

# PSYCHOLOGISCHE ASPECTEN VAN TAALWETENSCHAPPELIJK ONDERZOEK

## Samenvatting

De taalwetenschap doet de laatste 20 jaar steeds vaker en steeds intensiever een beroep op de inhoud, de methoden, en technieken der sociale wetenschappen. In deze voordracht zal ik schetsen dat op deze gebieden in de psychologie de problemen huizehoog opgestapeld liggen.

\* Wat het inhoudelijke aspect betreft moet o.a. op de naïef-psychologische contra de tegen-intuitieve wetenschappelijke psychologische theorieën gewezen worden.

\* De onderzoeksmethoden der psychologie hebben een alleszins aanvechtbare status: zowel wat betreft de test-procedure, het experiment, als het veldexperiment. De statistische technieken die gehanteerd worden leveren zelden of nooit verklaringen voor psychologische processen (die in het vak: toch centraal dienen te staan), maar geven slechts produktinformatie.

De bovengenoemde kritische aspecten zal ik met onderhanden zijnd onderzoek belichten. Vooral zal daarbij ingegaan worden op de deugdelijkheid van tests, enquêtes, vragenlijsten, rating e.d. en zal geprobeerd worden te wijzen op de consequenties voor taalwetenschappelijk onderzoek.

## Inleiding

Het heeft mij altijd verbaasd waarom toch zovele psychologen die zich met taal bezighielden en houden zich zo intensief in taalwetenschappelijke theorieën inwerkten en blijkens hun onderzoek zoveel vertrouwen in deze theorieën hadden. Dit is des te merkwaardiger wanneer men zich realiseert, dat in deze theorieën nauwelijks psychologische inzichten, vooronderstellingen en implicaties aan bod komen. Integendeel eerder is het zo, dat daar hiervan geabstraheerd wordt. Nemen we bij wijze van voorbeeld de transformationele theorie (een der door psychologen meest gehanteerd theorieën), dan wordt weliswaar met

explanatory adequacy en dus met de psychologische relevantie gekoketteerd, maar wordt tegelijkertijd met een aangeboren LAD (als een taalontwikkelingsverklaring die geen verklaring is) en de onafhankelijke status der syntaxis en een grammatica die weliswaar op een of andere wijze een rol speelt in de waarneming en produktie van zinnen maar geen model daarvoor is<sup>1</sup>, de psychologische vraagstelling weer min of meer buiten de deur gezet.

Het zou me eigenlijk net zo bij zoveel linguïsten moeten verwonderen-ware het niet dat ik mezelf tot die groep moet rekenen- die te rade gaan in de psychologie en die vaak voor psychologische vraagstellingen t.a.v. hun taaldata meer interesse hebben dan voor de taalkundige analyse van hun materiaal. Maar dat is niet zo. Dit ligt waarschijnlijk wel daaraan dat binnen mijn voorstellingen een taalkundige analyse, een taalkundige theorie nooit een doel op zich kan zijn, maar dat deze altijd een -overigens belangrijk- onderdeel van dat soort onderzoek dient te zijn, dat taal, taalgedrag en taalanalyse in een maatschappelijk en psychologisch kader plaatst. Ik zou zover willen gaan met te veronderstellen dat een taalkundige analyse die niet aansluit bij sociologische en psychologische theorieën en vraagstellingen -of daar tenminste niet compatibel mee is- nauwelijks aanspraak kan maken iets met de werkelijkheid van doen te hebben. Ik beseft dat ik hierbij niet elke linguïst aan mijn zijde zal hebben.

Het probleem van de taalkundige die zich aan de menswetenschappen oriënteert is dat de denkwijzen van de letteren en die der sociale wetenschappen vaak hemelsbreed van elkaar verschillen. Zo valt bijvoorbeeld bij de taalwetenschap de deductief descriptieve analyse van taal op (waarbij het materiaal overigens eerder voorbeeldkarakter heeft). Alle onderdelen van dit materiaal worden hierbij diepgaand verklaard en geïnterpreteerd. Deze activiteit leidt tot een hypothese, tot een bevestiging van een theorie of een verbetering daarvan. Kernpunt is dat bijvoorbeeld alle uitingen/zinnen/woorden van de te beschrijven klasse met een voor te stellen uitspraak beschreven kunnen worden

Deze eis leidt tot subclassificatie, differentiatie en uiteindelijk tot grote belangstelling voor het individuele geval. Een goed voorbeeld voor deze denkwijze leveren Labov & Fanshel (1977) in hun analyse van een kort onderdeel uit een psychoanalytisch interview. In de sociale wetenschappen valt aan de andere kant eerder een tegengestelde tendens vast te stellen met betrekking tot de doelstelling van onderzoek. Hier gaat het meestal om een statistische analyse van slechts bepaalde aspecten in het materiaal (waarin overigens ook best tegenvoorbeelden voor de te onderzoeken hypothesen mogen voorkomen) om inductief tot conclusies voor de gehele populatie te komen. Tegenvoorbeelden zijn vervelende uitzonderingen die storend voor de statistische berekening zijn maar verder niet belangrijk zijn om tot globale, generaliserende, uitspraken te komen.

In dit artikel wil ik proberen aan te tonen dat, hoewel -zoals gezegd- taalonderzoek noodzakelijkerwijs sociaal-wetenschappelijk georiënteerd dient te zijn om zinvol te zijn, de sociale wetenschappen zoals die inhoudelijk en methodisch functioneren, zelf echter vol voetangels en klemmen zitten die de niet-ingewijde linguïst en zelfs de ingewijden (vanwege een zekere hoeveelheid beroepsdeformatie?) meer nadeel dan voordeel kunnen opleveren wanneer hij zich van inhouden en methoden der sociale wetenschappen in zijn taalonderzoek bedient. Het is nu eenmaal zo, dat, wanneer men een andere discipline als hulpdiscipline invoert, men veel meer vertrouwen in het attractieve nieuwe dan het oude vertrouwde heeft.

De punten die ik hier achtereenvolgens zal behandelen zijn:

1. naiefpsychologische tegenover wetenschappelijke theorieën;
2. statistische technieken en de kracht van psychologische uitspraken ten aanzien van psychologische processen;
3. de proefpersonenproblematiek in zowel experiment als test;
4. de problemen bij experimenten;
5. de problemen bij testen, vragenlijsten, rating e.d.

## 1. Naieve psychologie en wetenschappelijke psychologie.

In *Knowledge without foundations* verdeelt Paul Feyerabend (1962) wetenschappelijke theorieën onder in intuïtieve en counter-intuïtieve theorieën en komt onder andere tot de slotsom dat vruchtbare theorieën in het begin zeer vaak als counter-intuïtieve beoordeeld zijn.

De intuïtieve theorieën hebben het schijnbare voordeel dat zij veel mensen, vooral ook niet-ingewijden onmiddellijk aanspreken. De onderzoeksresultaten hebben echter helaas meestal het teleurstellende "maar-dat-wist-ik-toch-allang" resultaat. Vooral in de sociale wetenschappen treden zulke herkenningseffecten veelvuldig op. Het kan echter nauwelijks de bedoeling van de wetenschap zijn, steeds weer reeds voorhanden kennis te bevestigen, zodat alleen daarom al minder intuïtieve theorieën nodig zijn, om onze kennis te vergroten.

