

Het unieke menselijke taalvermogen: van PAUS naar [pɑus] in een halve seconde

Peter Hagoort

Communicatie door middel van natuurlijke taal is uniek voor de mens. Het wordt uitgevoerd in verschillende vormen, zoals praten, luisteren, schrijven, lezen, en in gebarentaal. In al deze gevallen hebben we te maken met complexe cognitieve vaardigheden bestaande uit een cascade van verwerkingsprocessen. Gedurende deze verwerkingsprocessen worden relevante bronnen van informatie uit het geheugen opgehaald. Deze bronnen van informatie omvatten conceptuele kennis, kennis over de syntaxis van woorden en zinnen en kennis over spraakgeluiden of de orthografische patronen in lezen en schrijven. De verschillende 'ophaal'-processen vinden met hoge snelheid plaats en worden temporeel georkestreerd met een precisie in de orde van milliseconden. Recente 'brain-imaging' technieken maken onderzoek naar de neurale fundamenteën van menselijke taalvaardigheden mogelijk. Het blijkt dat zelfs voor de relatief simpele taalkaak van het produceren van éénwoorduitingen de gezamenlijke actie van een aantal gebieden in met name de linkerhersenhelft vereist is. In de laatste paragraaf worden de implicaties voor een theorie over de mens kort besproken.

Inleiding

Geen andere soort dan homo sapiens heeft in de loop van zijn evolutionaire geschiedenis een communicatiesysteem ontwikkeld, waarin een eindig aantal symbolen samen met een reeks van regels voor het combineren daarvan een oneindig aantal uitdrukkingen mogelijk maken. Dit natuurlijke taalsysteem stelt leden van onze soort in staat om gedachten uiterlijke vorm te geven en uit te wisselen met de sociale groep en, door de uitvinding van schriftsystemen, met de gehele samenleving. Spraak en taal zijn effectieve middelen voor het beheer van sociale cohesie in samenlevingen waar dit door de groepsgrootte niet langer kan door middel van 'vlooiën', de wijze waarop onze genetische buren, de primaten van de oude wereld, zich bij voorkeur aan elkaar binden (Dunbar, 1996; Levelt, 1999).

De generatieve kracht van het menselijk taalsysteem rust op zijn driedelige architectuur. (Jackendoff, 1997; 1999). In deze architectuur wordt taal-relevante informatie gecodeerd in tenminste drie afzonderlijke representaties: een voor betekenis, een voor syntaxis en een voor de klankstructuren van woorden en uitingen. Door deze representatie-structuren met elkaar in verband te brengen (op elkaar af te beelden), worden de conceptuele

structuren die de inhoud van de spreker's boodschap specificeren uitgedrukt als een lineaire reeks van spraakklanken (spreken). Anderzijds worden tijdens het luisteren naar spraak conceptuele structuren afgeleid van een lineaire aaneenschakeling van spraakklanken. In dit afbeeldingsproces is het combineren van eenheden tot hiërarchische constituentenstructuren (syntaxis) de noodzakelijke link tussen conceptuele structuren en fonologische (of klank-) structuren.

In de verdeling van werk tussen de wetenschappen die het menselijke taalvermogen onderzoeken, is het de taak van de linguïst de betrokken representatie-structuren te specificeren; en is het de taak van de psycholinguïst de toegang tot en het gebruik van deze structuren tijdens luisteren en spreken te onderzoeken. De cognitief neurowetenschapper gaat de uitdaging aan te specificeren hoe de hersenen menselijke taal mogelijk maken en het spatieel-temporeel profiel van neurofysiologische activiteit vast te stellen dat ten grondslag ligt aan spreken (schrijven) en luisteren (lezen).

Om een preciezer beeld te vormen over wat deze verschillende wetenschappen onthullen over het menselijk taalvermogen, moeten we eerst verduidelijken waar de allesomvattende term 'menselijk taalvermogen' voor staat. Deze verwijst naar de verzameling van de volgende complexe en verwante, maar tegelijkertijd afzonderlijke, vaardigheden: praten en luisteren, lezen en schrijven, en in dovensculturen het gebruik van gebarentaal. Elk van deze vaardigheden vereist dat afzonderlijke representatie-structuren in het geheugen worden opgehaald en gebruikt in 'real time'. De cognitieve architecturen voor deze vaardigheden specificeren welke representatie-structuren hierbij betrokken zijn en hoe deze worden bewerkt. De neurale architecturen specificeren de wijze waarop deze vaardigheden worden geconcretiseerd in the 'wetware' van het menselijk brein. Men moet echter in gedachten houden dat het onderscheid tussen de cognitieve architectuur en de neurale architectuur een idealisering is. Voor een eerste benadering is het nuttig om een onderscheid te maken tussen processpecificaties in symbolische termen (cognitieve architectuur) en in neurofysiologische termen (neurale architectuur), maar voor een algehele cognitieve neurowetenschap van taal zouden deze niveaus bij elkaar gebracht moeten worden.

