

4

Objectieve en subjectieve geografische factoren, informanten- en veldwerkerskenmerken

Een modelmatige aanpak van t-deletie in Nederlandse dialecten¹

4.0. Inleiding

In de twee vorige hoofdstukken werden achtereenvolgens de validiteit en de betrouwbaarheid van dialectgegevens behandeld. In dit hoofdstuk willen we nader bepalen welk deel van de variabiliteit van dialectgegevens door *geografische* factoren valt te verklaren. Enerzijds wordt aangetoond dat a) de “Grundmundart”, het lokale dialect, voornamelijk geografisch bepaald is. Dit impliceert dat andere sociale kenmerken naar verhouding weinig van de variabiliteit in dat stratum van taalgedrag verklaren.² Anderzijds wordt aangetoond dat b) afstand tot de standaardtaal een deel van de variabiliteit in de Grundmundart mee bepaalt, en dat c) mogelijk ook de afstand tot naburige dialecten, zoals die door dialectsprekers gepercipiëerd wordt, een rol speelt. daaruit volgt dat we ons in de rest van dit boek kunnen beperken, wat de taalexterne factoren betreft, tot de geografische factoren.

Om een en ander aan te tonen kiezen we een lineair model waar de bepalende factoren vooraf, op grond van theoretische overwegingen en bestaande kennis, in hun onderling verband worden vastgelegd. Vervolgens wordt dit model aan de feiten getoetst. We houden er al van te voren rekening mee dat we met een zuiniger model toekunnen, met minder factoren dan het oorspronkelijke, dat de gegevens ook optimaal verantwoordt. We proberen dus twee modellen te schatten. Op grond van de veronderstelling dat geografische factoren het belangrijkste zijn, verwachten we dat er uiteindelijk volstaan kan worden met een model waaruit de sociale factoren verwijderd zijn, maar dat niettemin vrijwel even goed de aanwezige variabiliteit verklaart als het eerste, uitgebreide model. We kiezen een causaal standpunt zoals bij het schatten van vergelijkingen via multiple regressie.³ De

¹ Bewerkte versie van een lezing: “De afstand van de dialecten tot de standaardtaal”, in het kader van een serie over Minderheden en Taal, op 15 november 1985, Rijksuniversiteit Groningen.

² Hiermee worden de conclusies van Goeman (1986 (hier hoofdstuk 3.2 e.v.)) nader getoetst.

³ Deze procedure is verwant met de in de sociologie meer gebruikelijke pad-analyse. Een toepassing van deze laatste techniek op taalsociologische gegevens is te vinden in Geerts, Nellemans en Jaspaert (1985). Voor een vergelijking van pad-analyse en van econometrische modellen met de meer algemene Lisrel-modellen zie Saris en Stronkhorst (1984, appendix).

feiten zijn in dit verband dialectgegevens die verzameld werden in het kader van het GTProject en ze betreffen Zuid-Holland en het Rivierengebied. De afhankelijke variabele is t-deletie aan het wordeinde.

Zoals in hoofdstuk 2, wordt er ook hier vanuitgegaan dat de “Grundmundart” niet een volstrekt homogeen stratum is. De aandacht voor variatie-dialectologie is betrekkelijk nieuw, maar we moeten er wel op wijzen dat in Europa het traditionele dialectologische standpunt een expliciete belangstelling voor variatie en externe, bijv. sociale, factoren, niet uitsloot.

Het schoolvoorbeeld in de Nederlandse dialectologie is Kloeke’s (1927) behandeling van de diftongering van wgm. *û* en de “Hollandse Expansie” ervan. Als er tegenwoordig in de dialectologie systematischer rekening wordt gehouden met variatie, dan moeten we niet vergeten dat dit vroeger in zekere mate ook al gebeurde, althans in de Nederlandse dialectologie.⁴

Naast de externe factor “geografische ligging” die de gegevens beïnvloedt is er nog een tweede externe factor die met de geografie samenhangt: de invloed van de standaardtaal op de dialecten of van dialecten die ten nauwste verwant zijn met die standaardtaal. Dat zijn dan met name de dialecten in Holland. Daar gaat Kloeke’s Hollandse Expansie over: het afnemen van de westelijke invloed naar gelang men verder verwijderd raakt van dát geografisch zwaartepunt in Holland waar de dialecten het meest lijken op de spreektaal van de algemene standaardtaal (Kloeke 1927).

Een derde geografische factor bestaat uit de mate van overeenkomst die men ervaart met naburige of meer verwijderde dialecten. In Holland waarden sprekers hun dialecten niet direct als dialect, maar veeleer als een platte variant van de standaardtaal (Goeman 1984, 7-8).

Een bepaald dialect kan dus minder of meer verschillen van een ander dialect maar het kan ook een grotere of kleinere afstand vertonen tot de standaardtaal met betrekking tot bepaalde verschijnselen. Zo zijn de verschillende afstanden van dialecten tot de standaardtaal globaal gevisualiseerd op een indelingskaart van de Nederlandse dialecten door Daan (Daan en Blok 1970). Deze kaart is voor Nederlands gebied gebaseerd op de standpunten van dialectsprekers over die dialecten in hun omgeving, die ze beoordelen als min of meer lijkend op hun eigen dialect.

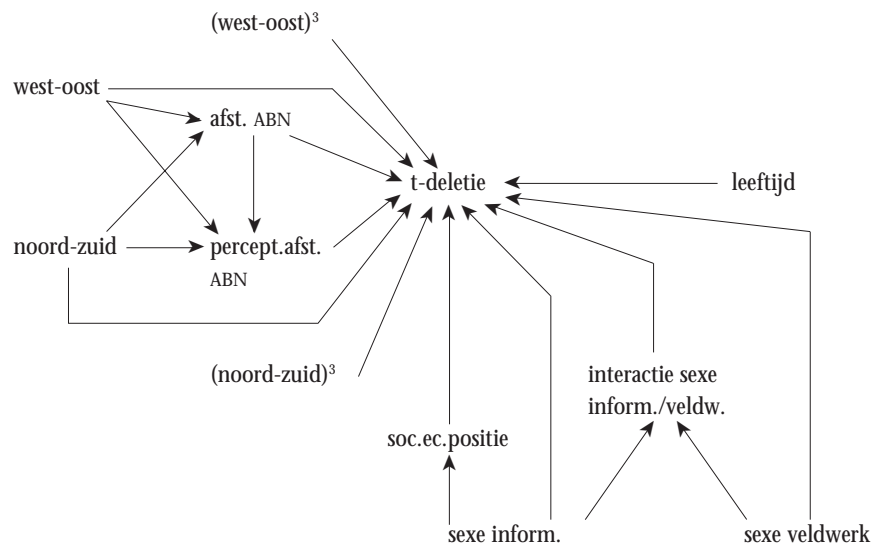
In dit hoofdstuk wordt in sectie 4.1 eerst een korte beschrijving van de externe, factoren (of onafhankelijke variabelen) gegeven, en van de onderlinge verbanden daartussen. Vervolgens worden twee modellen gespecificeerd (sectie 4.2). In sectie 4.3 worden twee aspecten van die modellen apart behandeld. De toetsing van de twee modellen wordt besproken in sectie 4.4. Vervolgens wordt in sectie 4.5 de discrepantie tussen uiteindelijk model en gegevens besproken. Dat is het gedeelte van de gegevens waarvoor het model geen verklaring biedt oftewel de residuen. Op dit aspect kom ik overigens nog terug in hoofdstuk 5. Tenslotte volgt in sectie 4.6 een samenvatting van de conclusies.