Een van de problemen in de psychologie is dat de grens

tussen de naieve en de wetenschappelijke theorieën dwars door de partijen loopt. Globaal genomen kan men stellen dat ze en in het wetenschappelijk-theoretische bedrijf en in de toepassing beide veelvuldig voorkomen. Dat niet-ingewijden meestal uitsluitend over de naieve theorieën beschikken, mag niet tot de conclusie verleiden dat daarom deze niet in het wetenschappelijke bedrijf thuis horen.

Ten eerste zegt de kwalificatie naief of wetenschappelijk niet per se iets over de juistheid van de theorie. Ten tweede moet de wetenschap zich met de naieve theorieën bezighouden zelfs voor het geval als aangetoond kan worden dat ze fout zijn. Dit in eerste instantie wel daarom, daar zulke naieve theorieën in de hoofden der mensen vaak de leidraad voor hun gedragingen, beoordelingen van zichzelf en anderen en voor hun beslissingen in het dagelijkse leven zijn.

Een voorbeeld uit de persoonlijkheidsleer moge dit verduidelijken. Oudere opvattingen over persoonlijkheid gaan ervan uit dat persoonlijkheid als een bundel min of meer stabiele eigenschappen gedefinieerd kan worden. In feite komt deze theorie met de alledaagse naieve opvatting der persoonlijkheid overeen: mensen schrijven personen eigenschappen toe (attributie) en nemen aan dat deze eigenschappen de verklaring voor bepaalde gedragingen zijn (bijvoorbeeld: "Dat doet ie omdat ie zo snel jaloers is"; causale attributie). Nu is in onderzoek vastgesteld, dat gedrag en gedragsmogelijkheden voor een zeer belangrijk deel met context en situatie verbonden zijn. In taalgedragsonderzoek is bijvoorbeeld vastgesteld dat bepaalde kinderen (bijvoorbeeld meisjes, lager milieu kinderen e.d.) in taaltestsituaties of in bepaalde onderwijsleersituaties minder taalvaardig zijn dan in andere hen minder belastende situaties (Cazden, 1970; Labov, 1970; Van der Geest, 1978). Ook kan vastgesteld worden dat stabiele eigenschapstoeschrijving tot foute waarneming en beoordeling van gedrag voert. Als voorbeeld moge Herr Kraus, een leerkracht aan een Duitse lagere school dienen. Als beoordeling na 6 weken onderwijs in de eerste klas gaf hij van een gastarbeiderskind te kennen, dat alleen de taalvaardigheid van het kind te wensen over liet en wel zou de gemiddelde zinslengte niet in orde zijn (GZL). Bijna drie jaar later staat als commentaar onder een opstel van dit kind dat inmiddels 9 jaar is: zeer fantasievol, spannend en origineel: alleen je zinnen zijn nog te kort; bijna geen fouten; cijfer: tevredenstellend. Een zulk blijvend oordeel is een van de nadelen dat kinderen 4 jaar lang dezelfde leerkracht hebben. De feiten waren de volgende: de GZL bedroeg 5,6. De langste uiting bedroeg 14 woorden, de kleinste 3. Zoals u weet is een van de mogelijkheden spanning te bereiken, om daar waar dit nodig is de uitingen kort te houden. Met andere woorden: 1. de waarneming van de heer Kraus was niet helemaal juist. Een uiting was zelfs zeer lang; 2. de beoordeling was niet geheel korrekt. Korte uitingen zijn niet per se slecht. Men kan niet tegelijkertijd zeggen dat iets spannend is en tegelijkertijd het stijlmiddel dat dit bewerkt afkeuren.

Onder andere om deze twee tekortkomingen (te weten de situatieafhankelijkheid en de foutieve waarneming en beoordeling) tegen te gaan, wordt persoonlijkheid in wetenschappelijke theorieën als de verzameling van die gedragsalternatieven beschreven waarover de persoonlijkheid

per situatie beschikt. De deelnemers aan menselijke interactie zouden eigenlijk om in de communicatie beter te kunnen functioneren volgens deze theorie moeten handelen. Vandaar dat de wetenschappelijke theorie in dezen niet zozeer de werkelijkheid van alledag beschrijft, maar veel eerder voor de interventie een beter alternatief aanbiedt. Ter verklaring voor wat in de werkelijkheid gebeurt zijn naïef-psychologische verklaringen bitter noodzakelijk. Zij blijven dan ook de beste verklaring voor hoe mensen zich (fout) gedragen ten opzichte van anderen en hoe zij (fout) over deze mensen denken.

Om een voorbeeld te geven dat een wetenschappelijke theorie niet noodzakelijkerwijze beter moet zijn dan een naïeve theorie, en tevens om te laten zien dat we in taalbeheersingsonderzoek en -praktijk wel degelijk met deze tweedeling en haar consequenties te rekenen hebben, volgt hier iets over taalverwerving.

In een onderzoek heb ik eens niet-ingewijde moeders twee alternatieve taalverwervingstheorieën voorgelegd: Chomsky's innateness-theorie en mijn, op interactie moeder-kind gebaseerde, theorie waarbij de omgeving enige invloed op de verwervingsprocessen kan uitoefenen. De proefpersonen waren vrijwel unaniem van mening dat Chomsky's theorie belachelijk is. Dit wil echter nog niet zeggen dat Chomsky's opvatting fout is. Dat kan alleen wetenschappelijk onderzoek aantonen. Dit soort onderzoek heeft echter de mening van mijn proefpersonen bevestigd, zodat we hier niet mogen koketteren met Feyereabend en zeggen: deze theorie is tegen-intuïtief en dus goed, zoals Elffers-Van Ketel, De Haan & Klooster (1974) dit lijken te doen.

Wat belangrijker is, is dat deze theorieën andere consequenties voor de praktijk hebben. Uit Chomsky's theorie moet men afleiden dat het taalverwervingsproces nauwelijks te beïnvloeden is en dat interventie dus nauwelijks wat uithaalt. Met andere woorden we kunnen naar huis gaan omdat taalverbetering blijkbaar niet geleerd kan worden door anderen. Uit heel wat onderzoek, bijvoorbeeld met doven, blijkt dat deze zienswijze god zij dank fout is. Tenslotte zou ik in deze paragraaf nog op een aantal punten willen wijzen die belangrijk zijn voor de latere behandeling van het proefpersonenprobleem en het test en experimenteerprobleem.

Bovendien zijn ze alle van belang voor (taal) gedragstheorieën en de opvattingen omtrent persoonlijkheid. Ze demonstreren bovendien alle hoe naïeve en wetenschappelijke opvattingen uit elkaar lopen.