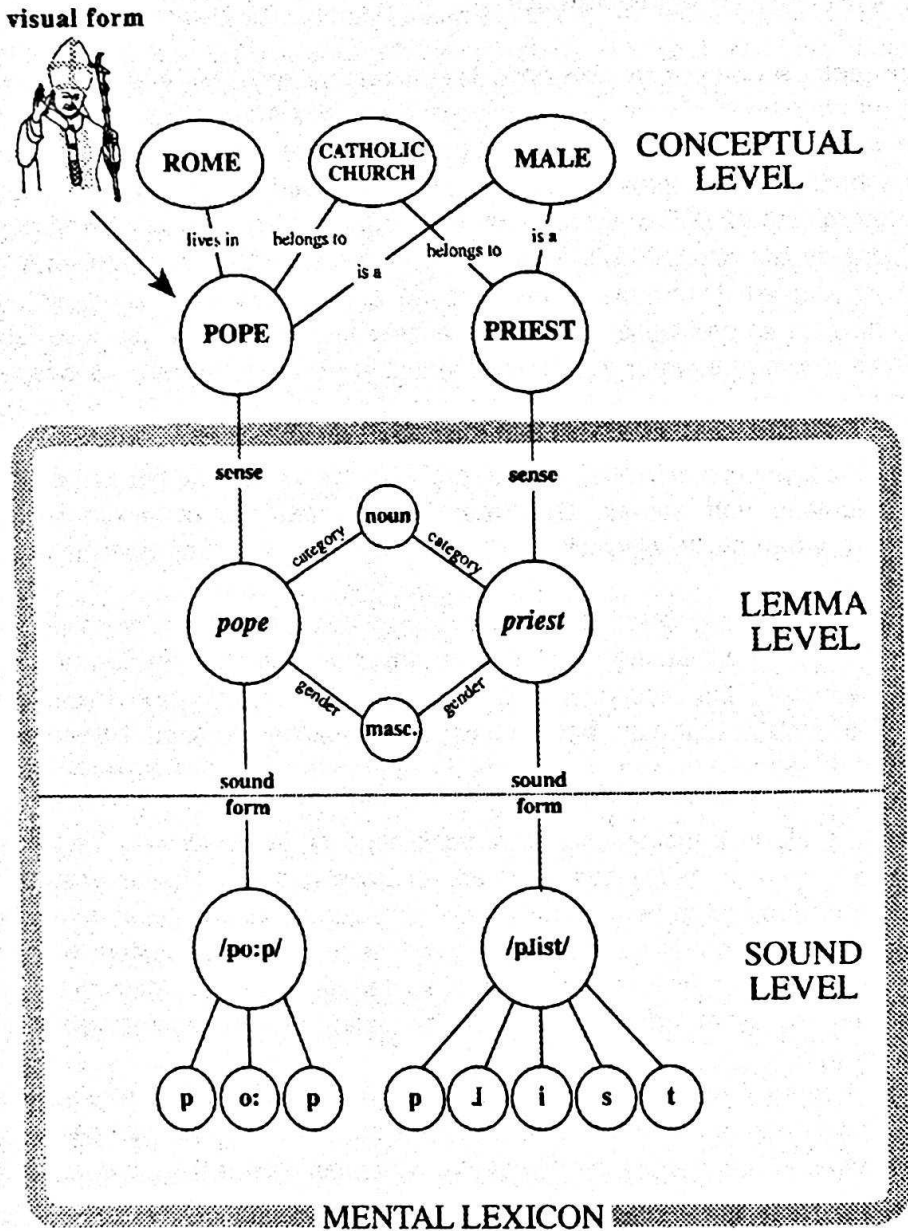
In het vervolg van dit hoofdstuk bespreek ik eerst de cognitieve architectuur van een van deze vaardigheden in detail en daarna komt de neurale architectuur aan bod. Tenslotte zal ik enkele implicaties schetsen voor een theorie van de mens.

De cognitieve architectuur

Een centrale component van onze taalvaardigheden is het mentale lexicon. Het mentale lexicon is het deel van het declaratief geheugen dat de kennis die een taalgebruiker heeft van de woorden in zijn/haar moedertaal specificiert. De schatting is dat sprekers van een taal een actieve woordenschat hebben van tenminste 40.000 woorden (voor een algemene inleiding van het mentale lexicon, zie Aitchison, 1987). Sprekers weten wat deze woorden betekenen en hoe ze klinken. Daarnaast kennen ze de syntactische eigenschappen van woorden als woordklasse (zelfstandig naamwoord, werkwoord, etc.). Al deze lexicale informatie wordt zeer snel opgehaald uit het geheugen. Gemiddeld produceert een spreker twee of drie woorden per seconde. Dit vereist niet alleen het ophalen van verschillende bronnen voor woordinformatie, maar ook de gecoördineerde activatie van een groot aantal spieren dat betrokken is bij de articulatie van spraak. Ongeveer 100 spieren zijn betrokken bij het spreken, waarvan de activiteit op de milliseconde nauwkeurig gecoördineerd moet worden.

Ondanks de complexiteit en snelheid van deze cognitieve activiteit zijn sprekers zeer nauwkeurig en maken ze gemiddeld minder dan één fout per 1000 woorden. De incidentele spreekfouten zijn niettemin zeer informatief voor de architectuur van het spreekproces. Klanken kunnen bijvoorbeeld verwisseld worden tussen verschillende woorden zoals in 'heft /emisphere' (in plaats van 'left hemisphere'; Fromkin, 1973), of kunnen te vroeg geproduceerd worden zoals in anticipaties ('it's a meal mystery' in plaats van 'it's a real mystery'; Fromkin, 1973). Wat deze voorbeelden van spreekfouten illustreren is dat woorden niet in het geheugen zijn opgeslagen als eenheden, maar dat ze, elke keer wanneer we een woord produceren, gevormd moeten worden uit de constituerende fonemen. Dit proces van samenvoegen gaat zo nu en dan verkeerd, waardoor klanken in de omgeving van het verkeerde woord terechtkomen.

Figuur 1 toont een blauwdruk voor het spreken van éénwoorduitingen. Het verduidelijkt wat er gebeurt tussen het moment dat we een bepaald beeld op het netvlies herkennen, bijvoorbeeld van Johannes Paulus II, en het moment dat we de feitelijke articulatie van de klankenstroom 'paus' produceren.



