 1981. Approaches to translation. Oxford, New York, Toronto, Sydney, Parys, Frankfort: Pergamon Press.

NEWMARK, P. 1988. A textbook of translation. New York, Londen, Toronto, Sydney, Tokio: Prentice Hall.

NIRENBURG, S. (*ed.*). 1987. Machine translation: theoretical and methodological issues. Cambridge University: Cambridge.

NIRENBURG, S. en NIRENBURG, I. 1988. A framework for lexical selection in natural language generation, Proceedings of Twelfth International Conference on Computational Linguistics, Budapest, Hungary, p. 471 - 475.

O'GRADY, W. 1987. Principles of grammar and learning. University of Chicago Press.

OSGOOD, C.E. 1953. Method and theory in experimental psychology. New York: Oxford University Press.

OSGOOD, C.E. 1957. Contemporary approaches to cognition. Cambridge, Mass.: Harvard University Press.

PARKER, K. 1989. Learnability theory and the acquisition of syntax. *University of Hawaii Working Papers in ESL* 8: 49-78.

PIENEMAN, M. 1980. The second language acquisition of immigrant children. (*In* Felix, S. (*ed.*). Second language development. Tubingen: Gunter Narr. p. 41-56).

PIENEMAN, M. 1984. Psychological constraints on the teachability of languages. *Studies in second language acquisition* 6(2): 186-214.

PIENEMAN, M. 1987. Determining the influence of instruction on L2 speech processing. *Australian review of applied linguistics* 10(2): 83-113.

PIENEMAN, M. en JOHNSTON, M. 1987. Factors influencing the development of language proficiency. (*In* Nunan, D. (*ed.*). Applying second language acquisition research. Adelaide, NSW: National

Curriculum Resource Centre. p. 45 - 141).

PINKER, S. 1984. Language learnability and language development. Cambridge, Mass.: Harvard University Press.

PINKER, S. 1987. Resolving a learnability paradox in the acquisition of the verb lexicon, Lexicon working papers 17. Cambridge, Mass.: MIT Center for Cognitive Science. p. 1- 100.

POLITZER, R. 1978. Errors of English speakers of German as perceived and evaluated by German natives. *Modern Language Journal*, 62, 253-261.

RAVEM, R. 1968. Language acquisition in a second language environment. *International review of applied linguistics* 6: 165-185.

RAVEM, R. 1970. The development of Wh-questions in first and second language learners. (*In* Richards, J. (*ed.*). 1974. Error analysis. Longman. p. 134-155).

RICHARDS, J. 1971. Error analysis and second language strategies. *Language sciences* 17: 12-22.

RICHARDS, J.C en RODGERS, T.S. 1986. Approaches and methods in language teaching. Cambridge: Cambridge University Press.

RICHARDS, J.C. (*ed.*). 1974. Error analysis: Perspectives on second language acquisition. Londen: Longman.

RIVERS, W.M. 1968/1981. Teaching foreign-language skills. 2de uitgawe. Chicago en Londen: University of Chicago Press.

RUMELHART, D.E. en McCLELLAND, J.L. 1986. Parallel distributed processing. Vol. I. Cambridge, Mass.: MIT Press.

RUTHERFORD, W. 1982. Markedness in second language acquisition. *Language learning* 32(1): 85-108.

SCHACHTER, J. 1974. An error in error analysis. *Language learning* 27: 205-214.

SCHACHTER, J. 1988. Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics* 9: 219-235.

SCHMIDT, R. 1983. Interaction, acculturation and acquisition of communicative competence. (*In* Wolfson, N. en Judd, E. (*eds.*). Sociolinguistics and second language acquisition. Rowley, Mass.: Newbury House. p. 137-174).

SCHUMANN, J. 1987. The pidginization process: a model for second language acquisition. Rowley, Mass.: Newbury House.

SELINKER, L. 1972. *Interlanguage.* IRAL 10: 219-231. (*Ook in* Richards, J.C. (*ed.*). 1974. Error analysis: Perspectives on second language acquisition. Londen: Longman. p. 31-54).

SELINKER, L., SWAIN, M. en DUMAS, G. 1975. The interlanguage hypothesis extended to children. *Language Learning* 25: 139-152.