1. Eigenschappen die men iemand toedicht komen niet systematisch met zijn gedragswijzen overeen: bijvoorbeeld hoeven (goede) taalkundigen niet goed in taalgedrag te zijn;
2. De waarneming van een ander of zichzelf is vaak zo selectief dat men alleen maar die gedragswijzen "ziet" die de overtuiging die men t.a.v. de toegedichte eigenschappen had, versterken.
3. Vaak worden tegengestelde gedragswijzen met een en dezelfde eigenschap verklaard. Eigenschapstoekenning is blijkbaar essentiëler voor de mens dan gedrag: gedrag en eigenschapstoekenning hebben blijkbaar niet veel met elkaar te maken;

4. Wanneer het om zelfattributie gaat: gedragsintenties en uiteindelijk gedrag stemmen niet noodzakelijkerwijze met elkaar overeen. De omgeving en de situatie beïnvloeden in belangrijke mate welke gedragsmogelijkheid gerealiseerd wordt. Weliswaar is men in staat uit persoonlijkheidseigenschappen gedragsnormen af te leiden, maar dit leidt niet automatisch tot de uitvoering van dit gedrag. Bijvoorbeeld, van een lerares die in de klasse-interactie in tegenstelling met haar normale gedrag nu eens alle kinderen even vaak aan de beurt wilde laten komen, stelden we vast dat haar beurtgeefgedrag in dit lesuur in niets afweek van haar gedrag in de andere lessen. (Van der Geest, 1978) Dit betekent bijvoorbeeld voor de interventie, dat in de communicatietraining na een cognitief omstructureringsprogramma een gedragstrainingscomponent moet volgen om tot de beoogde communicatieve veranderingen te komen.
5. Op grond van bepaalde arbitraire maar opvallende eigenschappen, bijvoorbeeld sexe, sociaal milieu, ethnische afstamming: e.d., worden vaak (voor) oordelen omtrent iemands gedragsmogelijkheden gevormd. Hierbij moet men vooral op het zgn. halo-effect letten, dwz. op grond van een bepaalde waarneming worden conclusies over niet waargenomen gedrag dat hier (volgens vooroordeel) mee samenhangt getrokken. Dit punt heeft vooral betrekking op het rating-principe. De consequentie voor deze onderzoeksmethode is dat men hier mogelijkwijze alleen maar vooroordelen verzamelt. Deze onderzoeksmethode wordt steeds vaker gebruikt onder andere in tekstbeoordelingsonderzoek (bijvoorbeeld Jansen, 1980) en in opstelbeoordelingen (bijvoorbeeld Westorp, 1974). De beoordelingscriteria onttrekken zich hierbij bijna geheel aan onze waarneming. Dit is mogelijkwijze de reden waarom Jansen (1980) niet in staat is gedetailleerd mee te delen welke criteria voor de beoordeling uiteindelijk verantwoordelijk zijn<sup>2</sup>).

Een ander aardig (of eigenlijk onaardig) voorbeeld der eigenschapstoekenning t.a.v. taalgedrag is gelegen in de GZL. Uit de literatuur is bekend dat mensen die langere zinnen maken voor intelligenter, taalvaardiger, afstandelijker, beleefder, tot een hoger sociaal milieu behorend, beter argumenterend, meer gelijk hebbend, en wat al nog niet meer gehouden worden (Garfinkel, 1973; Schönbach, 1974; Bernstein, 1971, Bereiter & Engelmann, 1967). Er is voldoende literatuur die overtuigend aantoonde dat dit niet noodzakelijkerwijze het geval hoeft te zijn (Labov, 1970; Cazden, 1970; Appel, 1972). Voor de taalontwikkeling bleek bijvoorbeeld dat de gemiddelde kwantitatieve prestatie niets zegt over de kwaliteit van het taalgebruik (Van der Geest, e.a., 1973). Dat we wel degelijk met halo-effectachtige vooroordelen en vaak zeer gevaarlijke vooroordelen te maken hebben, blijkt uit een sociaalpsychologisch onderzoek van Müller-Eckhard (1974). Hij hield juristen en aankomende juristen teksten in twee versies voor die inhoudelijk identiek waren maar wat de GZL betreft systematisch verschilden. In deze teksten zijn wetsovertreders aan het woord, die vertellen wat ze misdreven hadden. Toen de juristen naar de adequate strafmaat gevraagd werd, bleek dat de bereidheid hogere

straffen te geven met de teksten met kleine GZL significant correleerden. Deze rating maakt duidelijk dat het expert zijn (hier in dit geval de juristerij) niet tot betere beoordelingen voert, maar hooguit tot systematische vooroordelen. Geldt dit ook voor tekstbeoordeling? Het is, lijkt me, zeer wel mogelijk dat tekstbeoordelingen door geschetste halo-effecten tot stand komen. Het is de taak van de onderzoeker de bron van deze beoordelingen op te sporen. Dit soort onderzoek is echter tot op heden niet zo succesvol verlopen.

Resumerend kan gezegd worden dat de naieve psychologie met zijn vooroordelen een groter maar liever te vermijden rol speelt in taalvaardigheidsonderzoek.

## 2. Statistische technieken en de kracht van psychologische uitspraken

De psychologie is een empirische wetenschap waarin er naar gestreefd wordt met empirische methoden en technieken causaal te interpreteren samenhangen te ontdekken. Met andere woorden is de psychologie er alles aan gelegen processuele veranderingen te ontdekken en te beschrijven. Dit geschiedt normaliter door experimenten (waarover meer in 3) en de voor dit doel ontwikkelde statistische technieken. Door het systematisch variëren en controleren van bepaalde variabelen in het experiment is het mogelijk door bepaalde statistische berekeningswijzen causaal relationele verbanden tussen de afhankelijke en de ingestelde variabelen te vinden.

Het probleem hierbij is natuurlijk dat men bij de berekening op een of andere manier van gemiddelden en spreiding moet uitgaan.

Een gemiddelde prestatie van een proefpersoon wordt omgewerkt naar een gemiddelde prestatie van een groep. Hierbij gaat men gemakshalve -in elk geval in de praktijk- aan het probleem voorbij dat de verschillende meetpunten in de test t.a.v. een persoon mogelijk niet gelijkwaardig zijn, niet tot de zelfde categorie behoren, zodat hij op sommige aspecten goed en op andere fout scoort. Zijn gemiddelde score zegt mogelijk iets over een artefact, namelijk niets over twee grondverschillende vaardigheden die de experimentator niet bekend zijn. Op een hoger niveau worden dan al deze gemiddelde prestaties op een grote hoop gesmeten, alsof die paar variabelen t.a.v. de in het experiment meedraaiende groepen die gecontroleerd werden (bijvoorbeeld: leeftijd, IQ, sexe, sociale klasse, prestatie in een andere test) de verklaring zijn voor alle (kunstmatige?) verschillen in de groep of alle overeenkomsten in de groep (die ook weer kunstmatig kunnen zijn). Met andere woorden er worden t.a.v. de te meten vaardigheden en de deelnemende groepen veel te weinig variabelen gecontroleerd en onderzocht, om van gemiddelden uit te mogen gaan. Blijven we bij wijze van voorbeeld bij de GZL als prestatie maat dan blijkt dat deze GZL zeer van de situatie afhangt (Bakker & Hoefnagel, 1974). Bij dezelfde kinderen kan men zeer verschillende GZL's vaststellen afhankelijk van de verschillen in de alledaagse situatie. Wanneer we nu deze verschillen middelen dan

krijgen we een situatieloze GZL, die niet per se iets over iemands taalvaardigheid in het algemeen hoeft mede te delen (wanneer zo iets al bestaat), aangezien door de middeling de situatie als bepalende factor van het gemiddelde onder de tafel verdwijnt. Wat kan men bijvoorbeeld beginnen met de informatie dat de Nederlander 1,999 benen heeft. Dit geldt ook voor de gemiddelde taalprestatie van een groep kinderen van bijvoorbeeld dezelfde leeftijd, sexe en sociale klasse. Allerlei verschillende gespreksituaties, die zich mogelijk ook nog verschillend op de kinderen in de onderzochte groep kunnen uitwerken, allerlei verschillen in gespreksonderwerpen en gesprekspartners om een paar storende aspecten te noemen, werken zich op een niet gecontroleerde manier op de meetresultaten uit. Ook kunnen allerlei verschillen in ontwikkelingsstadium tot foute conclusies voeren. Een voorbeeld van onderzoek waar men aan deze problematiek geheel voorbijgaat is Van Ierland (1980).