⁴ Zie overigens hoofdstuk 2 voor vroege ontkenningen van de homogeniteit van dialecten.

4.1. De modellering

Zoals gezegd ontwikkelen we hier een causaal model. Het bevat maximaal de volgende factoren (figuur 1), maar we zullen zien dat niet alle factoren significant zijn:

Figuur 1. Modelfactoren



- afhankelijke variabele:
t-deletie aan het wordeinde;
- onafhankelijke variabelen:
 - 1) de coördinaat “west-oost” van een plaats;⁵
 - 2) de coördinaat “noord-zuid” van een plaats;
 - 3) idem “west-oost” tot de derde macht;
 - 4) idem “noord-zuid” tot de derde macht;⁶
 - 5) geografische afstand tot standaardtaalcentrum;
 - 6) gepercipiëerde afstand tot standaardtaalcentrum;
 - 7) leeftijd informant;

⁵ Het gebruikte coördinatensysteem loopt op van West naar Oost en van Noord naar Zuid. Ieder coördinatensysteem is in zekere zin arbitrair, maar voor deze vorm werd gekozen om negatieve correlaties met de afhankelijke variabele te vermijden; er zou dan iedere keer gewezen moeten worden op het omgekeerde verband.

⁶ Juist bij geografische variabelen kunnen niet-lineaire verbanden optreden; via hogere machten kunnen deze niet-lineaire verbanden toch meegenomen worden in een lineair model. Omdat de twee coördinaten onafhankelijk van elkaar zijn is een interactie “west-oost x noord-zuid” niet als aparte factor in het model opgenomen.

- 8) sociaal-economische positie van informant;
- 9) sekse informant;
- 10) sekse veldwerker;
- 11) interactie tussen sekse informant en veldwerker;

Sommige factoren zullen wel met elkaar samenhangen, sommige zelfs sterk, andere minder of niet. Maar samenhang tussen factoren is op zichzelf nog geen causaal verband. Als A samenhangt met B wil dat nog niet zeggen dat A de oorzaak is van B of omgekeerd. Het verband tussen A en B zou veroorzaakt kunnen worden door een derde factor: C.

Een voorbeeld van het verschil tussen samenhang en causaal verband: blijkt dat er een samenhang (correlatie) bestaat tussen geslacht en inkomen, dan zegt dat op zichzelf niets over een causaal verband. Maar als we de richting van het verband vastleggen spreken we van een causaal verband. Op grond van kennis en ervaring weten we dat het inkomen van iemand niet zijn geslacht bepaalt, maar omgekeerd. We weten dat het feit of iemand man of vrouw is maakt dat men gemiddeld hoger of lager op de inkomensschaal zit.⁷ Het causale verband loopt dus van geslacht naar inkomen.

Van causale verbanden binnen het model is pas sprake als men door veranderingen in de onafhankelijke variabelen effect kan sorteren in de afhankelijke variabele. Dit veronderstelt een experimentele opzet van het onderzoek, die bij dialect-survey onderzoek, net als bij survey-onderzoek in het algemeen, principieel uitgesloten is. Toch is het mogelijk causaliteit voor het model te claimen op voorwaarde dat men een relevant theoretisch uitgangspunt heeft. Op grond van theoretische overwegingen en bestaande kennis kan men vóóraf de richting van de relaties aangeven tussen elk paar variabelen waartussen men samenhang postuleert en veronderstelt. Als de causale effecten vooraf op theoretische gronden gespecificeerd zijn, dan kan men zich de situatie voorstellen waarbij alle andere relevante variabelen constant gehouden worden terwijl men één onafhankelijke variabele manipuleert. Als dan een verandering in de onafhankelijke variabele leidt tot een verandering in de afhankelijke variabele, kan men toch van een causaal verband spreken. Men moet dus vooraf een theorie ontwikkeld hebben over de verbanden en die theorie is geformuleerd in het model. Dat model bevat dus de causale verbanden. Zo'n theorie kan overigens vrij globaal zijn. De procedure van het toetsen van het model leidt tot het schatten van het relatief gewicht van het effect van een bepaalde factor ten opzichte van alle andere. Tevens leidt het tot het bepalen van niet alleen de significantie van het model in zijn geheel, maar ook van de afzonderlijke factoren die van dat model deel uitmaken.

⁷ We weten dat het zo ligt in onze maatschappijvorm. Het constateren van maatschappelijke ongelijkheid betekent natuurlijk niet de legitimatie ervan.

4.1.1. *De afhankelijke variabele t-deletie*

De afhankelijke variabele is: t-deletie aan het wordeinde: bijv. /knech/ in plaats van /knecht/. Over het algemeen komt t-deletie vaker voor in lopende spraak dan in geïsoleerde woorden of in pauza. Toch is het gerechtvaardigd om t-deletie te onderzoeken aan geïsoleerde woorden, omdat:

- a) we in het hoofdstuk over validiteit gezien hebben dat dialectdata verkregen door geïsoleerde elicitatie niet invalide zijn, en
- b) het opportuun is: het is immers buitengewoon moeilijk om t-deletie te beschrijven voor die dialecten waar in pauza al een zeer hoog percentage t-deletie gerealiseerd wordt.

Dit laatste gebeurt bijvoorbeeld in de Betuwe met percentages deletie van rond de 90%. In die dialecten zit t-deletie in pauza dus al bijna tegen het plafond van het maximaal mogelijke. Als daar de invloed van het volgconsonantisme bijgevoegd wordt zit men direct tegen de bovengrens van 100% aan en wordt de invloed verluierd van andere intern-talige factoren als: voorafgaand consonantisme, morfologische klasse, token frequentie etc., die alle eigenschappen van het woord zijn.⁸ Daarom baseren we ons op de GTP-enquête die voornamelijk uit geïsoleerd aangeboden woorden bestaat.

Voor een globaal idee van zowel de invloed van morfologische klasse als van volgconsonantisme wordt verwezen naar Diagram 1 in hoofdstuk 6. In wat volgt blijft naast het volgconsonantisme ook de superlatief buiten beschouwing vanwege het bijzondere karakter.⁹

Gegeven deze inperking blijven er dan per plaats 116 gevallen van t-deletie over.

4.1.2. *De onafhankelijke variabelen*

Voor elf onafhankelijke variabelen wordt hun mogelijk effect op t-deletie bekeken. In sectie 4.1.2.1 zijn dat variabelen die net als in het model van hoofdstuk 3 te maken hebben met het specifiek geografische karakter van dialecten: 1) de coördinaat "west-oost" van een plaats; 2) de coördinaat "noord-zuid" van een plaats; 3) idem "west-oost" tot de derde macht; 4) idem "noord-zuid" tot de derde macht.

In sectie 4.1.2.2-4.1.2.3 komt de verhouding van dialecten tot de standaardtaal aan de orde. In de eerste plaats is dat in sectie 4.1.2.2 5) de objectieve geografische afstand tot het gebied waar men de standaardtaal geografisch kan situeren; in sectie 4.1.2.3 komt de subjectieve, gepercipiëerde afstand tot dat standaardtaalcentrum aan de orde. Vervolgens komen een aantal bio-sociale kenmerken van

⁸ Zie hoofdstuk 6 (Goeman en Van Reenen 1985). Tevens is er overeenkomst tussen de resultaten van een andere analyse van RND-materiaal en onze gegevens (cf. Van Hout 1980; Goeman 1986 (hier hoofdstuk 3) en hoofdstuk 5). Het RND-materiaal betreft voornamelijk taal in verbonden spraak.

