

Het lineariseringsprobleem van de spreker*

Willem J. M. Levelt

Traditioneel wordt het proces van spreken opgevat als het omzetten van gedachten (intenties, gevoelens, etc.) in natuurlijke taaluitingen. Eén eis waaraan deze omzetting moet voldoen is dat er een strikte rangorde wordt opgelegd aan de uit te drukken informatie. Gesproken taal sluit goedgevoelens gelijktijdige expressie van meer proposities uit zodat de spreker een z.g. „lineariseringsprobleem” heeft: er moet een volgorde worden bepaald over een te formuleren kennisstructuur. Dit kan betrekkelijk eenvoudig zijn wanneer een kennisstructuur zelf lineair geordend is, zoals vaak het geval is voor temporele structuren, maar het vereist speciale procedures wanneer de kennisstructuur complexer is, zoals bij twee- en driedimensionale ruimtelijke structuren. Hoe moet een spreker bijvoorbeeld beschrijven hoe zijn woning in elkaar zit, of zijn stad?

Er liggen twee belangrijke typen beperking op lineariseringsbeslissingen van de spreker. De ene stamt uit wat „gedeelde kennis” kan worden genoemd, de andere uit eigenschappen van het menselijk werkgeheugen. Gedeelde kennis speelt een rol in zoverre de spreker kan verwachten dat de hoorder verschillende interpretaties zal toekennen aan verschillende volgorden van expressie (vgl. ze trouwde en werd zwanger met ze werd zwanger en trouwde). Determinanten van dit type zijn pragmatisch en cultureel van aard, en afhankelijk van het onderwerp van gesprek. Beperkingen van het werkgeheugen kunnen van invloed zijn op de linearisering omdat de spreker zal proberen de geheugenlast zo klein mogelijk te houden tijdens het proces van formuleren. Een meerdimensionale kennisstructuur zal bij voorkeur zó worden doorlopen dat het aantal te onthouden „terugkeeradressen” wordt geminimaliseerd. Dit gebeurt onder andere

* Dit is een Nederlandse bewerking van „The speaker's linearization problem’, verschenen in de *Philosophical Transactions of the Royal Society London*, B 295, 305-315, 1981.

door de connectiviteit van het gesprokene te maximaliseren, en door onthouden terugkeeradressen in omgekeerde volgorde af te werken. Omdat deze beperkingen van het werkgeheugen zeer algemeen en biologisch bepaald zijn, zullen de effecten ervan voor het lineariseringsproces betrekkelijk onafhankelijk zijn van de inhoud van het gesprokene.

Ten slotte wordt besproken of de eis tot linearisering slechts voortkomt uit de orale modaliteit van spraak, of dat we te doen hebben met een diepere modaliteits-onafhankelijke eigenschap van taalgebruik.

Inleiding

Telkens wanneer een spreker iets wil uitdrukken dat meer is dan een eenvoudige bewering, vraag, enz., moet hij of zij oplossen wat ik het „lineariseringsprobleem” heb genoemd: de spreker moet beslissen wat er eerst gezegd moet worden, wat er dan gezegd moet worden, en zo vervolgens. Ofschoon het lineariseringsprobleem het onderwerp is geweest van pedagogische en filosofische verhandelingen van vóór Aristoteles tot de huidige dag, wordt het geheel veronachtzaamd in de moderne psycholinguïstische literatuur.

De zestiende-eeuwse humanist-filosoof Pierre de la Ramée (1555, 122) demonstreerde het lineariseringsprobleem aan het voorbeeld van een leraar die een cursus grammatica gaat geven: „Posons que toutes les définitions, distributions, reigles de Grammaire soyent trouvées et chacune soyent jugée véritablement, et que tous ces enseignemens soyent escriptz en diverses tablettes lesquelles soyent toutes ensembles pesle mesle tournées et brouillées en quelque cruche, comme au jeu de la blanque. Icy je demande quelle partie de Dialectique me pourroit enseigner de disposer ces préceptes ainsi confus et les réduire en ordre.”

Ramus maakt zorgvuldig onderscheid tussen dit ordeningsprobleem en twee andere aspecten van (gesproken) tekstgeneratie. De eerste is het uit het geheugen ophalen en beoordelen van de uit te drukken zaken. Ramus zegt het zó (122-123): „Premièrement ne sera besoin des lieux d'invention car tout est déjà trouvé: chacune énonciation particulière est prouvée et jugée. Il ne faudra ny premier jugement de l'énonciation ny deuziesme du syllogisme.” Met andere woorden, het aanbrenge van volgorde veronderstelt volgens Ramus reeds het ter beschikking hebben van geverifieerde informatie. Anderzijds moet het ordeningsproces worden onderscheiden van het formuleerproces zelf, de keuze van linguïstische middelen voor de expressie van die gedachten-volgorde. Op dit punt breekt Ramus met de traditionele behandeling van linearisering. Sinds de klassieke retorische teksten werd linearisering behandeld in het hoofdstuk over *dispositio*.¹ Maar de rhetorica gaat over *oratio*, het vermogen tot spreken. Daartegenover benadrukt Ramus herhaaldelijk dat zowel oordelen als op volgorde brengen onafhankelijk zijn van dat vermogen, en

exclusief behoren tot het domein van de *ratio* (zie Nuchelmans 1980, 17). Het lineariseringsprobleem moet dus niet als een taalprobleem, maar als een denkprobleem worden behandeld, en Ramus deelt het gevolgelijk, net als de logica, in bij de *dialectica* in plaats van bij de *rhetorica*.