Om de realiteit van de problematiek te schetsen zou ik graag een voorbeeld uit de eigen onderzoekservaring willen geven. In ons schoolonderzoek (Van der Geest, 1978) werd o.a. door een aantal zelf ontworpen taaltests gemeten in hoeverre bij eerste klassers en kleuters de taalvoorwaarden voor succesvol rekenonderwijs aanwezig waren. In deze tests hadden de eerste klassers significant slechtere prestaties dan de kleuters. Zulke tendensen werden ook vastgesteld tussen meisjes en jongens en tussen hoger en lager milieu (Bolten, 1975; Drubbel, 1976). Alleszins interessante gegevens om mee voor de dag te komen: bijvoorbeeld dat de school een slechte rol op de taalrekenprestatie van kinderen uitoefent, iets wat overigens al eens "vastgesteld" is. (Koster, 1975). Bij nadere beschouwing van de data bleek echter dat mogelijkerwijze twee factoren voor deze merkwaardige uitkomst verantwoordelijk kunnen zijn. De eerste berust op de vaststelling dat de prestaties al bij de oudere kinderen op de kleuterschool duidelijk afnamen. Een gegeven dat door middelen verdween. Het kan dus zeer goed zijn dat de prestaties der lagere schoolkinderen door een ontwikkelingsfactor in plaats van door een schoolfactor te verklaren zijn<sup>3</sup>. Bijvoorbeeld biedt zich hier de overgang van het preoperationele naar het concreet operationele stadium als verklaringsmogelijkheid aan. Het is bekend uit de piagetiaanse literatuur dat kinderen in het begin van een bepaald stadium, bij het gebruik van een nieuwe strategie, fouten gaan maken die zij voordien niet gemaakt hebben. Door niet alleen te testen, maar door bijvoorbeeld na te vragen zoals dat in Piaget's klinische methode gebruikelijk is -en zoals wij dat ook deden- komt men achter deze verklaring. Zo ontdekt men op die manier ook dat voor kinderen vragen, die ons als gelijkwaardig voorkwamen, vaak tot geheel verschillende categorieën behoren.

Een andere tendens die ook door navragen te verifiëren is, is de afname der prestatie aangezien het kind met de door de school aangeboden probleemoplossingsmethode nog niet voldoende uit de voeten kan. Met andere woorden is het mogelijk dat de strategie van het kind en die van de school nog niet perfect harmoniëren, en dat door dit feit meer fouten voor zullen komen.

Fouten op zich zeggen dus niets over de ontwikkelingsstand en de prestatie en vaardigheid van het kind; maar, wanneer we de oorzaken van deze fouten kennen en dus ook beter deze fouten kunnen categoriseren, dan zijn we pas in staat een essentiële bijdrage tot de kennis van ontwikkelingsprocessen te leveren. Ook de (gemiddelde) goede prestatie zegt ons niets definitiefs over de ontwikkelingsstand van het kind. Ook hier is de vraag hoe het kind tot dit goede antwoord gekomen is beslissend voor de categorisering der goede antwoorden en voor de kennis van de ontwikkelingsprocessen die op een bepaald moment aflopen. Samenvattend: een test meet een gemiddelde prestatie en is dus als een produktmaat te beschouwen en niet als procesmeetpunt. Processuele uitspraken zijn uitspraken der interpretatie van produktmetingen en zijn daafom noodzakelijkerwijze subjectief.

Doorvraagmethoden zoals de klinische methode leveren ons inzicht in wat voor processen zich op een bepaald moment ontwikkelen. De statistische berekening is daarbij nog oninteressant. Zij wist de fouten van de onderzoeker in zijn experiment (zijn methode, technieken, opvattingen, interpretaties e.d.) niet uit, maar berekenen verfijnder, de mogelijke foute uitkomsten van zulk onderzoek. Het lijkt me dan ook zinvol dat de waarde van statistische technieken, psychometrische methoden, e.d. en het denk-kader dat daar achter ligt in de psychologie nog eens (kritisch) overdacht wordt.

### 3. De proefpersonenproblematiek

Uit het voorafgaande is reeds min of meer duidelijk geworden dat de mens -de proefpersoon- mogelijk de storendste factor van vruchtbaar psychologisch onderzoek is. Het is niet overdreven te stellen dat de proefpersoon zich vaak niet zo gedraagt als de experimentator dat graag zou willen. Dit wordt bijvoorbeeld zeer uitvoerig en goed gedocumenteerd in Bungard (1980) "*Die "gute" Versuchsperson denkt nicht; Artefakte in der Sozialpsychologie*". Helaas blijkt dat proefpersonen wel denken en dat daarom de waarde van veel proefpersooninformatie in test en experiment van niet geheel betrouwbare waarde is. Met dit feit zal men moeten leren leven.

In onderzoek werd vastgesteld dat men proefpersonen moet onderverdelen in hen die wel of niet mee willen werken. Vooral wanneer we beseffen dat meestal proefpersonen gebruikt worden die in het kader van hun psychologiestudie verplicht zijn als proefpersonen op te treden, of die voor de deelname een financiële vergoeding ontvangen, zal duidelijk zijn dat niet bij voorbaat van een goede arbeidsmotivatie uitgegaan mag worden en dat het punt medewerking niet automatisch als voldoende geregeld opgevat mag worden.

Wanneer de proefpersoon besluit mee te werken dan bouwt hij hypothesen op over zichzelf, de experimentator en het experiment, bijvoorbeeld t.a.v. de afgevraagde stof, cq. de gestelde opgaven. Zo werden bijvoorbeeld systematische verschillen in de beantwoording van een persoonlijkheidstest vastgesteld, afhankelijk van de reeds

volbrachte studieduur en afhankelijk van de kennis die men had van de theoretische uitgangspunten van de experimentator. De proefpersonen waren hier natuurlijk psychologie-studenten. Met andere woorden, proefpersonen proberen voor zich de vraag te beantwoorden wat de zin van het onderzoek is, wat de experimentator van hem wil en wat hij in dit geval van zichzelf denkt of wat anderen van hem denken. Wanneer de overwegingen van de pp. in de richting van de hypothese gaan dan heeft de experimentator -wat betreft zijn uitkomsten- tenminste geluk gehad. De vraag is echter of deze overwegingen met de werkelijkheid van de pp. overeenstemmen en of de uitkomsten correct zijn.

Vaak komt daar dan nog de onderzoeksopbouw bij. Het maakt blijkens onze ervaringen voor de testresultaten veel uit of men eerst een angsttest en daarna een IQ-test aanbiedt of omgekeerd.