Figur 1: Het model van Levelt en Roelofs voor het spreken van éénwoorduitingen (zie Levelt, 1989, Roelofs, 1992). Conceptuele knopen (POPE) worden geactiveerd op basis van sensorische en/of conceptuele input. Activatie van een conceptuele knoop spreidt naar zijn lemmaknoop (pope) in het mentale lexicon. Activatie van POPE leidt ook tot verhoogde activatie van gerelateerde concepten in het semantisch geheugen, zoals PRIEST. Elke conceptuele knoop is verbonden met precies één lemma in

het lexicon. Op het lemmaniveau wordt syntactische woordinformatie gespecificeerd, zoals grammaticaal geslacht en woordcategorie. Bijvoorbeeld, in het Italiaans is het geslacht van het lemma pope mannelijk (il papa), terwijl het geslacht van het Italiaanse lemma voor kerk vrouwelijk is (la chiesa). Nadat het lemma is geselecteerd, wordt de woordvorminformatie opgehaald en voorbereid voor articulatie.

Spreeken begint met het specificeren van de conceptuele inhoud van de uiting. Dit specificeren kan bepaald worden door visuele input, zoals in ons voorbeeld. Maar in veel gevallen wordt de conceptuele specificatie bepaald door interne input, bijvoorbeeld de spreker's intentie om een bepaald idee uit te drukken. Om het even wat de conceptuele specificatie van de uiting in gang zet, de spreker moet een bepaald concept of een reeks concepten selecteren uit het kennisbestand in het geheugen, en hij/zij moet een uitingwijze kiezen en daarover een beslissing nemen (bijvoorbeeld, een boodschap kan worden uitgedrukt als een bewering, maar ook als een vraag, met of zonder ironie, etc.). In ons voorbeeld kan de spreker bijvoorbeeld beslissen 'Johannes Paulus de Tweede' of 'de paus' te zeggen, om maar twee mogelijke alternatieven te noemen.

De conceptuele selectie en het specificatieproces gaan vooraf aan het feitelijke proces van formulering, waarin preverbale conceptuele structuren het ophalen van linguïstische structuren in gang zetten, wat nodig is om het idee uit te drukken als een reeks spraakklanken. Hier moeten twee volledig verschillende typen linguïstische informatie worden opgehaald, één die eigenschappen van de woordklank specificeert, de ander betreft de grammaticale eigenschappen van een woord.

Elke woordvorm in het mentale lexicon is geassocieerd met syntactische woordinformatie (Levelt, 1989, 1999; Roelofs, 1992, 1993). Deze laatste vorm van informatie wordt lemma-informatie genoemd. Lemma's specificeren de syntactische eigenschappen van woorden, zoals hun woordklasse (zelfstandig naamwoord, werkwoord, bijvoeglijk naamwoord, etc.). Bij zelfstandige naamwoorden in geslachtsgemarkeerde talen is het grammaticale geslacht eveneens gespecificeerd (bijvoorbeeld, *paard* in het Frans is mannelijk, in het Nederlands is het onzijdig). Werkwoordlemma's bezitten informatie over syntactische frames (de argumentstructuren) en over de thematische rollen van de syntactische argumenten (de thematische structuur). Bijvoorbeeld, het lemma voor het werkwoord *schénken* specificeert dat het een nominale constituent ('noun-phrase', NP) nodig heeft als grammaticaal subject, een NP als grammaticaal object, met de optionele toevoeging van een voorzetselconstituent ('prepositional phrase', PP) als indirect object (bijvoorbeeld, *Johannes* <subject-NP> *schénkt een boek* <direct object-NP> *aan de bibliotheek* <optioneel indirect object-PP>). Daarnaast wordt gespecificeerd hoe dit syntactische frame wordt afgebeeld op thematische

rollen). Bij *schenken* is het subject de agens, het direct object het *thema*, en het indirect object het *doel* of de *ontvanger* van de actie uitgedrukt door het predikaat.

In de volgende fase van het formuleringsproces zet de selectie van de juiste lemma's het ophalen van klankpatronen van de uiting in gang (zie figuur 1). Tijdens deze fase komen de spraakklanken (fonemen) van het woord beschikbaar. Naast de fonemen wordt ook de metrische informatie van een woord opgehaald. Deze specificiert het aantal syllaben en het leggen van de klemtoon (niet getoond in figuur 1). Gedurende een verwerkingsstap die fonologische codering wordt genoemd, worden de fonemen toegewezen aan hun syllabeposities, in volgorde van links naar rechts. Het resultaat van fonologische codering is een 'fonologisch woord' met woordklankinformatie als een opeenvolging van syllaben met het juiste klemtoonpatroon. Syllaben zijn de codes die de basis vormen voor de articulatorische bewegingen van de stembanden, het velum, de tong, de kaak en de lippen. Het zijn abstracte codes omdat ze onafhankelijk zijn van de beginposities van, bijvoorbeeld, de lippen en de tong. Spreken met of zonder een pijp in de mond resulteert in verschillende articulatorische bewegingspatronen, welke niettemin geïnstrueerd worden door dezelfde abstracte syllabencodes. Het uiteindelijke resultaat van deze hele cascade van ophaal- en activatieprocessen is een akoestisch signaal dat de luisteraar gebruikt om de bedoelde boodschap af te leiden.

Los van het experimentele bewijs voor een onderscheid tussen het ophalen van lemma- en van woordvorminformatie, zijn we allen bekend met een fenomeen dat dit onderscheid ondersteunt. Dit is de zogenoemde 'tip-of-the-tongue state', wat verwijst naar de vaak gênante situatie waarin we weten dat we het woord kennen, we weten zelfs dat het, bijvoorbeeld, een zelfstandig naamwoord is met een bepaald grammaticaal geslacht (bijvoorbeeld, onzijdig), maar om de een of andere reden wordt het ophalen van de klankvorm belemmerd. Het feit dat we toegang hebben tot bepaalde aspecten van woordinformatie, maar andere aspecten van woordinformatie niet kunnen ophalen, illustreert het idee dat verschillende aspecten van woordinformatie differentieel worden opgeslagen en opgehaald. Het simplistische idee dat woorden eenheden zijn die ergens in de hersenen te vinden zijn is onjuist. 'Woord' is niets meer dan een handige kreet voor een diversiteit aan informatie. Ondanks verschillen in de details van diverse modellen voor spreken, zijn onderzoekers naar taalverwerking het er in het algemeen over eens dat een temporeel georkestreerd ophaalproces van de verschillende typen informatie, zoals hiervoor besproken, vereist is.

