

ONDER WOORDEN BRENGEN
BESCHOUWINGEN OVER HET
SPREEKPROCES

HAARLEMSE
VOORDRACHTEN

VOORDRACHTEN GEHOUDEN
IN DE HOLLANDSCHE MAATSCHAPPIJ DER
WETENSCHAPPEN TE HAARLEM

LIV

Dr W. J. M. Levelt

HAARLEM
1994

ONDER WOORDEN BRENGEN
BESCHOUWINGEN OVER HET
SPREEKPROCES

door

Dr W. J. M. Levelt

*Voordracht gehouden in de Algemene Vergadering
van 28 mei 1994*

HAARLEM
1994

TZ
c
2842

S

EX LIBRIS
UNIVERSITATIS
NOVIOMAGENSIS

S. 509. 080

Met het spreken is het ongeveer zoals met de lucht of de zwaartekracht. Die zijn absoluut essentieel voor ons menselijk bestaan, maar we hoeven er nooit op te letten. Ze zijn er gewoon, ze werken, en daarmee basta. Ons vermogen tot spreken ligt zonder twijfel aan de basis van onze cultuur, ons sociale en maatschappelijke bestaan, onze kennisoverdracht, onze opvoeding. Maar we hoeven nooit na te denken over hoe we spreken. De meesten van ons weten niet eens dat het vermogen tot spreken de meest complexe mentale en motorische vaardigheid is die alle mensen met elkaar delen. Het is ook soortspecifiek, de mens is het enige sprekende dier. Het spraakvermogen is een geschenk van de evolutie; het is in onze genen voorgeprogrammeerd. En we gebruiken dat vermogen de godganselijke dag. Wanneer we niet met anderen kwebbelen, praten we wel met onszelf.

Maar hoe werkt het nu eigenlijk, spreken? Tot voor kort was daar buitengewoon weinig over bekend. Het is pas sinds een jaar of vijf en twintig dat dit mechanisme (zo zal ik het noemen, want dat is het) op systematische wijze wordt onderzocht. Vandaag wil ik de resultaten van dat onderzoek met u delen door u in vogelvlucht te schetsen wat we inmiddels over dat systeem te weten zijn gekomen. Onze huidige kennis erover is voornamelijk gebaseerd op experimenteel laboratoriumonderzoek, en dat zit verschrikkelijk ingewikkeld in elkaar. Het is echter wel mogelijk de grote lijnen van het spraakmechanisme te schetsen zonder diep in te gaan op dat onderzoek. Althans, dat zal ik in het volgende proberen.

Spreeksnelheid 1

Spreken is een proces dat zich in de tijd afspeelt. Hoe snel? De volgende tekst is een transcriptie van een vakantieverslag, een mondelinge beschrijving van een reis door Spanje in mei:

De bedoeling was om vooral in de bergen wat rond te trekken. Dan is het erg lekker in het binnenland van Spanje. Maar we zijn ook naar steden wezen kijken. We zijn naar Cordoba geweest en

we zijn naar Granada geweest. En eigenlijk was de combinatie -uh- een hele goeie. Want als je een paar dagen in de bergen hebt gelopen tussen de bloemetjes, ach dan wil je ook wel weer eens -uh- cultuur zien en dan zijn natuurlijk Granada en Cordoba fantastische plaatsen om te bezoeken.

Dit stukje tekst duurt 36 seconden. Wat doet de spreker allemaal in die tijd? Ten eerste bedenkt hij wat hij zal gaan zeggen. De tekst is een serie mededelingen, zoals wat de bedoeling was, dat het dan lekker is in het binnenland, etc. Hoeveel van die elementaire mededelingen doet de spreker in die 36 seconden? Dat hangt ervan af hoe je ze telt. Even aannemende dat elke mededeling zich rond een hoofdwerkwoord groepeert, bevat deze tekst 10 mededelingen, gemiddeld één elke 3.6 seconden. Je kunt het ook anders tellen, maar een gemiddelde "mededelingstijd" van circa 4 seconden is normaal voor de lopende spraak. Dat is niet erg snel, maar het gaat hier ook om een beslissingsproces dat veel aandacht vergt.

Ten tweede produceert de spreker woorden. Het zijn er hier 85, ofwel 2.36 per seconde. Het is geen geringe prestatie om 2 tot 3 woorden per seconde op te halen uit een mentaal lexicon dat vele tienduizenden woorden bevat. Dat ophaalproces verloopt vrijwel automatisch; het vergt weinig aandacht.