Ook is de problematiek der sociale wenselijkheid met betrekking tot de antwoorden bekend. Pp's tenderen er kennelijk naar zich van een geidealiseerd beeld van zichzelf te bedienen. Dit zou men nog als een deel van de persoonlijkheid kunnen zien, ware het niet dat dit ideale zelfbeeld mede door de proefopstelling veroorzaakt wordt. Zo veranderden we in een test de tekst van de items zodanig dat telkens de mening over het tegenovergestelde gevraagd werd (bijv.: wanneer de deur opengaat schrik ik/ schrik ik niet; met 7 antwoordmogelijkheden tussen "klopt helemaal" en "helemaal niet"). Tevens werd de suggestie door deze verandering gewekt, dat de neergeschreven tekst blijkbaar de normale gang van zaken is. Op basis van de feitelijke inhoud der beide tests zou er een grote negatieve correlatie tussen de antwoorden op beide tests moeten bestaan. Dit was echter niet het geval: we vonden matige maar positieve correlaties. De formulering van de testvragen veroorzaakt dus dat proefpersonen zich anders gaan zien, beter nog, dat proefpersonen vragen anders beantwoorden. Het is dan ook de vraag of men op onderzoekshygiënische gronden pp's aan zelfwaarneming en zelfoordeel mag onderwerpen. De Deze beide aspecten spelen in bijna elk psycholinguïstisch experiment op een of andere manier een rol.

Een voorbeeld moge dit verduidelijken. In veel onderzoek naar de uitwerking van reclame worden vragen gesteld of stimulus a of b de proefpersoon iets doet, opwindt e.d. De antwoorden op deze vragen zijn bijna altijd in systematische tegenspraak met wat fysiologisch (bijv. door huidweerstand, transpiratie, warmte en doorbloeding te meten) vastgesteld wordt en wel zegt de pp. dat hij van een bepaalde reclame-aankleding (die vaak neerkomt op uitkleding) niet of nauwelijks aangedaan is, terwijl fysiologisch eerder het tegendeel vastgesteld wordt.

Wanneer men nu onderzoek naar tekstbeoordeling doet en bijv. rating invoert, of wanneer men onderzoek naar de uitwerking van bepaald taalgebruik (woorden/ uitingen / argumentaties e.d.) doet zoals dat in de semantische differentiaal het geval is, waarbij men ( in deze test) naar de (emotionele ?) indrukken t.a.v. iets vraagt, dan dient men m.i. de uitkomsten van zulk onderzoek met een gezonde dosis scepsis op te nemen <sup>4</sup>.

in veel experimenten gaat het om geautomatiseerde processen waar motivatie niets toe doet

Ook grammaticaliteitsoordelen zijn van een twijfelachtige kwaliteit zoals o.a. door Levelt (1971), Snow (1975) en Van der Geest (1980) uiteengezet is.

Zelfs wanneer men proefpersonen niet naar meningen vraagt maar ze opdrachten laat uitvoeren, zoals in zoveler psycholinguïstische experimenten gebeurt, moet men er rekening mee houden dat de experimentele situatie de pp. aan het denken zet, of dat de experimentator hem d.m.v. de experimentele situatie aan het denken zet, en hem zodoende dingen laat doen die met zijn gedrag in het dagelijkse leven niet te verenigen zijn. Een voorbeeld hier zijn de experimenten waarin pp.'s homonieme zinnen krijgen voorgezet. Deze en andere experimenten zullen in het volgende deel onder de loep genomen worden.

#### 4. Het psychologische experiment.

*Waar is het anders?* De psychologie is een vrolijke, blije en optimistische wetenschap waarin alles dat lukt gepubliceerd wordt en wat niet lukt doodgezwegen wordt. Wanneer we ervan uitgaan dat in een experiment een aantal variabelen gecontroleerd worden en een aantal andere onderzocht worden en dat alle andere variabelen als storend te boek staan - wanneer ze al opgemerkt worden - en waarbij men hoopt dat de storende invloed bijvoorbeeld door een grote hoeveelheid proefpersonen geëlimineerd kan worden, dan zal het duidelijk zijn dat het vaak een gok is of de hypothese bevestigd kan worden. Wanneer de hypothese niet bevestigd kan worden is men methodisch gedwongen het tegendeel (de nulhypothese) aan te nemen. Dit gebeurt echter zelden, want het geloof (de theorie) is sterker dan de feiten. Men zoekt dan ook gemeenlijk naar de oorzaak van het falen van het experiment, verandert wat in de proefopstelling en start zo een "verbeterd" experiment. Dit herhaalt men zo vaak tot de beoogde significante resultaten optreden. Deze worden dan meestal zonder vermelding der vele tegenslagen gepubliceerd. Men kan overigens uitrekenen hoe vaak men een onderzoek herhalen moet tot het beoogde resultaat optreedt. Als voorbeeld voor deze gang van zaken zou men de zgn. klick-tests kunnen aanvoeren (o.a. Fodor & Bever, 1965). Smolenaars heeft in Amsterdam eind jaren '60 en begin jaren '70 onderzocht of men voor het Nederlands gelijke regelmatigheden kan vinden als voor het Engels werden vastgesteld. Dit alles zonder resultaat. Bij mijn weten zijn deze negatieve resultaten nooit gepubliceerd. Het kan dus best zijn dat met de klickexperimenten iets niet klopt. In elk geval werden ze na kort en intensief gebruik in de cognitieve psychologie nooit meer ingezet. In plaats van de syntactische verklaring der uitkomsten deden daarna semantische verklaringen opgeld (Fodor & Garrett, 1966; Hörmann, 1977), waartussen deze experimenten niet beslissen kunnen. Ook schijnen triviale gegevens als de zinslengte ter verklaring aangevoerd kunnen worden.

Uit linguïstische hoek wordt de laatste tijd de belangstelling voor het klickexperiment weer groter (o.a. Piennemann, 1980), zonder dat deze linguïsten met de methodische voetangels rond deze experimenteertechniek vertrouwd zijn.

Een belangrijk punt van kritiek op het experiment blijft de vraag in hoeverre de experimentele uitkomsten uit de experimentele situatie voortvloeien en in hoeverre ze op de alledaagse situatie geprojecteerd kunnen worden. Indien hier de nodige beperkingen vastgesteld kunnen worden (en veel onderzoek toont aan dat dit daadwerkelijk het geval is), dan betekent dit dat het experiment en de experimentele situatie een ernstig storende variabele voor psychologisch valide onderzoek is. Dat beide situaties inderdaad niet identiek zijn lijkt mij ook voor het psycholinguïstische onderzoek duidelijk. Ten eerste is taalwaarneming in het dagelijkse leven in de context ingebed en voor de interpretatie van deze context afhankelijk. In de experimentele situatie moet de proefpersoon echter deze context mentaal zelf opbouwen en hij moet zolang verder zoeken tot hij een context gevonden heeft waarin de uiting zinvol geïnterpreteerd kan worden. In de realiteit van elke dag moet men bijvoorbeeld een imperatief uitvoeren of tenminste aangeven dat men hem verstaan heeft maar (alsnog) niet uitvoeren wil of kan. In het experiment, daarentegen moet van een imperatief alleen maar gezegd worden waar de klick gevallen is.

Ten tweede worden pp.'s vaak met opgaven geconfronteerd die zij in het dagelijkse leven t.a.v. taalproblemen nooit voorgeschoteld krijgen. Zij moeten snel reageren, uit lijsten equivalenten uitzoeken, woordrijen van buiten leren, etc. De vraag schijnt dan ook gerechtvaardigd wat men nu meet, taalperceptie of geschiktheid voor bepaalde experimentele opgaven, of naieve taalkundige opvattingen die men in het experiment als een soort ezelsbruggetje gebruikt om bepaalde beslissingen te kunnen bereiken. Ten derde is de "taalsituatie-taalopdracht" combinatie vaak verschillend van die van alledag. Zo wordt de pp. in het experiment bijvoorbeeld systematisch tot het zoeken van homoniemen gedwongen, terwijl hij in het dagelijks leven deze homonymie niet eens gemerkt zou hebben, en wanneer hij die zou merken door contextuele interpretatie weer overboord gewerkt zou hebben.