Het spreekproces heb ik tot dusver besproken als een 'feedforward' proces, van intentie naar articulatie (Levelt, 1989). Echter, introspectief bezien hebben we vaak het gevoel dat de manier waarop we onze gedachten uiten, onze intenties verscherpt en vormt. Dat wil zeggen, spraak begint met een intentie, maar intenties zijn ook (deels) afgeleid van spraak. Deze intuïtie heeft tot

kritiek geleid op een visie die taalgebruik beschouwd als een eenvoudige informatiestroom van intentie naar articulatie (cf. Dennett, 1991). Echter, mijns inziens kan met deze kritiek eenvoudig afgerekend worden als we ons realiseren dat we als sprekers tegelijkertijd ook luisteraars zijn. Dat wil zeggen, we luisteren naar onze eigen spraak, gebruikmakend van dezelfde mechanismen waarmee we het spreken van anderen analyseren. Gedurende het luisterproces leiden we de intentie af uit de spraakklanken die ons oor bereiken. Spreken is een hoogst incrementeel proces, hetgeen betekent dat we nog niet alle details van onze preverbaal boodschap hebben gespecificeerd voordat we beginnen met het formuleringsproces. De incrementele aard van spraakplanning biedt ons de gelegenheid om al luisterend naar onze eigen spraak onze intenties te verscherpen en te vormen en via deze route het almaar voortschrijdende formuleringsproces te beïnvloeden. Aangezien ons cognitieve taalapparaat zowel een productie- als een begripscomponent heeft, instantieert het een ononderbroken interne dialoog tussen 'spreker' en 'luisteraar', resulterend in het introspectieve gevoel dat intenties niet alleen de bron maar ook het bij-product van spreken zijn.

In het voorbeeld hierboven heb ik een grove schets gegeven van de cognitieve architectuur van het spreken, hoofdzakelijk het spreken van éénwoorduitingen. Een volledig model van spreken specificeert ook hoe woorden gecombineerd worden tot langere uitingen, hoe de intonatiecontouren van meerwoorduitingen worden bepaald, etc. Soortgelijke modellen kunnen worden gemaakt voor luisteren, lezen en schrijven (voor gedetailleerde voorbeelden, zie Brown & Hagoort, 1999). In al deze gevallen is het vaststellen van de details van de cognitieve architecturen voor de verschillende taalvaardigheden gebaseerd op een combinatie van conceptuele analyse, computationeel modelleren en slimme experimentatie. Aan de hand van deze cognitieve architecturen kunnen we zinnige vragen stellen over de neurale concretisering van de verschillende taalvaardigheden. Zonder dergelijke expliciete modellen is het onderzoek naar de neurale mechanismen van taal gedoemd te mislukken.

Ter illustratie van wetenschap in actie zal ik een simpel experiment in enig detail bespreken. Als lexicale concepten zoals PAUS en PRIESTER zijn opgeslagen in een netwerk zoals getoond in figuur 1, en als de activatie van een bepaald concept gedeeltelijk spreidt naar naburige conceptknopen in het netwerk, dan worden bepaalde gevolgen voorspeld voor de verwerking van woorden die worden voorafgegaan door semantisch gerelateerde woorden. Dit wordt op de volgende manier getest. Proefpersonen die meedoen aan het experiment zien woordparen die kort worden getoond op een computerscherm. Als eerste wordt één woord gedurende een halve seconde op het scherm getoond, gevolgd door een blanco scherm gedurende een paar honderd milliseconden. Daarna wordt het doelwoord gedurende een halve seconde getoond. De proefpersonen worden geïnstrueerd om het tweede

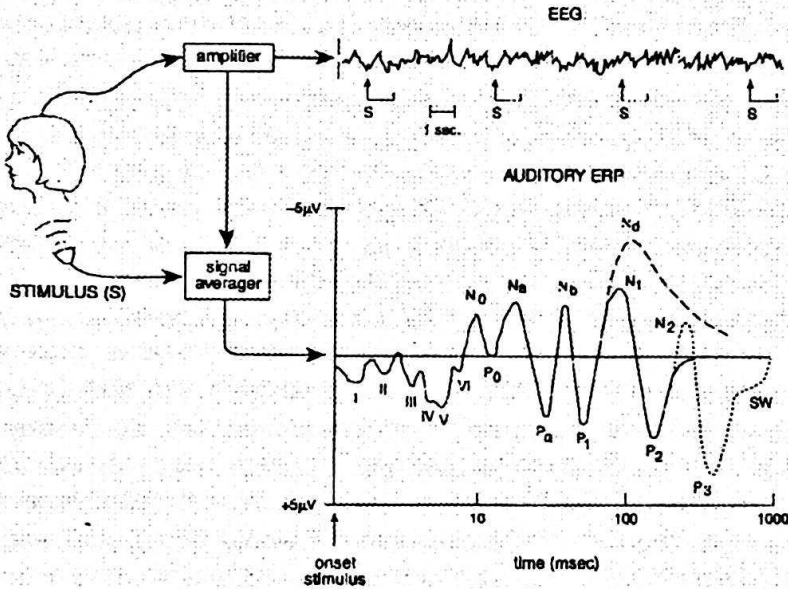
woord hardop te lezen zodra het op het scherm verschijnt. De aanbidding van het tweede woord start een klok die gestopt wordt door de verbale respons van de proefpersoon. Zodra de proefpersoon het woord op het scherm hardop leest, stopt de klok. Hiermee wordt de reactietijd van de proefpersoon gemeten. In één conditie van het experiment zien de proefpersonen voor het doelwoord (bijvoorbeeld, 'priester') een semantisch gerelateerd woord ('paus'). In de andere conditie is het eerste woord niet in betekenis gerelateerd aan het tweede woord (bijvoorbeeld 'paard'). De voorspelling van het netwerkmodel is dat het zien van het woord 'paus' tot gedeeltelijke activatie van het woord 'priester' leidt door de verbinding tussen de concepten PAUS en PRIESTER. Wanneer 'priester' direct na 'paus' op het scherm verschijnt, zou het lezen van 'priester' sneller moeten gaan dan wanneer het geïsoleerd wordt aangeboden, aangezien het al deels geactiveerd is door het voorafgaande woord 'paus'. Maar het woord 'paard' spreidt niet een deel van zijn activatie naar 'priester', aangezien de concepten PAARD en PRIESTER te ver van elkaar af liggen in het semantisch netwerk om elkaar te beïnvloeden. Dus als 'priester' direct na 'paard' wordt gelezen, zou dit het leesproces niet versnellen. De resultaten van dit soort experimenten komen overeen met de voorspellingen van het netwerkmodel. Proefpersonen zijn een aantal tienden van milliseconden sneller in het lezen van 'priester' voorafgegaan door 'paus' dan in het lezen van 'priester' voorafgegaan door 'paard'. Dit zogenoemde priming-effect suggereert dat informatie over de betekenis van woorden in het geheugen is opgeslagen als een netwerk van verbonden delen van informatie en niet als geïsoleerde pakketjes individuele woordbetekenissen.

