

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/19561>

Please be advised that this information was generated on 2017-12-05 and may be subject to change.

N en nog meer

INAUGURELE REDE DOOR DR. A.C.M. RIETVELD

Radboud Universiteit Nijmegen



N EN NOG MEER

N en nog meer

3

Rede uitgesproken bij de aanvaarding van het ambt van hoogleraar Methodologie van Onderzoek op het terrein van de Taal- en Spraakpathologie aan de Faculteit der Letteren van de Radboud Universiteit Nijmegen, op vrijdag 17 december 2004.

door dr. A.C.M. Rietveld

INLEIDING

Mijnheer de rector, dames en heren,

De titel van mijn openbare les is ambigu: *N en nog meer*. Fonetici en fonologen denken waarschijnlijk: nu volgt een verhandeling over nasalen, en wel in het bijzonder over de 'ng' van 'bang', en de spreker hanteert hierbij het transcriptiesymbool van het systeem SAMPA. Een behandeling van de 'ng' suggereert echter niet direct toegang tot linguïstische vergezichten. Statistici en methodologen veronderstellen mogelijk dat nu een betoog over de steekproefgrootte volgt, gewoonlijk aangeduid met de hoofdletter N. Dat is een, statistisch gezien, nogal afgezaagd onderwerp. Taal- en spraakpathologen denken dat ze bij de verkeerde voorstelling zijn beland en proberen nu ongemerkt de zaal te verlaten. Kortom, de verwachtingen zijn op dit moment niet erg hoog gespannen. In feite is deze enigszins confuse situatie een goede afspiegeling van de situatie in het land van de taal- en spraakpathologie: mensen van allerlei disciplines werken op dit terrein en allen hebben hun eigen onderzoekstraditie en ook hun eigen verwachtingen: KNO-artsen, logopedisten, neurologen, orthopedagogen, psychologen, psycholinguïsten, fonetici, neurolinguïsten, taalwetenschappers en niet te vergeten: taal- en spraakpathologen.

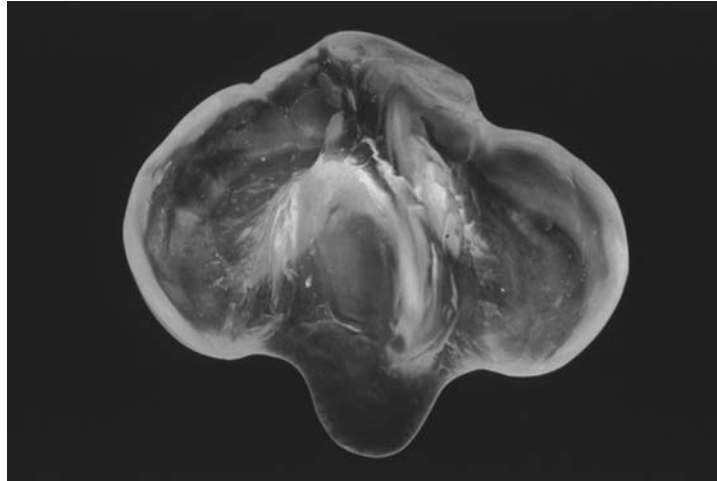
Om gestommel in de zaal te vermijden zou ik taal- en spraakpathologen willen aanraden nog even te blijven zitten, en ook de fonetici en methodologen komen mogelijk nog aan hun trekken. Over biologen, accountants, beheerders en andere zeer nuttige mensen in onze universitaire gemeenschap heb ik helemaal niets te zeggen.

Ik zal twee boodschappen aan de orde stellen:

1. Boodschap 1: N en de noodzaak om gestandaardiseerde afhankelijke variabelen te gebruiken in de taal- en spraakpathologie.
2. Boodschap 2: Meer meta-analyses moeten worden uitgevoerd, en er kan een groter onderscheidend vermogen worden bereikt door standaardisatie van variabelen in de Taal- en Spraakpathologie en vermindering van onnodige variatie.

1. BOODSCHAP 1: N EN DE NOODZAAK OM GESTANDAARDISEERDE
 AFHANKELIJKE VARIABELEN TE GEBRUIKEN

Een van de twee boodschappen van mijn les is mede gebaseerd op de ervaring die ik heb opgedaan in de stuurgroep van een longitudinaal onderzoeksproject dat door de Ziekenfondsraad is gefinancierd en waarop Mw. E. Konst in 2002 is gepromoveerd (*The effects of infant orthopaedics on speech and language development in children with unilateral cleft lip and palate*). Het betrof een project waarin werd nagegaan of het aanbrengen van een gehemelteplaatje bij heel jonge kinderen, alvorens het gehemelte operatief werd gesloten, zinvol was.



Figuur 1: Gehemeltesplaatje, met dank aan mw. dr. E. Konst

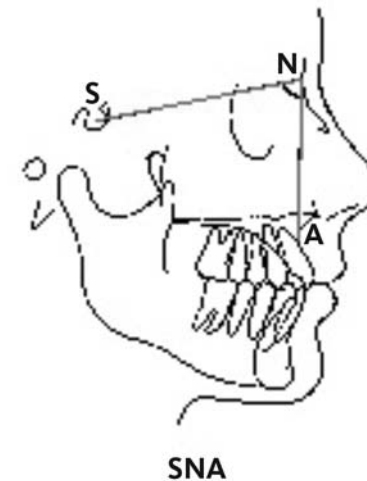
De zinvolheid van de behandeling werd geoperationaliseerd in orthodontische variabelen, zoals de breedte van de spleet, maar ook in voedingsgedrag en in variabelen op het terrein van de spraak- en taalontwikkeling, zoals nasaliteit (dat is het perceptieve effect van het bijschakelen van de neusholte als resonator bij de mond-keelholte; bij schisis kinderen gebeurt dat continue, door de gehemeltespleet), verstaanbaarheid, realisatie van specifieke medeklinkers zoals /p/ en /k/, gemiddelde zinslengte enzovoort. Een - tegenwoordig belangrijke - variabele, de kosteneffectiviteit, laat ik voor het gemak nu even buiten beschouwing; op dat belangrijke aspect kom ik echter later terug.

