

## DE TAALGEBRUIKER IN DE MENS

### Schets van zijn bouw en functie, toepassingen op moedertaal- en vreemdetaalverwerving\*

GERARD KEMPEN

**Inleiding.** — Slechts zelden treft men in de taalpsychologische literatuur modellen van de taalgebruiker aan, pogingen tot een totaalvisie op het systeem dat de mens in staat stelt tot taalgedrag in al zijn variëteiten. De oorzaken hiervan zijn niet moeilijk aan te geven. Ten eerste, de theorievorming in de psycholinguïstiek heeft altijd nauw contact onderhouden met de linguïstische theorievorming, en vanuit die hoek, waar de klemtoon ligt op beschrijving en verklaring van *taalintuïties*, zijn geen baanbrekende ideeën te verwachten aangaande de machinerie achter *primair taalgedrag*. Zulke ideeën zijn wel te verwachten — en vele werden er reeds geleverd — vanuit een specialisme binnen de computerwetenschappen dat “kunstmatige intelligentie” heet. Daar gaat het immers juist om het ontwerpen van “kunstmatige taalgebruikers” die zinnen uit een natuurlijke taal kunnen begrijpen en produceren. Opvallend is nu dat opbouw en werking van deze taalgedragsmodellen steeds sterker gaan contrasteren met de taalmodellen in de linguïstiek. Neem als een concreet vergelijkingspunt bijvoorbeeld de wijze waarop grammaticale kennis is geformaliseerd in enerzijds de regels van de transformationeel-generatieve grammatica en anderzijds de regels van “augmented transition networks” (zie par. 1.3 en 1.4). Ten tweede, de theorievorming binnen de experimentele psychologie munt bepaald niet uit door het streven de deelsystemen waaruit het cognitieve systeem is opgebouwd, in hun onderling verband te plaatsen.

In het bestek van dit artikel wil ik allereerst een schets van de menselijke taalgebruiker ontvouwen — een schets die, en dat moet van meet of aan duidelijk zijn, ten dele speculatief is. Op nogal wat plaatsen heb ik me meer laten leiden door de ervaringen van bouwers van kunstmatige taalgebruikssystemen dan door de empirische gegevens van taalpsychologen. Over vele onderdelen van de menselijke taalgebruiker zijn overigens in het geheel geen data beschikbaar! Met de data die er zijn raakt het hierna te schetsen taalgebruikersmodel nergens in conflict. De vrije opstelling die ik me hier veroorloof ten opzichte van empirische gegevens wordt m.i. gerechtvaardigd door de voordelen die een geïntegreerd model van de taalgebruiker biedt boven een onsamenhangend stel theorieën over onderdelen ervan. Deze voordelen zijn niet alleen theoretisch (de werking van de onderdelen is pas goed te begrijpen vanuit hun onderlinge interacties) maar vooral ook praktisch: als de taalpsychologie zich op toegepast gebied wenst te begeven moet zij ermee rekening houden dat ze “totaal taalgedrag” zal aantreffen. In de tweede sectie zal de aandacht gericht zijn op moedertaal- en vreemdetaalverwerving vanuit het perspectief van het taalgebruikersmodel

\* Tekst behorende bij een voordracht voor de Studiedag van de Nederlandse Vereniging voor Toegepaste Taalwetenschap op 23 november 1974 te Amsterdam.  
Adres van de auteur: Psychologisch Laboratorium, Eramuslaan 16, Nijmegen.

### 1. *Bouw en werking van de menselijke taalgebruiker*

Niemand zal moeite hebben, met de opsplitsing van de (mondellinge/ auditieve) taalgebruiker in de volgende vijf hoofdsystemen (zie Fig. 1): (a) spraakherkenner, (b) zinsontleder, (c) conceptueel systeem, (d) zinsgenerator, (e) articulator.

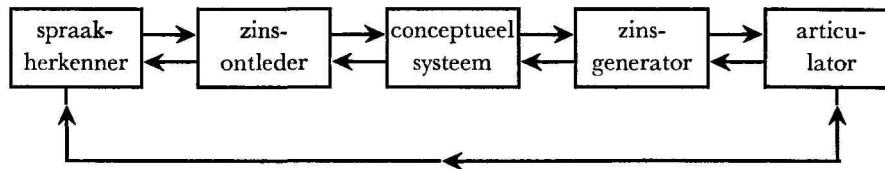


Fig. 1. De vijf hoofdsystemen van de menselijke taalgebruiker.

Wat deze systemen doen is het gemakkelijkst aan te duiden met vijf werkwoorden: (a) luisteren, (b) ontleden, (c) denken, (d) formuleren, (e) spreken. Onderdelen (a), (b) en (c) zijn nodig voor het *verstaan* van gesproken taaluitingen; (c), (d) en (e) zijn werkzaam tijdens het *produceren* van gesproken taal.

Wat veel minder voor de hand ligt en meer stof geeft tot nadenken en tot theoretische conflicten is de wijze waarop deze deelsystemen afzonderlijk en in onderling verband functioneren.

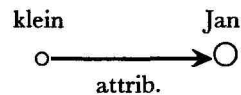
#### 1.1. Het conceptuele systeem

Het hart van de menselijke taalgebruiker is het conceptuele systeem.<sup>1</sup> Men kan het zich het beste voorstellen als een reusachtig netwerk waarvan de knopen corresponderen met concepten en de verbindingsdraden (pijlen) tussen knopen met relaties tussen concepten. Sinds ongeveer 10 jaar zijn er minstens een dozijn van dergelijke systemen ontwikkeld (zie literatuurlijst van Levelt & Kempen, 1994b), maar hier beperk ik me tot het verst ontwikkelde: Schank's "conceptual dependency network" (1972, 1973).

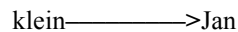
De concepten (knopen) in dit netwerk zijn van vierderlei soort: nominale concepten, actieconcepten, bepalers (modifiers) bij acties, en bepalers bij nominale concepten — resp. N, A, B<sub>A</sub>, B<sub>N</sub>. Deze conceptuele categorieën kunnen op welomschreven wijzen met elkaar in verband

<sup>1</sup> Het conceptuele systeem bevat zowel "semantische" als "pragmatische" kennis.

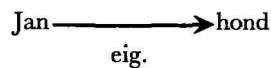
staan. Schank noemt deze verbanden “afhankelijkheden” (dependencies) want bij de meeste relaties tussen twee concepten is sprake van een asymmetrie. Bv.  $B_A$ 's kunnen niet zelfstandig voorkomen, ze staan altijd in attributieve relatie tot een A-concept. Dit in tegenstelling tot A-concepten die best zonder een  $B_A$  in het netwerk kunnen figureren. A- en N-concepten zijn van nature “regenten” van een afhankelijkheid,  $B_A$  en  $B_N$  zijn van huis uit “dependenten”. Pijlen in het netwerk wijzen steeds naar de regent. Bv. *kleine Jan* zou gerepresenteerd worden als:



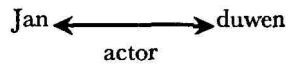
of, vereenvoudigd, als



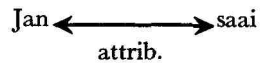
De labels van de pijlen geven het type afhankelijkheidsrelatie aan. Bv. in *Jan's hond* schuilt het relatietype “eigendom van”:



In bepaalde gevallen is er sprake van wederzijdse afhankelijkheid, bv. tussen actor en act,



en wanneer een eigenschap of lidmaatschap wordt toegekend aan een N.



Tot hiertoe zou men de indruk kunnen hebben dat er een één-op-één relatie bestaat tussen woorden en concepten. Niets is echter minder waar: Schrank is er, meer dan andere ontwerpers van conceptuele netwerken, op uit te laten zien dat de rijkdom van de woordenschat van een natuurlijke taal te vangen is onder een beperkt aantal elementaire concepten, mits men beschikt over een gedifferentieerd systeem van middelen om deze basisconcepten tot “conceptualisaties” aan elkaar te smeden. Bij wijze van illustratie volgt hier de conceptualisatie van zin (1).

## (1) Jan gooit stenen naar Piet.

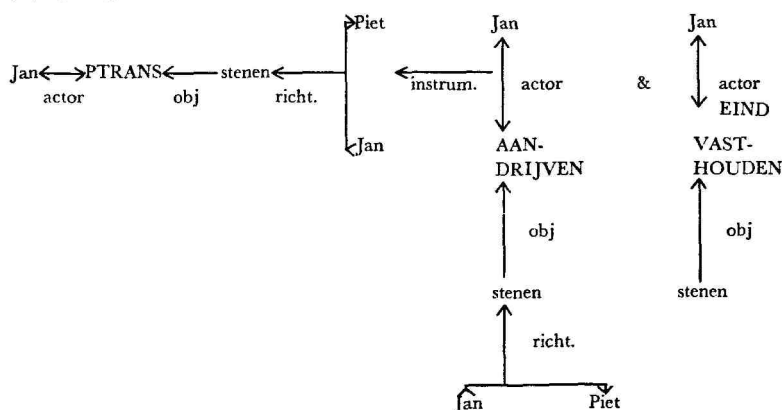


Fig. 2. Conceptualisatie behorende bij zin (1).