In dit opzicht ben ik een ramist: wanneer men het produceren van gesproken tekst onderzoekt dient men zorgvuldig te onderscheiden tussen de selectie van uit te drukken informatie, het op volgorde brengen van die informatie, en de linguïstische formulering ervan. In een behandeling van het lineariseringsprobleem moet, meer in het bijzonder, worden vermeden het probleem te stellen in termen van het ordenen van zinnen, constituenten, of andere linguïstische eenheden. Het gaat juist om elementen in de „taal van het denken”, en die kunnen we helaas nog aanzienlijk minder goed definiëren dan linguïstische elementen (zie bijvoorbeeld Fodor 1975). In de terminologie die Ramus uiteindelijk gebruikte waren deze eenheden *argumenten* – ongeveer equivalent met concepten – en *axioma's*, die op hun beurt tot beweringen samengevoegde argumenten waren.

Deze strikte scheiding van taal en denken houdt natuurlijk geenszins in dat men interactie tussen de twee systemen zou ontkennen. De grammatica van een taal kan bepaalde grensvoorwaarden stellen aan de volgorde waarin gedachten uitdrukbaar zijn; omgekeerd zal een eenmaal gekozen volgorde een zekere beperking kunnen inhouden ten aanzien van welke grammaticale constructies geschikt zijn voor de uitdrukking van die informatie. Niettemin kunnen de twee systemen wezenlijk verschillen in hun belangrijkste organisatieprincipes.

Welke principes liggen er ten grondslag aan de linearisering die sprekers aanbrenge(n)? Zoals vermeld is er in de moderne psycholinguïstische literatuur haast niets over dit onderwerp te vinden, ofschoon het een centrale plaats zou dienen in te nemen in theorieën over de sprekers. In het volgende zal ik enkele lineariseringsprincipes schetsen die volgens mij een belangrijke rol spelen. Ze zijn, deels, nog van speculatieve aard, maar voor een ander deel zijn ze gebaseerd op empirisch onderzoek. Ik zal beginnen met het speculatieve gedeelte dat betrekking heeft op principes die samenhangen met de inhoud van het gesprokene. Daarna stappen we over naar het meer empirische gedeelte dat betrekking heeft op principes die voortkomen uit algemene beperkingen van de aandacht in het proces van spreken. Een afsluitende paragraaf gaat over de relatie tussen linearisering en de orale modaliteit van het spreken.

Inhoudsgebonden principes

De kunst der rhetorica is ontwikkeld om hoorders effectief te beïnvloeden of te overtuigen. Elk onderdeel van die kunst, of het nu inventie, elocutie of uitspraak is, is op dat doel gericht. Linearisering heeft eveneens als hoofddoel om

bepaalde inferenties op te wekken in de geest van de luisteraar. Volgorde is een van de determinanten van zulke inferenties („implicatures”), zoals Lyons (1977), Gazdar (1979, 1980) en anderen hebben duidelijk gemaakt. Een voorbeeld van Kempson (1975) kan dit voldoende demonstreren. Als de context niet al te ver gezocht is, klinkt de volgende tekst raar:

(1) The Lone Ranger rode off into the sunset. He mounted his horse.

Volgorde van benoeming impliceert voor de hoorder blijkbaar volgorde van gebeurtenissen, en er is dus een pragmatische beperking die een volgorde als in (2) vereist:

(2) The Lone Ranger mounted his horse. He rode off into the sunset.

Het beproefde retorische principe dat hieraan ten grondslag ligt staat bekend als *ordo naturalis*.² Natuurlijke volgorde, kan men zeggen, is de ongemarkeerde wijze van linearisering. Alle andere volgordes, de zogenaamde artificiële, zijn er om bepaalde aandachts- of esthetische effecten te bewerkstelligen. Een psycholinguïstische beschouwing over linearisering kan het beste beginnen bij het ongemarkeerde geval.

Daarbij dienen drie vragen aan de orde te komen. De eerste is waarom natuurlijke volgorde voordeel kan hebben voor spreker en luisteraar. De tweede is hoe men natuurlijke volgorde kan definiëren voor verschillende onderwerpen van gesprek. En de derde vraag is wat er gebeurt wanneer zo'n onderwerp geen unieke natuurlijke volgorde oplevert. Laten we deze drie vragen achtereenvolgens kort in beschouwing nemen.

Het principe van natuurlijke volgorde heeft twee verwante, maar toch verschillende psychologische bronnen. Beide worden uitvoerig besproken in Osgood's recente boek (1980), waarin een interessant hoofdstuk voorkomt over „naturalness”. Een eerste bron is de organisatie van de spreker's eigen – prelinguïstische – ervaring. Een complexe gebeurtenis, bijvoorbeeld, wordt in onze cultuur bij voorkeur zó onthouden dat opvolgende gebeurtenissen directer met elkaar worden geassocieerd dan niet-opvolgende gebeurtenissen. Wanneer deze informatie voor expressie moet worden opgehaald uit het geheugen ligt het het meest voor de hand om elke zojuist opgehaalde gebeurtenis te laten werken als „retrieval cue” voor de erop volgende gebeurtenis. Experimentele evidentie voor deze veronderstelling is te vinden in Mandler en Johnson (1977). We komen op dit gebruik van connectiviteit weer terug in de volgende paragraaf, die betrekking heeft op procesdeterminanten van linearisering.