De taal- en spraakcomponent is wat later toegevoegd aan het project. De benodigde omvang van de steekproeven (deze zouden bestaan uit een groep bij wie wél een plaatje zou worden aangebracht, en een controlegroep bij wie dat niet gebeurde) was vastgesteld door collega's van MIES, de afdeling Medische Informatiekunde, Epidemiologie en Statistiek van het UMCN, nu Epidemiologie en Biostatistiek geheten. Zij hadden de benodigde steekproefgrootte bepaald aan de hand van de vier klassieke ingrediënten:

- Welk effect wil men detecteren als het inderdaad in de populatie aanwezig is,
- Wat is de spreiding van de betreffende afhankelijke variabele in de populatie(s),
- Welke kans op het ten onrechte aantonen van een effect neemt men voor lief (de zogenaamde type I fout; 'alpha'),
- Welk risico wordt geaccepteerd om een in de populatie(s) wel aanwezig effect in de steekproef niet aan te kunnen wijzen (type II fout, 'beta'). Het vermogen van de test (ook wel onderscheidingsvermogen of power genoemd) is $1 - \text{de kans op het maken van een type II fout}$.

7
 Samengevat: $N = f(\text{grootte van het effect in de populatie dat men wil detecteren, standaarddeviatie(s) van de afhankelijke variabele(n), type I fout, type II fout})$

De statistici gebruikten in overleg met de onderzoeksleiders als afhankelijke variabele de cephalometrische hoek SNA (sella-nasion angle), uitgedrukt in graden (de standaarddeviatie werd geschat op 4.1° , de effectgrootte die men wilde detecteren op 3°). De onderzoekers hadden kennis van de variatie in de SNA die je bij de kinderen kon verwachten, en de type I en II fouten mochten ze zelf bepalen (zij hanteerden de veel gebruikte waarden van respectievelijk 0.05 en 0.20). De klassieke formules leidden tot $N = 23$ in beide groepen, vermeerderd tot 25 om eventuele uitval op te vangen (Prah, 1993).



Figuur 2: Sella-nasion angle

Toevoeging van de spraak- en taalcomponent aan het onderzoek leidde niet tot bijstelling van de steekproefomvang. Zelfs als Dr. Herman Peters van het UMCN en ik assertiever waren geweest of eerder bij de opzet van het onderzoek betrokken zouden zijn geweest, zouden we weinig gegevens hebben kunnen bijdragen om de omvang van de steekproeven te bepalen en eventueel te wijzigen. Immers het opgeven van een effectgrootte en de standaarddeviatie van de betrokken afhankelijke variabele vooronderstelt het bestaan van gestandaardiseerde variabelen en ervaring met die variabelen. Helaas, met name voor spraakvariabelen zijn nog weinig gestandaardiseerde maten beschikbaar, vooral als het om perceptieve maten gaat, maten waarbij luisteraars het meetinstrument

8

vormen. Zo gebruikte een invloedrijke Britse onderzoeksgroep op het terrein van gespleten gehemelte-spraak (cf. Grunwell, 1993) een maat voor waargenomen nasaliteit die uitgedrukt werd op een schaal van 0 tot 2, en gebruikte een Nederlandse groep een schaal van 1 tot 7 (Konst, 2002), terwijl voor verstaanbaarheid door Jansonius (1999) en Konst (2002) een schaal van 1 tot 10 (in aansluiting bij de Nederlandse schoolcijfers) werd gehanteerd. Peng et al. (2004) hanteerden daarentegen weer een schaal van 1 tot 5 bij hun experiment over onder andere de verstaanbaarheid van spraak van kinderen met een cochleair implantaat, een instrument waarmee mensen die geen functionerend binnenoor meer hebben toch klanken kunnen analyseren.

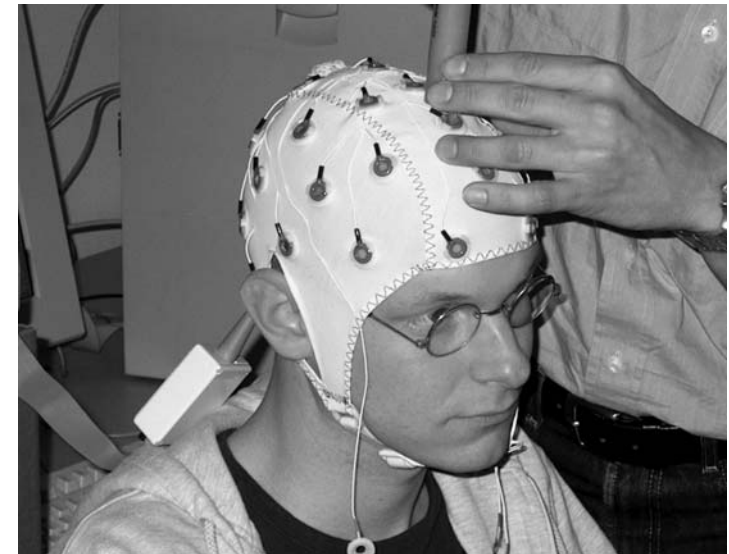
Kortom, als spraak- en taalpatholoog respectievelijk methodoloog stonden dr. Herman Peters van het UMCN en ik met lege handen toen het erom ging de grootte van de steekproeven vast te leggen om een nader te bepalen effect van de prechirurgische plaatsing van een plaatje bij kinderen met een gespleten gehemelte te kunnen vaststellen. Deze omstandigheid lijkt de spraak- en taalpathologie in het voorwetenschappelijke stadium te brengen, een stadium waarin medisch onderzoek niet meer verkeert. Ik weet niet precies wat daarvan de oorzaak is. Er is een aantal factoren dat hierbij een rol kan spelen:

- 1 Spraak- en taalpathologisch onderzoek is veelal interdisciplinair van karakter, en vaak weten de betreffende onderzoekers – met hun psychologische dan wel fonetische achtergrond – meer van statistiek dan hun medische collega's. De laatsten zoeken dan ook vaker hulp bij echte statistici/methodologen, en komen zo eerder bij de vraag: hoe groot moeten de steekproeven zijn?
- 2 Medisch onderzoek is vaak belastend: om die reden willen de onderzoekers op ethische gronden niet méér proefpersonen (in het Angelsaksische taalgebruik nu 'participants' genoemd) gebruiken dan strikt noodzakelijk. Taal- en spraakpathologisch onderzoek is vaak ook belastend, denk aan tongdrukmetingen, metingen met Elektromagnetische Midsagittale Articulografie (EMMA), metingen aan Event Related Potentials (ERP) enzovoort, maar wordt vaak niet als zodanig geafficheerd.
- 3 De afhankelijke variabelen zijn vaak niet gestandaardiseerd; ik heb daar hierboven al op gewezen en ik denk dat deze factor de belangrijkste is. Deze factor zorgt er bijvoorbeeld voor dat gegevens die worden geleverd door onderzoekers die een schaal van 0 tot 2 gebruiken, niet direct te gebruiken zijn door onderzoekers die op basis daarvan hun steekproefgrootte willen bepalen, maar wel een schaal van 1 tot 7 willen hanteren. Natuurlijk, we weten dat de effectgrootte ook kan worden gestandaardiseerd (cf. Cohen 1977). Echter, het bereik van de schaal zal zeker ook effect hebben op de spreiding (een ingrediënt van de standaardisatie). Uiteinden van schalen worden vaak niet gebruikt. Aldus zal een schaal met waarden van 0 tot 2 dreigen te krimpen tot een schaal met 1 waarde. De bijbehorende standaarddeviatie zal dus ook wel klein worden.

9



Figuur 3: EMMA-meting, met dank aan de firma Carstens te Göttingen



Figuur 4: ERP-meting, met dank aan mw. dr. M. Lamers van de afdeling Taalwetenschap

10

De volgende vraag is natuurlijk: waarom zijn de variabelen niet gestandaardiseerd? Ik heb daarover vermoedens, helaas niet meer dan dat.

- a) Metingen in de taal- en spraakpathologie hebben betrekking – het woord zegt het al – op spraakuitingen. In de ene taal is het realiseren van specifieke toonhoogtebewegingen cruciaal voor de verstaanbaarheid, voor de andere taal is het ‘slechts’ (dit laatste met opzet tussen aanhalingstekens) van belang voor de structurering van het discours. In de ene taal is nasaliteit betekenisonderscheidend, in een andere taal is dat niet zo (denk aan het Frans en het Nederlands); om die reden wordt er bijvoorbeeld in de handleiding van de Nasometer – een instrument om de relatieve nasale geluidsdruk te meten – op aangedrongen normwaarden op basis van eigen teksten te bepalen. De manier waarop afwijkingen van taalspecifieke elementen worden gemeten, wordt dan ook vaak overgelaten aan de onderzoekers van het betreffende taalgebied, zie ook Okalidou, Karathanasi & Grikoraki (2003). Op die wijze wordt het moeilijk om algemeen geaccepteerde en gestandaardiseerde afhankelijke variabelen te introduceren.
- b) De onderzoekers hebben een verschillende achtergrond; mensen met een taalwetenschappelijke achtergrond richten zich vaak op andere aspecten van de talige communicatie dan gedragswetenschappers; de eisen die ze stellen aan de afhankelijke variabelen zijn vaak ook verschillend. Taalwetenschappers gebruiken meer discrete variabelen (een segment wordt al-dan-niet gedeleerd), terwijl fonetici meer geïnteresseerd zijn in gradiënte verschijnselen (een segment uitgesproken in de experimentele groep verschilt gemiddeld n milliseconden van een segment uitgesproken door een controlegroep).
- c) Een verbijzondering van het vorige punt is dat technici en linguïsten vaak niet dezelfde taal spreken. Een technicus bedenkt bijvoorbeeld een formule voor de meting van de variatie in toonhoogte, of, om meer precies te zijn, in de lengte van de periodetijd P van de stembandtrillingen (*jitter*):

$$jitter = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} P_i - P_{i+1}}{\frac{1}{N} \sum_i P_i}$$

Figuur 5: Formule voor de bepaling van variatie in stembandtrillingen; P is de periodetijd van een stembandtrilling (uitgedrukt in ms), i = periodenummer, N = het aantal perioden van het spraakfragment.

11

U kunt zich voorstellen wat in deze formule kan worden gevarieerd:

- het aantal perioden (overeenkomend met het aantal keren dat de stembanden tegen elkaar slaan) waarover gemiddeld wordt: N
- de omvang van het venster (kijken we alleen naar de periode i en de volgende periode $i+1$, of kijken we ook naar achteren ($i-1$);
- hoe ver kijken we naar achteren en naar voren: 1, 3, of 5 perioden;
- het middelingsproces; middelen we gewoon (delen door N), of schatten we de gemiddelde waarde van de periodetijden op een andere manier (mediaan, modus, via regressie-analyse enzovoort)?

Afhankelijk van het teksttype (aangehouden klinkers, losse woorden, zinnen met veel en weinig toonhoogtebewegingen) is de ene formule of maat geschikter om pathologische jitter te detecteren dan de andere. Van belang is dan wel dat de onderzoeker zich bewust is van de tekstsoort en de daarmee samenhangende eisen aan de betreffende maat.

Mijn pleidooi voor standaardisatie moet niet misverstaan worden. Voor men het weet is een verkeerde benadering standaard of mode geworden. Een voorbeeld van een verkeerde standaardisatie is die welke nu gebruikt wordt bij het rapporteren van psycholinguïstisch en soms ook taal- en spraakpathologisch onderzoek.

Van Hout, Schreuder en ik hebben een paar jaar geleden enige jaargangen van de invloedrijke tijdschriften *Journal for Memory and Language*; *Journal for Psycholinguistic Research*; *Journal for Speech and Hearing Research*; *Language and Cognitive Processes* en *Memory and Cognition* gescreend; het betrof onderzoek waarin psycholinguïstische data afkomstig uit repeated measures designs waren geanalyseerd, designs waarin proefpersonen herhaaldelijk worden gemeten (het betrof de jaargangen 1994 en 1995: 145 artikelen). In dit onderzoek bleek vaak een analysemethode gehanteerd te zijn die echt fout is; bij die foute analysemethoden werden ook niet de toetsen gebruikt die nodig zijn om vast te stellen of aan de assumpties voldaan wordt waarop die, overigens foute, analyses gebaseerd zijn. Het gaat hier om de zogenaamde F_1 en F_2 analyse; bij de eerste wordt gedaan alsof de proefpersonen random uit de populatie getrokken zijn en de woorden niet (de laatste noemt men dan een fixed factor), bij de tweede alsof de proefpersonen een fixed factor vormen en de woorden een random factor (in de meeste gevallen zijn de proefpersonen een quasi-random factor en de woorden een fixed factor). Bij de eerste analyse hoort een Huynh-Feldt correctie om te corrigeren voor een eventuele schending van de sphericiteits-assumptie, bij de tweede analyse een test van Levine om vast te stellen of voldaan is aan de assumptie van homogeniteit der varianties (cf. Rietveld & van Hout, 1993). Niemand voerde de benodigde toetsen uit, niemand

12

die in Nijmegen een proefschrift schrijft – buiten de Faculteit der Letteren, op enige uitzonderingen na – doet dat. Het is alsof de heren Greenhouse & Geisser en Huynh & Feldt niet hebben bestaan (dertig jaar geleden waren ze al zeer actief) en de statistiekboeken niet op deze diagnostische toetsen hebben gewezen (Rietveld & van Hout, 1993).