PTRANS (fysische plaatsverandering), AANDRIJVEN en VASTHOUDEN zijn drie van ongeveer een dozijn "primitieve actieconcepten" die Schank ten grondslag legt aan een zeer groot aantal werkwoorden. De plaatsverandering (PTRANS) van Jan naar Piet (richting) waaraan de stenen (object) onder invloed van Jan's optreden onderhevig zijn, wordt bewerkstelligd door (instrument) twee acties van Jan (actor): (1) door te fungeren als krachtbron (AANDRIJVEN) die de stenen naar Piet voortstuwt, en (2) door het VASTHOUDEN van de stenen te beEINDigen, d.w.z. door ze los te laten. We zien tevens dat aan het casusmechanisme (object, richting, instrument) de belangrijke en gecompliceerde taak is toegedacht de primitieve acties op de juiste wijze samen te voegen, hier tot het complexe idee *gooien*.

Dit demonstreert dat het conceptuele systeem in wezen niet-talig is. De relaties tussen de concepten en de woorden van de natuurlijke taal kunnen zeer ingewikkeld zijn; aan de andere kant zijn de concepten het gemakkelijkst op te vatten als codenamen voor de constanten die de wereld van het waarnemen en van het denken bevolken: personen, objecten, eigenschappen, handelingen, gebeurtenissen.

Het is van belang het laatste punt verder uit te werken. Stel een onderwijzer ziet een van zijn leerlingen, Jan, een steen gooien naar een klasgenootje, Piet. Hoe zal hij deze episode in zijn lange-termijn-geheugen coderen? Waarschijnlijk zal onze onderwijzer na een paar dagen nog over een voorstellingsbeeld van de episode beschikken. Betekent dit dat hij de episode

gedetailleerd heeft geregistreerd, net zoals een filmcamera elk detail op het filmpje vastlegt? Nee, het is voldoende als hij:

- (a) de waarnemingsbeelden die hij tijdens de episode binnenkrijgt onderbrengt in zijn systeem van concepten, d.w.z. herkent (in deze episode herkent hij o.m. PTRANS, STEEN, AANDRIJVEN, JAN etc),
- (b) de afhankelijkheidsrelaties tussen de herkende concepten invult zodat een netwerk ontstaat analoog aan Fig. 2,
- (c) dit netwerk opslaat in zijn lange-termijn-geheugen.

Maar hoe kan vanuit zo'n netwerk een voorstellingsbeeld geconstrueerd worden? Hiertoe is een belangrijke extra assumptie nodig, nl. dat het cognitieve systeem beschikt over een "voorstellingengenerator", een systeem dat conceptualisaties tot invoer neemt en deze vertaalt in voorstellingsbeelden. Zo'n systeem zou analoog kunnen werken aan een computer: de conceptualisaties fungeren als programmeertaal; voor elk concept (JAN, VASTHOUDEN, PTRANS etc.) kent het systeem een subroutine; de compiler schakelt de subroutines aaneen tot een serie van instructies die, bij uitvoering, beelden doen verschijnen op een "mentaal beeldstation".

Deze speculatieve uitwijding moge illustreren dat de concepten, hoewel aangeduid (vaak) met alledaagse woorden (JAN, AANDRIJVEN etc), niet-talige codes zijn die voor allerlei doeleinden gebruikt kunnen worden. Eén ervan hebben we gezien: de *voorstellingengenerator* vat conceptualisaties op als zinnen van een programmeertaal. Het *inferentie-mechanisme* ziet de conceptualisaties als databestand: uit de gegevens in bv. Fig. 2 kan dit mechanisme o.m. een antwoord afleiden op de vraag *Wie gooide een steen naar Piet*; door Fig. 2 in contact te brengen met een ander stukje netwerk zal het concluderen dat Jan (waarschijnlijk) agressieve bedoelingen had etc. Hoe de *taalgebruiker* het conceptuele systeem hanteert bij het verstaan en produceren van taaluitingen vormt het onderwerp van de volgende paragrafen.

## 1.2. De spraakherkenner.

Het spraakverstaan begint met de analyse van het akoestische signaal met als gewenst resultaat een sekwentie van woorden die het gehele signaal dekt. Het staat momenteel wel vast dat dit doel onbereikbaar is zonder syntaktische en semantische (conceptuele) aanwijzingen (vgl. Newell et al. 1973). De spraakherkenner zal van het spraaksignaal in eerste instantie slechts gedeelten kunnen thuisbrengen. Omtrent de identiteit van de niet-erkende passages

kunnen de zinsontleder en het conceptuele systeem wellicht nadere informatie verschaffen: de herkende woorden suggereren een bepaald syntactisch schema met bijbehorende funktiewoordjes, of het conceptuele systeem van de luisteraar — dat vele conceptualisaties met de spreker deelt — voorspelt woorden met een bepaalde conceptuele inhoud. Deze syntactische en conceptuele kennis brengt een enorme beperking aan in wat de spraakherkenner aan woorden kan verwachten, en leidt mogelijk zelfs tot correctie van fouten die de spraakherkenner in eerste aanleg gemaakt heeft.

Hier zien we een uiterst belangrijk kenmerk van de architectuur van de taalgebruiker: de diverse deelsystemen zijn niet in serie achter elkaar geschakeld maar ze werken in wederzijdse afhankelijkheid. In Fig. 1 duiden de dubbele pijlen op het tweerichtingsverkeer in de informatieoverdracht en de procesbesturing.

Overigens is nog heel weinig bekend van de manier waarop de spraakherkenner het akoestische signaal verwerkt. De bekendste theorieën (o.a. de motor-theorie, de analyse-door-synthese-theorie) zijn slechts fragmentarisch en bovendien controversieel (zie Campbell, 1974, voor een overzicht en voor verdere literatuur).

### 1.3. De zinsontleder.

Bij het percipiëren van een taaluiting staat syntactische ontleding geheel in dienst van het ontdekken van de conceptualisatie die eraan ten grondslag ligt. Het conceptuele systeem beschikt over een uitgebreide inventaris van de typen conceptualisaties waarin elk bekend woord kan voorkomen: het lexicon. Door de lexicale gegevens van de afzonderlijke woorden van de taaluiting te combineren is vaak een onbetwistbare beslissing mogelijk omtrent de conceptualisatie die de spreker bedoeld heeft: dan namelijk wanneer de conceptuele omgevingen die de woorden vereisen, precies in elkaar passen. In zulke gevallen, en die zijn eerder regel dan uitzondering, is slechts een minimum aan syntactische ontleding nodig. De uitingen hoeven niet eens grammaticaal te zijn: neem bv. de woordsekwentie *stormen-bomen-voetgangers-doden* waarin onmiddellijk een conceptualisatie wordt herkend zonder dat — afgezien van de volgorde — syntactische middelen gebruikt zijn. Zelfs staat van minstens twee woorden de woordsoort niet vast.

Het is vooral Schank geweest (1972, 1973) die dit principe naar voren heeft gebracht: de zinsontleder van de menselijke taalgebruiker is sterk woord-

en gaat pas na uitputting van zijn lexicale kennis over tot onderzoek van syntaktische verbanden. Duidelijk spreekt het belang van dit principe als men zich realiseert hoe men een zin als

(2) Jan gooit Piet met stenen.

puur syntactisch zou moeten ontleden zonder raadpleging van conceptuele informatie. Een grote verscheidenheid aan conceptuele relaties kan schuilgaan onder het zinschema NC-W-NC-mei-NC:

(3) Ik bezocht het museum met mijn vrouw.

(4) Ik sneed het vlees met een mes.

(5) Ik zag de man met de snor.

(6) Ik zag Arnold met Ineke.

(7) Ik haalde de eindstreep met moeite.

(8) Ik vergeleek de theorie met de werkelijkheid.

Bij deze zinnen is vaststelling van de constituentenstructuur en van de syntaktische relaties tussen de constituenten niet mogelijk zonder conceptuele kennis. Met name de conceptuele casussen die het werkwoord vereist zijn van groot belang (alhoewel niet uitsluitend: vgl. (5) en (6)).

De kunstmatige zinsontleder die Schank (1973) heeft gebouwd begint niettemin met een rudimentair stukje syntaktische analyse: hij zoekt de eerste naamwoordsgroep op die niet wordt geregeerd door een voorzetsel, d.w.z. die zou kunnen dienen als onderwerp, en het hoofdwerkwoord (eventueel in combinatie met hulpwerkwoord). Hierna geeft de ontleder het roer onmiddellijk over aan het conceptuele systeem. In geval van zin (2) zal het de met *gooien* geassocieerde conceptuele structuur aanreiken. Deze is gelijk aan die van Fig. 2 afgezien van de "casushouders" (Jan, Piet, stenen, etc.) die nog leeg zijn. Raadpleging van het lexicon levert voorts de informatie op dat het object van PTRANS (en van AANDRIJVEN etc.) het eerste FYSISCH OBJECT is dat na de persoonsvorm wordt aangetroffen, eventueel voorafgegaan door *met*. Het doelwit van het *gooien*, zo zegt het lexicon, is een FYSISCH OBJECT dat volgt op *naar*. Is het doelwit een MENS of DIER, dan is een andere volgorde mogelijk: doelwit-mei-object. M.b.v. deze regels kan de conceptualisatie voltooid worden door de concepten corresponderend met Jan, stenen, e.d. te deponeren in de casushouders.