Wanneer de spreker de hoorder tot bepaalde gevolgtrekkingen wil brengen op grond van de volgorde waarin de dingen gezegd worden, moet hij rekening houden met wat – onder die omstandigheden – kan doorgaan voor gedeelde

kennis. In onze cultuur bestaat er gedeelde kennis ten aanzien van algemene regels of wetmatigheden, zoals bijvoorbeeld dat oorzaken voorafgaan aan gevolgen, dat intenties kunnen leiden tot handelingen, dat planning voorafgaat aan uitvoering, enz. Door de ongemarkeerde volgorde te gebruiken kan de spreker zeker stellen dat de hoorder zulke algemene kennis correct zal toepassen en de beoogde gevolgtrekking zal maken. Bovendien zijn er in onze cultuur meer specifieke stereotiepe volgorderelaties tussen gebeurtenissen zoals het eerst-eten-dan-betalen in (niet-zelfbedienings)restaurants (zie Schank en Abelson 1977), stereotiepe ruimtelijke relaties zoals die tussen vertrekken in een woning (zie Linde en Labov 1975) of tussen meubels in een woonkamer (zie Ullmer-Ehrich 1979 en 1982). Van zulke stereotypen kan de spreker aannemen dat hij die kennis met de hoorder deelt. Door zulke voorgegeven schema's „af te lopen" in de volgorde van uitdrukking kan de spreker het de hoorder gemakkelijk maken. De interesses van spreker en luisteraar kunnen hier samenvallen: wanneer het ophalen van de kennis bij de spreker wordt gestuurd door wat stereotiep is in de cultuur, dan kan de hoorder hetzelfde stereotiep gebruiken om de informatie te interpreteren en op te slaan. De vraag wat natuurlijke volgorde zo natuurlijk maakt heeft dus kortweg als antwoord dat het de spreker helpt kennis beschikbaar te maken voor expressie, en de hoorder helpt gevolgtrekkingen te maken, wanneer beiden zich kunnen baseren op algemene of meer specifieke gedeelde kennis in de taalgemeenschap. Het kan overigens niet genoeg benadrukt worden dat wat voor natuurlijke volgorde doorgaat cultuur-afhankelijk is. Becker (1980) laat bijvoorbeeld zien dat aristotelisch temporeel-causaal verloop van gebeurtenissen geen basis vormt voor linearisering in Javaans Wayang; de volgorde van expressie wordt daar veeleer bepaald door iets wat hij „coïncidentie" noemt en wat blijkbaar een diep handelingsmotief is in de Wayang-cultuur.

De tweede vraag, wat de natuurlijke volgordes zijn voor verschillende gespreksonderwerpen, kan nooit volledig worden beantwoord. Er zijn net zoveel natuurlijke volgordes als er dingen zijn waarover men kan praten, en voor dat men het weet belandt men in het moeras van de inhoudsanalyse. Een enkele meer algemene observatie kan echter wel gemaakt worden: wanneer de te formuleren kennisstructuur zelf lineair geordend is dan is de ongemarkeerde of natuurlijke manier om die uit te drukken het aflopen van die ordening. Voor niet-simultane gebeurtenisstructuren komt de ongemarkeerde linearisering overeen met de temporele volgorde van de gebeurtenissen. Voor lineaire ruimtelijke structuren geldt dat de spatiële connectiviteit moet worden gevolgd. Dat is met name het geval bij het wijzen van de weg – onderzocht door Klein (1979 en 1982) en anderen in ons instituut, en door Munro (1977) en Wunderlich en Reinelt (1982) –. De taak om iemand de weg te wijzen houdt ten eerste in het zoeken van een kortste of snelste weg door een interne representatie van een meerdimensionale configuratie. Ten tweede moet die kortste of snelste route worden geformuleerd. Sprekers doen dat altijd door de to-

pologische connectiviteit van de route te volgen, ofschoon dat geenszins de enige logische mogelijkheid is.

Soms zijn kennisstructuren gedeeltelijk georganiseerd in termen van subordinatie. De ongemarkeerde linearisering volgt dan een weg van bovengeordende naar ondergeordende items. Dat is, bijvoorbeeld, het geval in een computerprogramma ontworpen door Collins e.a. (1975) dat aardrijkskunde onderricht.

De auteurs beweren dat zij deze super-naar-subordinatie-volgorde van behandeling in het programma hebben geïmplementeerd op grond van systematische observaties die tijdens het gewone aardrijkskundeonderwijs gemaakt werden. Het klinkt als een verre echo van de genus-naar-species-ordering die in retorische teksten uit alle tijden werd aanbevolen.

De derde vraag is wat te doen wanneer er geen unieke natuurlijke volgorde is voor een bepaalde kennisstructuur. Dat kan het geval zijn voor temporele structuren, nl. wanneer er sprake is van simultane gebeurtenissen. Het is het normale geval voor twee- of driedimensionale ruimtelijke structuren, en voor talloze andere kennisstructuren. De enige empirische evidentie over wat er in zulke gevallen met de linearisering gebeurt betreft de beschrijving van meerdimensionale ruimtelijke structuren. Sinds het verschijnen van het stimulerende artikel van Linde en Labov (1975) over flatbeschrijvingen is er op ons instituut veel aandacht besteed aan de manier waarop proefpersonen ruimtelijke beschrijvingen geven (Klein *op cit*, Ullmer-Ehrich *op cit*, Ullmer-Ehrich en Koster 1981, Levelt en Maassen 1981, Levelt 1982a en 1982b). Een belangrijk resultaat van dit werk is dat proefpersonen het lineariseringsprobleem oplossen door een gebeurtenisstructuur te projecteren op de ruimtelijke structuur. Sprekers maken bij voorkeur een mentale rondgang door de ruimte, hetzij een soort wandeling (bijvoorbeeld in flat- en wegbeschrijvingen), of een bliktoer (in woonkamer- en netwerkbeschrijvingen). Als gevolg daarvan worden ruimtelijke relaties veelal uitgedrukt door middel van temporele connectieven („en dan”), met overvloedig gebruik van bewegingswerkwoorden en -adverbia. Op deze manier wordt de natuurlijke-volgorderegels voor temporele structuren uitgebreid naar ruimtelijke structuren, die op hun beurt slechts „indirect” worden beschreven via de gebeurtenisstructuren die erover worden geprojecteerd. Het is niet geheel duidelijk waarom sprekers hun toevlucht nemen tot deze toer-oplossingen. Een voordeel van de strategie is wel dat spreker en hoorder een aanzienlijke hoeveelheid kennis delen met betrekking tot het zich verplaatsen door steden, woningen, kamers, enzovoorts; een temporele beschrijving wekt gemakkelijk de bedoelde inferenties ten aanzien van de ruimtelijke relaties. Toch moet onder ogen worden gezien dat, ofschoon een temporele strategie een lineaire oplossing garandeert, zo'n oplossing helemaal niet uniek hoeft te zijn. Gedeelde kennis kan één determinant zijn voor de oplossing die dan in feite wordt gekozen, andere beperkingen komen voort