⁹ Goeman en Van Reenen (1985), hier hoofdstuk 6, waar de speciale kenmerken van de superlatief en de overige interne condities aan de orde komen.

informanten en veldwerkers ter sprake: die betreffen in sectie 4.1.2.4 7) de leeftijd van de informant; in sectie 4.1.2.5 8) de sociaal-economische positie van informant; en in sectie 4.1.2.6 9) de sekse van de informant; 10) de sekse van de veldwerker; en tot slot 11) de interactie tussen sekse informant en veldwerker. Deze laatste groep van persoonskenmerken is in hoofdstuk 2 besproken omdat ze mogelijk een systematische invloed hebben op de dialectgegevens. We willen hier nagaan wat hun relatieve belang is in de data betreffende t-deletie.

4.1.2.1. ONAFHANKELIJKE VARIABELEN 1-4: GEOGRAFISCHE POSITIE

Geografische positie wordt in de traditionele Europese continentale traditie van dialectologie opgevat als de bepalende factor voor het lokale karakter van dialecten. Dialecten worden meestal opgevat als het dialect van een bepaalde plaats. Dit dialect wordt nogal eens als relatief homogeen opgevat: varianten binnen een plaats worden dan weer eerst geografisch gedefinieerd, bijv. naar wijk (zo bijv. Ausems (1953) voor Culemborg, Taeldeman (1985, 33 en vooral 233-234) voor Gent)¹⁰ vervolgens stelt men dan vast dat dit geografische verschil binnen een gemeente gepaard gaat met sociaal verschil. In deze situaties lijkt de sociale klasse waar men toe behoort te bepalen waar men gehuisvest is. Dit wordt ook in subjectief opzicht ervaren. De negatieve waardering blijkt dan uit de vernoemingsnamen voor de buurten. Rentenaar (1985², 261-262) wijst op een hele rits van “Lombok”, via “De Balkan” en “Gomorra” tot “De Vodbuurt”. Positieve waardering blijkt dan bijvoorbeeld uit een vernoeming als “Goudkust”.

Het wekt dan ook geen verbazing dat de nadruk die er ligt op het geografisch karakter van dialecten er toe geleid heeft de term topolinguïstiek te gebruiken in plaats van dialectologie; zo bijvoorbeeld Weijnen (1975).¹¹

Deze inperking van het probleemveld gaat toch wat te ver, omdat dan socio-dialectologisch onderzoek wordt uitgesloten. Het blijft evenwel van belang te benadrukken dat de geografische factor een relatief groot gewicht heeft, maar dat houdt niet in dat het stratum “Grundmundart” homogeen is, zoals in hoofdstuk 2 is aangetoond.

De geografische positie is gemeten langs een West-Oost en een Noord-Zuid as. Uitgangspunt voor de vaststelling van de ligging van iedere plaats was de vakkenindeling van de grondkaart van Grootaers-Kloeke. Deze vakken werden genummerd vanaf een nulpunt. De vakken zijn geen zuivere vierkanten en daardoor is de Noord-Zuid schaal korter dan de West-Oost schaal. De verschijnselen zitten

¹⁰ Taeldeman geeft de geschiedenis van de opvattingen over de onderlinge verhouding van geografische en sociale variatie binnen het Gents, te beginnen met Winkler (1874). Schatz (1986) kon voor het Amsterdams van de vorige eeuw aantonen dat de verwijzingen naar de vele buurtgebonden dialecten in wezen als sociolecten opgevat moeten worden. 19e - eeuwers als Hoeufft, Winkler en Alberdink Thijm geven al blijk van een sociale waardering van dialect. Dit blijkt onder andere ook uit de wijze van vraagstelling in hun enquêtes: “Hoe spreekt het volk te uwent...” (mond. meded. J. Berns).

¹¹ Iets soortgelijks gebeurt als men voor dialectologie de term: taalgeografie gebruikt, in het Duits: Sprachgeographie.

op de Noord-Zuid schaal dus dichter bij elkaar in de buurt dan op de West-Oost schaal. Bij de uiteindelijke resultaten speelt dit -overigens kleine- verschil geen rol omdat we de resultaten ook met gestandaardiseerde gewichten kunnen weergeven. Dan zijn West-Oost en Noord-Zuid herleid tot een zelfde schaal, overigens net als de andere factoren en zijn ze qua relatief gewicht alle onderling vergelijkbaar.

4.1.2.2. ONAFHANKELIJKE VARIABELE 5: DE AFSTAND VAN HET DIALECT TOT DE STANDAARDTAAL

Om de afstand tot de standaardtaal te meten zou men voor ieder dialect afzonderlijk een contrastieve studie met de standaardtaal kunnen maken. De mate van contrast zou dan een maat voor die afstand zijn. Dit is evenwel een werk van zeer lange adem, en dan nog blijft er het fundamentele probleem bestaan hoe men het grote aantal verschillen onderling zou moeten wegen. Deze procedure lijkt vooralsnog praktisch niet uitvoerbaar.¹²

Via een andere invalshoek is het toch mogelijk om de relatieve afstand dialect-standaardtaal te operationaliseren, namelijk via geografische afstand. Het uitgangspunt hiervoor is een eerdere vaststelling (zie Goeman 1986, hier hoofdstuk 3) dat ogenschijnlijke geografische verbanden, die een storing in het materiaal leken te zijn, geïnterpreteerd konden worden als verbindingen tussen dialecten, geordend op één uniform gerichte dimensie met een eveneens geografisch karakter: namelijk afstand tot de standaardtaal gemeten ten opzichte van die groep dialecten die het meest op de standaardtaal lijkt.

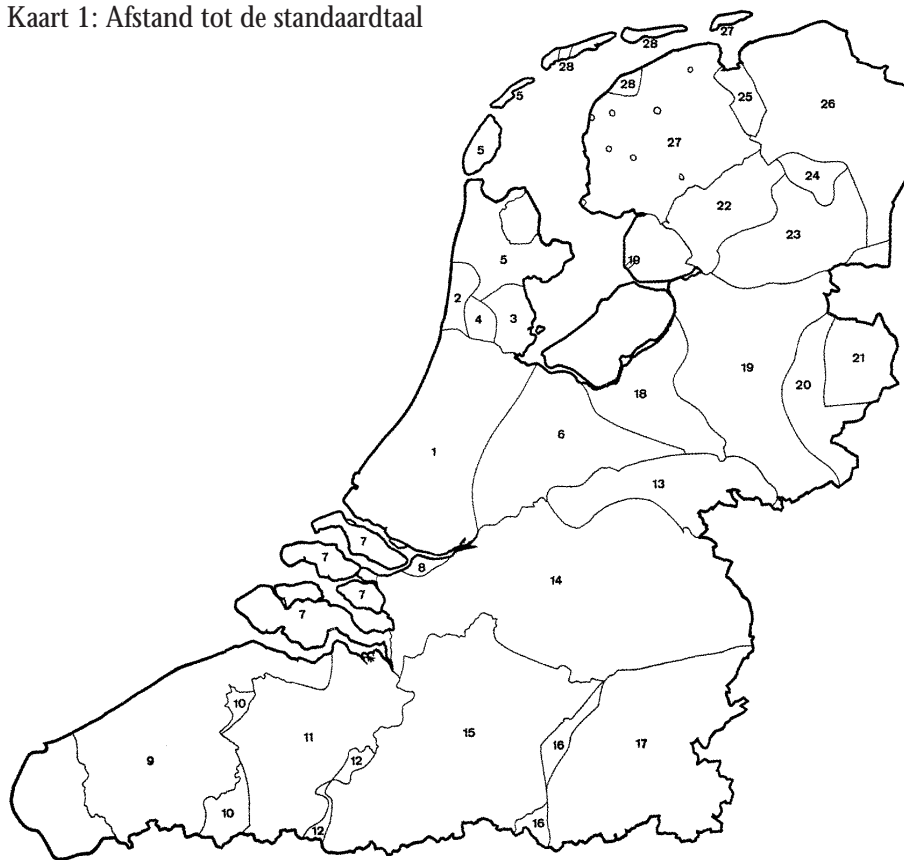
Winkler beschouwde de Hollandse dialecten al als gesproken varianten van de standaardtaal (Berns 1981), waarbij het echte Zuid-Hollands ten zuiden van de Oude Rijn gezocht moet worden. Dit standpunt werd in de achttiende eeuw ook verwoord door Ten Kate (Daan 1992, 178; 184): Ten Kates opkomende standaardtaal wordt bepaald door a) de spelling en b) het Hollands dialect dat in zijn uitspraak met deze spelling overeenstemt en dat is in Ten Kates visie niet het dialect van Amsterdam of van Amstelland-Rijnland, maar het zijn de zuidelijker gelegen Hollandse dialecten. Ten Kate is dus op zoek naar een neutrale variant van de standaardtaal, een variant die hij niet in de Hollandse steden kan vinden. In die optiek vertegenwoordigen stadsdialecten als Amsterdam, Leiden of Den Haag een niet-neutrale vorm van spreektaal.¹³ De indelingskaart van de Neder-