Blijkt uit dit voorbeeld dat prioriteit van syntaktische boven conceptuele analyse kan leiden tot verregaande inefficiëntie, in sommige situaties, nl. als conceptuele richtlijnen ontbreken, moet van de nood een deugd gemaakt wor-

den. Dit is bv. het geval als de spreker woorden gebruikt die de luisteraar niet kent. Een extreem voorbeeld is Lewis Carroll's "wauwelwok" (ontleend aan Reedijk & Kossman's vertaling van *Through the looking glass*):

*Wauwelwok*

‘t Wier bradig, en de spiraments  
Bedroorden slendig in het zwiets:  
Hoe klarm waren de ooiefents,  
Bij ‘t bliuften der beriets.

Een zeer aantrekkelijk model voor niet-conceptueel gestuurde zinsontleding bieden de zog. "augmented transition networks" (Woods, 1970; Kaplan, 1972). Bij wijze van voorbeeld schets ik hier een summiere (en onvolledige) grammatica voor de ontleding van zin (9).

(9) De aardige dokter belde de patiënt op.

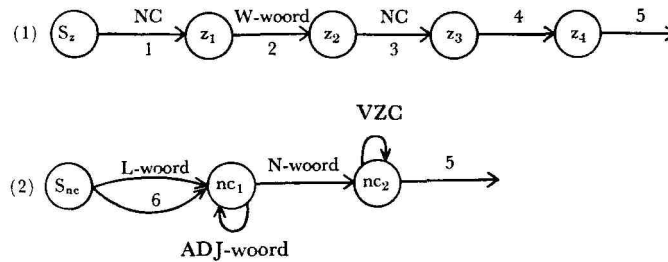
De grammatica (zie de drie deelnetwerken van Fig. 3) bestaat uit toestanden (cirkels) en overgangen (pijlen) tussen toestanden. De ontleder start in toestand  $S_z$  (start-zin) en probeert deelnetwerk (1) van links naar rechts te doorlopen. Een overgang mag alleen dan gemaakt worden als aan alle voorwaarden die bij die overgang opgesomd staan, voldaan is. Er zijn meerdere soorten voorwaarden:

- de cijfers 1, 2, 3, etc. verwijzen naar de condities die in het lijstje onder de figuur vermeld staan;
- L-woord, N-woord, W-woord etc. betekenen dat het eerstvolgende woord van de zin moet behoren tot de woordsoort lidwoord, zelfstandig naamwoord, werkwoord etc;
- een voorwaarde zoals NC of VZC omdat de opdracht om een constituent van een bepaald type, hier een naamwoordsgroep of een voorzetselgroep, te gaan opzoeken in het eerstvolgende gedeelte van de zin. Om bv. de opdracht (subroutine) NC te vervullen springt de ontleder rechtstreeks over naar toestand  $S_{nc}$  in het tweede deelnetwerk van Fig. 3. Hij werkt zich vooruit naar het einde van dit deelnetwerk en keert met de gevonden naamwoordsgroep terug naar de plaats waar hij de NC-opdracht had ontvangen. Omdat nu aan de NC-voorwaarde is voldaan mag de betreffende overgang gemaakt worden.



Verdere bijzonderheden:

- de ontleder leest de woorden van de zin in volgorde van links naar rechts;
- als uit een toestand meerdere pijlen vertrekken worden ze klokgewijs beproefd;
- elk deelresultaat wordt weergegeven m.b.v. haakjesnotatie; bv. als een L-woordovergang wordt geprobeerd met *de* als in te lezen woord, zal het succesvolle resultaat aangeduid worden als (L *de*).



*Conditie*

1. Onvoorwaardelijk
2. Onvoorwaardelijk
3. W-woord is overgankelijk
4. Het eerste te ontlenen woord vormt samen met het woord in (W \_\_\_\_\_) een scheidbaar samengesteld werkwoord
5. Onvoorwaardelijk
6. Onvoorwaardelijk

*Operatie*

- Plaats de ontlede NC in (SUBJ\_\_\_\_)
- Verander het W-woord in de onbepaalde wijs
- Plaats de ontlede NC in (OBJ\_\_\_\_)
- Plaats dat scheidbare werkwoord in (W\_\_\_\_)
- Maak haakjesontleding. Keer terug \_\_\_\_\_

Fig. 3. Een eenvoudige netwerkgrammatica.

Laten we de ontleder stap voor stap volgen. Op overgang 1 komt hij de opdracht “zoek een NC” tegen en springt naar toestand  $S_{nc}$ . De eerste overgang die hij hier beproeft is L-woord. Dit lukt, en met als eerste resultaat (L *de*) gaat hij over op toestand  $nc_1$ . Het eerstvolgend woord is geen N-woord maar een ADJ-woord zodat de lus wordt doorlopen en (ADJ *aardige*) resulteert. *Dokter*

is een N-woord en wordt als (N *dokter*) bewaard. Vervolgens wordt een voorzetselconstituent geprobeerd. Omdat *belt* echter geen voorzetsel is moet de ontleder onverrichterzake terugkeren van toestand  $S_{vzc}$  en gaat nu operatie 5 uitvoeren. Het eerste deel hiervan, “Maak haakjesontleding”, voegt de resultaten van het deelnetwerk waarin de ontleder momenteel bezig is, in elkaar tot een haakjesontleding van een bepaald type. Binnen deelnetwerk (2) bv. heeft operatie 5 tot gevolg dat de L-, ADJ-, N- en VZC-resultaten (voor zover aanwezig) geplaatst worden in (NC\_\_\_\_), en wel in de aangegeven volgorde. Kwamen meerdere adjectiva voor, dan worden deze eenvoudigweg aan elkaar vastgeknoopt; bv. *mooi oud huis* wordt (NC (ADJ *mooi*) (ADJ *oud*) (N *huis*)). In deelnetwerk (1) heeft operatie 5 een heel andere uitwerking. De haakjesontleding krijgt nu de vorm

$$(Z (NC\_\_\_\_) (C (W\_\_\_\_) (NC\_\_\_\_))),$$

waarbij de eerste NC afkomstig is uit (SUBJ\_\_\_\_), de tweede uit (OBJ\_\_\_\_).

Het tweede deel van operatie 5, “Keer terug”, heeft tot gevolg dat de ontleder met de zojuist gefabriceerde haakjesontleding terugkeert naar het punt van oorsprong. Nu is dat overgang 1, en de haakjesontleding is

$$(NC (L\ de) (ADJ\ aardige) (N\ dokter)).$$

Conditie 1 wordt nu opgezocht in de lijst. Deze is zonder meer vervuld, maar eerst moet nog de onder 1 vermelde operatie uitgevoerd worden: de ontleder NC wordt ingevoegd in (SUBJ\_\_\_\_), d.w.z. wordt opgevat als onderwerp. Via *belt*, het eerste te ontleden woord, komt hij in  $z_2$  na eerst de standaardvorm *bellen* ingevuld te hebben in (W\_\_\_\_). Omdat aan de conditie 3 voldaan is (*bellen* is bruikbaar als overgankelijk werkwoord) gaat de ontleder opnieuw aan de slag met een NC-opdracht. We slaan nu een paar stappen over en gaan naar toestand  $nc_2$ . Hier zal de ontleder eerst VZC beproeven en springen naar  $S_{vzc}$ . *Op* is onder meer een voorzetsel, dus de overgang naar  $vzc_1$  wordt gemaakt. Omdat de zin nu af is mislukt de VZC-poging. De ontleder laat *op* nog even zitten, keert terug naar  $nc_2$  en neemt nu de overgang 5. Het resultaat,

$$(NC (L\ de) (N\ patiënt)),$$

wordt (zie conditie/operatie 3) geplaatst in (OBJ\_\_\_\_). Vanuit de huidige toestand,  $z_3$ , wordt conditie 4 getoetst. Het eerste te ontleden woord, *op*, vormt samen met *bellen* een scheidbaar samengesteld werkwoord zodat (W *bellen*)

wordt vervangen door (W *opbellen*). Operatie 5 zet nu de haakjesontleding

$$(Z (NC (L de) (ADJ aardige) (N dokter)) (VC (W opbellen) (NC (L de) (N patiënt))))$$

in elkaar, die er als zinsdiagram uit ziet als Fig. 4.