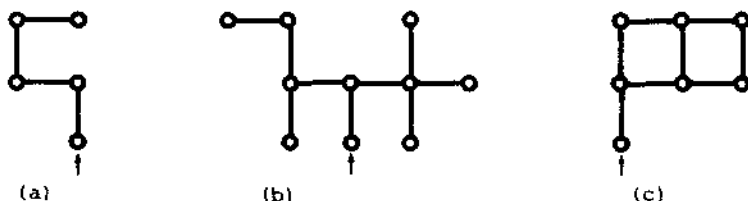
uit eigenschappen van werkgeheugen en aandacht. Ik zal nu enkele van deze procesgebonden principes bespreken.

Procesgebonden principes

Voor de geordende uitdrukking van complexe informatie moet de spreker een zekere „boekhouding” voeren. Het is, in de eerste plaats, nodig om uiteen te houden wat al gezegd is, en wat nog gezegd moet worden. Dit kan een eenvoudige aangelegenheid zijn voor lineaire structuren, maar het vraagt om bepaalde geheugenvoorzieningen bij de uitdrukking van meer complexe informatie zoals dadelijk zal blijken. In de tweede plaats zal een coöperatieve spreker ook rekening houden met de boekhouding van de hoorder, die moet proberen de complexe informatiestructuur te reconstrueren uit de gelineariseerde tekst. De belasting voor de hoorder zal onder andere afhangen van de mate waarin de spreker met modale, anaforische, deictische en andere middelen aangeeft hoe de gelineariseerde informatie weer samengesteld dient te worden. De belasting zal ook samenhangen met de volgorde van expressie die de spreker gekozen heeft.

In het volgende zal ik bespreken hoe deze boekhouding leidt tot beperkingen op het lineariseringsproces die van heel algemene aard zijn. Ze zijn onafhankelijk van de informatie-inhoud, en verschillen in dat opzicht wezenlijk van de inhoudsgebonden lineariseringsprincipes die zojuist ter sprake zijn geweest.

Ofschoon ze van heel algemene aard zijn zal ik de procesgebonden principes toelichten aan de hand van een heel beperkt tekstdomein dat we hebben gebruikt in het empirische onderzoek. Het gaat om beschrijvingen van ruimtelijke rasterachtige netwerken zoals in figuur 1, bestaande uit verschillend ge-



Figuur 1. Voorbeelden van lineaire (a), hiërarchische (b) en luspatronen (c) die aan proefpersonen ter beschrijving werden voorgelegd. De knooppunten waren in werkelijkheid verschillend gekleurd.

kleurde knooppunten die verbonden zijn door horizontale of verticale verbindinglijnen. De patronen konden een lineaire structuur (zonder vertakkingen) hebben, zoals in figuur 1a, een hiërarchische structuur met meervoudig vertakkende knopen, zoals in figuur 1b, of ze konden nog complexer zijn door

de aanwezigheid van lussen, zoals in figuur 1c. We lieten zulke patronen aan proefpersonen zien en vroegen hen de figuur zó te beschrijven dat een volgende proefpersoon het patroon correct zou kunnen tekenen op grond van de op de band opgenomen beschrijving. Verder vroegen we de sprekers hun beschrijving te beginnen bij de met een pijltje aangeduide knoop.

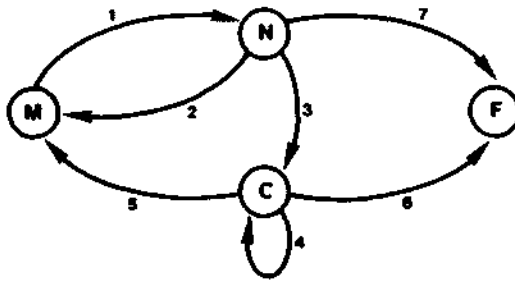
Hoe organiseert een spreker zijn boekhouding bij de beschrijving van zulke patronen? Ofschoon de patronen tijdens de beschrijving visueel gegeven zijn is boekhouding ook hier nodig, want de spreker moet onthouden welke delen van het patroon hij al beschreven heeft, en in sommige gevallen naar welke knopen hij nog moet terugkeren om de beschrijving te completeren. Ik heb twee lineariseringsmodellen gemaakt die voor elk patroon van dit type een sequentie oplevert van alle in het patroon voorkomende knopen en verbindingslijnen. De modellen voorspellen dus de volgorde van expressie voor elk patroon.

Voor de huidige discussie is het genoeg om een van de modellen nader te bekijken. Het model bestaat uit twee delen, een non-deterministisch gedeelte dat voor elk patroon een verzameling van mogelijke volgordes genereert, en een probabilistisch gedeelte dat iets zegt over de waarschijnlijkheidsverdeling binnen zo'n verzameling. Het non-deterministische gedeelte is een z.g. „aangevuld transitienetwerk” (ATN). Het is voorgesteld in figuur 2. De werking van dit ATN is elders beschreven (Levelt 1982b). Waar het hier om gaat zijn de principes waarop de constructie van dit ATN is gebaseerd; dat zijn namelijk de principes waarvan we aannemen dat ze algemeen en inhoudsonafhankelijk zijn.