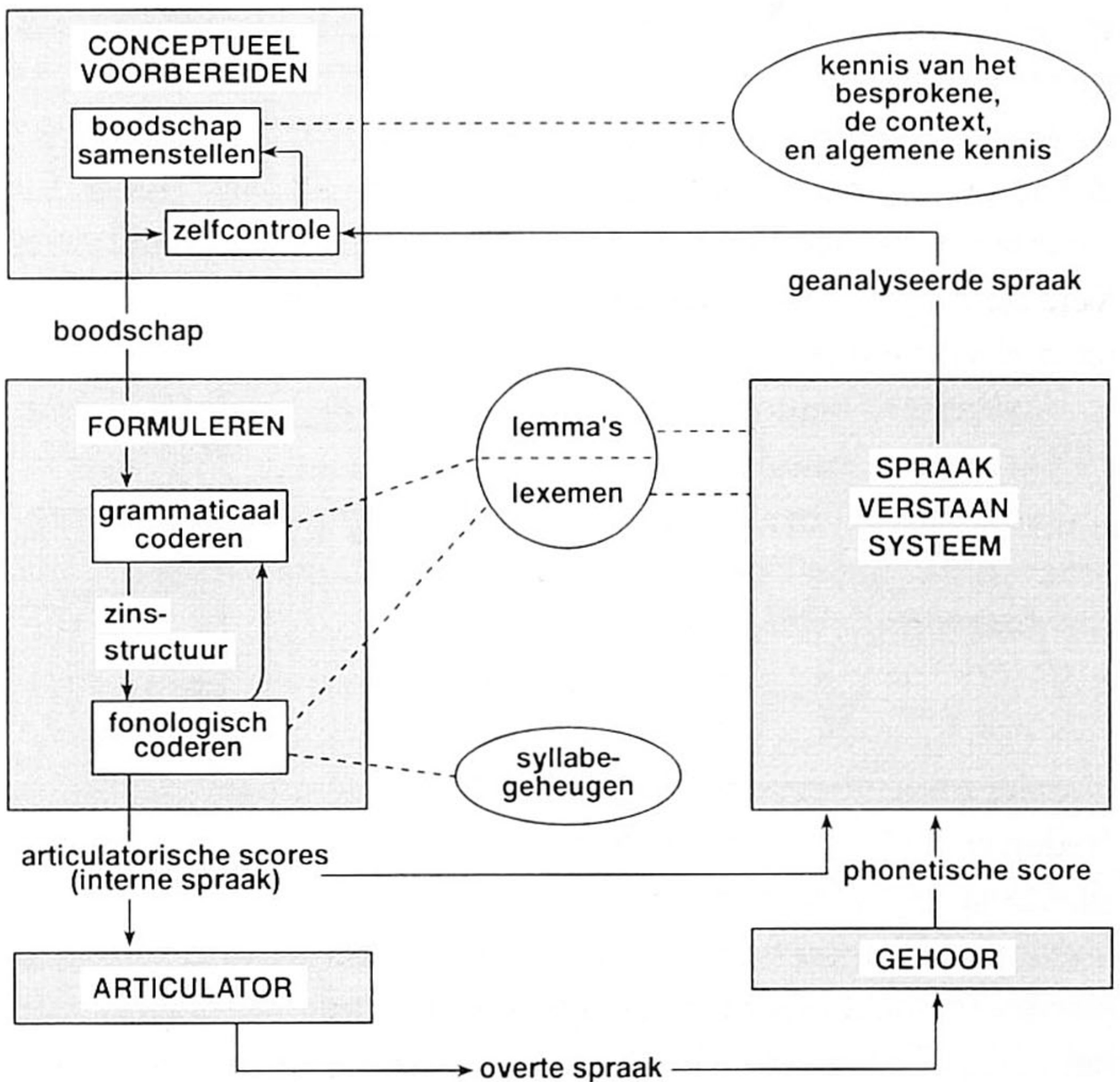
Ten derde produceert de spreker "uitspraakeenheden", syllaben en fonemen. Er zitten 128 syllaben in deze tekst, dat zijn er 3.6 per seconde. En de tekst bevat 334 fonemen (klinkers en medeklinkers); dat zijn er 9.3 per seconde. En dit was geen snelle spreker. Vier syllaben en 10-12 fonemen per seconde is heel gewoon. Er kan geen sprake van zijn dat een spreker voortdurend moet beslissen "nu ga ik een t zeggen, nu ga ik een a zeggen", etc. Vloeiend spreken is slechts mogelijk omdat die uitspraakeenheden geheel automatisch tot stand komen. We hoeven er inderdaad niet bij na te denken.

En al die processen verlopen simultaan. Terwijl je het ene woord uitspreekt wordt het volgende woord al opgehaald uit het lexicon en ben je al aan het bedenken wat je daarna zal gaan zeggen. Dat simultane verloop is een voorwaarde voor het produce-

ren van vloeiende spraak en het bestaat alleen bij gratie van het juist genoemde automatisme. De spreker hoeft alleen maar te letten op *wat* hij wil gaan zeggen; al het andere gebeurt vrijwel vanzelf.

Het systeem en zijn componenten

Hoe zit het spraakproductiesysteem in elkaar? Zoals bij de analyse van elk complex systeem, moeten we ook bij het spraakmechanisme proberen het in zijn componenten uiteen te leggen, onderzoeken hoe elk van die componenten werkt, en hoe die componenten met elkaar samenwerken. Op het ogenblik stellen we ons voor dat dat systeem in elkaar steekt zoals weergegeven in Figuur 1.



Figuur 1. Blauwdruk van de spreker

Links boven staat de "strategische" component weergegeven die zich bezig houdt met *conceptuele voorbereiding*, het beslissen *wat* er moet worden meegedeeld. Laten we dat de "boodschap" noemen. Wat zich hier afspeelt heeft een bewust aspect; het vraagt aandacht en het verloopt langzaam. De uit te drukken boodschap moet dan worden *geformuleerd*, d.w.z. in taal worden uitgedrukt. Dat gaat in twee stappen, *grammaticaal coderen* en *fonologisch coderen*. Bij het grammaticaal coderen haalt de spreker woorden op uit het lexicon die uitdrukking kunnen geven aan de begrippen waaruit zijn mededeling is samengesteld. En die woorden gebruikt hij om zinnen mee te bouwen. Bij het fonologisch coderen wordt de uitspraak van elk woord voorbereid, met name zijn samenstelling uit syllaben, klinkers en medeklinkers. Ook wordt de intonatie en het metrisch verloop van de uiting als geheel voorbereid. Het resultaat is een articulatoirisch programma, dat we ook kunnen ervaren als innerlijke spraak. Dat programma kan worden uitgevoerd door een zeer complexe *articulatorische component* en dan hebben we overte, hoorbare spraak. Er is, tenslotte, geen spreker naar wie we meer luisteren dan onszelf. En zoals we in de spraak van anderen versprekingen kunnen horen of andere onvolkomenheden, kunnen we dat ook bij het luisteren naar onze eigen spraak. We kunnen dan stoppen en onszelf corrigeren. Dit heet *zelfcontrole* en zelfcontrole is een aspect van elke complexe vaardigheid.