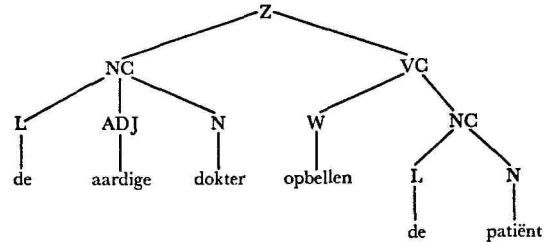


Fig. 4. Ontledingsresultaat van zin (9) m.b.v. de netwerkgrammatica in Fig. 3.

Deze gedetailleerde weergave van het functioneren van netwerkgrammatica's moge een aantal punten demonstreren. (a) De zinsanalyse verloopt, net als bij de mens, van links naar rechts, (b) De ontleder kiest bepaalde overgangen eerder dan andere en zal dus bij ambiguë zinnen slechts met één interpretatie voor de dag komen, net zoals mensen (maar zie ook Lackner & Garrett, 1972). (c) De operaties die bij sommige overgangen worden uitgevoerd hebben het karakter van "omgekeerde transformaties". Een voorbeeld was de samenvoeging van *bellen* met *op*. Als de zin passief was geweest dan zou de eerste NC aanvankelijk als "logisch" subjezt zijn aangemerkt (d.w.z. geplaatst in (SUBJ\_\_\_)) en pas na ontvangst van passief-signalen, bv. een passief hulpwerkwoord, zijn verhuisd naar (OBJ\_\_\_). Dit laatste vereist natuurlijk een uitbreiding van de grammatica van Fig. 3.

Deze paragraaf kort samengevat, in situaties die het conceptuele systeem goed kan overzien zal de menselijke taalgebruiker met een minimum aan syntactische analyse toekunnen; de behoefte hieraan wordt groter naarmate het conceptuele systeem er minder in slaagt verwachtingen op te bouwen aangaande de inhoud van de binnenkomende boodschap.

#### 1.4. De zinsgenerator.

Taalpsychologisch onderzoek naar het produceren van taaluitingen bestaat nauwelijks. Dit is des te betreurenswaardiger omdat bij het onderzoek naar

moedertaal- en vreemdetaalverwerving en naar pathologisch taalgedrag het spreken een centrale plaats inneemt.

Hoe moeten we ons het systeem voorstellen dat gedachte- en waarnemingsinhouden omzet in verstaanbare taaluitingen? Zou het transformatiemechanisme van de TGG model kunnen staan voor de zinsgenerator in de menselijke taalgebruiker? In de generatieve semantiek start het generatieproces met een zinsdiagram dat de uit te drukken inhoud weergeeft: de semantische representatie. Via transformatieregels wordt dit aanvankelijke zinsdiagram omgezet tot een zinsdiagram dat oppervlaktestructuur heet. De eindrij van dit laatste zinsdiagram bepaalt dan de uit te spreken zin.

Behalve (1) de strikte inhoud, d.w.z. de stand van zaken waarnaar de zin verwijst, zijn er nog andere factoren die de uiteindelijke vorm van de zin bepalen. In zijn dissertatie brengt van der Geest (1974) ze in drie groepen onder: (2) de attitude, het standpunt van de spreker m.b.t. de strikte inhoud, (3) de uiteindelijke bedoeling die de spreker met het uiten van de zin heeft, en (4) de linguïstische en situationele voorkennis die de spreker (naar hij aanneemt) met de luisteraar deelt. Van der Geest laat uitvoerig zien hoe deze factoren ingebouwd kunnen worden in de semantische representaties en hun invloed op de vormgeving van de zin doen gelden. (N.B. Onder de term “inhoud” van een zin begrijpen we verderop niet alleen punt (1), maar ook (2), (3), en (4).)

Niettemin kleven aan het transformatiemechanisme, indien het dienst zou moeten doen als zinsgenerator in een taalgebruikersmodel, twee hoofd-bezwaren.<sup>2</sup> Op de eerste plaats verloopt het generatieproces niet van links naar rechts: de laatste transformatie — die welke de oppervlaktestructuur als uitvoer geeft — kan de volgorde van de morfemen/woorden helemaal omgooien (bv. de transformatie “adjektiefverplaatsing” die de eindrij *een oud vrouwtje stierf* afleidt uit de volgorde *vrouwtje-oud-sterven*). Dit betekent dat de spreker zijn zin pas kan starten nadat het transformationele apparaat de oppervlaktestructuur-eindrij kant en klaar heeft afgeleverd. Hetgeen zou impliceren: veel pauzes tussen zinnen, binnen zinnen weinig of geen pauzes; weinig hernemingen en verbeteringen van half afgemaakte zinnen. De menselij-

2 Afgezien nog van het feit dat in alle versies van transformationele taalkunde het transformatiemechanisme non-deterministisch is. Dit bezwaar is gemakkelijk te ondervangen.

ke taalgebruiker daarentegen begint zijn taaluiting vaak, zo niet gewoonlijk, op een moment dat de woordkeus verderop in de zin nog niet vaststaat. De generator wordt dus op elk moment gestuurd door twee instanties: (a) de inhoud die de spreker wil uitdrukken, en (b) de syntactische beperkingen die het reeds geproduceerde zinsgedeelte oplegt aan het vervolg.

Ten tweede, de impliciete assumptie dat de zinsgenerator volledig onderworpen is aan de beslissingen die het conceptuele systeem neemt ten aanzien van de inhoud van de eerstvolgende zin, is onjuist. In vele gevallen moet aangenomen worden dat syntactische factoren de inhoud van de eerstvolgende zin medebepalen. Argumenten voor deze stelling zijn te ontlenuen aan zowel linguïstische observaties als aan experimenteel taalpsychologisch onderzoek. Hier geef ik er de voornaamste van; voor details zij verwezen naar Kempen (in voorbereiding).

De oudste observatie is ongetwijfeld van de denk- en taalpsycholoog Selz. Hij wees er al in 1922 op dat iemand die een definitie van een begrip wil geven uit een beperkt aantal syntactische schema's kan kiezen, zoals *een X is een Y die .....*; *een X is een .....* *Y*; *X bestaat hierin dat .....*. Selz draagt zelfs empirisch materiaal uit geheugenproeven aan: deze schema's bevorderen het vinden van een goede definitie voor moeilijke begrippen — een duidelijk geval van syntactische bepaaldheid van de inhoud van zinnen.

Er zijn grote verschillen tussen talen m.b.t. de genuanceerdheid van het getalsysteem. Het Japans maakt geen onderscheid tussen enkelvoud en meervoud zoals het Nederlands, Engels etc. Het Oud-Grieks kent bovendien een speciale dualisvorm en onderscheidt tussen de aantallen 1, 2, > 2. Het Arabisch gaat zelfs nog verder: 1, 2, 3-10, > 10.

Analoge contrasten vindt men op het gebied van beleefdheidsvormen (Ned.: *je/jij* vs. *U*; Eng.: alleen *you*; Japans: een uitgebreid stelsel), en werkwoordstijden (Ned.: onvoltooid en voltooid verleden tijd, tegenwoordige tijd, toekomstige tijd; Bahasa Indonesia: geen tijdsaanduidingen in het werkwoord). Wanneer men, zoals veel gebeurt, ervan uitgaat dat semantische representaties (dieptestructuren) van zinnen invariant zijn over talen, dan bevat een Japanse, Indonesische, Engelse zin informatie over resp. getal, tijd, en sociale relatie tussen spreker en toegesprokene — informatie die in het transformatieproces eenvoudigweg overboord gegooid wordt. Veel realistischer en efficiënter is de veronderstelling dat de syntaxis van de taal die men spreekt bepaalt welke semantische dimensies onderdeel gaan uitmaken van de “dieptestructuur” van de zinnen.

Een derde argument vloeit voort uit eigen onderzoek over het onthouden en reproduceren van zinnen (Kempen, 1974; Levelt & Kempen, 1974a). De gebruikte basistechniek heb ik “parafrastische reproductie” genoemd, d.w.z. de proefpersonen moesten zinnen die ze volledig uit het hoofd geleerd hadden, reproduceren in de vorm van een (partiële) parafrase. Bv. (12) is een volledige, (13) een partiële parafrase van (11).

- (11) ..... want die waakse honden beten fietsers.  
 (12) ..... omdat die waakse honden fietsers beten.  
 (13) ..... want die waakse honden beten.

Laat ik één experiment wat meer in detail bespreken. De ene helft van de proefpersonen had 12 zinnen van type (11), allemaal het voegwoord *want* bevattend, letterlijk van buiten geleerd, de andere helft 12 zinnen van type (12) (steeds met *omdat*). Daarna begon het eigenlijke experiment: meting van de tijd benodigd om na aanbieding van het onderwerp van één der zinnen zo snel mogelijk het predikaat van die zin hardop op te noemen. Deze reactietijd werd herhaaldelijk gemeten in vier verschillende condities:

- W<sub>2</sub>: het onderwerp werd voorafgegaan door *want*, bv. *want die waakse honden*, en de gevraagde responsie was het hele predikaat, bv. *beten fietsers*;  
 W<sub>1</sub>: idem, maar de gevraagde responsie was alleen het eerste woord van het predikaat, dus bv. *beten*;  
 O<sub>2</sub>: het onderwerp werd voorafgegaan door *omdat*, de responsie was het hele predikaat, dus bv. *fietsers beten*;  
 O<sub>1</sub>: idem, maar de gevraagde responsie was alleen het eerste woord van het predikaat, dus bv. *fietsers*.