Het eerste principe is de *maximalisering van connectiviteit*. Voor de onderhavige patronen houdt dit principe in dat, waar mogelijk, als volgende knoop er een wordt gekozen die een verbindingslijn heeft met het laatstgenoemde knooppunt. Of in andere woorden, het connectiviteitsprincipe stelt dat complexe informatie zoveel mogelijk wordt gelineariseerd „zonder het potlood op te tillen”; nieuwe concepten worden als het kan geïntroduceerd middels een voor de beschrijving relevante relatie met het meest recent uitgedrukte concept.

In allerlei experimenten die we met zulke patronen gedaan hebben (Levelt 1982a en 1982b) verkregen we een kleine vierduizend patroonbeschrijvingen. Schendingen van het connectiviteitsprincipe zijn hoogst zeldzaam. Ze doen zich nooit voor in patronen die zelf lineair zijn (zoals figuur 1a), en zijn verwaarloosbaar in de beschrijvingen van hiërarchische patronen (zoals figuur 1b); ze komen nu en dan voor in luspatronen (zoals figuur 1c) en hebben daar wellicht te maken met de moeilijkheden die een proefpersoon ervaart wanneer hij mentaal een volledige draaiing in de ruimte moet uitvoeren. Hij geeft er dan – nog steeds bij uitzondering – wel eens de voorkeur aan de lus aan de bovenkant „open te knippen”.

Connectiviteit is natuurlijk ook voordelig voor de hoorder omdat de nieuwe



Figuur 2. Lineariseringsmodel voor de patroonbeschrijvingen

| Overgang | Conditie | Actie |
|----------|--|--|
| 1 | ingang van patroon anders dan ingang | beschrijf ingang; reduceer valentie van ingangsknoop met 1 ga lus binnen indien herkend bij overgang 2; kies en beschrijf verbindingslijn, reduceer valentie van verlaten en binnengegane knoop met 1 |
| 2 | knoopvalentie = 1 knoop is ingang lus, en valentie > 1 | beschrijf knoop beschrijf knoop |
| 3 | knoopvalentie > 1 | beschrijf knoop |
| 4 | knoopvalentie > 1 | zoek M/F |
| 5 | knoopvalentie = 1 | - |
| 6 | knoopvalentie = 0 | stop |
| 7 | knoopvalentie = 0 | beschrijf knoop, stop |

informatie steeds kan worden verbonden met de meest recente, die veelal nog beschikbaar zal zijn in het werkgeheugen van de hoorder.

Het is niet altijd mogelijk voor de spreker nieuwe items connectief te introduceren zonder terug te keren naar reeds eerder vermelde items. Als het patroon hiërarchisch is, zoals in figuur 1b, kiest het ATN non-deterministisch de ene of de andere uitgang vanuit het keuzepunt; wanneer het vervolgens het einde van dat hele pad bereikt heeft springt de controle terug naar het keuzepunt om met de beschrijving van de andere uitgang en het daar ontspringende pad te beginnen. Deze terugsprong schendt natuurlijk het connectiviteitsprincipe. De enige manier om in deze situatie connectiviteit te handhaven zou voor de spreker zijn om stap-voor-stap weer terug te keren via alle reeds beschreven elementen tot aan het betreffende keuzepunt. Dit is precies de manier waarop terugkeer wordt georganiseerd in het andere, hier niet behandelde ATN-model, en er zijn proefpersonen die zich volgens dat model gedragen. De principiële kwestie is deze: wanneer de terugkeer naar keuzepunten sprongsgewijs plaatsvindt, zoals het ATN in figuur 2 voorspelt, hoe houdt de spreker dan bij

naar welke keuzepunten hij nog moet terugkeren, en in welke volgorde? Als keuzepunten „ingebod” zijn, zoals in figuur 1b, dat wil zeggen als de spreker een nieuw keuzepunt tegenkomt voordat hij is teruggekeerd naar het eerdere keuzepunt, hoe regelt hij dan zijn schema van terugkeer? Gaat hij eerst terug naar het oudste knooppunt, of juist naar het meest recente?

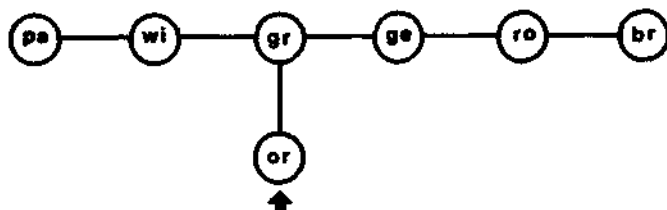
Het tweede principe zegt dat terugkeer naar keuzepunten in omgekeerde volgorde plaatsvindt. Het ATN van figuur 2 is uitgerust met een stapelgeheugen waarin terugkeeradressen op elkaar worden gestapeld in de volgorde van binnenkomst. Wanneer het eind van een lineair pad wordt bereikt verspringt de controle van het ATN (de aandacht van de spreker) naar het bovenste adres op de stapel. Dat betreft het keuze-element dat het laatst beschreven is. Dit gaat recursief zo verder tot het geheugen leeg is. Een andere manier om dit principe te formuleren is te zeggen dat het terugsprongen zo klein mogelijk maakt. Dit tweede procesgebonden principe blijkt eveneens zeer robuust te zijn. We vonden minder dan 3% schendingen in 297 patroonbeschrijvingen waarin sprake was van enkelvoudige of multipele inbedding. Het is aannemelijk, maar nog niet empirisch aangetoond dat deze wijze van terugkeren naar keuzepunten ook voordelig is voor de hoorder. Het corresponderende ATN-model voor de hoorder dat we hebben ontworpen heeft een stapelgeheugen met dezelfde volgorde-omkeringseigenschap.

Maximalisering van connectiviteit en inverse volgorde van terugkeer naar keuzepunten vormen de belangrijkste lineariseringsprincipes voor het non-deterministische gedeelte van het model. Samen geven ze de linearisering het „diepte-eerst”-karakter dat zo kenmerkend is voor menselijk probleem-oplossen in het algemeen (zie Newell en Simon 1972).