De neurale architectuur

De neurale architectuur specificereert de spatio-temporele dynamiek van de hersenprocessen die het netvliesbeeld van bijvoorbeeld Johannes Paulus II omzetten naar de spraakklank [pɑus]. We moeten dus specificeren welke gebieden van de hersenen geactiveerd worden gedurende deze processen en hoe de hiermee samengaande activaties temporeel georkestreerd worden.

Recentelijk heeft het meten van elektrische hersenactiviteit geresulteerd in een tamelijk gedetailleerde schatting van het tijdsverloop van de verschillende betrokken processen. Gewoonlijk wordt elektrische hersenactiviteit gemeten door middel van een serie elektrodes op de schedel. Wanneer deze metingen in de tijd gekoppeld aan sensorische, motorische of cognitieve processen worden uitgevoerd, resulteert dit in op de schedelgemeten 'event-related' hersenpotentialen ('event-related brain potentials', ERPs). Deze reflecteren het totaal van simultane postsynaptische activiteit van een grote verzameling

neuronen. De ERPs hebben een hoge temporele resolutie, in orde van grootte van enkele milliseconden. Op basis van de latentie van deze componenten kan relevante informatie verkregen worden over het tijdsverloop van de onderliggende processen (zie figuur 2).



Figuur 2 (naar Hillyard & Kutas, 1983): Geïdealiseerde golfvorm van een serie ERP-componenten die zichtbaar worden na middeling van het EEG van herhaalde presentaties van een korte auditieve stimulus. In deze figuur wordt het EEG gemeten met één elektrode, geplaatst op een centrale middenlijn locatie op de schedel. Gewoonlijk is het nodig te middelen over een aantal stimuli om een adequate signaal-ruis ratio te verkrijgen. Langs de logaritmische tijdsas worden de vroege hersenstampotentialen (Golven I-VI), de midlatentie componenten (No, Po, Na, Pa, Nb), de voornamelijk exogene componenten (P1, N1, P2), en the endogene, cognitieve ERP-componenten (Nd, N2, P300, Slow Wave) getoond. De componenten met een negatieve polariteit zijn naar boven uitgezet, de componenten met een positieve polariteit zijn naar beneden uitgezet. De exogene componenten geven vooral de fysieke stimulseigenschappen weer (bijvoorbeeld, intensiteit, grootte, duur). Met name de endogene componenten geven de gevolgen weer van stimulatie van (een van) de sensorische systemen voor cognitieve informatieverwerking.

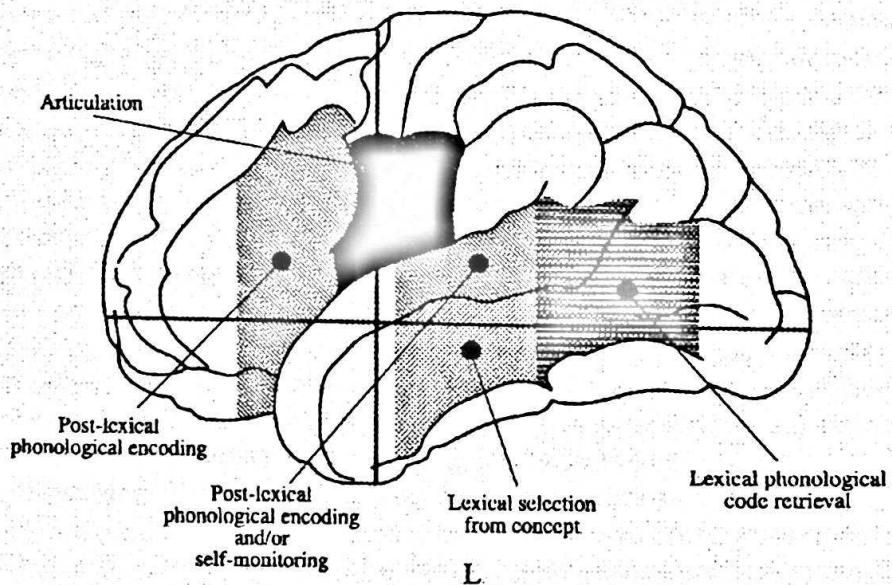
Hoewel om bepaalde redenen (cf. Hagoort & van Turenhout, 1997) ERPs vooral door taalwetenschappers gemeten werden in relatie tot aspecten van taalbegrip, zijn deze technieken in recent onderzoek succesvol toegepast bij het bestuderen van taalproductie (van Turenhout, Hagoort, & Brown, 1997, 1998). Zonder in details te treden kan er op basis van deze en andere studies een gefundeerde uitspraak worden gedaan over de temporele dynamiek van het omzetten van een visueel beeld naar een gearticuleerde reeks spraakklanken (zie Hagoort & van Turenhout, 1997, voor meer details). In het geval van ons voorbeeld zal het tenminste 150 milliseconden duren om het retinale beeld waar te nemen en te categoriseren als Johannes Paulus II (Thorpe, Fize, & Marlot, 1996). De navolgende activatie van het concept PAUS duurt minder dan 200 milliseconden. Het wordt gevolgd door een cascade van ophaalprocessen. Ongeveer 40 milliseconden voor het ophalen van het begin-foneem van het woord [pɑus] worden de syntactische eigenschappen van *paus* (de lemma-informatie) geactiveerd (van Turenhout, Hagoort, & Brown, 1998). Het is belangrijk op te merken dat de informatie over de fonologische vorm van een woord niet meteen beschikbaar is, maar zich opbouwt in volgorde van links naar rechts. Bij woorden van drie fonemen /p//ɑu//s/ duurt het nog maximaal 80 milliseconden om de resterende segmenten op te halen als eenmaal het begin-foneem beschikbaar is (van Turenhout, Hagoort, & Brown, 1997). Aangezien het ongeveer 600 milliseconden duurt voordat de articulatie van het woord [pɑus] aanvangt, is de resterende tijd nodig voor het prepareren (en deels uitvoeren) van het articulatorische motorprogramma op basis van de fonologische informatie.