De zinnen die de proefpersonen afmaakten in condities W<sub>2</sub>, O<sub>2</sub> en W<sub>1</sub> waren resp. van de typen (11), (12) en (13). In conditie O<sub>1</sub> bleven de zinnen syntactisch onvolledig (zie (14)).

- (14) ..... omdat die waakse honden fietsers.

De metingen onder elk der vier condities vonden bloksgewijs plaats<sup>3</sup>; bv. de O<sub>2</sub>-metingen over de 12 zinnen vonden vlak na elkaar plaats, dan kwamen alle W<sub>1</sub>-metingen etc. (Elke proefpersoon kreeg een andere volgorde van condities.)

<sup>3</sup> In contrast tot een eerder experiment (vermeld in Levelt & Kempen, 1974b) waar de metingen in de diverse condities door elkaar gehutseld waren.

De gemiddelde reaktietijden vertoonden het patroon  $W_1 < W_2$ ;  $O_1 > O_2$ . In een gangbare redenering (o.a. Johnson, 1970) zou men voorspeld hebben  $W_1 < W_2$ ;  $O_1 < O_2$ : een langere responsie (twee woorden) vergt meer zoektijd door het geheugen en meer responsie-planning dan een kortere responsie (één woord). Men zou ook de tegengestelde voorspelling kunnen opperen,  $W_1 > W_2$ ;  $O_1 > O_2$ . Tijdens het van buiten leren van de zinnen hebben de proefpersonen de twee woorden van elk predikaat tot één geheel verenigd en het kost extra moeite en tijd als die eenheid weer verbroken moet worden (in  $W_1$  en  $O_1$ ).

De gevonden uitkomst wordt echter voorspeld door een theorie die uitgaat van het groundbegrip "syntactische constructie". Een syntactische constructie definiëren we hier als een syntactisch grondscheema dat een samenstel van conceptuele relaties uitdrukt. Bv. het grondscheema  $N_1VN_2$  (naamwoord + werkwoord in bepaalde wijs + naamwoord) kan de conceptuele relatie actor-actie uitdrukken tussen  $N_1$  en  $V$ , en de relatie actie-object tussen  $V$  en  $N_2$ , zoals in *Jan sloeg Piet*. Hetzelfde syntactische grondscheema kan ook andere conceptuele relaties weergeven als  $N_2$  een tijdsduur aanduidt: *Jan sliep uren*. Deze voorbeelden maken duidelijk dat een syntactische constructie een combinatie is van een syntactisch grondscheema (dat men kan weergeven als een reeks van woordsoortaanduidingen zoals  $N$ ,  $VZ$ ,  $V$  etc., evt. ook met funktiewoordjes zoals *door* in de passief-constructie) en een samenstel van conceptuele relaties.

Ik neem nu aan dat de taalgebruiker voor elke frekwent voorkomende syntactische constructie beschikt over een programma dat twee dingen doet: (a) het zoekt in een vooraf aangeduid gedeelte van het conceptuele systeem naar een exemplaar van zijn karakteristieke samenstel van conceptuele relaties, en (b) het giet de gevonden informatie in de vorm van zijn karakteristieke syntactisch grondscheema. Het programma behorende bij het eerste voorbeeld van daarnet zoekt in een concrete conceptualisatie van de vorm actor-actie-object (met eventueel verdere casussen en bepalers) precies die informatie die zich laat uitdrukken in het syntactische patroon  $N_1VN_2$ .

Alvorens dit idee nader uit te werken wil ik even terugkomen op de uitkomst van het geheugenexperiment. Tijdens elk blok van reaktietijd-metingen (re)construeerden de proefpersonen steeds meerdere zinnen van dezelfde syntactische constructie. Bv. in conditie  $W_2$  vormden ze telkens zinnen van het type  $XN_1VN_2$  ( $X$  staat hier gemakshalve voor de woorden voorafgaande aan het onderwerpssubstantief). De proefpersonen waren hierdoor in de gelegenheid het "zoek/formuleer"-programma dat bij die constructie hoort, te mobi-

liseren en effectief te gebruiken: ze wisten dat dit programma feilloos de gevraagde sekwentie van responsiewoorden zou opleveren. In de condities  $W_1$  konden ze een eenvoudiger zoek/formuleer-programma  $XN_1V$  hanteren: alle werkwoorden waren onovergankelijk en de  $XN_1V$ -zinnen waren partiële parafrazen van de geleerde  $XN_1VN_2$ -zinnen. We mogen hier  $W_1 < W_2$  verwachten, want zowel de op te zoeken conceptuele informatie als de responsieplanning zijn eenvoudiger in het  $XN_1V$ -programma dan in het  $XN_1VN_2$ -programma.

Wat voorspelt deze theorie nu over  $O_1$  en  $O_2$ ? Omdat  $XN_1N_2$  (zie (14)) geen syntactische constructie is kent de taalgebruiker geen zoekformuleer-programma dat precies past op de eisen die de proefleider in de  $O_1$ -conditie stelt. De beste tactiek is het  $XN_1N_2V$ -programma toe te passen, dat responsies van het type *fietsers beten* oplevert, en deze responsies tot één woord in te korten. Dit laatste vergt een extra bewerking zodat de voorspelling luidt:  $O_1 > O_2$ ; hetgeen met de data overeenstemt.

De experimentele situatie is minder kunstmatig dan men misschien denkt. In het alledaagse spreken komen opsommingen veelvuldig voor, nl. wanneer men meerdere malen inhouden van hetzelfde type moet formuleren (bv. een reeks sportuitslagen). In dergelijke spreesituaties ligt de syntactische vorm van de zin grotendeels vast nog voordat de inhoud vaststaat.

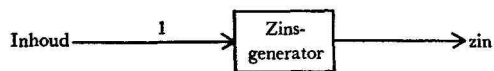
Drie argumenten heb ik de revue laten passeren die nopen tot de aanname dat bij het spontane spreken syntactische factoren soms de inhoud van de zin meebepalen. Waarmee ik overigens het bestaan van de "gewone" volgorde (conceptie van de inhoud éérst, daarna syntactische vormgeving) niet wil wegcijferen. Deze argumenten toevoegende aan de hogerop genoemde wenselijkheid van een generator die van links naar rechts werkt, kom ik tot de slotsom dat een transformatiemechanisme naar TGG-voorbeeld niet model kan staan voor de generator in de menselijke taalgebruiker.

Fig. 5 vat het contrast tussen TGG-achtige (5a) en taalpsychologisch gewenste (5b) zinsgeneratoren beknopt samen. In Fig. 5a gaat een eenmaal geselecteerde inhoud de generator in die er een grammaticale zin uit tracht te construeren. In Fig. 5b is alleen het globale thema vooraf geselecteerd, bv. het gespreksonderwerp; de gebeurtenis die ik nu waarneem; wat ik gistermiddag heb meegemaakt. De keuze die de inhouds-selector hieruit maakt voor een nieuw te formuleren zin (pijl no. 1) staat onder invloed van niet slechts inhoude-



lijke (pijl no. 2) maar ook syntaktische (pijl no. 3) criteria. Het eindproduct in Fig. 5a is een zin die “als één geheel” wordt afgeleverd; in Fig. 5b komt de zin “bij stukjes en beetjes” (van-links-naar-rechts) tot stand.

(a)



(b)

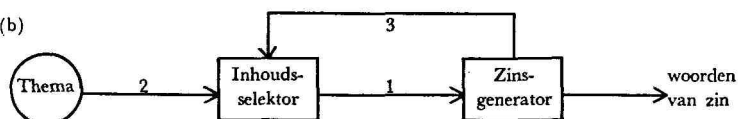


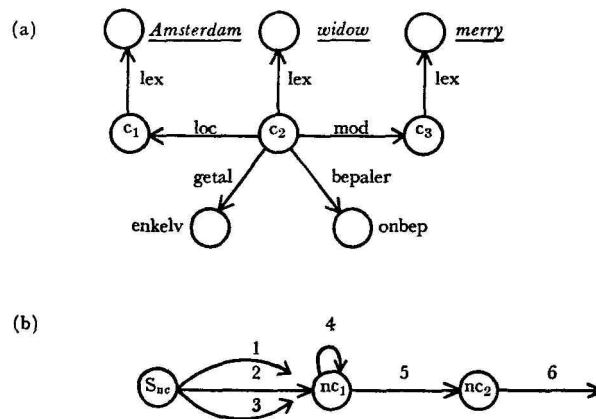
Fig. 5. Stroomdiagram van (a) een zinsgenerator, en (b) een zinsgenerator die aan de in tekst beschreven taalpsychologische eisen voldoet.