Raaijmakers, Schrijnemakers & Gremmen (1999) lieten in hun zeer lezenswaardige artikel 'The $F_1 \times F_2$ fallacy' gelijksoortige resultaten zien: zij bekeken zes delen van *Journal of Memory and Language* (1993-1997). Van de 124 relevante artikelen rapporteerden er 120 $F_1 \times F_2$ als criterium en slechts vier naast F_1 en F_2 , min F' . Door naar een langer tijdvak te kijken (vanaf 1974) stelden zij vast dat vanaf een standaard gebruik van min F' de standaard nu $F_1 \times F_2$ is, terwijl dat NIET de juiste procedure is (het $F_1 \times F_2$ criterium is positief gebiased (met dus een te hoge type I fout)) wanneer de factor items (woordmateriaal) een echt random factor voorstelt (cf. Forster & Dickinson (1976)). Raaymakers, Schrijnemakers & Gremmen (1999) zeggen dat in de praktijk de F_1 de correcte test statistic is. Helaas, dit artikel heeft de foute praktijk niet weten te wijzigen, hetgeen niet veel goeds belooft voor de effecten van mijn oratie.

Wat in die tijdschriften evenmin gedaan werd en wordt, is vaststellen hoe groot de steekproef zou moeten zijn om een bepaald effect te kunnen detecteren. Laat ik beginnen met enige spraak- en taalpathologische tijdschriften. Wij hebben de tijdschriften *Journal of Speech and Hearing Research* en *Clinical Linguistics & Phonetics* gescreend, en wel de jaargangen 2002 en 2003. Van de 253 artikelen die betrekking hadden op twee of meer onderzoeks- en/of controlegroepen vonden wij slechts twee artikelen waarin gemeld werd dat men op basis van poweranalyse (analyse van het onderscheidingsvermogen) de steekproefgrootte had bepaald. Dat is een ratio van 0.008, vrij gering om een understatement te gebruiken.

Tabel 1: Aantal keren dat in twee jaargangen van twee tijdschriften op het terrein van de Taal- en Spraakpathologie vermeld is dat men de steekproefgrootte a-priori heeft bepaald.

Tijdschrift:	2002		2003	
	aantal in aanmerking komende onderzoeken	aantal onderzoeken dat benodigde N a-priori bepaalde	aantal in aanmerking komende onderzoeken	aantal onderzoeken dat benodigde N a-priori bepaalde
Journal of Speech and Hearing Research	102	0	76	1
Clinical Linguistics & Phonetics	32	0	43	1

13

Het geringe aantal artikelen waarin a-priori is vastgesteld hoe groot de steekproefgrootte zou moeten zijn om een tevoren vastgesteld onderscheidingsvermogen te bereiken is verbazingwekkend, omdat er tegenwoordig prachtige software beschikbaar is om de steekproefgrootte te bepalen: *NQuery* en *SamplePower* (dr. van Limbeek heeft al lang geleden een handig programma voor poweranalyse gemaakt voordat de commerciële sector zich op deze problematiek stortte; wel toevallig: hij is een epidemioloog van de Sint Maartenskliniek). Ik ben geneigd om te wedden om iets dat ik nu niet wil noemen, gezien de tekstsoort (een oratie), dat deze programma's in de meeste multimediale studiecentra of computerzalen niet aanwezig zijn. In feite is deze omstandigheid een bevestiging van de eerder genoemde situatie: het is moeilijk eenmaal ingezette dan wel gevestigde modes te veranderen.

Bij de steekproefgrootte en het onderscheidend vermogen van toetsen denk ik aan een artikel dat in 1978 door Elstrodt & Mellenbergh is gepubliceerd: 'Een minus de vergeten fout'. De vergeten fout is de type II fout, waarmee elk inleidend boek in de statistiek begint, althans in hoofdstuk 2. De type II fout gaat om de kans die een onderzoeker wil accepteren om een in de populatie bestaand effect niet te detecteren. Niemand in de psycholinguïstiek, bijna niemand in de psychologie, niemand in de taal- en spraakpathologie, maar bijna iedereen in de medische wetenschap maakt zich daarover zorgen.

Elstrodt & Mellenbergh hadden ook een tijdschrift-survey doorgevoerd, al vele jaren geleden; het betrof de jaargangen 1971 tot en met 1974 van het *Nederlands Tijdschrift voor de Psychologie*. Zij vonden dat het onderscheidend vermogen van 503 toetsen waarover in de genoemde jaargangen was gerapporteerd over het algemeen zeer gering was, wat erop neerkomt dat de steekproefgrootte onvoldoende was geweest om een redelijk onderscheidend vermogen te waarborgen. In de volgende tabel reproduceer ik een deel van tabel 5 van Elstrodt & Mellenberghs artikel.

Tabel 2: De klasse waarin het mediane onderscheidingsvermogen van een aantal statistische toetsen valt voor drie verschillende effectgrootten en een significantieniveau van 0.05 (tweezijdig).

Effect: Toets:	Klein	Middelmatig	Groot
t- en Mann Whitney	0.00-0.09	0.20-0.29	0.50-0.59
Tekentoets	0.00-0.09	0.10-0.19	0.30-0.39
Chi-kwadraat toets	0.20-0.29	0.50-0.59	0.80-0.89
F-toets	0.00-0.09	0.20-0.29	0.60-0.69

14

De tabel laat duidelijk zien dat het onderscheidingsvermogen voor kleine en middelmatige effecten veel kleiner is dan de vaak gewenst geachte waarde van 80 procent. Ik vrees dat een soortgelijke tabel gemaakt op basis van gegevens in gepubliceerde (of geplande) artikelen op het terrein van de taal- en spraakpathologie ongeveer dezelfde gegevens zou opleveren.