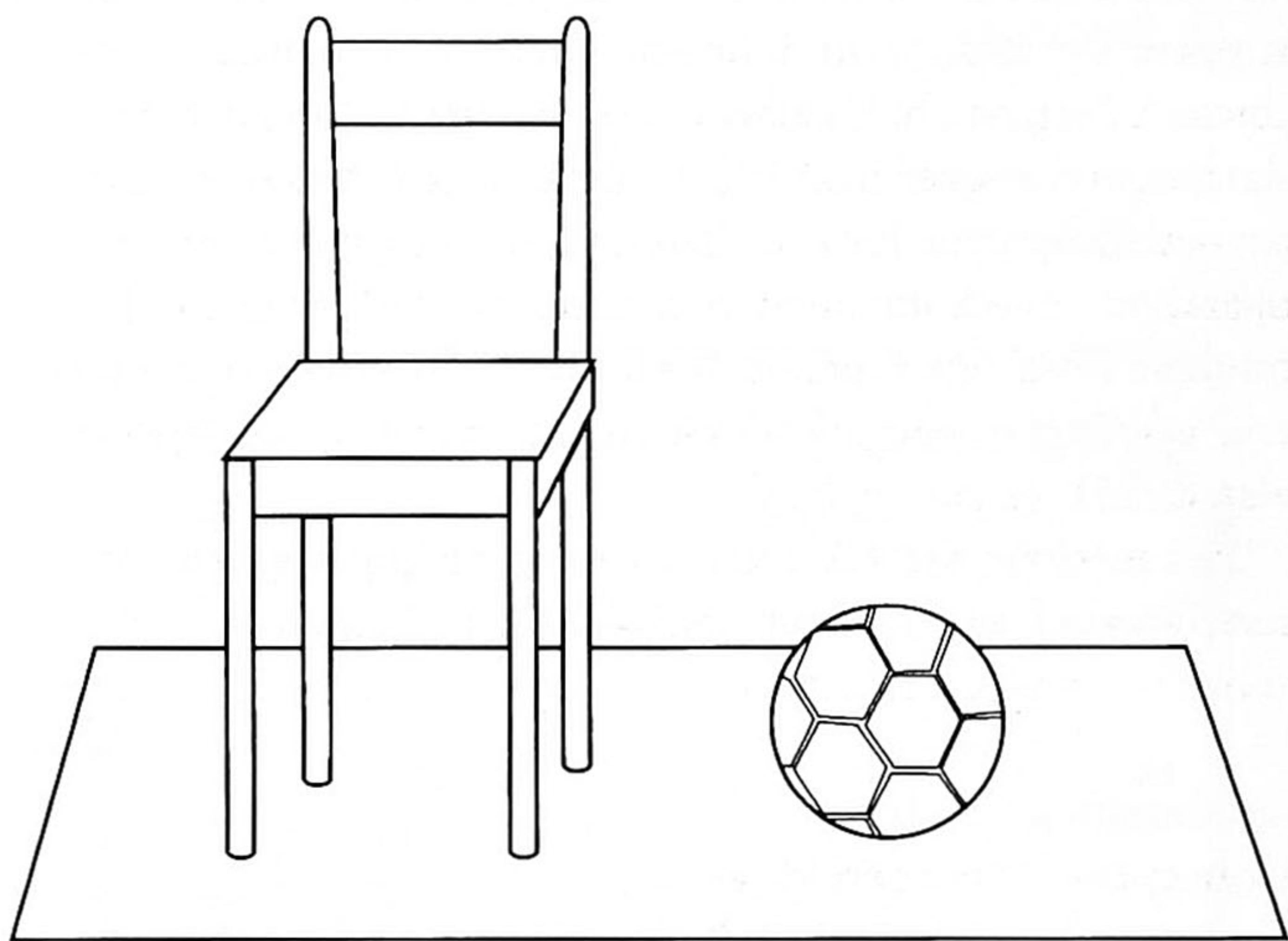
In het volgende zal ik over elk van deze componenten iets zeggen, echter zonder ook maar enigszins volledigheid na te streven (zie daarvoor Levelt, 1993).

Conceptuele voorbereiding

Wat de spreker gaat zeggen hangt ten nauwste samen met zijn bedoelingen. De boodschap die hij gaat formuleren moet tenslotte als effect hebben dat die bedoeling herkend kan worden door de hoorder. (Er moet ook wel eens iets worden versluierd, maar ook daarvoor moet weer een effectieve mededeling worden be-

dacht.) De stap van bedoeling naar boodschap is, ondanks zeer veel onderzoek, nog grotendeels *terra incognita* in de taalproductietheorie. Het gaat hier om het hele arsenaal van retorische middelen dat sinds Aristoteles (en eerder) is onderzocht en verrijkt. Over dat arsenaal zelf is veel bekend, maar hoe een "gewone" spreker onder welgedefinieerde omstandigheden een bedoeling omzet in een boodschap - daarover weten we bedroevend weinig.

Hier wil ik een ander aspect van conceptuele voorbereiding belichten. We kunnen de meest uiteenlopende soorten informatie in taal uitdrukken: muzikale impressies, emoties, geuren, ruimtelijke relaties, en nog veel meer. Maar wat daarbij altijd nodig is, is die informatie af te beelden op begrippen waarvoor we woorden hebben; een boodschap moet bestaan uit *lexicale begrippen*. Dit kan worden geïllustreerd aan de hand van Figuur 2.



Figuur 2. Ligt de bal links of rechts van de stoel?

Wanneer je iemand naar deze figuur laat kijken en vraagt "Waar ligt de bal?", dan is het antwoord meestal iets als, "Rechts van de stoel". Maar sommige mensen antwoorden, "Links van

de stoel". Hoe komt het dat dezelfde ruimtelijke relatie door de één als "rechts van" wordt opgevat en door de ander als "links van"? Dat ligt aan het perspectief dat de spreker kiest. In het eerste geval kiest de spreker zichzelf als basis van een coördinatensysteem ("deictisch perspectief"). Als de spreker zich recht naar de stoel richt, ligt de bal in zijn rechter halfveld. In het tweede geval kiest hij de stoel als basis ("intrinsiek perspectief"). De stoel heeft een "eigen" linkerkant en aan die zijde ligt de bal. Het gekozen perspectief bepaalt dus hoe één en dezelfde ruimtelijke relatie kan worden omgezet in twee tegengestelde lexicale begrippen, het begrip LINKS of het begrip RECHTS. En dat bepaalt welk woord er gebruikt gaat worden.