¹² Voor een onderzoek naar regionale verschillen in de standaardtaal van de maatschappelijke bovenlaag buiten de Randstad zie Voortman (1994).

¹³ Een alternatief model dat een getrapte beïnvloeding van 1) stad naar stad inhoudt met 2) vanuit die centra uitstraling over het omringende platteland is voor dialecten in Nederland bij mijn weten alleen door Goossens (1991, 117) gebruikt voor de rol van de Veluwe IJsselmeerstadjes op de noordelijke Veluwe en voor Twentse industriesteden (Goossens 1987, 155; 1991, 129). In het eerste geval betreft het een inverse expansie, of een mislukte, waardoor deze plaatsen in een "Reliktage" zitten. In het tweede geval is er volgens hem sprake van beperkte uitstraling van standaardtaalkenmerken. Een ander gebied waar steden deze rol hadden kunnen spelen is Friesland. De druk van het Nederlands heeft geleid tot het ontstaan van het Stadsfries in de Friese steden, maar van daaruit expandeert dat niet verder, het blijft tot die steden beperkt. Het

landse dialecten van Daan (Daan en Blok 1970, 9 en 32-33) kan gezien worden als de uitwerking van dit idee: de dialecten in de randstad (met uitzondering van de kuststreek) zijn met de neutrale kleur wit weergegeven en, naarmate de andere dialecten meer van de standaardtaal afwijken, werden steeds uitgesprokener kleuren gebruikt (zie kaart 1 voor een zwart-wit weergave).

Kaart 1: Afstand tot de standaardtaal



Dialectindelingkaart J. Daan (1969), bewerking H. Entjes (1974)

1=Zuidhollands, 2=Kennemerlands, 3=Waterlands, 4=Zaans, 5=Westfries-Noordhollands, 6=Utrechts-Alblasserwaards, 7=Zeeuws, 8=Westhoeks, 9=Westvlaams en Zeeuwsvlaams, 10=overgang West- en Oostvlaams, 11=Oostvlaams, 12=overgang Oostvlaams en Brabants, 13=Zuid-gelders, 14=Noordbrabants en Noordlimburgs, 15=Brabants, 16=overgang Brabants en Limburgs, 17=Limburgs, 18=Veluws, 19=Gelders-Overijssels, 20=Twents-Graafschaps, 21=Twents, 22=Stellingwerfs, 23=Zuiddrents, 24=Middendrents, 25=Kollumerlands, 26=Gronings en Noorddrents, 27=Fries, 28=Bildts, Stadsfries, Midlands, Amelands. Niet genummerd op de kaart is Frans-Vlaanderen.

De getallen coderen de kleur van de oorspronkelijke kaart en geven bij benadering een aanduiding van de afstand ten opzichte van de standaardtaal maar aangezien de nummering van de gebieden dient om dialectgebieden te *identificeren* en aangezien het aantal kleurklassen beperkter is dan het aantal dialecten wijkt de nummering af van de kleurcodering, met name in de hogere nummers.

De oorspronkelijke kleurklassen-indeling is dus in feite een *rangorde*-schaal. Zuid-Holland wordt op die manier beschouwd als de bakermat van het algemeen Nederlands (Daan 1963b).

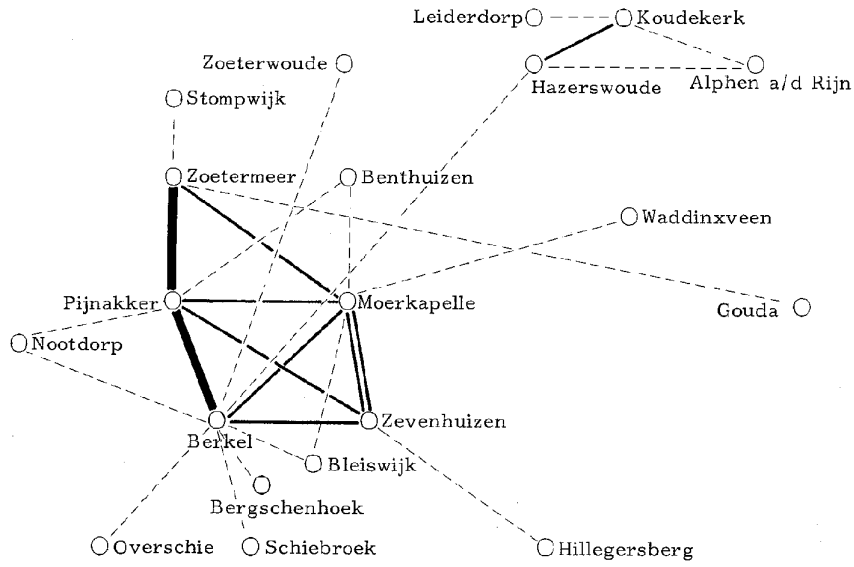
De lokalisatie van zo'n centrum is niet meteen evident. Leiden, Den Haag, Rotterdam of Delft als uitgangspunt te nemen is minder juist omdat deze stadsdialecten a) in het oordeel van sprekers als sterk geïsoleerd naar voren komen (kaart 2), b) sterk als sociolect gewaardeerd worden.¹⁴ De grote steden vallen er dus buiten. We situeren hier de kern in de door een cirkel omsloten plaatsen Zoetermeer, Berkel en Gouda. De eerste twee zijn opgenomen omdat ze, blijkens eerder onderzoek naar subjectief beoordeelde dialectverwantschap (Goeman 1984, 181) één cluster vormen (zie figuur 2), en de gegevens waarop die analyse gebaseerd was worden hierna gebruikt voor de afstandsbepaling. Gouda is opgenomen omdat volgens Heeroma (1968) dit dialect als een typische vertegenwoordiger van het Zuid-Hollands mag gelden, hetgeen een zekere historische dimensie inbrengt. Zoetermeer kan gezien worden als specimen van een dorp met toen reeds ontwikkelde industrialisatie. Het wordt later een stedelijk ontwikkelingsgebied. De kern bevat dus zowel enkele plattelandsdialecten, een 'kleinstedelijk' dialect (Gouda) en een 'nieuwstedelijk' dialect onder druk en bevat alle elementen van de ontwikkelingen sinds de jaren dertig.