Naast het beantwoorden van de vraag over het tijdsverloop, moeten we de hersengebieden specificeren die betrokken zijn bij de cascade van verwerkingsoperaties betrokken bij het spreken. Hiervoor moeten we vertrouwen op evidentie van laesiedata of op metingen van hersenactiviteit met behulp van moderne 'brain-imaging' technieken zoals 'Positron Emissie Tomography (PET)' en 'functionele Magnetische Resonantie Imaging (fMRI)'. Laesiedata zijn afkomstig van patiënten die een beroerte of een hersentumor hebben gehad, waardoor een taalstoornis is ontstaan. Op basis van precieze analyse van enerzijds de locatie en grootte van de laesie en anderzijds de specifieke aard van de taalstoornis wordt vastgesteld welke gebieden van de hersenen een bepaald aspect van taalverwerking voor hun rekening nemen. Bijvoorbeeld, laesies in de frontale cortex inclusief het gebied van Broca resulteren vaak in een verstoring van het produceren van de juiste klankpatronen van de bedoelde woorden. Dit suggereert dat onder normale omstandigheden dit hersengebied betrokken is bij onder andere het samenvoegen van de klankpatronen van een woord. Op deze manier wordt de relatie tussen de locatie van de laesie en de taalstoornis gebruikt om taalfuncties toe te schrijven aan hersenstructuren.

PET en fMRI meten hemodynamische signalen. Ze maken de detectie en visualisatie mogelijk van functioneel geïnduceerde veranderingen in locale doorbloeding van het brein (PET) of veranderingen in zuurstofvoorziening van het bloed (fMRI), welke verondersteld worden te correleren met de activatie van nabij gelegen neurale weefsel (voor een algemene inleiding, zie Raichle, 1994). Op deze wijze wordt de locus van neurale activiteit die gerelateerd is aan cognitieve processen vastgesteld door middel van een vasculair filter. Dit impliceert dat de temporele resolutie van deze technieken inherent beperkt wordt door de temporele dynamiek van veranderingen in bloedstroom of bloedzuurstofvoorziening. Deze heeft een ordegrrootte van honderden milliseconden tot enkele seconden. Dit staat in contrast met de electrofysiologische metingen die direct gerelateerd zijn aan neurale activiteit en een temporele resolutie hebben in de orde van grootte van milliseconden. De ERP-metingen echter lijden onder het zogenoemde inverseprobleem, wat het lokaliseren van de op de schedel gemeten elektrische potentialen bemoeilijkt. Voorlopig verschaffen alleen PET en fMRI metingen de vereiste spatiële resolutie.

De afgelopen jaren zien we een groeiend aantal PET- en fMRI-studies naar taalverwerking. Door afwezigheid van een diemodel voor taal zijn we sterk afhankelijk van deze nieuwe 'brain-imaging' technieken om de hersenen in actie te zien tijdens taaltaken. De volgende logica ligt ten grondslag aan de meeste 'brain-imaging' studies naar taal. De hersenactivatiepatronen die samenhangen met taken die een bepaald proces vereisen (bijvoorbeeld, het ophalen van lemma-informatie) in de cascade van processen betrokken bij het spreken worden vergeleken met activatiepatronen die samenhangen met taken waar dit specifieke proces niet bij betrokken is. Door deze vergelijking kan men vaststellen welke hersengebieden sterker geactiveerd zijn tijdens deze stap (i.c. lemma-activatie) in het algehele proces. Aangenomen wordt dat de gebieden die sterker geactiveerd worden de gebieden zijn die specifiek betrokken zijn bij dit aspect van spreken. Men kan bijvoorbeeld proefpersonen vragen om woorden en pseudo-woorden hardop te lezen. De laatstgenoemde zijn fonotactisch legale letterreeksen, die echter geen bestaande woorden vormen in bijvoorbeeld het Engels. Een voorbeeld van zo'n letterreeks is *floke*. Iedereen kan dit woord lezen, maar niemand weet wat het betekent. Wanneer men het Engelse woord *smoke* vergelijkt met het pseudo-woord *floke*, ontstaan de volgende twee verschillen in het zien van het woord en het hardop horen. Eén verschil is dat er bij *smoke* sprake is van toegang tot de betekenis gedurende het verloop van het proces, terwijl er van *floke* geen semantische representatie in het geheugen bestaat. Bovendien kunnen we geen fonologische code uit het geheugen ophalen voor het pseudo-woord *floke*. In plaats daarvan moeten we zo'n fonologische code vormen door omzetting van de individuele grafemen in de corresponderende fonemen. Wanneer we de hersenactivatiepatronen meten die geassocieerd zijn met het