Dit is geen toevallige bijzonderheid van "links" en "rechts", maar een volstrekt algemene eigenschap van conceptuele voorbereiding. Alles waar we naar verwijzen, elke relatie, elke toestand kan op meerdere wijzen worden afgebeeld op lexicale begrippen. Dezelfde persoon Jan kan ik aanduiden met "collega", "broer", "vriend", etc., afhankelijk van het gekozen perspectief (nu niet in ruimtelijke maar metaforische zin). Het gaat hier niet om een filosofische curiositeit (u weet wel, de morgenster en de avondster), maar om een afbeeldingsprobleem waarmee we als sprekers altijd worden geconfronteerd. Maar we weten alleen enigszins hoe sprekers dit doen voor ruimtelijke begrippen (Levelt, 1994).

Het uiteindelijke resultaat van conceptueel voorbereiden is een "boodschap", die bestaat uit lexicale begrippen.

Grammaticaal coderen

Deze component doet twee dingen. Voor elk lexicaal begrip in de boodschap selecteert hij het juiste woord uit het lexicon. En hij brengt die geselecteerde woorden in syntactische verbinding met elkaar, zó dat de boodschap als geheel juist wordt afgebeeld.

Zoals boven vermeld, halen we gemiddeld zo'n twee tot drie woorden per seconde op uit een zeer omvangrijk lexicon. Het

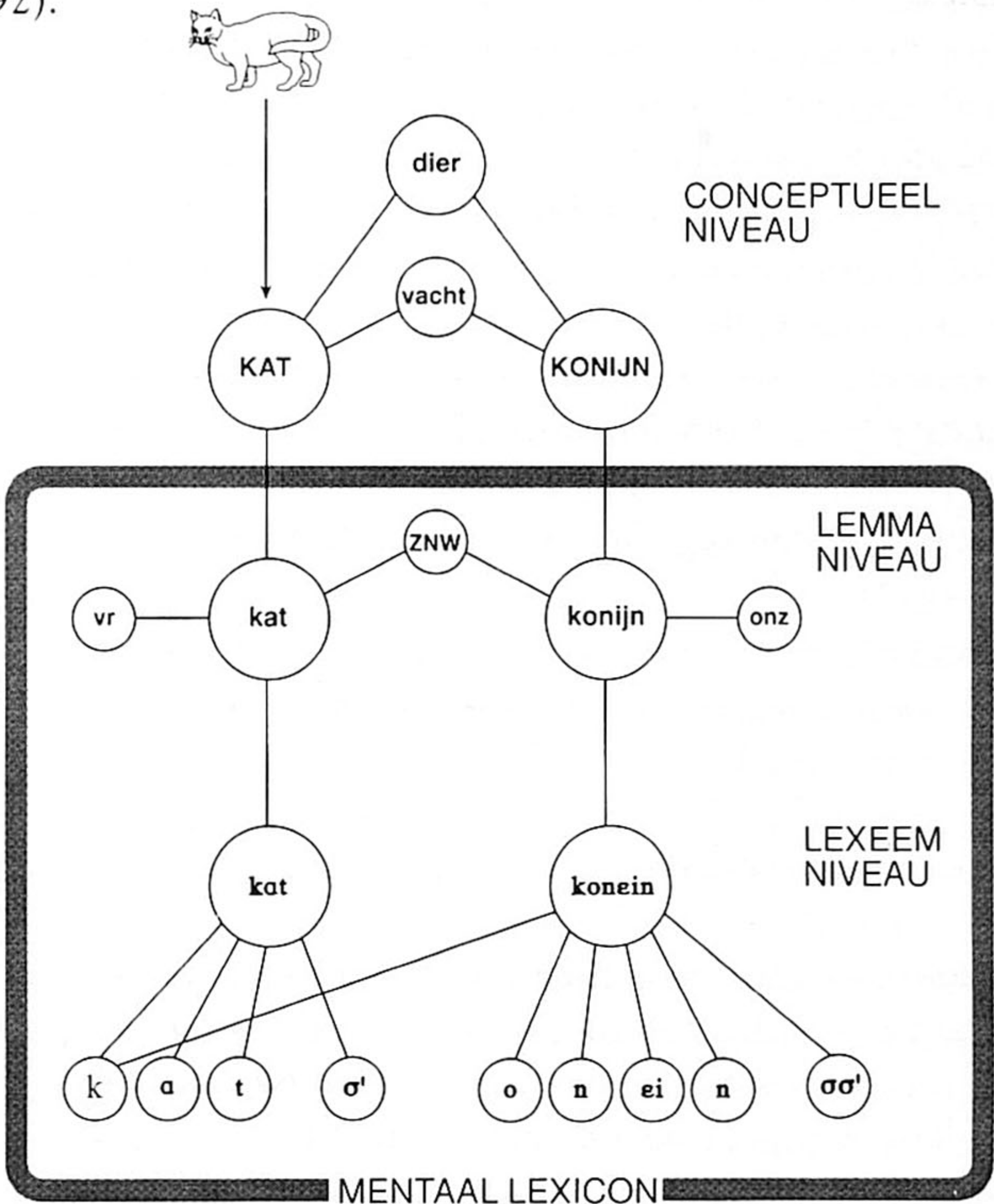
meest indrukwekkende is wel dat we daarbij opvallend weinig fouten maken, gemiddeld niet meer dan ongeveer één promille. Maar ze komen voor, selectiefouten. Hier zijn drie zelf opgetekende voorbeelden:

Ik hoorde Chris pi- viool spelen

We zijn naar het wrak toegelopen - gevaren

't Is prachtig herfst-uh lenteweer

Zulke fouten zeggen iets over de organisatie van het mentale lexicon. De selectiefout is meestal betekenisverwant met het doelwoord. Die organisatie stellen we ons voor als een netwerk waardoor activatie kan stromen. Figuur 3 laat een heel klein fragment zien van dat netwerk (zoals voorgesteld door Roelofs, 1992).