Voor Zoetermeer kon bovendien worden vastgesteld dat over ruim een eeuw gerekend de verhouding van het dialect tot de standaardtaal relatief constant was gebleven, al waren er wel schommelingen waar te nemen (Goeman 1984, 124-126). Daardoor is er, hoewel impliciet, toch een contrastief moment ingebouwd. De afstand van een dialect tot de standaardtaal operationaliseren we dan als de geografische afstand van dat dialect tot het hollands dialectcluster van drie dat we als de gesproken variant van de standaardtaal beschouwen. We passen dus een operationalisatie toe die vergelijkbaar is met die van Ten Kate. Een steden netwerk zou als aparte factor invloed kunnen uitoefenen. We hebben dat niet als aparte factor opgenomen en daarvoor zijn verschillende redenen. De eerste reden is van praktische aard: er zijn naar verhouding weinig steden in de database. De tweede reden is meer principieel: zoals zal blijken in de volgende sectie vormen de

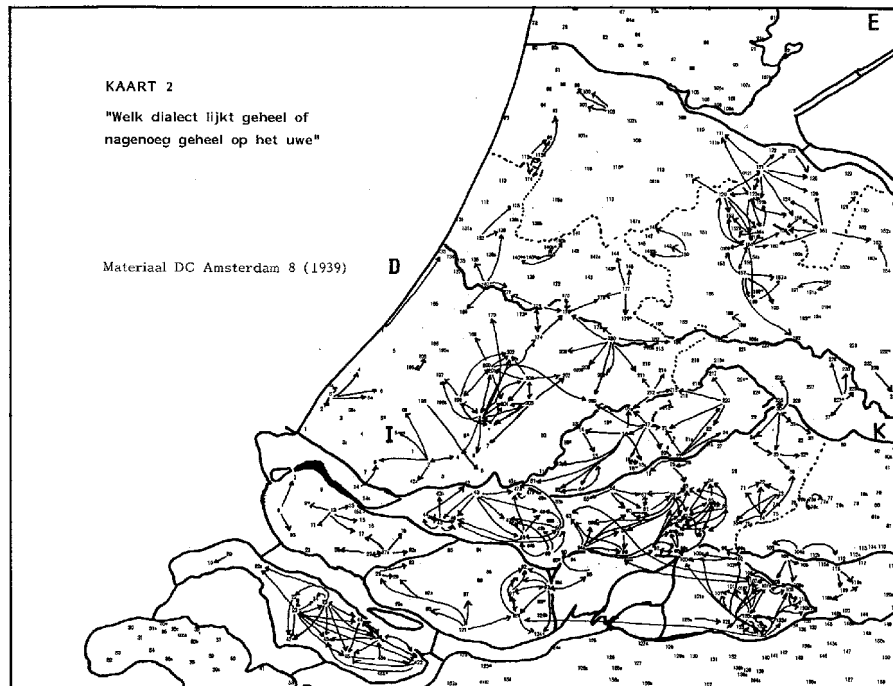
lijkt er op dat steden zeer vaak hun eigen netwerk hebben, onafhankelijk van en naast de algemene, globale druk die de standaardtaal uitoefent. Zie voor expansieve en relictmatige eigenschappen van dergelijke netwerken in Vlaams België Goossens (1992).

¹⁴ Van Hout en Münstermann (1981, 110-111) geven een ietwat andere operationalisering van de geografische afstand tot het standaardtaalcentrum, namelijk als het gemiddelde van de som van de afstand van een plaats tot Den Haag en van de afstand tot Utrecht: $0.5 * ((\text{dist}(i, \text{Den Haag}) + \text{dist}(i, \text{Utrecht})))$.

Figuur 2: Dialectcluster



Kaart 2: Subjectieve gelijkheidsoordelen



subjectieve oordelen van sprekers over gelijkheid van hun dialect geen clusters die interpreteerbaar zijn als steden-netwerk. Dit betekent niet dat een inzicht in het effect van een mogelijk steden-netwerk voor ons is afgesloten. Immers, het model dat we ontwikkelen is er juist op gericht om globale tendensen (de trend) in het dialectgedrag te scheiden van specifiek lokaal gedrag.¹⁵ Die specifiek lokale constellatie weerspiegelt zich in de residuen (vgl. hoofdstuk 3, kaart 1) en die residuen bevatten naast onsystematische variatie tevens alle effecten die we eventueel vergeten zouden zijn. Stel dat stedelijke netwerken wel een overwegende rol zouden spelen met betrekking tot t-deletie, dan zou, zelfs als niet alle steden in de database vertegenwoordigd zijn, het effect van die veronderstelde netwerken alsnog in een kaart van de residuen tot uiting moeten komen. Het kaartbeeld van de residuen (hoofdstuk 3, kaart1) valt evenwel moeilijk te interpreteren als steden-netwerk en daarom voelen we ons gerechtvaardigd de kwestie van de stedelijke netwerken terzijde te laten.¹⁶

4.1.2.3. ONAFHANKELIJKE VARIABELE 6: GEPERCIPIËERDE AFSTAND

Bij de indelingskaart van Daan, die hiervoor besproken werd, leiden a) subjectieve oordelen van dialectsprekers, b) de dialecten qua afstand tot de standaardtaal ingedeeld in groepen, en c) de objectieve indeling op basis van enkele dialectkenmerken tot hetzelfde resultaat: een globale indeling van de dialecten van Nederland.¹⁷ De aan de sprekers gestelde vraag was, welke dialecten in hun omgeving door hen als geheel of nagenoeg geheel gelijk werden beoordeeld. Uit de subjectieve oordelen van sprekers werden door Daan inductief de globale, grotere dialectgebieden geconstrueerd.

Mattheier (1985) ziet als een van de nieuwe opgaven voor de dialectologie dat ze ook rekenschap moet geven van het "Alltagserleben", de subjectieve ervaringen van dialectsprekers. Een deel van dit programma is dus in de Nederlandse dialectologie al betrekkelijk vroeg voorwerp van onderzoek geweest getuige Weijns onderzoek uit 1944.

¹⁵ Dit onderscheid is het thema van hoofdstuk 5.

¹⁶ Zoals zal blijken vertoont het beeld van de residuen uit de analyse in dit hoofdstuk en die van hoofdstuk 5 evenmin de effecten van zo'n steden-netwerk.

¹⁷ Voor de dialecten in België is niet het subjectieve oordeel van dialectsprekers gebruikt. Toch is het mogelijk voor Vlaanderen een zelfde kaart te tekenen op basis van het materiaal Willems, dat geldt voor de jaren tachtig van de 19e eeuw. In Goeman (1984, 12) is er op gewezen dat deze kaart tevens de veel geringere geografische mobiliteit in de vorige eeuw weergeeft. Bij de gelijkenisoordeelen uit de 19e eeuw is de gemiddelde afstand die men overziet tussen eigen en ander dialect, een uur gaans: dus gemiddeld 5 kilometer. Dezelfde vraag, gesteld aan het eind van de jaren dertig van de twintigste eeuw, waar Weijnen en Daan zich op baseerden, levert gemiddelde afstanden op van 15 kilometer. Onlangs is de vraag weer herhaald en de oordelen zijn nu nog globaler, doordat ze veel grotere afstanden bestrijken. Men ziet hier duidelijk het effect van eerst de fiets en dan de auto. Wie verder van huis komt hoort meer afwijkingen en oordeelt globaler, wie een kleine actieradius heeft hoort minder afwijkingen en kan meer verschillen "bijhouden" en belangrijk vinden.