Keren we nu naar het probabilistische gedeelte. Tot nu toe voorspelde het model niet *welk* van de keuzealternatieven er bij een keuzepunt eerst zal worden gekozen, *welk* vervolgens na terugkeer, enzovoorts. Het probabilistische gedeelte van het model voorspelt de keuzepreferenties van de spreker in dit opzicht. Ook hier weer hebben we te doen met een algemeen principe dat, naar wij aannemen, van toepassing is voor elk gesproken tekstdomein.

Dit derde principe zegt dat de spreker bij een meervoudig vertakt keuzepunt een preferentie zal hebben om zijn keuzen *zó te ordenen dat de geheugenbelasting voortkomend uit het onthouden van terugkeeradressen wordt geminimaliseerd*. Dit principe kan het beste worden toegelicht door het in twee gedeelten te splitsen; het eerste op de *duur* van de geheugenbelasting, het tweede op de *omvang* ervan.

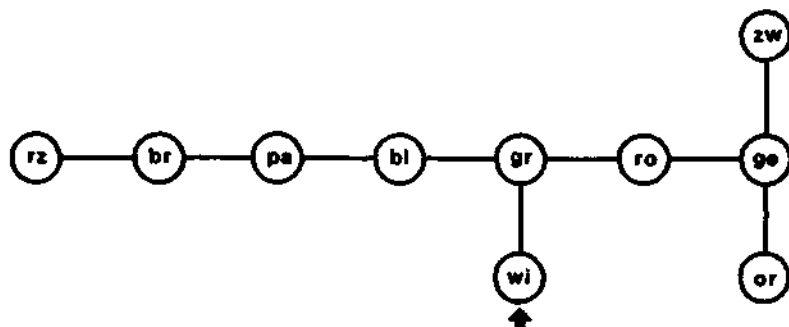
De minimalisering van belastingsduur kan worden verduidelijkt aan de hand van figuur 3. Wanneer de spreker in zijn beschrijving bij het groene keuzepunt terecht is gekomen moet hij kiezen tussen links- of rechtsaf gaan. In beide gevallen zal hij later moeten terugkeren naar het groene keuzepunt om het andere pad nog te beschrijven. Maar er is een verschil: wanneer het lange, rechter pad eerst wordt beschreven moet de spreker het terugkeeradres langer



Figuur 3. Patroon voor het testen van minimalisering van belastingsduur. (De kleuren zijn met letters aangegeven.)

beschikbaar houden dan wanneer het korte, linker pad eerst wordt beschreven. De voorspelling is dus dat in dit geval en in soortgelijke de waarschijnlijkheid het grootst is dat de spreker zijn keuzen ordent van korte naar lange paden. In de experimenten vonden we duidelijke statistische bevestiging van deze voorspelling (zie Levelt 1982b).

De minimalisering van belastingsomvang kan worden toegelicht met figuur 4. Hier is er weer een keuze te maken bij het groene punt, maar nu is het aantal te beschrijven elementen links en rechts van het keuzepunt gelijk. De twee



Figuur 4. Patroon voor het testen van minimalisering van belastingsomvang

vertakkingen verschillen echter daarin dat er zich rechts, maar niet links een tweede keuzepunt voordoet. Wanneer de spreker deze rechter vertakking eerst beschrijft komt hij op een gegeven moment in de situatie dat hij er zowel aan moet denken naar het groene keuzepunt terug te keren, als ook om naar het gele terug te gaan. In termen van het ATN: er zijn dan twee terugkeeraadressen op het stapelgeheugen. Dit kan niet gebeuren wanneer het linker pad eerst wordt beschreven; het is dan nooit nodig meer dan één terugkeeradres tegelijk in het werkgeheugen vast te houden. Men kan dus voorspellen dat voor deze en soortgelijke figuren de spreker ertoe zal neigen eenvoudige paden vóór complexe paden te kiezen. Ook hiervoor vonden we overtuigende statistische evidentie in de experimenten.

Het kan ook voor de hoorder van voordeel zijn wanneer de spreker zich aan deze minimaliseringstrategieën houdt. Wanneer de hoorder zich ervan be-

wust is dat de spreker van plan is later naar een bepaald knooppunt terug te keren dan kan de hoorder die punten voor zichzelf markeren. De hoorder kan zulke intenties vermoeden op grond van expliciete of impliciete aanduidingen van de spreker („we komen straks nog een keer naar dit kruispunt terug” versus „hier zou je rechtsaf kunnen gaan”). Wanneer de hoorder kan vermoeden wat de keuzepunten zijn en dus de terugkeeradressen, zal ook hij ermee gebaat zijn zo weinig mogelijk van zulke adressen te hoeven onthouden, en zo kort mogelijk.

De proces-gebonden principes kunnen nu als volgt worden samengevat. De spreker zal, ten eerste, maximaal connectief zijn bij het introduceren van nieuwe informatie-items. Voorts zal de spreker zijn terugkeer naar keuzepunten in omgekeerde volgorde programmeren, en tenslotte zal hij vanuit een keuzepunt de paden in die volgorde kiezen waarmee het onthouden van terugkeeradressen wordt geminimaliseerd.

Het is opmerkelijk dat de combinatie van deze principes voor hiërarchische structuren formeel equivalent is met Yngve's „dieptehypothese” (Yngve 1961). Yngve voorspelde dat syntactische structuren maximaal rechtsvertakkend zouden moeten zijn, gegeven de geheugenbeperkingen van de spreker (zie Levelt 1973 voor een bespreking van deze hypothese). De argumenten die Miller en Chomsky (1963) aanvoerden tegen Yngve's theorie, namelijk dat op grond van dezelfde beperkingen de hoorder linksvertakking zou prefereren, gelden niet voor de lineariseringstheorie. Zoals opgemerkt is elk principe mogelijk ook van voordeel voor de hoorder. Een belangrijke empirische toetssteen voor een algemene lineariseringstheorie is dus de voorspelling dat de structuur van gesproken tekst (NB niet van de syntaxis van de tekst) dominant rechtsvertakkend zal zijn. We konden deze dominantie aantonen voor de behandelde ruimtelijke beschrijvingen, maar beschikken slechts over anekdotische evidentie voor andere tekstdomeinen.