Figuur 3. Een fragment van het lexicale netwerk

Het bovenste niveau is het niveau van de concepten. Elke knoop stelt een begrip voor, en de betekenis van zo'n concept ligt opgeslagen in de verbindingen met andere concepten. Een kat, bijvoorbeeld, is een dier en heeft een vacht (de aard van de verbindingen staat in de figuur niet nader aangegeven). Een konijn is ook een dier en heeft ook een vacht. Dat een kat miauwt en een konijn knaagt zit ook in het netwerk (maar eenvoudigheidshalve niet in de figuur). Een begrip kan op allerlei manieren worden geactiveerd. De spreker ziet bijvoorbeeld een kat of herinnert zich wat zijn kat gisteren voor iets bijzonders deed. Dan kan het concept KAT actief worden. Die activatie spreidt zich door het conceptuele netwerk, zodat het concept KONIJN ook enige activatie krijgt.

Een lexicaal concept, zoals KAT, spreidt zijn activatie ook naar het mentale lexicon. De activatie van KAT spreidt naar het woord *kat*, of beter naar het "lemma" *kat*. Een lemma is niet het complete woord, maar het syntactische woord. De lemma-knoop is verbonden met syntactische informatie op het tweede niveau van het netwerk. Zo is de lemma-knoop *kat* verbonden met een knoop die aangeeft dat *kat* een zelfstandig naamwoord (ZNW) is. En *kat* is syntactisch van vrouwelijk geslacht; in het netwerk is er een verbinding met de knoop *vr*. Het lemma *konijn* is ook een zelfstandig naamwoord, maar het geslacht is onzijdig (*onz*).

Lexicale selectie is in het netwerkmodel niets anders dan het selecteren van een lemma, gegeven een actief lexicaal concept. En dat is een statistisch proces. Stel KAT is het actieve begrip. Op elk moment is de kans dat het lemma *kat* wordt geselecteerd gelijk aan de activatie van dat lemma gedeeld door de som activatie van alle actieve lemmas. Met die selectieregel kunnen we nauwkeurig voorspellen hoelang een selectie gemiddeld duurt. En het model is verrassend goed bevestigd in experimenten waarin we onder verschillende condities meten hoeveel tijd het kost om plaatjes van objecten (bijvoorbeeld van een kat of een konijn) te benoemen (zie Roelofs, 1992).

Hoe ontstaat nu zo'n selectiefout, *piano* voor *viool* of *herfst* voor *lente* of *konijn* voor *kat*? Dat komt door de activatiespreiding

op conceptueel niveau in het netwerk. KAT spreidt een klein deel van zijn activatie naar KONIJN, en beide spreiden een klein deel van hun activatie naar hun lemmas (*kat* en *konijn*). Omdat de selectieregel een probabilistische is, is er altijd een geringe kans dat *konijn* wordt geselecteerd, zelfs als *kat* het grootste deel van de activatie krijgt. En die activatie kon alleen maar zo spreiden omdat KAT en KONIJN conceptueel verbonden zijn, d.w.z. betekenisverwant zijn.

Zogauw het lemma is geselecteerd komen de syntactische eigenschappen van het woord beschikbaar. Op grond daarvan kunnen de woorden nu een zinsverband met elkaar aangaan. Elk woord vereist zijn eigen specifieke syntactische omgeving. Een werkwoord heeft een onderwerp nodig (in het Nederlands), soms ook een lijdend of medewerkend voorwerp. Een zelfstandig naamwoord vereist vaak een bepalend woord, bijvoorbeeld een lidwoord dat eraan voorafgaat, enz. Al die syntactische omgevingen moeten tegelijk gerealiseerd worden voor de geselecteerde lemmas. Het is een beetje zoals het oplossen van een stelsel simultane vergelijkingen. Dat proces heet *unificatie* (zie Kempen, 1993). Syntactische unificatie verloopt meestal foutloos, maar soms gaat het mis. Hier zijn twee voorbeelden:

Kunt u met uw hart op uw hand verklaren ...

Wij vinden dat dit verhaal maar bij de werkgevers neergelegd moeten worden.

Het eerste voorbeeld laat zien hoe twee geselecteerde woorden op de verkeerde plaats terecht komen. Maar het is opmerkelijk dat daarmee geen syntactische regels worden geschonden. Syntactisch is de zin in orde; alleen de beoogde betekenis gaat verloren. Syntactische unificatie is een efficiënt maar dom proces. Er wordt alleen maar op de syntaxis van woorden gelet, niet op hun betekenis. Zo'n proces dat alleen aan zichzelf denkt noemen we "modulair". Veel van de processen in Figuur 1 zijn modulair van aard.

In het tweede voorbeeld wordt de syntactische congruentie tussen onderwerp (*dit verhaal*) en persoonsvorm (*moet*→*moeten*) geschonden. Uit experimenten weten we dat dit gebeurt onder

invloed van een interveniërend zelfstandig naamwoord (hier: *werkgevers*). Zulke syntactische fouten zijn zeldzaam.

Maar dat modulaire syntactische proces kan een groot deel van zijn automaticiteit verliezen na een beroerte die de linker hersenhelft beschadigt. Een patiënt met Broca's aphasie heeft zijn syntactische kennis weliswaar niet verloren, maar hij kan een eenmaal opgebouwd syntactisch patroontje niet lang genoeg vasthouden om het verder te integreren in de voortgaande zin. Wat altijd automatisch verlopen is, vergt plotseling veel aandacht.