Een voorbeeld hoe experimentator, proefopstelling en theorievorming tot uitkomsten kunnen leiden, haal ik uit de cognitieve psychologie. In een onderzoek naar muziekperceptie stelt Stoffer (1981) voor muziek met een systeem volgens Chomsky te beschrijven. Dit wordt met zeer eenvoudige muziek (het volks- en kinderlied) gedemonstreerd. Wanneer nu een dergelijk systeem aanspreken op "explanatory adequacy" wil maken moet men, wanneer men de muziek stopt, in staat zijn de volgende toon te voorspellen (muziek wordt dus als volstrekt voorspelbaar gedefinieerd) of wanneer men een klick aanbiedt deze klick in de richting van de syntactische grens verschoven waarnemen. Stoffer onderzocht deze twee aspecten met behulp van zeer eenvoudige toonopvolgingen (die ik weiger muziek te noemen). Hij trainde voordien een deel van zijn pp.'s voor de voorspellingsopgave met behulp van zijn beschrijvingsysteem. Zij die na deze training een slechte prestatie leverden werden uit het experiment verwijderd. Na de uitvoering van het experiment bleek dat de getrainde proefpersonen betere prestaties in beide tests hadden. De onderzoeker deelt daarop mee dat zijn beschrijvingsysteem psychologische

realiteit bezit t.a.v. de waarneming en verwerking van muzikale eenheden<sup>5</sup>. Hij komt niet op de gedachte dat de resultaten van zijn experiment door zijn proefopstelling gemanipuleerd zijn. Ten eerst kan het zeer eenvoudige materiaal ertoe leiden dat men muziek volledig voorstellen kan. Ten tweede wijzen de uitkomsten erop dat de resultaten voor een belangrijk deel door de training bewerkt zijn. Hoe kan het nu voorkomen dat een onderzoeker niet op een zo voor de hand liggende verklaring voor zijn resultaten komt. De verklaring ligt daarin dat in bijna elk cognitief psychologisch experiment de proefpersonen getraind worden (zie bijvoorbeeld Neisser, 1976) en wel omdat men reactietijden meet en omdat procentueel te veel fouten in de test dan ook een storende interactie tussen reactietijd en testprestatie op zou leveren. Door vele psychologen wordt de cognitieve psychologie als het kwalitatieve kernstuk van psychologisch onderzoek gezien.

Op het gevaar af ervan beschuldigd te worden recepten te verspreiden waarmee men in elk geval tot significante uitkomsten kan komen volgt hier nog een tip: wanneer men in elk geval tot significante uitkomsten wil komen en men niet precies weet hoe bepaalde psychologische aspecten met elkaar samenhangen, kan men het aantal variabelen uitbreiden zodat in de correlatiematrix altijd wel een aantal cellen met significante samenhang opduiken. Ook hier kan men uitrekenen hoeveel cellen men nodig heeft om tot zulke resultaten te komen. Zulke praktijken komen zeer vaak voor (bijvoorbeeld Geppert, 1979, die taalcomplexiteitsmaten in hun samenhang met sexe, klasse, intelligentie en prestatiemotivatie onderzoekt).

Tot slot, resultaten van experimenteel onderzoek hangen helaas van vele factoren af die met de aard van het experiment en het experimenteren samenhangen en waarvan ik in vele gevallen niet weet hoe men ze anders verhelpen kan, dan door van het experimenteren verder maar af te zien.

## 5. Test, vragenlijst, rating e.d.

In het voorafgaande kwam telkens reeds kritiek aan de oppervlakte wanneer er gesproken werd over proefpersonen met betrekking tot meningen, weten en beoordelingen; reden genoeg op deze problematiek in een aparte paragraaf nader in te gaan.

Voor alle vragenlijsten e.d. geldt, dat:

1. antwoorden op vragen niet zozeer de juistheid van bepaalde psychologische theorieën bevestigen maar eerder informatie geven over de naïef-psychologische processen waarvan de pp. zich in meningen en beoordelingen bedient;
2. sociale wenselijkheid, de opzet van de lijst, de opvattingen van de afvrager en constructeur van de test van grote invloed op het uiteindelijke resultaat is;
3. strategieën ter beantwoording der vragen tot bepaalde foute resultaten kunnen voeren;
4. halo-effecten, vooroordelen en extrapolaties tot bepaalde beoordelingen/antwoorden voeren die niet noodzakelijk de mening/het weten van de pp. representeren.

Zulke aspecten spelen een wezenlijke rol in de beantwoording van vragenlijsten en rating. We hoeven hierop in deze paragraaf niet meer in te gaan. Belangrijker is het hier op de zgn. gestandaardiseerde vragenlijsten, de tests, in te gaan. Het doel van tests is om iemands vaardigheid (bijv. intelligentie of taalvaardigheid) te meten of om iemands psychische toestand (bijv. angst, emotionaliteit) vast te stellen. Meet een test dat nu ook? Daarmee samenhangend: hoe wordt de vaardigheid of toestand gedefinieerd, zodat men controleren kan of de test meet wat hij meten moet? Het antwoord is hier van teleurstellende eenvoud: intelligentie/taalvaardigheid is datgene wat de test meet. Een dergelijke tautologische definitie t.a.v. de inhoudelijke validiteit is in de testtheorie bijna normaal geworden. Vandaar dat het voor elke test belangrijk is na te gaan volgens welke theorie en inhoudelijke desiderata een test is opgebouwd. Helaas moet men hier globaal stellen dat de testbouwers vaak goede methodologen zijn en een zekere creativiteit bezitten in het uitdenken van testitems, maar dat van enige theoretische diepgang t.a.v. het object nauwelijks sprake kan zijn. De bekendste taalvaardigheidstest bijvoorbeeld 'de Illinois Test of Psycholinguistic Abilities' (ITPA), in de praktijk zeer vaak gebruikt, zowel in Amerika, Duitsland als Nederland, heeft als theoretische ondergrond de theorie van Osgood. Deze theorie is in de wetenschap reeds lang geleden terzijde gelegd vanwege zijn onbruikbaarheid en naar ik meen ook vanwege zijn inadequaatheid. Wat verder dan in de test uiteindelijk aan deze theorie nog terug te vinden is, is minimaal. In elk geval is mij het taalvaardigheidsconcept niet geheel duidelijk.

Een ander punt dat bij de testconstructie grote aandacht krijgt, is vanuit het inhoudelijke aspect gezien als extern criterium te karakteriseren. Het gaat hierbij om het discriminerend vermogen van een test. Zo zijn de constructeurs van eerder genoemde ITPA er trots op dat de test in staat is systematische verschillen tussen verschillende ethnische groeperingen in de VS vast te stellen. Zo zou men ook klasseverschillen, sexeverschillen, leeftijdverschillen kunnen inbouwen. De vraag is en blijft echter, of dat vanuit de taalvaardigheidstheorie gemotiveerd kan worden.