2. BOODSCHAP 2: MEER META-ANALYSES MOETEN WORDEN UITGEVOERD, EN ER KAN EEN GROTER ONDERSCHIEDEND VERMOGEN WORDEN BEREIKT DOOR STANDAARDISATIE VAN VARIABELEN IN DE TAAL- EN SPRAAK-PATHOLOGIE EN VERMIJDING VAN ONNODIGE VARIATIE.

We hebben nu laten zien dat er relatief weinig aandacht is voor steekproefgrootte en standaardisatie van variabelen in de taal- en spraakpathologie. Ik heb ook vermoedens geuit omtrent de oorzaken van deze betreurenswaardige situatie. We kunnen nu bij de pakken neerzitten, of er wat aan gaan doen. Ik denk dat er minstens twee acties mogelijk zijn: de ene heeft betrekking op meta-analyses (dat komt neer op roeien met de riemen die je hebt, maar dan op een verstandige manier), de andere probeert de problemen wat meer bij de wortel aan te pakken, en het onderscheidingsvermogen te vergroten door variabelen te standaardiseren en onnodige variatie te verkleinen.

2.1 *meta-analyses*

Ik heb in het voorafgaande gepleit voor poweranalyses en bepalingen van de benodigde steekproefgrootte om vooraf vastgestelde effectgrootten te kunnen detecteren. Nu is dat makkelijker gezegd dan gedaan. Proefpersonen of patiënten zijn vaak niet in voldoende aantallen beschikbaar. Soms is dat heel erg prettig, omdat in een bepaald jaar bijvoorbeeld relatief weinig kinderen met een specifieke aandoening worden geboren. Men moet dan werken met een relatief kleine steekproef – bijna even groot als de populatie in dat jaar, waardoor de beslissing tussen het bedrijven van ‘inductieve’ dan wel ‘descriptieve’ statistiek interessant wordt – en daarover rapporteren. Het volgende jaar is er misschien weer een kleine groep proefpersonen te vinden waarover gerapporteerd kan worden. Elke keer zal de power klein zijn omdat de steekproefgrootte – N – te klein is. We kunnen dan onze toevlucht nemen tot meta-analyses.

Meta-analyses zijn analyses van de resultaten van meer dan een experiment. Een voorbeeld: het is denkbaar dat vijf onderzoeken zijn uitgevoerd naar de effecten van twee stottertherapieën, A en B; geen van die onderzoeken liet zien of A beter was dan B, in de zin dat de effecten significant waren, maar in vier van de vijf gevallen bleek A het beter te doen. Het gebrek aan power (onder meer als gevolg van een te kleine N) was waarschijnlijk oorzaak van de niet-significantie. Vier van de vijf niet-significante resultaten in een bepaalde richting lijken wel op iets te wijzen; misschien is A toch wel beter dan B. Meta-analyse maakt het mogelijk om subresultaten te gebruiken voor een

meer algemene uitspraak van het type ‘A’ is significant beter dan ‘B’. Meta-analyses in de spraak- en taalpathologie zijn nog niet usance, maar zouden dat wel moeten worden. Echter, omdat we niet alleen willen vaststellen of x onderzoeken aanleidingen geven tot de veronderstelling dat inderdaad behandeling A beter is dan B, maar ook willen weten in welke mate A beter is dan B, de effectgrootte, en tegen welke prijs – hier komt de kosteneffectiviteitanalyse (Severens, Konst & Prah, 1998) om de hoek kijken –, moeten de onderzoeken waarop meta-analyse wordt uitgevoerd gebruik maken van gestandaardiseerde afhankelijke variabelen. Bovendien moeten dan wel de voor de bepaling van de effectgrootte in meta-analyses benodigde gegevens geleverd worden (cf. Korting & Nouri, 2000).

Er moeten dus meer meta-analyses worden uitgevoerd, maar er moet meer gebeuren:

- 1) We moeten proberen overeenstemming te bereiken over de wijze waarop afhankelijke variabelen worden gehanteerd (bijvoorbeeld: door af te spreken dat we zevenpuntsschalen of tienpuntsschalen gaan gebruiken wanneer we Equal Appearing Interval schaling toepassen, of misschien wel Visual Analog Scales of Magnitude Estimation in perceptieve testen (cf. Chen & Rietveld, 2004).
- 2) We moeten de variatie in de afhankelijke variabelen gering houden; op die wijze is het mogelijk het vermogen van de gebruikte toetsen te vergroten. Het is daarom van groot belang rekening te houden met de factoren die ongewenste variatie in de afhankelijke variabelen teweeg kan brengen.

2.2 *de rol van de fonetiek in de taal- en spraakpathologie.*

Naar mijn stellige overtuiging kan de fonetiek een belangrijke rol spelen bij de opzet van spraak- en taalpathologisch onderzoek. Ik begin daarom over de situatie in Nederland enigszins ongerust te worden. De taal- en spraakpathologie wordt op steeds ruimere schaal in klinische en academische contexten beoefend (ik zie ook steeds meer visitekaartjes met als beroep: clinical linguist), en ik zie op de fonetische congressen een toename van het aantal bijdragen in secties die betrekking hebben op spraak- en taalpathologie.

Ondanks dat feit neemt het aantal mensen dat fonetische kennis heeft, eerder af dan toe. Dit jaar (2004) hebben afscheid genomen: Dr. Tj. de Graaf (RUG, de enige Nederlandse foneticus met een eredocraat), Prof. Nooteboom (UU), dr. van Beinum (UvA), Prof. Pols van de UvA stopt binnenkort, andere fonetici zijn technoloog geworden; kortom het aantal fonetische specialisten neemt af.