SINCLAIR, J.M. 1987. Looking up: an account of the Cobuild project in Lexical Computing. New York en Londen: Collins.

SKINNER, B.F. 1938. Behavior of organisms: an experimental analysis. New York: Appleton-Century-Crofts.

SKINNER, B.F. 1954. The science of learning and the art of teaching. *Harvard educational review* 24: 86-97.

SKINNER, B.F. 1957. Verbal behavior. New York: Appleton-Century-Crofts.

SKINNER, B.F. 1958. Teaching machines. *Science* 128: 969-977.

SKINNER, B.F. 1961. Teaching machines. *Scientific American*, November, 1-13.

SPOLSKY, B. 1989. Conditions for second language learning. Oxford University Press.

STENSON, N. 1974. Induced errors. (*In* Schumann, J. en STENSON, N. (*eds.*). New frontiers in second language learnings. Rowley, Mass.: Newbury House).

STERN, H.H. 1974. Retreat from dogmatism: toward a better theory of language teaching. *Canadian modern language review* 30: 244-254.

STERN, H.H. 1983. Fundamental concepts of language teaching. Oxford University Press.

STERNBERG, R.J. 1987. Most vocabulary is learned from context. (*In* McKeown, M.G. en Curtis, M.E. (*eds.*). The nature of vocabulary acquisition. Londen en Hillsdale, N.J.: Lawrence Erlbaum Associates. p. 89-105).

STOCKWELL, R., BOWEN, J. en MARTIN, J.W. 1965. The grammatical structures of English and Spanish. Chicago: University of Chicago Press.

TARONE, E. 1980. Communication strategies, foreigner talk and repair in interlanguage. *Language learning* 30: 417-431.

THORNDIKE, E.L. 1921. The teacher's word book, New York: Teachers College.

THORNDIKE, E.L. 1969. Human nature and social order. New York: Maple Press Corporation.

TSUJII, J. en FUJITA, K. 1991. Lexical transfer based on bilingual signs: towards interaction during transfer, Proceedings of the European Chapter of the Association for Computational Linguistics, Berlin, Germany, p. 275 - 280.

UFERT, D. 1988. Computers in lexicology and lexicography. (*In* Jung, U.O.H. (*ed.*). Computers in applied linguistics en language teaching. New York, Parys, Bern: Verlag Peter Lang.)

WAGNER-GOUGH, J. en HATCH, E. 1975. The importance of input data in second language acquisition studies. *Language learning* 25: 297-308.

WARDHAUGH, R. 1970. The contrastive analysis hypothesis. *TESOL Quarterly* 4: 123-130.

WATSON, J.B. 1909. The beginnings of the teaching of modern subjects in England. Londen: Pitman.

WATSON, J.B. 1913. Psychology as the behaviorist views it. *Psychological Review* 20: 158 - 177.

WATSON, J.B. 1914. Behavior: an introduction to comparative psychology. New York: Holt.

WATSON, J.B. 1919. Psychology from the standpoint of a behaviorist. Philadelphia: Lippincott.

WHITELOCK, P. 1992. Shake-and-bake translation, Proceedings of Fourteenth International Conference on Computational Linguistics, Nantes, France, p. 784 -791.

WHITMAN, R. en JACKSON, K. 1972. The unpredictability of contrastive analysis. *Language learning* 22: 29-41.

WIDDIG, W. en ESSER , R. 1988. New possibilities in Computer-assisted learning of foreign language vocabulary using the Apple Macintosh. (*In* Jung, U.O.H. (*ed.*). Computers in applied linguistics en language teaching. New York, Parys, Bern: Verlag Peter Lang. p. 265-278).

WODE, H. 1976. Developmental sequences in naturalistic L2 acquisition. *Working papers on bilingualism* 11: 1-31.

WONG FILLMORE, L. 1976. The second time around: cognitive and social strategies in second language acquisition. Ongepubliseerde Ph.D. verhandeling, Stanford University.

WOODWORTH, R.S. 1964. Contemporary schools of psychology. Great Brittain: John Dickenson & Company.

ZOBL, H. 1982. A direction for contrastive analysis: the comparative study of developmental sequences. *TESOL Quarterly* 16: 169-183.