Een ander punt betreft de item-selectie. Wanneer men een keer de taalvaardigheid geoperationaliseerd heeft (op wat voor basis en (voor)-oordelen dan ook), dan worden vele aspecten van taalvaardigheid gemeten, vele echter ook niet. Het is ook mogelijk dat een aantal items die verondersteld worden hetzelfde aspect te meten, in feite verschillende aspecten representeren. Na berekening der item-test correlatie worden o.a. die items verwijderd waarbij de 'minder taalvaardigen' (op basis van de gebruikte theorie) betere prestaties verrichten dan de goeden. Wanneer men beseft dat de geoperationaliseerde vaardigheden meestal op wankele benen staan, dan zal duidelijk zijn dat deze vorm van goed-slecht discriminatie een riskante onderneming is.

Een laatste punt der testconstructie betreft de vertaalbaarheid der theorie in testitems. Het is namelijk zo, dat niet dat wat getest moet worden, afgevraagd wordt, maar

wat men door middel van een test afvragen kan. Daarom worden bijv. rekenopgaven gesteld maar geen vragen over hoe men iets oplossen moet. Niet de reden van iets maar dit iets wordt gevraagd. Dit leidt tot systematische verschillen tussen bijvoorbeeld wat men taalvaardigheid zou willen noemen en dat wat gemeten wordt.

Tot zover de inhoudelijke en constructionele aspecten van een test. Er zijn echter andere aspecten aan een test die om onze aandacht vragen, te weten de objectiviteit van een test, haar reliabiliteit, en haar validiteit.

Men kan stellen dat op de projectieve tests na - de tests in het algemeen objectief zijn. Men levert een prestatie die objectief met punten gemeten kan worden. Daarmee wil niet gezegd zijn dat datgene wat objectief gemeten wordt ook inderdaad met taalvaardigheid gekarakteriseerd kan worden.

Over het algemeen zijn tests ook reliabel. Testherhaling of zgn. split half onderzoeken leveren globaal genomen dezelfde waarden op. Dit voert echter tot een merkwaardig probleem in de differentiële psychologie (waartoe de testtheorie behoort). Men gaat er in de testtheorie namelijk vanuit dat de persoonlijkheid als een bundel stabiele eigenschappen opgevat moet worden. Dat is de naieve theorie, zoals gezien, waarvan gevonden is dat zij fout is. Op grond van deze foute naieve opvatting wordt een test opgebouwd, waarbij o.a. die items moeten uitvallen die de temporele reliabiliteit in gevaar brengen. Zulke tests worden echter ontworpen in de differentiële psychologie om te kunnen diagnosticeren en om op grond daarvan wanneer dit nodig blijken te intervensiëren door therapie, training, e.d. Met andere woorden, op dit punt aangekomen zijn de eigenschappen der persoonlijkheid plotseling weer veranderlijk geworden.

Storend voor een test is volgens de motivatietheorie van Atkinson de wisselende motivatie van de persoon. Ook de situatie is een andere erkende storende variabele van de temporele reliabiliteit. Met name is hiervan veel bekend in het taalvaardigheidsonderzoek, zoals hierboven al is aangegeven. Men moet zich dan ook ernstig afvragen of temporele reliabiliteit niet een artefact der testresearch is en dit met name in de taalvaardigheidstests en, of niet juist dit gegeven een test eerder zwakker maakt dan sterker. Met deze opvatting vervalt echter de gehele basis aan de psychologische testtheorie: waarom zou ik op tijdstip a iets meten dat voor tijdstip b geen waarde meer heeft, wanneer ik in tijdstip b geïnteresseerd ben.

De validiteit van een test kan men wat mij betreft, maar ook wat de uitkomsten van onderzoek betreft, geheel vergeten. Wel is het zo, dat er een zekere validiteit t.a.v. andere verwante tests bestaat (interne validiteit) maar dat is niet zo verwonderlijk daar deze meestal op hetzelfde stramien geconstrueerd zijn en dus eigenlijk twee keer het zelfde gemeten wordt. Met de externe validiteit, daar waar het om gaat als we over validiteit spreken, is het echter treurig gesteld. Een IQ-test moet bijvoorbeeld school succes prognosticeren. Dit houdt in dat dat wat gemeten wordt met schoolsucces moet correleren. Nu is voor IQ-testen vastgesteld - en dit na bijna 100 jaar re-

search-ervaring op dit gebied - dat de validiteit voor de verschillende IQ-tests nog steeds maar tussen .20 en .40 ligt. Omgerekend wil dit zeggen dat door een IQ-test slechts .04 (=  $0.20^2$ ) tot .16 (=  $0.40^2$ ) van de variantie wordt opgehelderd. Met andere woorden: 90 percent wordt bij schoolsucces niet door wat IQ-tests meten verklaard. De prognostische capaciteit van IQ-tests is dus miniem. Het is dan ook een kleine ramp mee te moeten delen dat IQ-tests de beste tests zijn die de differentiële psychologie heeft voortgebracht. Dit houdt in dat de validiteit van taaltests nog veel lager liggen. Waarom testen we eigenlijk nog?

Voor taalvaardigheid en het testen daarvan kan men het volgende stellen:

1. meestal wordt competence of weten afgevraagd, hoewel men iets over gedrag (performance) en gedragsmogelijkheden wil weten;
2. wordt taalvaardigheid gemeten dan hebben deze opgaven weinig met de dagelijkse praktijk te maken; bijv.:
  - a. maak van 2 zinnen 1 zin;
  - b. maak van de volgende rij woorden een zin;
  - c. zoek een synoniem voor x;
  - d. welk woord betekent dat .... (en omgekeerd);
  - e. het antoniem van a is b;
  - f. transformeer vorm a in vorm b; etc.
3. temporele reliabiliteit en taalvaardigheid hebben niets met elkaar te maken; gegevens hierover in de verantwoording van de test zeggen iets over de constructeur, niets over taalvaardigheid en dus niets over de kwaliteit van een test;
4. taalvaardigheid is zo complex dat hij daarom alleen al niet door een test te meten is. Wanneer men de linguïstische en psychologisch relevante variabelen gevonden heeft, moet men ze nog in elke situatie onderzoeken, waarbij de testsituatie er slechts een is. Wanneer men beseft dat taalvaardigheid voor een belangrijk onderdeel bepaald wordt door met wie men in welke situatie spreekt, dan zal duidelijk zijn dat de aangesproken variabele het aantal te onderzoeken situaties ongelofelijk uitbreidt;
5. de validiteit van een taaltest is gering. Van objectiviteit is alleen maar sprake voor zover de score zich in een cijfer laat uitdrukken. Niet, echter, voor de keuze van de items.

## 6. Conclusies

Tot nu toe heb ik alleen vrij negatief gezegd wat men niet zou moeten doen. Ik heb de indruk, dat ik in deze mening niet op te veel medestanders moet rekenen. Dit komt mogelijkwijze daardoor, dat ik niet langzaamaan met de methoden van de psychologie vertrouwd ben gemaakt, maar, door mijn biografie bepaald, plotseling om mijn mening over allerhande dingen in de psychologie gevraagd ben. Op deze wijze ontbrak een gewenningseffect. Langzaam maar zeker, echter, merk ik dat er ook een aantal psychologen tot ongeveer de gelijke slotsom lijken te komen als ik. Ik noem hier alleen Wottawa (1979) die op eerder externe niet-psychologische, namelijk



mathematisch-statistische gronden tot de opvatting komt dat de psychologie van de gemiddelden niet een essentiële bijdrage tot onze kennis van psychologische processen kan leveren en Fiske (1978) die in zijn boek *Strategies for personality research* met als ondertitel *The observation versus interpretation of behavior* t.a.v. de testpsychologie tot vergelijkbare conclusies komt en ook tot dezelfde oplossing hoe het beter kan en moet.