Om te illustreren hoe zinnen van links naar rechts gegenereerd kunnen worden vanuit een vooraf gegeven thema volgt nu een (sterk vereenvoudigd) onderdeel van een zinsgenerator die geprogrammeerd werd door Simmons & Slocum (1972). De grammatica die erin gebruikt wordt is een augmented transition network (zie par. 1.3 waar deze netwerken dienst deden als syntaktische ontleders).

Fig. 6b geeft een stukje netwerk weer dat naamwoordsgroepen genereert; Fig. 6a is de semantische representatie van de naamwoordsgroep *a merry widow in Amsterdam*. Ik zal laten zien hoe het genereerprogramma een deel van deze inhoud uitdrukt als *a merry widow*. De generator kan het semantische netwerk doorlopen; laten we aannemen dat hij zich bevindt bij knoop  $c_2$  op het moment dat hij in toestand  $S_{nc}$  verkeert.  $c_2$  is een concept dat luistert naar de naam *widow*: dit wordt uitgedrukt door de pijl  $c_2 \text{---} lex \text{---} \rightarrow widow$ . In toestand  $S_{nc}$  onderzoekt de generator de wijze van bepaaldheid van de knoop waarbij hij staat. Hiertoe zoekt hij naar een als “bepaler” gelabelde pijl die uit  $c_2$  vertrekt. Aan het eind van die pijl wordt de waarde “onbep(aald)” aangetroffen, zodat grammatica-weg 2 ingeslagen mag worden. De generator produceert nu het eerste woord van de naamwoordsconstituent: *a* (zie conditie/operatie 2). Vanuit toestand  $nc_a$  gaat de generator na of een adjektief geproduceerd kan worden. Dit vereist een wat ingewikkelder zoekproces want er komen geen “adj.”-pijlen in het semantische netwerk voor. Voorwaarde 4 stelt dat uitgaande van de knoop waarop de generator momenteel staat (X) een mod(ifier)-pijl doorlopen moet

worden en dat de dan bereikte knoop (Y, in dit geval  $c_3$ ) te boek staat als adjektief. Aan deze voorwaarden is voldaan, en *merry* is het tweede woordje dat wordt uitgesproken. De generator blijft nu in toestand  $nc_1$ . Om te voorkomen dat hij doorgaat met alsmear opnieuw *merry* te zeggen, zullen we afspreken dat een eenmaal verwoord stukje semantisch network wordt uitgeveegd.

De rest van de grammatica wijst zich vanzelf. Opgemerkt zij alleen nog dat de operaties die bij de diverse overgangen uitgevoerd moeten worden, vaak complexe functies zijn die uitgebreid beroep doen op lexicale informatie (bv. in overgang 5 moet opgezocht worden of het naamwoord misschien een onregelmatige meervoudsvorm heeft).



<i>Condities</i>	<i>Operaties</i>
1. X—bepaler—>bep	Zeg <i>the</i>
2. X—bepaler—>onbep en X—getal—>enkely	Zeg <i>a</i>
3. Onvoorwaardelijk	————
4. X—mod—>Y—lex—>Z, waarbij Z een adj. is	Zeg <i>Z</i>
5. X—lex—>Y, waarbij Y een substantief is	Zoek X—getal—>Z; zet, al naar gelang de waarde van Z, subst. Y in enkelvoudsvorm- of meervoudsvorm; spreek het geconstrueerde woord uit.
6. Onvoorwaardelijk	Keer terug.

Fig. 6. Een eenvoudige netwerkgrammatica voor het genereren van naamwoordsgroepen (b), en een voorbeeld van het type semantische representatie dat als invoer kan fungeren (a).

### 1.5. De articulator.

Het uitspreken van de woordreeksen die de zinsgenerator heeft samengesteld verloopt via een uiterst ingewikkeld proces waarvan ik slechts enkele aspecten wil belichten die betrekking hebben op de *sturing* van de articulatieorganen. Povel (1974) besteedt hieraan in zijn dissertatie veel aandacht, vooral met het oog op de wijze waarop men leert articuleren, mét of zónder gerichte training.

Empirische gegevens maken het, aldus Povel, waarschijnlijk dat het foneem de eenheid van spraakproductie is, d.w.z. dat fonemen de elementen zijn die de articulatoren achtereenvolgens trachten te realiseren (en niet lettergrepen of woorden). Vervolgens trekt Povel een zeer illustratieve analogie tussen articuleren en fluitspelen: zowel mond als fluit produceren hun diverse geluiden door de resonantiekenmerken van een holte te variëren. Op bv. een blokfluit gebeurt dit door met de vingers de gaten open en dicht te maken, bij het spreken door de stand van de articulatieorganen te veranderen.

Hoe stuurt de spreker zijn spraakmotoriek ter realisatie van een gewenste fonemenreeks? In beginsel zou de articulator voor elk te realiseren foneem een vast neurale commando kunnen doorseinen naar de spraakmusculatuur. In de fluit-analogie vertaald: elke toon wordt gemaakt m.b.v. een vaste beweging met de vingers. Dit zou juist zijn als de speler elke toon zou maken vanuit een vaste uitgangspositie van zijn vingers. Maar dit is evident niet het geval: om bv. van f naar g te gaan maakt de blokfluit een andere vingerbeweging dan wanneer hij een g speelt na een a. Dit brengt ons bij een tweede mogelijke stuurprincipe. Bij elk foneem hoort een vaste doelpositie van de articulatieorganen, en de neurale commando's nodig om deze positie te bereiken zijn verschillend voor elke uitgangspositie (voorafgaand foneem). De werkelijkheid is echter iets ingewikkelder want uit fonetisch onderzoek blijkt dat één en hetzelfde foneem kan worden geproduceerd m.b.v. meerdere configuraties der articulatoren (McNeilage, 1970) — net zoals op een blokfluit één noot vaak met meer dan één configuratie van open en gesloten gaten gemaakt kan worden.

Dit betekent dat de perceptuele invariantie van het foneem niet verklaard kan worden vanuit articulatorische invariantie. Omdat ook in de geproduceerde klanken zulk een invariantie niet te vinden blijkt, moet geconcludeerd worden dat het foneem een nogal abstracte entiteit is. Een soortgelijke situatie treffen we overigens aan op andere terreinen van patroonherkenning en -productie, niet

alleen binnen het taaldomein (o.a. zinsperceptie en -productie) maar ook daarbuiten. In totaal verschillende foto's herkennen we hetzelfde gezicht, één en hetzelfde figuurtje tekenen we met zeer uiteenlopende bewegingspatronen. Op elk van deze terreinen is waarschijnlijk een systeem nodig bestaande uit een ontleder en een generator dat de complexe relaties beregelt tussen de externe invoer- en uitvoersignalen en de interne eenheden/categorieën/concepten die het cognitieve systeem hanteert. (Een goed voorbeeld van ontleding in het visuele domein staat bekend als "scene analysis". Zie Sutherland, 1974.)

## 2. Toepassingen op moedertaal- en vreemdetaalverwerving

Deze sectie is gewijd aan het ontstaan van taalgebruikers.

### 2.1. Moedertaalverwerving: het leren van grammaticaregels.

De moedertaalverwerving kan beginnen zodra het conceptuele systeem concepten bevat voor de objecten, eigenschappen, handelingen, gebeurtenissen die het kind in zijn omgeving waarneemt. Dan kan het associaties gaan vormen tussen concepten en de spraaklabels die personen uit zijn omgeving gebruiken voor objecten, handelingen, etc. waarnaar de concepten verwijzen. Op grond van (aangeboren? vlg. Povel, 1974) kennis van de mogelijkheden der articulatieorganen weet het kind welke de doelposities zijn die het ter imitatie van de spraaklabels zal moeten realiseren. De imitatiepogingen zullen aanvankelijk weinig succesvol zijn omdat de fijne neurale sturing die nodig is om de achtereenvolgende doelposities vlot en vloeiend te bereiken nog ontbreekt.

Rond de tijd van zijn eerste verjaardag beschikt het kind niet alleen over "losse" concepten maar ook over conceptualisaties, d.w.z. netwerken waarin de afhankelijkheidsrelaties tussen meerdere concepten worden gespecificeerd (zie 1.1). We maken nu de grondassumptie dat het kind beschikt over een leermechanisme dat regelmatigheden ontdekt en registreert in paren bestaande uit (a) een conceptualisatie die het kind in een bepaalde situatie maakt, en (b) de reeks woorden die het kind herkent in een taaluiting die binnen dezelfde situatie tot hem gericht is. Een voorbeeld kan dit verduidelijken. Jessica heeft net een theekopje laten vallen en staat beduusd naar de scherven te kijken. Vader komt binnen en zegt: *Heb jij het kopje laten stukvallen?* Stel dat Jessica nog alleen de

woorden *kopje* en *stuk* met de bijbehorende begrippen heeft geassocieerd. Het kent dan in vaders uiting alleen de sekwentie *kopje-stuk* terug. Omdat op dat moment Jessica de conceptualisatie KOP<—attrib—>STUK maakt, heeft het leermechanisme gelegenheid te registreren dat in de uiting de attribuutnaam onmiddellijk volgt op de naam van het concept waaraan het attribuut wordt toegekend. Als Jessica meer van deze woordsekwentie–conceptualisatieparen tegenkomt kan haar leermechanisme tot een algemene regel besluiten. Deze regel gaat dan deel uitmaken van de syntaktische ontleder en de syntaktische generator. Als de generator later de conceptualisatie JESSICA—eig.—>POP<—attrib—>LIEF aangereikt krijgt (“Jessica’s pop is lief”), dan kan Jessica zeggen *Pop lief*.