Bij elke verandering van context kan natuurlijk de vraag worden gesteld: is dat allemaal zo erg? In ons geval: Is die fonetische kennis dan noodzakelijk? Ja, die is noodzakelijk, is mijn stelling, om minstens drie redenen:

- 1 Pathologische spraak – tengevolge van dysartrie, dyspraxie, afasie, stotteren, schisis, stempathologie, velofaryngaal disfunctioneren, – is zeer variabel. Niet-pathologische

15

spraak – van mannen, vrouwen, kinderen, sprekers uit Amsterdam, Leiden, Nijmegen, Utrecht en Groningen – is ook zeer variabel. Kennis van die variabiliteit is noodzakelijk om de pathologische variabiliteit goed te kunnen inschatten. De omgang met normale spraak geeft de onderzoeker ook een idee van de effecten die specifieke taken kunnen hebben op de gerealiseerde spraak. Wat doet een spreker als hij/zij accenten moet leggen op gemarkeerde posities? Kan men de locaties van accenten voorspellen? Is co-articulatie naast taal- ook sprekerspecifiek? Wat doen sprekers als ze niet-frequente of niet-bestaande woorden moeten uitspreken: “*ik moet karaBONdra zeggen*”?

- 2 Mensen die alleen op het terrein van de taal- en spraakpathologie werken, hebben vaak onvoldoende mogelijkheden om de spraak van ‘normale’ sprekers te bestuderen. Zij moeten dus hun kennis afleiden uit de spraak van de gemarkeerde groepen, de sprekers met een spraak- of taalpathologie.
- 3 Het onderzoeksdesign, met name wat betreft de samenstelling van woordmateriaal in het onderzoek naar pathologische spraak, is cruciaal. Kennis van de fonetiek (en fonologie) is bij het samenstellen van woordmateriaal daarom onontbeerlijk.

Ik geef drie voorbeelden om mijn stelling te illustreren dat kennis van de fonetiek van groot belang is in de onderzoekspraktijk. Die voorbeelden zijn niet vrolijkmakend. Ze laten zien dat een brede kennis van de fonetiek noodzakelijk is wanneer men onderzoek wil doen op het terrein van de spraakpathologie. Deze laatste uitspraak lijkt ingegeven te zijn om het in de gevarenzone verkerende beroep van foneticus te redden, net nu ik die beroepsstand lijk te verlaten. Laat ik U geruststellen, dan wel alarmeren: eens een foneticus, altijd een foneticus. Nu de voorbeelden:

Een eerste voorbeeld:

Voor een onderzoek moest woordmateriaal worden samengesteld om de vraag te beantwoorden of een spreker met epilepsie bij het voorkomen van specifieke verschijnselen in het EEG ook afwijkend prosodisch gedrag vertoont als gevolg van zogenaamde *transient cognitive impairment* (TCI): een korte verslechtering van het cognitief functioneren. Je kunt dat spraakmateriaal op een naïeve manier samenstellen – zoals men deed –, bijvoorbeeld met alleen bisyllabische woorden, zoals *manen*, *paarden*; je kunt ook gebruik maken van de veel grotere prosodische variatie die 4-syllabische woorden bieden, zoals ‘rododendron’ maar dan moet je wel weten dat er 2- en 3-syllabische voeten bestaan, en dat die verschillende duurpatronen opleveren (cf. Rietveld, Kerkhoff & Gussenhoven, 2004). Woorden van het type *ballen*, *manen* en *katers* leveren minder en eigenlijk onvoldoende mogelijkheden op vergeleken met woorden van het type *rododendron*, of *piano*.

Een tweede voorbeeld:

Er zijn mensen die onderzoek doen naar de vraag of woorden met een zinsaccent vaker tot het optreden van stotteraars leiden dan woorden zonder zinsaccent. Zinsaccenten worden dan vaak gedefinieerd als pieken in het toonhoogteverloop. Daarbij wordt buiten beschouwing gelaten dat de meeste talen ook accenten kennen die met een toonhoogtedal (de L*H) worden gerealiseerd. Daarbij werd buiten beschouwing gelaten dat het realiseren van zinsaccenten op plekken waar dat niet default is, ook spanning kan opleveren. Aldus kunnen stotteraars en niet-stotteraars beide spanning ervaren bij de realisatie van het betreffende materiaal, en aldus eventueel aanwezige effecten versluieren.

Een derde voorbeeld:

Spreeknelheid als maat voor leeftijd bepaalde spraakontwikkeling. Zinsfinale syllaben, of zelfs zinsfinale voeten zijn langer dan niet-finale syllaben en voeten met de gelijke segmentele opmaak. Hoe langer de zin, hoe minder zinnen er in vijf minuten worden gerealiseerd, en hoe minder zinsfinale voeten er worden geproduceerd, en hoe minder verlenging er op die manier optreedt. Dus: in het zinsmateriaal ‘Dit is zin een’ en ‘Dit is zin twee’ is de spreeknelheid lager dan in de bijna even lange zin ‘Dit is zin een en dat is zin twee’. Dat leidt dus tot een bias in de schatting van de spreeknelheid. Dat moet je weten, dat weten fonetici, maar ja, dan moet men deze wel als adviseur bij de hand hebben (cf. Haselager, Slis & Rietveld, 1991).

3. WAT GAAN WE DOEN?

Een oratie is een wat bijzondere tekstsoort. Aan de ene kant wil men aan een breed publiek laten zien wat het vak inhoudt, soms aan de hand van een specifiek probleem, soms ook aan de hand van een overzicht van problemen, maar tegelijkertijd wil het publiek ook wel weten wat de plannen zijn van de nieuwe functionaris.

Wel, mijn plannen zijn de volgende; N is hier 5.

- Wat betreft het onderwijs, en ik bedoel het onderwijs in de minoren Taal- en Spraakpathologie van de BA-opleiding Taalwetenschap en in de MA-opleiding Taal- en Spraakpathologie van de Radboud Universiteit, wil ik studenten bewust maken van de noodzaak om onderzoek op het terrein van de taal- en spraakpathologie niet zomaar te starten, om “te kijken wat eruit komt”, maar allereerst na te gaan of de condities waaronder het onderzoek gaat plaatsvinden zich wel lenen voor de toetsing van de hypotheses. Ik zal proberen invloed op het curriculum van de MA-opleiding Taal- en Spraakpathologie uit te oefenen om deze doelstelling te bereiken.