Weijnen heeft als eerste van deze subjectieve dialectgrenzen een gedetailleerde kaart voor Noord-Brabant gepubliceerd (1944; later ook in Weijnen (1966)) en bovendien voor Limburg (Weijnen 1966). Rensink (1955) bevat slechts een globale kaart met begrenzingslijnen voor Nederland die als uitgangspunt heeft gediend voor de indelingskaart van Daan (Daan en Blok 1970). Kremer (1984) karteert op basis van een latere gegevensverzameling met dezelfde vraagstelling Twente, Achterhoek en het aangrenzende deel van Duitsland; Goeman (1984, 155) publiceert een kaart voor Zuid-Holland en de Vechtstreek op basis van het materiaal uit eind jaren dertig¹⁸ en daarbij wordt bovendien voor een kleine groep dialecten in centraal Zuid-Holland de rangorde in samenhang tussen de dialecten uit de sprekersoordelen afgeleid (ibid. 180-181); voor het gebied beneden de grote rivieren, daarin begrepen de Vlaamse dialecten is er nu een kaart op basis van oordelen uit het eind van de 19e eeuw (Goeman 1989).

Op grond van methodologische overwegingen heeft Schirmunski (1962, 145 e.v.) grote bezwaren tegen een indeling op basis van subjectieve oordelen. Ten onrechte, want hij interpreteert dit soort gegevens (de oordelen van sprekers) als waren het factoren die rechtstreeks de vorming van taalgemeenschappen mede bepalen en meent dat de procedure tot een volstrekt subjectieve en willekeurige vastlegging van dialectgrenzen leidt. Dezelfde mening is bij verschillende gelegenheden door Grootaers geventileerd.¹⁹ Volgens hem kan er geen overeenstemming bestaan tussen subjectieve en objectieve dialectgrenzen. Deze tegenstelling zien Schirmunski en Grootaers echter om twee redenen verkeerd.

De eerste reden is dat op cruciale punten Schirmunski en Grootaers zich -zoals zoveel dialectologen- de zaken deterministisch voorstellen en doordat ze het dialect als homogeen zien. Overeenstemming wordt opgevat in absolute zin: als "iets komt voor", of "iets komt niet voor"; er is overeenstemming, of er is geen overeenstemming. Dat is een kwalitatieve of deterministische aanpak. Realistischer is een niet-deterministisch of kwantitatief standpunt: iets is in meerdere of mindere mate het geval, iets komt in meerdere of mindere mate voor.

De deterministische opvatting is eigenlijk een bijzonder geval van het kwantitatieve standpunt, en beslaat uitsluitend de *uiteinden* van de kwantitatieve schaal, de punten 0 en 1 onder verwaarlozing van wat daartussen ligt.

We hebben al gezien dat dialecten heterogeen kunnen zijn. Bijgevolg zullen ook de grenzen tussen de objectief vastelbare dialectkenmerken vloeiend verlopen omdat de verschijnselen variabel zijn.²⁰ Men kan verwachten dat deze variabiliteit zich in de subjectieve oordelen van sprekers zal weerspiegelen.

¹⁸ Zie hoofdstuk 10 voor een kaart van oostelijk Nederland, die vergeleken kan worden met de kaart van Kremer (1984).

¹⁹ Grootaers (1959 en 1964). Weijnen (1968) heeft hierop gereageerd. Een meer uitgebreide bespreking van deze controversie is te vinden in Goeman (1989). Een enigszins anders gerichte opzet voor een perceptuele dialectologie is Preston (1989).

²⁰ Voorbeelden hiervan met betrekking tot t-deletie komen aan de orde in de hoofdstukken 6 en 10; voorbeelden uit de werkwoordsmorfologie komen aan de orde in de hoofdstukken 7 en 8.

De tweede reden waarom de eis van een absolute overeenkomst tussen objectief vaststelbare dialectkenmerken en subjectieve sprekersoordelen verkeerd gesteld is heeft te maken met het feit dat we subjectieve perceptie niet als rechtstreeks constituerend voor een dialectgemeenschap moeten zien, ook al blijft de subjectieve perceptie van sprekers mede bepalend voor het dialectgedrag van sprekers. Immers, de vraag die aan de informanten gesteld wordt levert geen directe toegang tot het complex van dialectkenmerken, maar levert wel een antwoord op dat betrekking heeft op gepercipiëerde gelijkens. Dit soort vragen moet op zijn eigen perceptuele merites bekeken worden en is als zodanig van even groot belang als de vraag naar de gebiedsvormende distributie van objectieve dialectkenmerken; perceptieoordelen leveren keuzepatronen op die tot een bepaalde ordening leiden. Die keuzepatronen hebben hun eigen intrinsieke variabiliteit. En het is zinvol ze aan andere verwante ordeningen te relateren. Zeker als men een theorie over die relaties heeft. Relaties die samenhang of, sterker, causaal verband weergeven, hoeven daarom niet noodzakelijkerwijs een alles of niets karakter te hebben. Dat heeft niets met subjectiviteit of willekeur van doen.

Er zijn verschillende manieren om perceptieoordelen van dialectsprekers die, gecarteerd, clusters van dialecten vormen, in afstand te relateren aan het standaardtaalcentrum dat we verondersteld hebben. We kunnen gebruik maken van de vorm en de richting van het cluster. Zo staat de hoofdrichting van het cluster met Overflakkeese dialecten loodrecht op de richting van dat cluster naar het standaardtaalcentrum (kaart 2). Bij andere clusters ligt de hoofdrichting daarop georiënteerd (bijv. het kleine Westlandse cluster). De *vorm* van een cluster kan bepaald zijn door het feit dat een informant zijn oordelen sterk in een bepaalde richting afgeeft, hetgeen resulteert in een *kleine spreiding* in de keus van als gelijk beoordeelde dialecten. Zo'n kleine spreiding laat daardoor een bepaalde hoofdrichting zien. Dit levert een schaal gebaseerd op de mate van spreiding. Die hoofdrichting binnen dat spreidingsgebied kan min of meer naar het standaardtaalcentrum wijzen, maar kan ook in de totaal andere richting wijzen, van het standaardcentrum af. Als we de afwijkingen in richting van de gelijkheidsoordelen in een cluster als basis nemen ten opzichte van de richting vanuit het cluster naar het standaardtaalcentrum dan krijgen we een schaal van de gemiddelde afwijking in oriëntatie ten opzichte van de standaardtaal. Deze twee operationalisaties van het begrip perceptuele afstand tot de standaardtaal zijn gebruikt in Goeman (1987a) maar daar moest geconcludeerd worden dat deze spreidings- en afwijkingsmaat uiteindelijk weinig gewicht in de schaal werpen. Dat komt omdat deze maten sterk lokaal bepaald blijven. Daarom geven we de voorkeur aan een andere maat, waarin uitsluitend de oordelen gebruikt worden die georiënteerd zijn op het standaardtaalcentrum. Het bleek dat deze maat (Goeman 1987a) beter functioneert. De scoring gaat als volgt: we construeren een keten van ieder cluster naar het cluster van het standaardtaalcentrum. De lege ruimte tussen clusters/afzonderlijke plaatsen wordt gezien als wederzijds neutraal en wordt gewaardeerd als het grootste verschil; wederzijdse erkenning van gelijkheid wordt gehonoreerd

met de laagste score: 0.5; geen 0, omdat er natuurlijk altijd wat verschil is. De scores tussen 0.5 en 3 worden gesommeerd en leveren dan de uiteindelijke afstandsmaat op.