Toegegeven, dit voorbeeld is zeer eenvoudig, maar ik zou willen poneren dat zulk een regel-leermechanisme, indien tot in details uitgewerkt, een plausibel taalleersysteem kan opleveren. (N.B. Het leggen van associatieve verbanden tussen spraaklabels en concepten (zie hogerop) is niet het werk van een apart mechanisme maar kan worden opgevat als de eerste manifestatie van het regel-leermechanisme).

Zo’n systeem heeft een aantal belangrijke theoretische consequenties.

(1) Kinderen spreken in geen enkele fase van hun taalverwerving in “dieptestructuren” of semantische representaties. Met name McNeil (1970) heeft verdedigd dat vroege taaluitingen van kinderen vrij directe weerspiegelingen zouden zijn van dieptestructuren. Sinds de orthodoxe Chomskiaanse dieptestructuur goeddeels verlaten is en, met name in de generatieve semantiek, vervangen door het begrip “semantische representatie” zou men in het voetspoor van McNeill kunnen opperen dat vroege uitingen heel dicht staan bij semantische representaties. Voorspellingen van dit type zijn op geen enkele wijze afleidbaar uit het hier getekende taalgebruikersmodel. Ons model kent geen semantische representaties maar conceptualisaties en “spreken in conceptualisaties” is een zinloze gedachte: (a) over de concepten die in conceptualisaties optreden is geen volgorde van links naar rechts gedefinieerd (wel over de eindknopen in semantische representaties à la de generatieve semantiek) ; (b) de concepten zijn met opzet taalvrij (niet-talig) gekozen en hun relaties met de woorden van het lexicon zijn uiterst complex. Ook de vroege taaluitingen, hoe eenvoudig van structuur ze ook mogen lijken, worden gefabriceerd door de zinsgenerator.

(2) Van der Geest snijdt in zijn dissertatie (1974) het probleem aan hoe kinderen, nog slechts over een eenvoudige grammatica beschikkend, in staat

zijn uitingen die volwassenen tot hen richten, te verstaan. Deze uitingen stijgen immers qua syntaktische complexiteit vaak uit boven het niveau van het kind. Dat dit laatste juist is wordt bevestigd door een fraai onderzoek van van der Geest, Snow en Drewes-Drubbels (1974) naar de wijze waarop moeder en kind met elkaar converseren. Zij vonden dat de uitingen van de moeders in syntaktisch opzicht complexer waren dan de uitingen van het kind. Maar een tweede uitkomst is minstens even interessant. De kinder-uitingen waren in semantisch-cognitief (ik zou zeggen: conceptueel) opzicht meer geavanceerd dan de moeder-uitingen. Het lijkt erop, aldus de auteurs, of het kind aan de moeder de inhouden aanreikt waarover ze zou kunnen praten en dat de moeder dan geschikte syntaktische middelen verschaft om deze inhouden uit te drukken.

Als we nu, gewapend met deze data, teruggrijpen naar hetgeen in par. 1.3 gezegd is over zinsontleding, dan blijkt dat de moeder-kind-conversatie een voorbeeld bij uitstek is van een situatie waarin syntaktische ontleding tot een minimum beperkt kan blijven. Syntaktische niveauverschillen tussen moeder en kind hoeven dan geen storing van de communicatie in de hand te werken. Voorts, omdat het kind beschikt over de conceptualisaties die tijdens de conversatie gewoonlijk aan de orde komen vormt dit type gesprekken voor het regel-leermechanisme een ideale vindplaats van conceptueel-syntaktische correspondenties.

Dit is m.i. een afdoende oplossing van het probleem dat in dit punt ter sprake is. De oplossing die van der Geest voorstelt en waarin hij een beroep doet op de beperkte onmiddellijk-geheugen-capaciteit van het kind acht ik dan ook overbodig. Hierover meer in punt (4).

(3) In de literatuur zijn verschillende mogelijkheden geopperd ter verklaring van de discrepantie tussen de relatief rijke conceptuele (semantische, diepte-) structuur van kindertaaluitingen en de eenvoud van hun syntaktische (oppervlakte-) structuur. Bloom (1970) stelt voor dat tot de aanvankelijke grammatica een aantal reductietransformaties behoren die later echter verdwijnen. Dit heeft iets tegen-intuïtiefs: binnen de TGG-stroming wordt taalontwikkeling gezien als het bijleren van transformaties, niet als het afleren ervan. Voor verdere kritiek zie Schaerlaekens (1973). Brown & Fraser (1963) en van der Geest (1974) leggen de oorzaak van de syntaktische eenvoud van kindertaaluitingen in een meer perifeer mechanisme, het onmiddellijk geheugen. Ze gaan ervan uit dat het transformatiemechanisme dit geheugen nodig heeft om tussenresultaten op te bergen. Als dit klein is zullen bij het uitdrukken van

complexe inhouden stukken moeten wegvallen (de minst beklemtoonde stukken het eerst, aldus van der Geest). In punt (4) wordt dit idee geëvalueerd.

Een derde verklaringmogelijkheid wordt naar voren gebracht door Schaerlaekens in haar “semantische-relatiemodel”. Sprekend over tweewoordzinnen oppert zij de gedachte dat semantische relaties een vast syntactisch uitdrukingspatroon bij zich hebben. Bv. een bezitsrelatie (“fixed allocation”) wordt weergegeven door de naam van het bezit onmiddellijk te laten volgen door de naam van de bezitter. Zo bezigde een van de door haar onderzochte kinderen de uiting *bed Joost* voor *dit is het bed van Joost*. Ter aanduiding van de plaats waar iemand zich bevindt hanteerde ditzelfde kind de omgekeerde volgorde (*Joost bed* voor *Joost is in bed*).

Dit laatste voorstel past het beste in de zinsgenerator van ons taalgebruikersmodel omdat het geen beroep doet op een transformatiemechanisme. Integendeel, de bedoelde uitdrukingspatronen vertonen verwantschap met de in par. 1.4 ontwikkelde zoek/formuleer-programma’s die eveneens een vaste combinatie vormen van (een samenstel van) conceptuele relaties en een syntactisch constructieschema. Concretisering en formalisering van Schaerlaekens’ voorstel in de vorm van netwerkgrammatica’s liggen dan ook voor de hand.

Een voorbeeld moge verduidelijken hoe m.b.v. netwerkgrammatica’s het aanvankelijke contrast tussen conceptuele rijkdom en syntactische eenvoud van kindertaaluitingen verklaard kan worden en hoe het geleidelijk aan wordt opgeheven. In de fase van de tweewoordzinnen is het kind in staat, nemen we aan, tot conceptualisaties waarin behalve de actie ook een actor en een objekt figureren. De syntactische netwerken die het kind tot op dat moment heeft opgebouwd m.b.v. zijn regel-leer-mechanisme, zijn echter nog beperkt. Bv. het beschikbare syntactische netwerk zoekt in conceptualisaties van de vorm *actor* <—> *actie* <—> *objekt* alleen actor- en objektconcepten op en plaatst hun lexicale benamingen achter elkaar in de volgorde actor-objekt. In een latere fase van het taalleerproces zal het syntactische netwerk uitgebreid zijn en ook voor het actie-concept een plaats in de taaluiting inruimen. Merk op dat in deze beschrijving van syntactische ontwikkeling alleen sprake is van uitbreiding van het regelsysteem en dat geen beroep nodig is op “extra” verklaringsmiddelen zoals een afnemend aantal reductietransformaties of een groeiend onmiddellijk geheugen.

(4) Mijn terughoudendheid ten aanzien van de verklaringen in termen van groei van het onmiddellijk geheugen (short-term memory, STM) is geba-

seerd op de uitkomsten van recent onderzoek. Deze suggereren dat de capaciteit van het volwassen STM niet, zoals gewoonlijk wordt aangenomen, in de buurt van 7-8 items ligt maar slechts ongeveer 3 items bedraagt. Zie voor details het overzicht van de experimentele literatuur door Thomassen & Kempen (1974). Met items bedoel ik goed geïntegreerde eenheden ("chunks") zoals cijfers, letters, woorden, korte zegswijzen e.d. De *geheugenspanne* die een normale volwassene weet te behalen is weliswaar 7-8 items, maar dit getal weerspiegelt niet rechtstreeks de STM-capaciteit. Ofschoon hierover het laatste woord nog niet gezegd is wordt steeds waarschijnlijker dat de geheugenspanne het resultaat is van toepassingen van allerlei mnemotechnische hulpmiddelen. Eén ervan is "hercodering": gebruikmakend van eerder verworven kennis slaagt men erin meerdere items tot een hecht geïntegreerde groep te verenigen die slechts één van de  $\pm 3$  STM-plaatsen bezet. Bv. het 6-stellige getal 492536 kan m.b.v. worteltrekken worden teruggebracht tot het 3-stellige getal 756. Een ander, machtig, middel tot hercodering is ... gebruikmaking van syntaktische structuur. Bv. de nonsens-reeks *twier bradi spiramen droor slend zwiets* lijkt veel moeilijker te onthouden dan de eerste twee regels van Carroll's Wauwelwok (par. 1.3).