Fonologisch coderen

De tweede stap van het formuleren, het fonologisch coderen (zie Figuur 1) is het voorbereiden van de uitspraak. Daarvoor is het nogmaals nodig om informatie uit het lexicon op te halen, nu over de *vorm* van de geselecteerde woorden. Dat brengt ons bij het derde niveau van het netwerk in Figuur 3. We hebben experimenteel kunnen aantonen dat zogauw een lemma is geselecteerd, de sluis naar beneden wordt opengezet, zodat de activatie van het geselecteerde lemma verder spreidt naar de bijbehorende knoop op het *lexeem*-niveau. Deze knoop staat voor de fonologische vorm van het geselecteerde woord. In het netwerk is die vorm gepresenteerd door verbindingen met andere knopen op dat lexeem-niveau. De lexeemknoop /kat/, bijvoorbeeld, is enerzijds verbonden met *foneem*-knopen, namelijk met de knopen /k/, /a/ en /t/. En anderzijds is er een verbinding met een *metrische* knoop, die aangeeft uit hoeveel syllaben het woord bestaat en waar de klemtoon valt; /kat/ is verbonden met de knoop σ' en dat betekent dat het woord één syllabe heeft, waarop de klemtoon valt. De lexeemknoop /konɛin/ is verbonden met de foneemknopen /k/, /o/, /n/, /ɛi/, /n/ en met de metrische knoop $\sigma\sigma'$. Dat laatste betekent dat het woord uit twee syllaben bestaat, met accent op de laatste.

De stap van lemma naar lexeem gaat niet altijd over rozen.

Wat we allemaal af en toe meemaken is het tipje-van-de-tong verschijnsel. We zijn aan het spreken en opeens lopen we vast omdat een bepaald woord niet wil komen. Vooral eigennamen zijn notoir in dit opzicht. We weten uit allerlei onderzoek dat dat probleem ontstaat bij de activatie van het lexeem. Het lemma is er gewoon (we weten dat omdat een spreker in die toestand bijna altijd het geslacht weet van het woord dat hem niet te binnen wil schieten - hij weet of het een *de*-woord is of een *het*-woord en dat is informatie op lemma-niveau). Maar het lexeem is onvoldoende geactiveerd. Soms komt er in die toestand nog wel eens iets beschikbaar; we weten dan bijvoorbeeld met welke klank het woord begint of wat voor metriek het heeft - hoeveel syllaben en met wat voor woordaccent. Er zijn afatische patiënten voor wie de toegang tot de woordvorm een permanent probleem is. Dat zijn *anomische* patiënten. Ook die patiënten hebben normaal toegang tot de lemmas. Ze hebben weinig moeite met het construeren van zinnen; ze weten ook meestal het syntactische geslacht van het woord dat niet komen wil.

Maar als de toegang tot de woordvorm normaal, ongestoord verloopt, gaat het toch nog steeds niet altijd even gemakkelijk. Woorden die we niet vaak gebruiken worden duidelijk trager opgehaald dan woorden die we veel gebruiken. Een plaatje van een mond benoemen duurt ongeveer 0.8 seconde (d.w.z. van aanbieding plaatje tot begin articulatie). Een plaatje van een mand benoemen duurt al gauw 100 milliseconden langer. Dat heet het "woordfrequentie-effect". Het verschil wordt niet veroorzaakt doordat het meer tijd kost het plaatje van een mand te *herkennen* dan het plaatje van een mond. Het ontstaat pas bij de lexicale toegang zelf. Experimenteel hebben we kunnen aantonen dat het gaat om de toegang tot de woordvorm, het lexeem. Die is trager bij laagfrequente woorden dan bij hoogfrequente woorden.

Zover over de toegang tot die woordvorm. Maar als we die bereikt hebben, wat doen we dan als spreker met die informatie? Laten we eerst constateren dat die woordvorm niet als geheel wordt opgehaald om dan vervolgens als geheel te worden uitgevoerd. Nee, de onderdelen van een woordvorm, met name de

klinkers en medeklinkers waaruit die is samengesteld, worden apart opgehaald en dan weer op de goede plaats gezet. Dat blijkt onder andere uit versprekingen, zoals deze:

patijten (in plaats van *tapijten*)

dan kan ik dat pasj - dat tasje ook pakken

er gaat een heleboel door je heen (i.p.v. *heleboel*)

In het eerste voorbeeld worden twee medeklinkers (*p* en *t*) verwisseld binnen hetzelfde woord. In het tweede komt de *p* van *pakken* veel te vroeg en neemt de plaats in van *t* in *tasje*. Dat heet "anticipatie". In de derde verspreking wordt een klinker geanticipeerd; de *oo* van *door* komt te vroeg en wel op de plaats in *heleboel* waar de klinker *oe* had moeten komen. Normaal positioneren we zo'n 10 tot 12 fonemen per seconde, en het mechanisme is bij die hoge snelheid vrijwel feilloos. Fouten zoals deze gebeuren niet veel vaker dan één op de duizend gesproken woorden.

Fonemen worden niet zomaar op volgorde gezet; waar het om gaat is dat ze in een *metrisch frame* worden ondergebracht. En daarmee bouwen we syllaben, de belangrijkste eenheden van articulatie. Waar komt dat metrische frame vandaan, en hoe zit het in elkaar? Boven vermeldde ik al dat er tweërlig woordvorm-informatie beschikbaar komt, segmentele en metrische. In Figuur 3 worden bij activatie van *kat* de foneemknopen (*/k/*, */a/* en */t/*) geactiveerd en de metrische knoop (σ'). Voor *konijn* is het metrische frame $\sigma\sigma'$. Die metrische frames gebruiken we vaak om weer grotere samen te stellen. Als ik bijvoorbeeld zeg, "werk eraan!", dan voeg ik de frames van *werk* en van *eraan* samen tot het drie-syllabige frame $\sigma\sigma\sigma$, een dactylus. En pas dan begin ik de fonemen (*/w/*, */e/*, */r/*, etc.) op hun plaats te zetten in dat grotere frame. Het is dus niet zo dat ik eerst het één-syllabige frame voor *werk* invul en dan het twee-syllabige frame voor *eraan*. Hoe weten we dat?