- Wat betreft het onderzoek: de situatie is daar wat ingewikkelder. De Faculteit der Letteren heeft momenteel geen eigen onderzoeksprogramma op het terrein van de taal- en spraakpathologie. Dat moet veranderen, omdat er voldoende competentie aanwezig is. Wat betreft de Sint Maartenskliniek, waar ik een dag per week gedetacheerd ben: daar wordt momenteel ook niet zoveel onderzoek gedaan op het terrein van de taal- en spraakpathologie, maar daar komt nu verandering in (dat komt niet door mij, dat was men al langer van plan). Dat onderzoek moet zo veel mogelijk in overeenstemming met de normale medische standaarden worden uitgevoerd; daarvoor zal ik mijn best doen.
- Er is in Nederland veel deskundigheid aanwezig op het terrein van de taal- en spraakpathologie: op de universiteiten van Amsterdam, Groningen, Maastricht, Rotterdam, Utrecht en natuurlijk Nijmegen, maar ook op een aantal Hogescholen. Soms moeten we elkaar tot de orde roepen; een mooi voorbeeld vond ik het congres over effectmeting van stem-, spraak- en taalstoornissen dat op 3 oktober 1997 op de – toen nog – KUN werd gehouden. Vele onderzoekers en practici hebben aan dat congres deelgenomen. Ik ben voornemens om in juni 2006 een congres te organiseren dat beoogt te komen tot een grotere standaardisering van afhankelijke variabelen in de taal- en spraakpathologie.
- De Hogeschool Arnhem en Nijmegen biedt een opleiding Logopedie aan en de Radboud Universiteit Nijmegen een MA-opleiding Taal- en Spraakpathologie. Samen kunnen we daarvan een voortreffelijke opleiding maken waarbij de klinische invalshoek van de Logopedie-opleiding en de wetenschappelijke invalshoek van de opleiding TSP elkaar zeer goed kunnen aanvullen. Wij zijn voornemens in het studiejaar 2006/2007 deze gecombineerde opleiding te starten.
- Nijmegen biedt unieke mogelijkheden om de spraak- en taalpathologie en de spraak- en taaltechnologie dichterbij elkaar te brengen. Immers, de Faculteit der Letteren heeft twee opleidingen op dit terrein op dezelfde etage gehuisvest (onder de naam: sectie Taal en Spraak). Zonder intensieve samenwerking met de Sint Maartenskliniek en het UMC Sint Radboud zal deze samenwerking steriel blijven.

DANKWOORD

Geachte toehoorders, aan het slot van een oratie past een woord van dank. Ik doe dat graag want ik ben veel mensen veel dank verschuldigd. Ik begin met de universiteit en haar hoogste vertegenwoordiger, de rector; ik dank hem zeer voor het in mij gestelde vertrouwen.

Voordat ik verder ga met mijn dankwoord, wil ik eerst een mogelijk misverstand wegnemen. Voor het gemak neemt men aan dat ik een taal- en spraakpatholoog ben: dat is niet het geval. Echte taal- en spraakpathologen zijn op de Radboud Universiteit mijn collega's prof. dr. Herman Kolk van het NICI, dr. Ben Maassen, drs. Bert deSwart,

mw. dr. Lianne Nijland, mw. drs. Wendy Huijnck, mw. drs. Mariette Embrechts en vele andere vaste en soms tijdelijke medewerkers van het UMCN en mijn vele collega's op de Sint Maartenskliniek. Deze laatsten zijn overigens voor een groot deel afgestudeerd aan de opleiding Spraak- en Taalpathologie van de toen nog Katholieke Universiteit Nijmegen.

Ik dank degenen die mijn leerstoel hebben mogelijk gemaakt, de directie van de Sint Maartenskliniek en het bestuur van de Faculteit der Letteren van de Radboud Universiteit Nijmegen. Daarbij waren de dynamische prof. dr. Hans Bots en de even dynamische mevrouw mr. Petri Holtus en mijn directe baas, de zeergeleerde dr. Jacques van Limbeek van de Sint Maartenskliniek drijvende krachten. Ik hoop noch de Maartenskliniek, noch de faculteit teleur te stellen in de tien jaren die mij nog resten (ik bedoel in mijn professionele leven).

Hooggeleerde professor van Hout, beste Roeland, met jou en met de zeergeleerde dr. Oud, ons aller Han heb ik vele genoeglijke uren doorgebracht, al speurende naar de complotten van mixed models. Anderen (behalve Jacques, Frans en Hubert) hebben nooit goed begrepen, zo vermoed ik, wat voor daadwerkelijk plezier we beleefden en beleven aan statistiek en methodologie en het is beter dat zo te laten.

Met professor Carlos Gussenhoven – sinds september van dit jaar verbonden aan de University of London – heb ik een prachtig researchleven van ongeveer 25 jaar achter de rug, resulterend in tientallen gemeenschappelijke publicaties. Zijn en mijn onderwerp was: experimentele taalkunde. Ik zat samen met hem op een rijdende trein. Carlos blijft op dezelfde trein, maar hij rijdt nu met een andere maatschappij – Thames Link –; ik zoek nu een andere trein, een trein die beter past bij de Maartenskliniek en bij het belang dat de Faculteit der Letteren nu hecht aan de Taal- en Spraakpathologie. Mijn hart blijft voor een deel bij de experimentele taalkunde; we zijn voornemens daar op de maandagavond samen mee verder te gaan.

Met professor Vincent van Heuven uit Leiden – een volbloed foneticus – heb ik nog een lang leven voor de boeg. Ons fonetiekboek gaat een derde druk tegemoet, mede dankzij de grote belangstelling vanuit de logopedie-opleidingen. Ik bereid me al weer voor op de intellectuele schermutselingen met die uiterst kritische maar oh zo gewaardeerde vriend en collega.

Drs. Joop Kerkhoff is nu studietoecoördinator TSP. Hij was vroeger fysicus, hij werkte vroeger aan de Nijmeegse spraaksynthese, helpt nu studenten door hun practica, lost op een uiterst vriendelijke wijze ingewikkelde roosterproblemen op. Zonder Joop (en mijn gevulde koeken) zouden vele studenten al lang naar de onderwijsdirecteur zijn gegaan.

Dr. Iman Slis nam 23 jaar geleden het initiatief – door mij als opportunist snel gevolgd – om een opleiding Spraak- en Taalpathologie te starten (“STP”). In 1983, 21 jaar geleden, is die opleiding aan de toen nog Katholieke Universiteit gestart. Dat

was een goed initiatief, zoals wij in de visitatie-rapporten hebben kunnen lezen. Anderen hebben ons initiatief opgevolgd, aan de universiteiten van Groningen, Utrecht en Amsterdam. Zonder deze opleiding is de leerstoel waarover ik nu spreek niet denkbaar. Beste Iman: nogmaals dank voor je initiatief, je collegiale steun en je vriendschap.