Linearisering en de orale modaliteit van het spreken

Waaruit komt het lineariseringsprobleem van de spreker voort? Er is een voor de hand liggend, maar nogal triviaal antwoord op deze vraag: de oorsprong van het probleem is de orale modaliteit van het spreken; we kunnen met onze stem geen simultane uitdrukking geven aan complexe informatie. Zelfs als we het erover eens zijn dat er een zekere gelijktijdigheid van uitdrukking mogelijk is door coarticulatie van fonemen, of nog duidelijker door de parallelle expressie van segmentale en suprasegmentale informatie, moeten we accepteren dat het onmogelijk is twee of meer zinnen tegelijk te formuleren.

Niettemin is er reden om aan te nemen dat de lineariseringsvoorwaarde modaliteitsonafhankelijk is, en meer te maken heeft met onze beperkte vermogens om propositionele informatie te verwerken dan met technische beperkin-

gen van onze stem. Er zijn twee gronden voor deze veronderstelling. De eerste is dat wanneer een spreker plotseling met twee of meer stemmen tegelijk zou kunnen spreken (er zijn tenslotte ook vogels die met twee paar stembanden tegelijk tweestemmig kunnen zingen), de hoorder er geen raad mee zou weten. Het werk van Broadbent en vele anderen (zie Broadbent 1971) laat er geen twijfel over bestaan dat een aantal simultaan gesproken teksten door een hoorder niet tegelijk kunnen worden verstaan. Dit geldt zelfs nog wanneer één tekst aan het ene oor wordt aangeboden en een andere tekst, goed van de eerste gescheiden, aan het andere oor. Stellen dat zulke tweestemmige sprekers niet bestaan en de hoorder er dus geen ervaring mee kan hebben is onjuist. Mensen verkeren vaak in de situatie dat er twee of meer sprekers tegelijk aan het woord zijn. Wat ze in zo'n „cocktail party"-situatie doen is, op elk moment selectief naar de een of de andere spreker luisteren.

De tweede reden is mijns inziens nog overtuigender. Er zijn natuurlijke talen die geen gebruik maken van de orale modaliteit, en die technische gelijktijdigheid van uitingen zouden toestaan. Dit zijn de natuurlijke tekentalen van de doven. In American Sign Language (ASL), zowel als vermoedelijk in alle andere tekentalen, kan men een teken naar believen met de rechter of de linker hand maken: er is vrije alternatie. Er is ook motorisch niets wat simultaan gebruik van beide handen in de weg staat. Ik heb inderdaad verschillende observaties kunnen doen van simultaan tekengebruik in ASL. In de meeste van deze gevallen houdt de ene hand een statische „classifier" (een anaforisch element dat voor een bepaalde semantische klasse kan staan), terwijl de andere hand de focale informatie geeft die een bepaalde relatie onderhoudt met dat anaforische element. Het lijkt een beetje op de deictische functie van een wijsgebaar dat simultaan gegeven wordt bij de uiting „daar ligt het boek". In één geval heb ik het spontane gebruik van een gewoon zelfstandig naamwoord („het meisje") in de ene hand waargenomen, met tegelijk in de andere hand een betrekkelijke bijzin. Opheffing van de technische beperking heeft dus wel degelijk een zeker effect. Maar het staat eveneens buiten twijfel dat het ongrammaticaal is in ASL om twee zinnen tegelijk te uiten met de twee handen (zie Levelt 1980 voor een verdere bespreking van deze kwestie). Er zijn nog geen experimenten gedaan over het vermogen van ASL-gebruikers om simultane uitingen te verstaan. Maar ofschoon het visuele systeem zagezegd ontworpen is voor de parallele verwerking van informatie, moeten we van zulke experimenten toch niet veel positiefs verwachten. De situatie zal vermoedelijk niet veel beter zijn dan voor simultane leestaken. Daarvan weten we dat er wel incidentele parallele registratie plaatsvindt van enkelvoudige woorden (Willows en McKinnon 1973), maar er is geen enkele aanwijzing dat er simultaan lezen mogelijk is van twee of meer zinnen tegelijk.

Met deze opmerkingen kom ik weer terug bij het ramistische uitgangspunt van de inleiding. Het lineariseringsprobleem moet worden behandeld onder het hoofd *ratio*, het denkvermogen, niet onder *oratio*, het vermogen tot spre-

ken. Dit zijn relatief autonome systemen (zie Levelt 1981, en voor verdere empirische evidentie Levelt en Maassen 1981). De lineariseringsvoorwaarde heeft te maken met beperkingen op ons vermogen onze aandacht te verdelen, en geldt derhalve evenzeer voor natuurlijke talen zonder spraak als voor elk beslissingsproces waarbij propositionele informatie is betrokken, zoals rekenen, schaken, of een artikel schrijven over lineariseringsproblemen.

Dankbetuiging. Gaarne dank ik Else Barth, Gerard Kempen, William Marslen-Wilson en Gabriël Nuchelmans voor discussies en commentaar bij de voorbereiding van dit artikel.

Noten

1. Zie bijvoorbeeld de Latijnse rhetoricatekst die traditioneel maar foutiefelijk wordt toegeschreven aan Marcus Tullius (Cicero), circa 85 v. C. Kempen en Van Wijk (1981) brachten deze tekst expliciet in verband met het lineariseringsprobleem.
2. Ofschoon ze vaak worden verward moeten natuurlijke-volgordetheorieën van de *dispositio* zorgvuldig worden onderscheiden van natuurlijke-volgordetheorieën in de *grammatica*. De laatste betreffen de woordvolgorde in zinnen. Zie voor een bespreking daarvan, Levelt (1967).