Hier dreigt dus het gevaar van verwisseling van oorzaak en gevolg: misschien is de groei van de syntaktische vaardigheid van het kind niet het *gevolg* van (o.m.) STM-groei, zoals van der Geest en anderen aannemen, maar juist oorzaak (althans één der oorzaken) ervan. Deze mogelijkheid, die in het licht van de recente geheugenpsychologie heel plausibel is, maakt verklaringen zonder STM gewenst. In punten (2) en (3) heb ik gepoogd hiervoor de grondslag te leggen.

## 2.2. Vreemdetaalverwerving.

In de slotparagraaf wil ik uit het taalgebruikersmodel een drietal gevolgtrekkingen voor vreemdetaalverwerving afleiden.

*De hypothese van de contrastieve analyse* stelt dat contrasterende structuren (structuren van de vreemde taal die in de moedertaal geen parallel hebben) de leerling voor grotere problemen zullen plaatsen dan parallelle structuren. Uit wat in par. 1.2 en 2.1 gezegd is over syntaktische ontleding valt te voorspellen dat deze hypothese niet zal gelden voor het verstaan (luisteren, lezen) van de



vreemde taal, tenminste indien het conceptuele systeem beschikt over de concepten waarnaar wordt verwezen door de afzonderlijke woorden van de tekst. Het conceptuele systeem kan dan het ontleedproces in belangrijke mate sturen zodat het er nog weinig toe doet of de lezer/luisteraar de contrasterende structuur al dan niet beheerst. Ulijn's onderzoek (1974), dat bij mijn weten het enige is waarin de contrastieve hypothese getoetst is aan tekstbegrip, levert gegevens op die met onze voorspelling kloppen.

*Fouten in de uitspraak* van fonemen, woorden, zinnen in de vreemde taal worden gewoonlijk — zonder nadere motivering — toegeschreven aan een onvermogen van de articulatieorganen om de vereiste standen en bewegingen uit te voeren. Ofschoon ik dit niet in zijn algemeenheid wil ontkennen, wijs ik, mede naar aanleiding van Povel's studie (1974), op de mogelijkheid dat uitspraakfouten voortkomen uit een perceptieprobleem. In par. 1.2 is gesteld dat spraakverstaan vaak mogelijk is zonder een nauwkeurige, foutloze analyse van het spraaksignaal omdat syntactische en conceptuele informatie de tekortkomingen van de spraakherkenner kunnen compenseren. Om het gechargeerd uit te drukken: de (volwassen) taalgebruiker is een goede verstaander maar een slechte luisteraar. Uitspraakfouten zijn dan een onvermogen om de karakteristieke eigenschappen van de vreemdetaalklanken op te pikken. Kinderen, die minder dan volwassenen kunnen vertrouwen op syntactische en conceptuele kennis, zijn ook bij het verstaan van *moedertaalzinnen* sterker aangewezen op pure klankinformatie. Ze zullen dan voor het klankaspect van de vreemde taal een scherper oor hebben.

*Zich uitdrukken in de vreemde taal* vergt van de syntactische generator méér dan alleen het aanleren van nieuwe constructieschema's. Van minstens even groot belang zijn "koppelingen" tussen conceptualisaties en constructieschema's — koppelingen die misschien de vorm kunnen aannemen, van programma's à la Fig. 6b. Het aanleren van zulke programma's vereist minstens twee leerstadia: de syntactische schema's zelf moeten ingeoeffend worden, en ze moeten aan conceptualisaties geassocieerd worden. Structuuroefeningen (pattern drills) lijken voor het eerste stadium een goede leervorm. Het tweede stadium eist dat de leerling geconfronteerd wordt met telkens wisselende conceptuele inhoud met als taak om uit zijn arsenaal van syntactische schema's snel het best passende te kiezen. Om te vermijden dat de leerling hierbij moet vertalen is het noodzakelijk die conceptualisatie op een niet-talige manier aan te bieden. Wil men dit leerproces in drill-formaat gieten, dan is men aangewezen op "situationele" of "communicatieve" drills (vgl. het overzicht van Extra, 1973). Een goed voorbeeld hiervan treft men aan in Lipson (1971).

## REFERENTIES

- Bloom, L., *Language development, form and function in emerging grammars*. Cambridge, Mass., 1970.
- Brown, R., & Fraser, C., The acquisition of syntax. In: C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York, 1963.
- Campbell, H. W., *Phoneme recognition by ear and by eye*. Dissertatie, Nijmegen, 1974.
- Carroll, Lewis, *Through the looking glass*. Nederlandse vertaling door C. Reedijk & A. Kossman: *Alice in wonderland en in spiegelland*. Rotterdam: Donkers, 1972.
- Extra, G., *Taalverwerving en vreemdetalenonderwijs*. Groningen: Wolters, 1973.
- Geest, A. J. M. van der, *Some aspects of communicative competence and their implications for language acquisition*. Dissertatie, Amsterdam, 1974.
- Geest, T. van der, Snow, C., & Drewes-Drubbels, A., Developmental aspects of mother-child conversation. Paper, Amsterdam, 1974.
- Johnson, N. F., The role of chunking and organization in the process of recall. In: G. H. Bower (ed.), *The psychology of learning and motivation, Vol. 4*. New York: Academic Press, 1970.
- Kaplan, R. M., Augmented transition networks as psychological models of sentence comprehension. *Artificial Intelligence*, 1972, 3, 77-100.
- Kempen, G., Syntactic constructions as retrieval plans. Rapport 74 FU 05, Psychologisch Laboratorium, Nijmegen, 1974. [In druk, *British Journal of Psychology*.]
- Kempen, G., Formuleren. Rapport in voorbereiding.
- Lackner, J. R., & Garrett, M., Resolving ambiguity: effects of biasing context in the unattended ear. *Cognition*, 1972, 1, 359-372.
- Levelt, W. J. M., & Kempen, G., Semantic and syntactic aspects of remembering sentences. In: Kennedy, R. A. (Ed.), *Studies in long term memory*. New York: Wiley, 1974.
- Levelt, W. J. M., & Kempen, G., Algemene taalpsychologie. In: J. A. Michon, L. F. W. de Klerk & E. G. Eijkman (Eds.), *Handboek der psychonomie*. Deventer: Van Loghum Slaterus, 1975 (rapport 74 FU 01, Psychologisch Laboratorium, Nijmegen, 1974).
- Lipson, A., Some new strategies for teaching oral skills. In: Lugton, R. C., & Heinle, C. H., *Towards a cognitive approach to second language learning*. Philadelphia: The Center for Curriculum Development, 1971.
- MacNeilage, P. F., Motor control of serial ordering of speech. *Psychological Review*, 1970, 77, 182-196.
- McNeill, D., *The acquisition of language*. New York: Harper, 1970.
- Newell, A. et al., *Speech understanding systems*. Amsterdam: North-Holland, 1973.
- Povel, D. J. L., *Articulation correction of the deaf by means of visually displayed acoustic information*. Dissertatie, Nijmegen, 1974.
- Schaerlaekens, A. M., *The two-word sentence in child language development*. Den Haag: Mouton, 1973.
- Schank, R., Conceptual dependency: a theory of natural language understanding. *Cognitive Psychology*, 1972, 3, 552-631.

- Schank, R., Identification of conceptualisations underlying natural language. In: Schank, R., & Colby, K., *Computer models of thought and language*. San Francisco: Freeman, 1973.
- Selz, O., *Über die Gesetze des geordneten Denkverlaufs, Zweiter Teil*. Bonn: Cohen, 1922.
- Simmons, R. F., & Slocum, J., Generating English from semantic networks. *Communications of the ACM*, 1972, 15, 891-905.
- Sutherland, N. S., Intelligent picture processing. In: Sutherland, N. S. (Ed.), *Tutorial essays in psychology, Vol.1*. Washington: Erlbaum, 1974 (in druk).
- Thomassen, A., & Kempen, G., De experimentele psychologie van het geheugen. In: Michon, J. A., Klerk, L. F. W. de, & Eijkman, E. G. (Eds.), *Handboek der psychonomie*. Deventer: Van Loghum Slaterus, 1975 (rapport 74 FU 06, Psychologisch Laboratorium, KU Nijmegen, 1974).
- Ulijn, J. M., *Contrastieve analyse en leesvaardigheid in een vreemde taal*. Voordracht ANELA-Studiedag, Maart 1974.
- Woods, W. A., Transition network grammars for natural language analysis. *Communications of the ACM*, 1970, 13, 591-606.