Beste professor Loe Boves, dr.Ir. Bert Cranen, dr. Helmer Strik en vele andere collega's van de sectie Taal en Spraak. Vroeger lag jullie belangstelling vooral bij de stembandfunctie, nu bij de spraaktechnologie. Ik blijf jullie achtervolgen, want de spraaktechnologie wordt steeds relevanter voor de spraak- en taalpathologie. Ik vind het zeer bemoedigend dat jullie al lange tijd interesse voor de interactie tussen pathologie en technologie aan de dag leggen en voor mij actieve raadgevers zijn.

Een apart woord van dank wil ik richten aan de promovendi, zowel zij die dat waren als zij die dat nog zijn. In de afgelopen jaren zijn zij het vaak geweest die het voor mij de moeite waard maken om aan de universiteit te werken. Promovendi zijn onverbidde gericht op de inhoud van onderzoek en met hen spreek ik dus vanzelf over de echte problemen.

Ik kom tot slot van mijn dankwoord bij mijn verwanten – waarvan ik er niet zoveel heb – en wel bij mijn vader en moeder. Ik vrees dat mijn ouders wel eens gedacht hebben dat ze niet wisten wat er uiteindelijk van mij terecht zou komen. Aan de andere kant geloof ik niet dat ze zich ooit echt zorgen hebben gemaakt. Ze dachten: die krijgt een mooie baan ter zee, de vraag is gewoon: bij de Koninklijke Marine, ('koleke mrine') zoals dat in het fonetiekboek van Rietveld en van Heuven wordt getranscribeerd, of bij de Koopvaardij.

[3ko+lkχ m3rinχ] (koleke mrine)

Het werd een compromis: de zeilerij, en dan wel 'nebenbei'. Ze hebben zich nooit direct willen bemoeien met mijn keuzes in studie en onderzoek, maar heel af en toe denk ik wel eens dat ze liever gezien hadden dat ik leraar Frans zou zijn gebleven. Beste vader en moeder, ik ben dankbaar voor jullie steun en voor het feit dat mijn vader vandaag hier kan zijn en deze feestelijke dag nog volop kan meemaken.

Tot slot zou ik willen zeggen: N nu niets meer, maar ik zeg uiteraard:

Ik heb gezegd.

DANKBETUIGING

Ik dank Mw. drs. M. Wegman voor het uitvoeren van de survey van de tijdschriften inzake de bepaling van de vereiste steekproefgrootte.

LITERATUUR

- Chen, A. & Rietveld, T. (2004). *Obtaining perceptual judgements in research of intonational meaning*. Lezing gegeven op de Workshop Experimental Prosody Research, 7-9 oktober. Leipzig, <http://www.uni-leipzig.de/~prosody/>.
- Cohen, J. (1977). *Statistical power analysis for the behavioural sciences* (2nd edition). Orlando, FL.: Academic Press.
- Cortina, J.M. & Nouri, H. (2000). *Effect sizes for anova designs*. London: Sage Publications.
- Elstrodt, M. & Mellenbergh, G.J. (1978). Eén minus de vergeten fout. *Nederlands Tijdschrift voor de Psychologie*, 33, 33-49.
- Forster, K.I. & Dickinson, R.G. (1976). More on the Language-as-Fixed-Effect fallacy: Monte Carlo Estimates of Error Rates for F1, F2, F' and Min F'. *Journal of Verbal Learning and Verbal Behaviour*, 15, 135-142.
- Frattali, C.M. (1998). *Measuring Outcomes in Speech-Language Pathology*. New York: Thieme.
- Grunwell, P. (1993). *Analysing Cleft Palate Speech*. London: Whurr Publishers.
- Haselager, G.J.T., Slis, I.H. & Rietveld, A.C.M. (1991). An alternative way of studying the development of speech rate in Dutch children. *Clinical Linguistics & Phonetics*, 5, 53-63.
- Jansonius, K. (1999). *Twee jaar spraak en taal bij schisis*. Dissertatie UvA, Amsterdam.
- Jones, M., Gebski, V., Onslow, M. & Packman, A. (2002). Statistical Power in Stuttering Research: A Tutorial. *Journal of Speech, Language and Hearing Research*, 45, 243-255.
- Konst, E.M. (2002). *The effects of infant orthopaedics on speech and language development in children with unilateral cleft lip and palate*. Dissertatie KUN, Nijmegen.
- Maassen, B. & Groenen, P. (1999). *Pathologies of Speech and Language*. London: Whurr Publishers, Ltd.
- Okalidou, A., Karathanasi, A. & Grikoraki, E. (2003). Nasalance norms in the speech of Greek normal adults: Linguistic, phonetic, gender effects and crosslinguistic comparisons. *Proceedings 15th International Conference of Phonetic Sciences, Barcelona*, 1659-1662.
- Peng, S.-SH., Tomblin, J.B., Spencer, L.J. & Hurtig, R.R. (2004). Acquisition of Rising Intonation in Pediatric Cochlear Implant Recipients: A Longitudinal Study, *International Congress Series*, 1273, 336-339.
- Peters, H., Rietveld, A.C.M. & Vieregge, W.H. (1998). Themanummer 'effectmeting van stem-, spraak- en taalstoornissen'. *Stem-, spraak- en taalstoornissen*, 7 (3), 149-140.
- Prael, C. (1993). *A Study into the effects of presurgical orthopedic treatment in complete unilateral cleft lip and palate patients*. Protocol february 1993.
- Raaymakers, J.W.G., Schrijnemakers, J.M.C. & Gremmen, F. (1999). How to deal with the 'language-as-fixed-effect fallacy': Common misconceptions and alternative solutions. *Journal of Memory and Language*, 41, 416-426.

- Rietveld, Toni & van Hout, Roeland (1993). *Statistical Techniques in the Study of Language and Language Behaviour*. Berlin: Gruyter Mouton.
- Rietveld, T, Kerkhoff, J. & Gussenhoven, C. (2004). Word prosodic Structure and Vowel Duration in Dutch. Te verschijnen in *Journal of Phonetics*.
- Severens, J.L., Konst, E.M. & Prahl, C. (1998). Principes van kosten-effectiviteitsanalyse bij stem-, spraak- en taalstoornissen. Themanummer 'Meting van therapie-effecten.'. *Stem-, Spraak- en Taalstoornissen*, 7 (3), 158-167.