Literatuur

- Becker, A. L. 1980. Text-building, epistemology, and aesthetics in Javanese shadow theatre. In: A. Becker en A. Yengoyan (red.), *The imagination of reality*. Norwood: Ablex.
- Broadbent, D. E. 1971, *Decision and stress*. Londen: Academic Press.
- Collins, A., E. H. Warnock en J. J. Passafiume 1975, Analysis and synthesis of tutorial dialogues. In: G. H. Bower (red.), *The psychology of learning and motivation*, vol. 9. New York: Academic Press.
- Fodor, J. A. 1975, *The language of thought*. Hassocks: Harvester Press.
- Gazdar, G. 1979, *Pragmatics*. Londen: Academic Press.
- Gazdar, G. 1980, Pragmatic constraints on linguistic production. In: B. Butterworth (red.), *Language production*. Londen: Academic Press.
- Kempen, G. en C. van Wijk 1981, Leren formuleren. *Tijdschr. Taalbeheersing* 3, 32-44.
- Kempson, R. M. 1975, *Presupposition and the delimitation of semantics*. Cambridge: Cambridge University Press.
- Klein, W. 1979, Wegauskünfte. *Z. LitWiss. Ling.* 9, 9-57.
- Klein, W. 1982, Some aspects of route directions. In: R. J. Jarvella en W. Klein (red.), *Speech, place, and action. Studies in deixis and related topics*. Chichester: Wiley.
- Levelt, W. J. M. 1967, *Over het waarnemen van zinnen*. Groningen: Wolters.
- Levelt, W. J. M. 1973, *Formele grammatica's in linguïstiek en taalpsychologie, deel III*. Deventer: Van Loghum Slaterus.
- Levelt, W. J. M. 1980, On-line processing constraints on the properties of signed and spoken language. In: U. Bellugi en M. Studdert-Kennedy (red.), *Signed and spoken language: biological constraints on linguistic form*. Weinheim: Verlag Chemie.
- Levelt, W. J. M. 1981, Déjà vu? *Cognition* 10, 187-192.

- Levelt, W. J. M. 1982a, Cognitive styles in the use of spatial direction terms. In: R. J. Jarvella en W. Klein (red.), *Speech, place, and action. Studies in deixis and related topics*. Chichester: Wiley.
- Levelt, W. J. M. 1982b, Linearization in describing spatial networks. In: S. Peters en E. Saarinen (red.), *Processes beliefs, and questions*. Dordrecht: Reidel.
- Levelt, W. J. M. en B. Maassen 1981, Lexical search and order of mention in sentence production. In: W. Klein en W. J. M. Levelt (red.), *Crossing the boundaries in linguistics. Studies presented to Manfred Bierwisch*. Dordrecht: Reidel.
- Linde, C. en W. Labov 1975, Spatial networks as a site for the study of language and thought. *Language* 51, 929-939.
- Lyons, J. 1977, *Semantics*, vol. 2. Cambridge: Cambridge University Press.
- Mandler, J. M. en N. S. Johnson 1977, Remembrance of things parsed: story structure and recall. *Cogn. Psychol.* 9, 111-151.
- Miller, G. A. en N. Chomsky 1963, Finitary models of language users. In: R. D. Luce, R. R. Bush en E. Galanter (red.), *Handbook of mathematical psychology*, vol. 2. New York: Wiley.
- Munro, A. 1977, *Speech act understanding*. (Ph. D. dissertation, University of California, San Diego.)
- Newell, A. en H. A. Simon 1972, *Human problem solving*. Englewood Cliffs, N. J.: Prentice Hall.
- Nuchelmans, G. 1980, *Late-scholastic and humanist theories of the proposition*. Amsterdam: North-Holland.
- Osgood, C. E. 1980, *Lectures on language performance*. New York: Springer.
- Ramée, P. de la 1555, *Dialectique*. Opnieuw uitgegeven door M. Dassonville, bij Librairie Droz, Genève 1964.
- Schank, R. en R. Abelson 1977, *Scripts, plans, goals and understanding*. Hillsdale: Erlbaum.
- Tullius, M. circa 85 v.C., *Ad Herennium* (translated and edited by H. Caplan, 1954). Cambridge, Mass.: Harvard University Press.
- Ullmer-Ehrich, V. 1979, Wohnraumbeschreibungen. *Z. LitWiss. Ling.* 9, 58-83.
- Ullmer-Ehrich, V. 1982, The structure of living space descriptions. In: R. J. Jarvella en W. Klein (red.), *Speech, place, and action. Studies in deixis and related topics*. Chichester: Wiley.
- Ullmer-Ehrich, V. en C. Koster 1981, Discourse organization and sentence form. The structure of static arrangement descriptions. (Niet gepubliceerd, Max-Planck-Institut für Psycholinguistik, Nijmegen.)
- Willows, D. M. en G. E. MacKinnon 1973, Selective reading: attention to the „unattended” lines. *Can. J. Psychol.* 27, 292-304.
- Wunderlich, D. en R. Reinelt 1982, How to get there from here. In: R. J. Jarvella en W. Klein (red.), *Speech, place, and action. Studies in deixis and related topics*. Chichester: Wiley.
- Yngve, V. H. 1961, The depth hypothesis. In: R. Jakobson (red.), *Structure of language and its mathematical aspects*. (Proc. 12th Symposium in Applied Mathematics.) Providence, RI: American Mathematical Society.

Willem J. M. Levelt is directeur van het Max-Planck-Institut für Psycholinguistik te Nijmegen.