Uit allerlei experimenten is ons gebleken dat fonemen "van links naar rechts" aan het frame worden opgehangen. Om het driesyllabige frame van *werkeraan* op te vullen begint de spreker vooraan, bij de eerste syllabe (σ'_1). Daaraan hangt hij op de fonemen */w/*, */e/* en */r/*. Dan begint hij met de tweede syllabe (σ'_2)

en hangt daar de /k/ en de /ə/ aan op. Tenslotte gaat hij naar de derde syllabe (σ'_3) en hangt er de laatste drie fonemen aan, /r/, /a/, en /n/. Daarmee ontstaat dus deze opdeling in syllaben: *werke-raan*. Wanneer de spreker echter eerst het woord *werk* had samengesteld en dan pas het woord *eraan*, zou de syllabificering anders zijn uitgekapt, namelijk zo: *werk-e-raan*. Maar zo spreken we niet. We laten syllaben rustig over woordgrenzen heenlopen; in de syllabe *ke* komt *k* uit *werk* en *e* uit *eraan*. Met andere woorden, we halen geen kant en klare syllaben op uit het mentale lexicon. Syllaben worden tijdens het spreken steeds opnieuw aangemaakt; het hangt er maar helemaal vanaf welke woorden we achter elkaar gaan uitspreken. Vaak koppelen we een "klein" woord (lidwoord, voorzetsel, voornaamwoord) aan een "groot" woord (een werkwoord, een zelfstandig naamwoord). Het resultaat noemen we wel een "fonologisch woord"; *werkeraan* is een fonologisch woord. En het fonologische woord is het domein van syllabificatie.

Zover samenvattend kunnen we zeggen dat een spreker incrementeel (dus "van links naar rechts") fonologische woorden bouwt, die elk weer uit fonologische syllaben bestaan. Maar die vers gevormde syllaben moeten tenslotte worden uitgesproken, en daarvoor is een uitspraakprogramma nodig. Elke syllabe is een motorisch gebaar, een geste die we maken met onze articulators. Om *wer* uit te spreken, moeten we beginnen met de mondholte even naar buiten toe af te sluiten door de onderlip tegen de boventanden te drukken (Limburgers en Surinamers gebruiken daar ook de bovenlip bij). Tegelijk moeten we onze stembanden in trilling zetten en dan die onderlip weer terugtrekken, etc. Wat ik hier beschrijf is de "score" van dit syllabische gebaar, althans het begin van die score.

Onze recente experimenten (Levelt en Wheeldon, 1994) hebben het waarschijnlijk gemaakt dat wij als sprekers een "syllabegeheugen" (*syllabary*) in ons hoofd hebben (zie Figuur 1). Voor elke fonologische syllabe die we opbouwen (bijvoorbeeld voor *wer*, voor *ke*, voor *raan*) zit in dat geheugen de bijbehorende score. Zogauw een fonologische syllabe klaar is wordt de correspon-

derende score uit dat geheugen opgehaald en ter beschikking gesteld van de articulatorische component. Wij hebben gevonden dat een veel gebruikte syllabescore sneller uit dat geheugen wordt gehaald dan een weinig gebruikte syllabescore. En dat effect is onafhankelijk van het boven besproken woordfrequentie-effect.

Het uiteindelijke product van fonologisch coderen is een rij syllabische scores, gegroepeerd per fonologisch woord. Daarmee moet de volgende component zijn werk doen.

Articuleren

Ons articulatorische apparaat, longen, keel en neusholte, mond, tong, lippen, bestaat uit meer dan honderd verschillende spieren. Die zijn er oorspronkelijk niet om mee te praten, maar om mee te ademen, te eten en te drinken. Het is een gril van de evolutie dat diezelfde apparatuur voor een andere functie gebruikt is gaan worden, het spreken. En dat is dan ook nog de meest verfijnde motoriek waar we überhaupt over beschikken. Maar de andere functies zijn bewaard gebleven, en wonder boven wonder wordt die dubbelrol heel redelijk vervuld. Behalve wanneer we beide functies tegelijk willen uitvoeren, praten tijdens het eten. Dan verslikken we ons nogal eens, en alleen mensen doen dat.

Bij de uitvoering van de articulatorische gebaren (waarop ik hier niet verder in zal gaan) zijn verschillende delen van de hersenen betrokken, niet alleen de motorische schors, maar ook de premotorische schors, de kleine hersenen en enkele basale ganglia. Elk van deze gebieden heeft zijn eigen functie en er is dan ook helaas een grote scala van articulatorische spraakstoornissen te observeren bij cerebrale stoornissen. Vloeiende spraakmotoriek veronderstelt feilloze samenwerking tussen al deze gebieden. Er kunnen al hoorbare problemen ontstaan bij timingsfouten van zo'n 20 milliseconden.

Zelfcontrole

Rechts in Figuur 1 staat de component aangegeven die zich bezig houdt met taalverstaan. Die is niet in detail ingevuld, omdat dit artikel niet gaat over taalverstaan. Toch is die component ook bij het spreken enigszins betrokken. We kunnen naar onze eigen innerlijke of overte spraak luisteren en een beetje letten op wat we aan het doen zijn. Soms valt ons dan een probleem op, een fout of iets wat we niet duidelijk geformuleerd hebben. Het "valt ons op", d.w.z. we besteden er even aandacht aan. Ofschoon we als sprekers onze aandacht vooral besteden aan het voorbereiden van de "boodschap", is er soms toch even gelegenheid voor zulke zelfcontrole. Dat doen we vooral tegen het einde van zinnen en van grote zinsdelen.