hardop lezen van woorden en pseudo-woorden, zijn de verschillen in hersenactiviteit geassocieerd met woorden en pseudo-woorden toe te schrijven aan het ophalen van woordbetekenis en fonologie en/of het vormen van de fonologische code voor pseudo-woorden. Om de hersenactiviteiten gerelateerd aan betekenis en fonologie verder te scheiden zijn andere taakvergelijkingen nodig. Op deze manier kan men ontcijferen welke hersengebieden verschillend geactiveerd worden gedurende de verschillende stappen in het spreekproces. Natuurlijk zijn er ook gebieden die cruciaal zijn voor beide taken, aangezien er naast de verschillen ook overeenkomsten bestaan tussen het lezen van woorden en pseudoworden. In beide gevallen ziet men bijvoorbeeld activatie in de primaire visuele cortex, aangezien het gehele leesproces begint met de analyse van de visuele patronen die op het netvlies terechtkomen.



Figuur 3 (naar Indefrey & Levelt, 2000): De verschillende functionele rollen van corticale gebieden in de linker hemisfeer bij productie van éénwoorduitingen. De figuur geeft een lateraal zicht op de linker hemisfeer. Verschillende gebieden zijn betrokken bij verschillende aspecten van het productieproces van het gesproken woord.

In een recente meta-analyse van meer dan 50 'brain imaging' studies naar de productie van éénwoorduitingen hebben Indefrey en Levelt (2000) onze huidige kennis samengevat van het neurale schakelsysteem dat ten grondslag ligt aan de cognitieve activiteit die ik hiervoor beschreven heb. Alle kernstappen van het spreekproces worden uitgevoerd door gebieden in de linker hemisfeer, wat bij de meeste mensen de taaldominante hemisfeer is. Bij het selecteren van het juiste concept voor het uitspreken van (PAUS) lijkt de middentemporale winding in de linkerhersen helft betrokken te zijn (zie figuur 3 voor een overzicht). Van daaruit spreidt de activatie naar het gebied van Wernicke dat van groot belang is voor het ophalen van de fonologische code van een woord. Het gebied van Wernicke speelt een cruciale rol in het gehele netwerk van taalverwerking door de lexicale aspecten van een woordvorm te koppelen aan de wijdverspreide associaties die zijn betekenis definiëren. Deze rol heeft het gebied van Wernicke voor zowel taalproductie als taalbegrip (Mesulam, 1998). De lexicale woordvorminformatie wordt doorgegeven aan het gebied van Broca in de linker frontale cortex en/of het middelste deel van de bovenste winding van de temporaalkwab in de linker hemisfeer. Deze gebieden spelen een rol bij de omzetting van de fonologische codes in het geheugen naar fonologische woorden waarvan het abstracte articulatorische programma wordt afgeleid. In de laatste fase van voorbereiding op articulatie en de uitvoering van articulatie worden sensorische gebieden geactiveerd, met de mogelijk additionele inbreng van het supplementaire motorgebied en het cerebellum (de laatste twee gebieden zijn niet weergegeven in figuur 3).

De conceptuele kennis die we hebben van woorden van onze taal lijkt meer diffuus in het brein opgeslagen te zijn dan de lexicale lemma- en vormaspecten en is niet beperkt tot de linker hemisfeer. Bovendien wijzen 'brain-imaging' en laesiestudies naar de semantiek van concrete naamwoorden uit dat toegang tot perceptuele en functionele kenmerken van woordbetekenis vanuit verschillende delen van het brein plaatsvindt, de perceptuele kenmerken liggen dicht bij de primair-sensorische gebieden en de functionele kenmerken liggen dicht bij de motorische cortex (voor een overzicht, zie Saffran & Sholl, 1999). Transmodale cortex, inclusief de middelste winding van de temporaalkwab en het gebied van Wernicke zijn convergentiezones of cruciale knooppunten voor de toegang tot relevante informatie die op gedistribueerde wijze is gerepresenteerd (Mesulam, 1998). Op basis van de algemene organisatie van de cortex zijn deze gebieden uitstekend geschikt om onze gedistribueerde conceptuele kennis in één woordvorm te gieten. Dat alles wat we weten over Johannes Paulus II samenkomt in de enkele woordvorm [pαus] vereist de betrokkenheid van hersengebieden die gespecialiseerd zijn in het binden van verspreide fragmenten van kennis tot een enkele output, in dit geval één woordvorm.

Door kennis over de algehele organisatie van de hersenen te combineren met informatie over specifieke patronen van taalgerelateerde hersenactiviteit en de bijbehorende temporele dynamiek, kunnen we inzicht te krijgen in de neurale organisatie van de unieke menselijke capaciteit van communicatie door middel van natuurlijke taal. Om dit zeer complexe communicatiesysteem te doorgronden is veel vakkundig experimenteel onderzoek naar gedetailleerde deelaspecten nodig. Uiteindelijk heeft het begrip dat we hiermee krijgen bredere consequenties voor centrale vragen over de mens.