Wat ons aan onderzoeksmogelijkheden overblijft is de observatie van (taal)gedrag. De voorkeur verdienen zo volledig mogelijke beschrijvingen met zo min mogelijk theorie, of systematische observatie van een bepaald fenomeen zoals turn taking of "contact maken" e.d. Met de uitkomsten van zulk onderzoek zou men de bestaande psychologische theorieën kunnen evalueren en eventueel corrigeren. Op grond van zulke evaluaties zou men in een tweede stap tot doelgerichtere analyses van taalgedrag kunnen komen. Belangrijk is dat men taalhandlungsstrategieën en regelmatigheden opspoorde in individueel gedrag, dat men in ieder geval geen gemiddelden van een groep mensen berekent, maar gedragingen per geval beschrijft. Wanneer men weet hoe een bepaald individu zich in een bepaalde situatie gedraagt, en wat van deze gedragingen de daaronder liggende processen zijn en wat de belangrijkste variabelen, dan kan men hypothesen opbouwen over hoe dit er voor andere individuen uitziet. Mogelijkerwijze moet men dan de gedragsregels differentiëren, gedragsalternatieve regels opbouwen e.d. Voor een verdere uitwerking van deze werkwijze verwijs ik naar Fiske (1978)<sup>6</sup>.

## Noten

1. Kenmerkend is Chomsky (1965, p. 139): "Thus it seems absurd to suppose that the speaker first forms a generalized Phrase-marker by base rules and then tests it for well-formedness by applying transformational rules to see if it gives, finally, a well-formed sentence. But this absurdity is simply a corollary to the deeper absurdity of regarding the system of generative rules as a point-by-point model for the actual construction of a sentence."; of: (o.c., p. 140): "But a generative grammar as it stands is no more a model of the speaker than it is a model of the hearer (...) it can be regarded only as a characterization of the intrinsic tacit knowledge or competence that underlies actual performance."
2. Hiermee wil ik niet een negatief oordeel over dit werk uitspreken. Veel meer is mij er aan gelegen vast te stellen dat wanneer men zonder trucs rekent, men eenvoudig vindt dat de globale oordelen op niets berusten. Dit is een onverwachte verdienste ervan.
3. Dit gegeven zou ook tot interessante interpretaties t.a.v. de uitkomsten van de sexeverschillen en de sociale klasseverschillen kunnen leiden.
4. We laten hierbij nog buiten beschouwing dat de slechts als ordinaal te beschouwen schalen in dit soort onderzoek vrijwel uitsluitend als intervalschalen behandeld worden. Men kan zo wel lekkerder rekenen, maar de uitkomsten zijn met grote waarschijnlijkheid fout.

5. Waar heb ik dergelijke uitspraken toch meer gehoord t.a.v. Chomsky's beschrijvingsysteem? In dit licht is het interessant noot 1 er nog eens bij te halen.
6. Tot slot moet ik me bij mijn vriend en collega Müller-Eckhard bedanken, met wie ik over deze onderwerpen zo veel heb gesproken, dat niet meer op te maken is of ik of hij voor de hier geformuleerde gedachten verantwoordelijk is.

## Bibliografie

- Appel, R.: *Het meten van taalvaardigheid*, Inst. v. Alg. Tw. Amsterdam, 1972.
- Bakker, H. & M. Hoefnagels: *Situatieverschillen in taalgebruik*, Inst. v. Tw. Univ. van Amsterdam, 1974.
- Bereiter, C. & S. Engelmann: *Teaching disadvantaged children*, Englewood Cliffs, 1966.
- Bernstein, B.: *Class, codes and control*, vol. 1..., London, 1971.
- Bungard, W. (Hg.): *Die gute Versuchsperson denkt nicht; Artefakte in der Sozialpsychologie*, München, 1980.
- Cazden, C.: "The neglected situation in child language research and education" in Williams (ed.), 1970.
- Chomsky, N.: *Aspects of the theory of syntax*, Cambridge (Mass), 1965.
- Elffers-van Ketel, E., S. de Haan en W. Klooster: "Een (+FANTAS-TISCHE) macrostructuur", *Spektator*, 3(1973-1974)p. 581-600.
- Feyerabend, P. (1962): *Knowledge without foundations*, Oberlin (Ohio), 1962.
- Fiske, D.W.: *Strategies for personality research*, San Francisco, 1978.
- Fodor, J. & Th. Bever: "The psychological reality of linguistic elements", *JVLVB* 4(1965)p. 414-420.
- Fodor, J. & M. Garrett: "Some reflections on competence and performance". In: J. Lyons & R. Wales (eds): *Psycholinguistics papers*, Chicago, 1966, p. 135-154.
- Garfinkel, H.: "Das Alltagswissen über soziale und innerhalb soziale Strukturen". In Arbeitsgruppe Bielefelder Soziologen (Hg.): *Alltagswissen, Interaktion und Gesellschaftliche Wirklichkeit*, Reinbek, 1973.
- Geest, T. van der e.a.: *The child's communicative competence*, The Hague, 1973.
- Geest, T. van der: *Entwicklung der Kommunikation*, Bochum, 1978.

- Geest, T. van der: "How to become a native speaker: one simple way". In: Florian Coulmas (ed.): *A Festschrift for the native speaker*, The Hague, 1980.
- Geppert, U.: *Sprachmilieu und Leistungsmotivgenese*, Bochum (inaugural Diss.) 1979.
- Hörmann, H. (1977): *Psychologie der Sprache*, 2. überarb. Aufl., Berlin, 1977.
- Ierland, M. van: "Tussen vier en acht: ontwikkelingen in taalgebruik". *Toeg. Taalwetenschap in Artikelen* 7(1979)p. 86-101.
- Jansen, C.: "Oordelen over teksten". *Tijdschrift voor taalbeheersing* 2(1980)p. 324-343.
- Koster, K.: *De ontwikkeling van het getalbegrip*, Groningen, 1975.
- Labov, W.: *The logic of nonstandard English*, In: Williams (ed.) 1970.
- Labov, W. & Fanshel, D.: *Therapeutic discourse; Psychotherapy as conversation*, New York, etc. 1977.
- Levelt, W.: "Some psychological aspects of linguistic data", *Linguistische Berichte*, 17(1971)p. 18-30.
- Müller-Eckhard, E.: *Ausserungslänge und Eindrucksbildung*, Bochum, 1974.
- Neisser, U.: *Cognition and reality*, San Francisco, 1976.
- Pienemann, M.: *Nützt die Spracherwerbsforschung den ausländischen Arbeiterkindern?* Universität Passau, 1980.
- Schönbach, P.: Soziolinguistik. In H. Koch (Hg.) *Perspektiven der Linguistik*, Stuttgart, 1974.
- Snow, C.: "Linguists as behavioral scientists". In A. Kraak (ed.): *Linguistics in the Netherlands*, Assen, 1975.
- Stoffer, T.: *Wahrnehmung und Representation musikalischer Strukturen*, inaug. Diss. Bochum, 1980.
- Wesdorp, H.: *Het meten van de produktief-schriftelijke taalvaardigheid*, proefschrift Amsterdam, RITP. Amsterdam, 1974.
- Williams, F. (ed.) *Language and poverty*, Chicago, 1970.

Wottawa (1979)