Dat kan er toe leiden dat we even stoppen om een oneffenheid glad te strijken. We hebben kunnen aantonen (Levelt, 1983) dat een spreker vrijwel direct stopt wanneer hij een ernstig probleem detecteert. Hij doet daarbij niet de geringste poging om het zinsdeel waarmee hij bezig is netjes af te ronden; desnoods stopt hij midden in het probleemwoord, zoals hier:

Eerst een brui-, uh geel en een groen rondje.

En als hij toch nog een of meer syllaben doorpraat na het probleempunt, dan komt dat doordat hij de fout nog niet meteen ontdekt heeft. Dat gebeurt bijvoorbeeld hier:

En van groen linksaf naar roze, of uh, van blauw linksaf naar roze.

Hier is *groen* het probleemwoord.

Verder hebben we kunnen aantonen dat de spreker die onderbroken syntaxis nog even vasthoudt. Daar kom ik direct op terug.

Na de zelfonderbreking gebeurt het nogal eens dat de spreker een of ander tussenwerpsel produceert, zoals *uh, sorry, offe*, enz. Dat tussenwerpsel geeft vaak aan wat voor een probleem het was. Sprekers gebruiken *sorry* en *nee* en *of* bijna uitsluitend wanneer er een regelrechte fout gemaakt is. Maar *dus* wordt vooral gebruikt wanneer een nadere specificatie nodig is, zoals hier:

Gaan we weer terug naar dat zwarte bolletje, dus het tweede bolletje. (1)

En dan komt de correctie zelf. Zojuist merkte ik op dat de spreker de onderbroken syntaxis meestal nog even vasthoudt. Die gebruikt hij dan als startpunt voor zijn correctie. We kunnen meteen horen of de spreker dat goed doet. De vorige correctie zal bijvoorbeeld nooit zó worden gemaakt:

Gaan we weer terug naar dat zwarte bolletje, dus we weer terug naar het tweede bolletje (2)

Waarom niet? Wat de spreker doet is vrijwel hetzelfde als het maken van een nevenschikking. De geobserveerde zelfcorrectie in (1) komt overeen met de (goed klinkende) nevenschikking:

Gaan we weer terug naar dat zwarte bolletje of het tweede bolletje

De onmogelijke zelfcorrectie (2) komt echter overeen met een al even onmogelijke nevenschikking:

Gaan we weer terug naar dat zwarte bolletje of we weer terug naar het tweede bolletje.

Sprekers zijn dus syntactisch nogal principiëel bij het uitvoeren van een zelfcorrectie, en dat maakt het voor de hoorder weer een beetje makkelijker. Het is meteen duidelijk hoe de correctie verband houdt met de oorspronkelijke onderbroken uiting. Het maakt het ook makkelijk voor de radio-omroep. Het lukt heel vaak om zelfcorrecties "weg te lassen" in een bandopname. Voorbeeld (1) wordt dan:

Gaan we weer terug naar het tweede bolletje

met een goede kans dat niemand de las zal opmerken. En dit lukt alleen maar omdat de spreker zelf syntactisch zo netjes te werk is gegaan.

Spreeksnelheid 2

Na deze vogelvlucht over de componenten in Figuur 1, keren we tenslotte nog even terug naar waarmee we zijn begonnen, de spreeksnelheid. Hoe snel kan dat hele systeem functioneren? We

weten al dat conceptuele voorbereiding relatief traag verloopt, gemiddeld zo'n vier seconden per elementaire boodschap. Maar de rest van het systeem is razendsnel. Dat weten we uit zg. *shadowing* experimenten. We laten de proefpersoon dan luisteren naar een gesproken tekst, met de instructie die tekst op het gehoor mee te spreken. En we benadrukken dat hij er "zo dicht mogelijk op moet blijven zitten". We krijgen dan zóiets:

Band:

De bedoeling was om vooral in de bergen wat rond te trekken.

Dan is het erg lekker ...

Proefpersoon:

... De bedoeling was om vooral in de bergen wat rond te trekken.

Dan is het erg ...

Om zo mee te praten moet de proefpersoon de tekst beluisteren en enigszins begrijpen (de rechter componenten in Figuur 1). En hij moet de tekst weer zelf formuleren en articuleren (de linker componenten in Figuur 1). Dat dat begrijpen en weer formuleren echt plaatsvindt, weten we uit proefjes waarbij kleine semantische of syntactische fouten worden gemaakt in de tekst op de band. De na-apende proefpersoon herstelt zulke fouten bijna altijd. Kortom, het grootste deel van het spraakproductiesysteem is bij deze taak betrokken. En de vertraging van de proefpersoon is dan een slordige meting van hoe snel het systeem doorlopen kan worden.

De meeste proefpersonen hebben vertragingen van 1.5 tot 3 seconde. Maar er zijn ook snelle na-apers (*close shadows*), die niet méér achterkomen dan een kwart seconde. Dat is ongeveer de duur van een enkele syllabe. In die korte tijd kunnen alle componenten hun werkje blijkbaar afkrijgen. En 250 milliseconde is niet meer dan een stuk of twintig opeenvolgende synaptische schakelingen in de hersenen. Dit wijst er nogmaals op dat er heel veel *parallel* wordt verwerkt binnen en tussen componenten. Spreken is eigenlijk alles tegelijk doen. En niemand klaagt daarover.

REFERENTIES

- Kempen, G. A. M. (1993). *Spraakkunst als bouwkunst*. Oratie Leiden.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41-104.
- Levelt, W. J. M. (1993, tweede druk). *Speaking. From intention to articulation*. Cambridge, Mass.: MIT Press.
- Levelt, W. J. M. (1994). Perspective talking and ellipsis in spatial descriptions. Manuscript ter perse.
- Levelt, W. J. M. en Wheeldon, L. (1994) Do speakers have access to a mental syllabary? *Cognition*, 50, 239-269.
- Roelofs, A. (1992). A spreading activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142. Ook verschenen in W. J. M. Levelt (Ed.), *Lexical access in speech production*. Cambridge, Mass.: Blackwell.