De subjectieve gelijkheidsoordelen van de sprekers zoals die vanuit de plaatsclusters naar voren komen worden als volgt gescoord:

plaats A \approx plaats B; en plaats B \approx A:	0.5
plaats A \approx plaats B; maar B blijft neutraal:	1
plaats A \approx plaats C; plaats B \approx C; C neutraal m.b.t. A en B:	2
plaats A neutraal m.b.t. B; plaats B neutraal m.b.t. A:	3

“ $X \approx Y$ ” betekent: “X beschouwt Y als geheel of bijna geheel gelijk aan zichzelf”.

Sommatie van de scores over de ‘verbindingslijnen’ die de kortste keten van een bepaalde plaats naar de regio van Zoetermeer en omstreken in de randstad vormen levert de gepercipiëerde afstandscore voor die plaats t.o.v. die regio. We gebruiken de oordelen hier dus niet als indelingscriterium, maar we trekken uit de gepercipiëerde ruimte die oordelen die georiënteerd zijn op het “standaardtaalcentrum”.

4.1.2.4. ONAFHANKELIJKE VARIABLE 7: LEEFTIJD INFORMANT

Taalgedrag hangt vaak sterk samen met leeftijd. Taalgedrag is het resultaat van een socialisatieproces. In de adolescentiefase kan er een aanpassing aan een andere norm (die van de standaardtaal) plaatsvinden, die op latere leeftijd, als men zijn uiteindelijke plaats in het maatschappelijk leven gevonden heeft, weer tenietgedaan kan worden. Dan komt men weer op het niveau van de eerste socialisatie van thuis en peer-group terug. Dit doet zich naar gelang de leeftijd hoger is in sterkere mate voor.²¹ Een goed voorbeeld hiervan is een jonge informant uit Zoetermeer in 1932/1933 (Goeman 1984, 124-125). Deze vertoonde een veel hoger aantal standaardtaalvarianten dan zijn generatiegenoten, inmiddels 50 jaar ouder, in 1979 bleken te gebruiken. Leeftijd moet dus in rekening gebracht worden.

De leeftijd van de informanten in het materiaal van dit hoofdstuk ligt tussen de 48 en 74 jaar. Men kan derhalve verwachten dat het taalgedrag van informanten in deze leeftijdsklasse in hoge mate “Grunddialekt” is, en dat, wat de dialecticiteit in dit onderzoek betreft, de leeftijdsverschillen naar verhouding weinig invloed zullen hebben.

²¹ Een voorbeeld in apparent time is te vinden bij Chambers en Trudgill (1980, 91) met betrekking tot de realisatie van <ng> aan het wordeinde. Voor alle onderzochte ‘stijlen’: Casual speech, Formal speech, Reading passage style en Word list style gold, dat na ongeveer het 45ste levensjaar de standaardtaalvariant met de jaren minder voorkwam, terwijl deze vanaf 10-jarige leeftijd tot het 45ste levensjaar nu juist toegenomen was, zeker in Casual en Formal speech. Vergelijk overigens voor leeftijd en sociale mobiliteit Gregersen en Pedersen (1991, 21).

4.1.2.5. ONAFHANKELIJKE VARIABELE 8: SOCIAAL-ECONOMISCHE POSITIE VAN DE INFORMANT

Een van de hoofdthema's uit de sociolinguïstiek is het onderzoek naar de invloed van de sociale klasse waartoe men behoort op iemands taalgedrag. Een van de problemen die zich bij dat onderzoek voordoen is dat een schaal die uit een -al dan niet gewogen- sommatie van scores op opleiding, beroep, woonomgeving en inkomen bestaat meestal niet in zijn totaliteit gebruikt wordt. De schaal wordt meestal ingedikt tot een beperkte klasse-indeling: bijvoorbeeld Upper-middle, Lower-middle, Upper-working en Lower-working class. De breekpunten om tot deze ingedikte schaal te komen zijn vaak arbitrair zoals Davis (1985) heeft laten zien. Afhankelijk van de gekozen klasse-indeling kan men bij het aantonen van verband met taalkundige variabelen op hetzelfde taalkundige materiaal significante of niet-significante samenhangen constateren. Een betere benadering is om ofwel het totale bereik aan scores te gebruiken dat past bij de te onderzoeken maatschappij ofwel een onafhankelijke, al door anderen gevalideerde sociale klassenrangorde als uitgangspunt te nemen.

In dit hoofdstuk is voor de derde benadering gekozen, daarentegen werd de tweede methode -in de vorm van de beroepsprestigeschaal van Ultée en Sixma- in hoofdstuk 2 gebruikt. De criteria die ten grondslag liggen aan de rangorde, die men verkregen heeft uit de sommatie van scores zijn onderling afhankelijk: immers, opleiding bepaalt zowel beroep als inkomen, en inkomen bepaalt in sterke mate waar men kan gaan wonen. Deze onderlinge samenhang is al in die gevalideerde klassenrangorde verwerkt. Vanwege het feit dat in de eerste fase van het onderzoek de beroepsprestigeschaal van Ultée en Sixma nog niet beschikbaar was, wordt hier de beroepsklassificatie gebruikt van het ITS (Klapper ITS, 1972) waarbij het opleidingsniveau verdisconteerd is voorzover een bepaalde opleiding voor dat beroep noodzakelijk is. "Overeducatie" met betrekking tot een bepaald uitgeoefend beroep heeft dus geen invloed meer op de sociale status. Dit resulteert in een rangorde van zes klassen, waarbij 1 de laagste en 6 de hoogste is. Gehuwde vrouwen die zelfstandig een beroep uitoefenen vervullen de daaraan gekoppelde sociale rol en maken op grond daarvan deel uit van een sociale klasse. Gehuwde vrouwen die huisvrouw zijn ontlene hun sociale status meestal aan die van hun echtgenoot.²²

²² De aantallen informanten per klasse, in dit subsample, zijn als volgt:
 klasse 1: 7; klasse 2: 8; klasse 3: 8; klasse 4: 26; klasse 5: 3 en klasse 6: 0 informanten. Klasse 1 omvat handarbeiders zonder beroepsopleiding; klasse 2: handarbeiders met lagere beroepsopleiding; klasse 3: hoofdarbeiders zonder beroepsopleiding of met lagere of voortgezette opleiding, maar in een ondergeschikte positie; klasse 4: kleine zelfstandigen (inclusief boeren) zonder hogere beroeps- of academische opleiding; klasse 5: voortgezette opleiding in leidinggevende positie of hoger beroepsonderwijs in ondergeschikte positie; en klasse 6: hogere beroepsopleiding/hoger onderwijs als zelfstandige, of hogere functionaris leidinggevend én zelfstandigen zonder hogere beroepsopleiding of hoger onderwijs maar wel eigenaar van een groot bedrijf.

In het hier voorgestelde model wordt de sociale status-rangorde behandeld als een intervalschaal. De correlatie tussen sociaal-economische positie en t-deletie is dan 0.092 en niet significant: $p > 0.51$.