Taal en de theorie van de mens

Een volledige theorie van de mens vereist een specificatie van de wijze waarop signalen van zeer verschillende functionele systemen als taal, geheugen, emotie, motoriek etc., met hun elk eigen neurale netwerk, leiden tot zelfbesef en een geïntegreerde persoonlijkheid. Hoe dit precies gebeurt is nog grotendeels onbekend. Hoe de hersenen het probleem oplossen om signalen van deze verschillende systemen samen te voegen tot een geïntegreerd zelfbesef met continuïteit van verleden naar toekomst is vrijwel geheel terra incognita voor de huidige cognitieve neurowetenschap. Succesvolle modellen van de verschillende cognitieve systemen ten spijt, is binnen de cognitieve neurowetenschap nog geen alomvattende theorie van de persoon voor handen. Echter, zelfs wanneer een theorie van de menselijke persoon afwezig is, lijkt het redelijk te veronderstellen dat zo'n theorie er heel anders uit zou zien als we geen taal zouden hebben. In de context van zijn schema-theorie beweert Arbib (Russel, Murphy, Meyering & Arbib, te verschijnen) dat het zelf een schema-encyclopedie is die honderdduizenden schema's bevat die een persoon gebruikt om nieuwe informatie te interpreteren en toe te voegen aan het geheugen. De schema-encyclopedie wordt gebruikt om 'een verhaal te vertellen' om nieuwe gegevens in te passen. Niet alleen zijn de metaforen die gebruikt worden om het zelfbesef en de persoonlijkheid te beschrijven zeer vaak ontleend aan het taaldomein (zie ook Gazzaniga's voorstel voor de interpretator, 1992), het is ook duidelijk dat we met taal de omvang van onze schema-encyclopedie in verbazingwekkend tempo kunnen vergroten. Hoewel een cognitieve neurowetenschap van taal de inhoud van onze schema-encyclopedie niet verklaart, is het onmisbaar voor de uiteenzetting van de machinerie die het ons mogelijk maakt deze grote schema-encyclopedie op te bouwen. Ongetwijfeld zouden ons zelfbesef en persoonlijkheid aanzienlijk beperkter en saaier zijn zonder onze taalvaardigheid. In dit opzicht is taal een cruciaal onderdeel van een theorie van de persoon waarvoor de inbreng vanuit de cognitieve neurowetenschap zeer gewenst is. Een vergelijkbaar verhaal kan verteld worden over de relatie tussen taal en bewustzijn. Hoewel het fenomenale bewustzijn niet toereikend verklaard is in

temen van cognitieve neurowetenschap, ondanks beweringen van het tegendeel (zie Chalmers, 1996, voor een tot nadenken stemmende beschouwing), maakt toegansbewustzijn een betere kans. Toegansbewustzijn houdt verband met onze vaardigheid om waameembare oordelen te geven over en verbaal verslag te doen van onze gevoelens. Zoals PET en fMRI zicht bieden op neurale activiteit door een hemodynamisch filter, zo biedt toegansbewustzijn zicht op fenomenaal bewustzijn door een linguïstisch filter. Het begrijpen van de eigenschappen van het filter is ook in dit geval van cruciaal belang om het belangrijke maar vluchtige verschijnsel van fenomeneel bewustzijn te begrijpen.

Dankbetuiging

Ik ben Michael Arbib dankbaar voor zijn commentaar op een eerdere versie van dit hoofdstuk.

Referenties

- Aitchison, J., 1987
Words in the mind: An introduction to the mental lexicon. Oxford: Basil Blackwell.
- Brown, C.M. & P.Hagoort, 1999
Neurocognition of Language. Oxford: Oxford University Press.
- Chalmers, D.J., 1996
The conscious mind: In search of a fundamental theory. Oxford: Oxford University Press.
- Dennett, D.C., 1991
Consciousness explained. Boston, MA: Little, Brown & Company.
- Dunbar, R., 1996
Grooming, gossip, and the evolution of language. London: Faber & Faber.
- Fromkin, V.A. (ed.), 1973
Speech errors as linguistic evidence. The Hague: Mouton.
- Gazzaniga, M.S., 1992
Nature's mind: The biological roots of thinking, emotions, sexuality, language, and intelligence. New York: BasicBooks.
- Hagoort, P. & M. van Turenout, 1997
The electrophysiology of speaking: Possibilities of event-related potential research on speech production. In: W. Hulstijn, H.F.M. Peters & P.H. H.M. van Lieshout (eds), *Speech production: Motor control, brain research and fluency disorders.* (pp 351-361). Amsterdam: Elsevier.
- Hillyard, S.A. & M. Kutas, 1983
Electrophysiology of cognitive processing. *Annual Review of Psychology*, 34, 33-61.

- Indefrey, P., & W.J.M. Levelt, 2000
The neural correlates of language production. In: M. Gazzaniga (ed.), *The cognitive neurosciences*, 2nd edition. Cambridge, MA: The MIT Press.
- Jackendoff, R., 1997
The architecture of the language faculty. Cambridge, MA: The MIT Press.
- Jackendoff, R., 1999
The representational structures of the language faculty and their interactions. In: C.M. Brown & P. Hagoort (eds), *Neurocognition of Language* (pp. 37-81). Oxford: Oxford University Press.
- Levelt, W.J.M., 1989
Speaking: From intention to articulation. Cambridge, MA: The MIT Press
- Levelt, W.J.M., 1999
Producing spoken language: A blueprint of the speaker. In: C.M. Brown & P.Hagoort (eds), *Neurocognition of Language* (pp. 83-122). Oxford: Oxford University Press.
- Mesulam, M.M., 1998
From sensation to cognition. *Brain*, 121, 1013-1052.
- Raichle, M.E., 1994
Visualizing the mind. *Scientific American*, 36-42.
- Roelofs, A., 1992
A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.
- Roelofs, A., 1993
Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition*, 47, 59-87.
- Russel, R.J., N. Murphy, T.C. Meyering & M.A. Arbib (eds), te verschijnen
Neuroscience and the person: scientific perspectives on divine action. Vatican City State and Berkely, CA: Vatican Observaatory and CTNS
- Saffran, E. M. & A. Sholl, 1999
Clues to the functional and neural architecture of word meaning. In: C.M. Brown & P. Hagoort (eds), *Neurocognition of Language* (pp. 241-272). Oxford: Oxford University Press.
- Thorpe, S., D. Fize & C. Marlot, 1996
Speed of processing in the human visual system. *Nature*, 381, 520-522.
- Turenout, M. van, P. Hagoort & C.M. Brown, 1997
Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 787-806.
- Turenout, M. van, P. Hagoort & C.M. Brown, 1998
Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science*, 280, 572-574.