Omdat in ons sample klasse 6 ontbreekt en de schaal korter is geworden (1-5) is er voor de zekerheid ook op ordinaal niveau getoetst: Spearman's $r_s = 0.097$ (gecorrigeerd voor ties) en evenmin significant. De waarde van interval en ordinale maat zijn praktisch gelijk. Het is gerechtvaardigd om de scores als interval-gegevens te behandelen, omdat we veronderstellen dat de schaal onderliggend continu is; deze veronderstelling is gewettigd gezien de veel "langere" beroepsprestigeschaal (zie hoofdstuk 2). Bovendien heeft Monte Carlo simulatie aangetoond dat schending van de interval-eis alleen in zeer extreme gevallen gevolgen heeft (Asher 1976, 35; 66).

4.1.2.6. ONAFHANKELIJKE VARIABELEN 9-11: SEXE VAN INFORMANT, VELDWERKER EN DE INTERACTIE

De variabele *sexe* wordt opgenomen in het model omdat herhaaldelijk gebleken is dat het taalgedrag van vrouwen anders is als dat van mannen. Vrouwen richten zich eerder dan mannen naar de standaardtaalnorm. Bij mannen werkt als extra behoudende factor op het niet-standaardtaal-gedrag de behoefte deel te hebben aan de "covert prestige" van hun subcultuur.²³ Daarnaast moet rekening gehouden worden met het feit dat het taalgebruik van de informant kan verschillen naar gelang deze te maken heeft met een vrouwelijke of mannelijke veldwerker. Dit is niet een artefact van de situatie. Ten eerste kan de gesprekspartner zelfstandig door eigen, mogelijk dominant taalgedrag de toon zetten, en ten tweede kan de informant uit eigen beweging het taalgebruik afstemmen op dat van de gesprekspartner of aangesprokene (Brouwer 1982; Jansen 1985). In beide gevallen is het taalgebruik bepaald door de interactie tussen informant en veldwerker.

Gerritsen (1985) heeft er op gewezen dat, althans bij ons, die interactie tot op zekere hoogte cultuurgebonden is. Het is de weerslag van rolverwachtingen en komt tot uiting in mannen- en vrouwengedrag dat op elkaar is afgestemd. Omdat de interactie als afzonderlijke variabele in het model werd opgenomen, is het overbodig een onderscheid te maken tussen enerzijds *sexe* als rolgedrag en anderzijds geslacht als biologische categorie.²⁴ *Sexe* van informant en veldwerker zijn kwalitatieve, nominale variabelen. In beide gevallen is de waarde 1 toegekend als het om een man ging.²⁵

Onder de informanten zijn 12 vrouwen en 40 mannen. Bij de opnamen was in 25 gevallen de veldwerker een vrouw en in 27 gevallen een man. De verdeling van manlijke en vrouwelijke veldwerkers naar gelang de *sexe* van de informanten is als volgt:

²³ Zie voor het relatief grotere adaptief gedrag van vrouwen bijvoorbeeld Schatz (1985) en voor het covert prestige Trudgill (1972).

²⁴ Zoals voorgestaan door Brouwer (1984) op grond van sociologische theorievorming.

²⁵ Ze zijn opgenomen in het model als z.g. dummy-variabelen.

$$\begin{aligned}
 M_{inf} \times M_{ve} &: 18; \\
 V_{inf} \times M_{ve} &: 9; \\
 V_{inf} \times V_{ve} &: 3; \\
 M_{inf} \times V_{ve} &: 22.
 \end{aligned}$$

Deze verdeling wijkt niet af van wat men volgens het toeval zou verwachten: $c^2=3.327$, $p> 0.05$ bij 1 vrijheidsgraad. De verdeling is dus niet van invloed op mogelijke interactie-effecten.

4.2. Specificatie van twee modellen

In enkele gevallen werd hierboven al impliciet aangegeven hoe de onderlinge relaties tussen de variabelen lagen. Bijvoorbeeld: geografische afstand van het standaardtaalcentrum is bepaald door de geografische positie en mogelijk zelf weer bepalend voor de gepercipiëerde afstand. Het is belangrijk om die relaties in hun onderling verband te beschouwen, omdat het gezamenlijke effect van de onderscheiden factoren een tweezijdig geconstateerd verband, dat in isolatie belangrijk lijkt, volledig naar de achtergrond kan dringen. De verbanden die significant aanwezig zijn tussen telkens twee variabelen staan in de tabel hieronder, waarbij de volgorde van links naar rechts die van hogere naar lagere correlatie is, en waar dus nog geen rekening gehouden is met het feit of zo'n geconstateerd verband een causaal karakter heeft.

Tabel 1: Significante correlaties

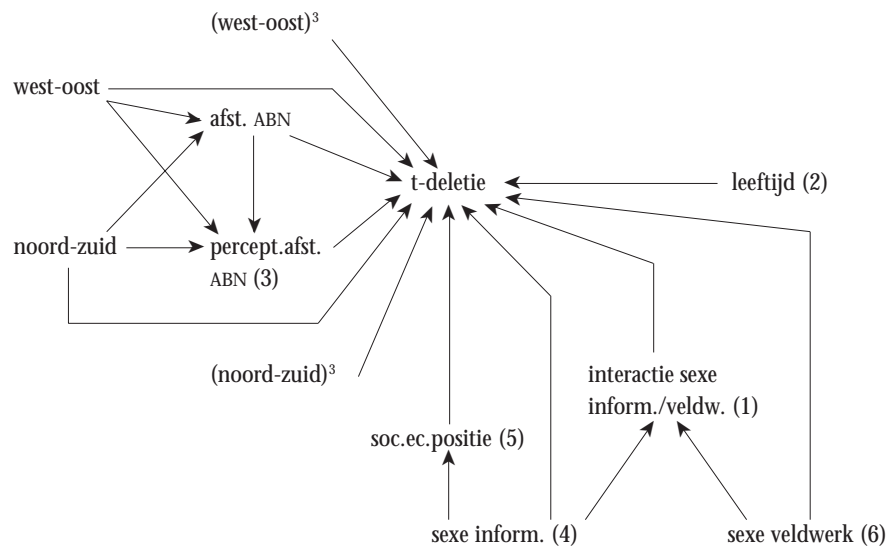
tdel:	woost	wocub	afst	percaf	sexveld
woost:	wocub	afst	percaf	isexen	
nzuid:	nzcub	percaf	afst	sexho	
sexinf:	isexen				
sexho:	isexen	lftijd	wocub		
nzcub:	percaf	afst			
wocub:	afst	percaf	isexen		
percaf:	afst	isexen			
tdel=t-deletie; woost=west-oost; nzuid=noord-zuid; afst=geografische afstand tot de standaardtaal; percaf=gepercipiëerde afstand tot de standaardtaal; sexinf=sexe informant; sexveld=sexe veldwerker; isexen=interactie sexen; lftijd=leeftijd informant; wocub=woost ³ , nzcub=nzuid ³ (woost en nzuid tot de derde macht; de verklaring hiervan volgt verderop).					

Stel dat we constateren dat er een significant verband bestaat tussen t-deletie en gepercipiëerde afstand, en stel dat we vastgesteld hebben dat perceptie van een grote afstand tot de standaardtaal tot veel t-deletie leidt. Dan zou dit effect helemaal te niet gedaan kunnen worden vanwege de geografische positie op de as west-oost (de variabele woost). Deze hangt natuurlijk sterk samen met de gepercipiëerde afstand, en als de factor geografische positie op zichzelf al een grote invloed heeft op t-deletie, dan zou het effect van de gepercipiëerde afstand, in vergelijking daarmee, in het niet kunnen zinken.

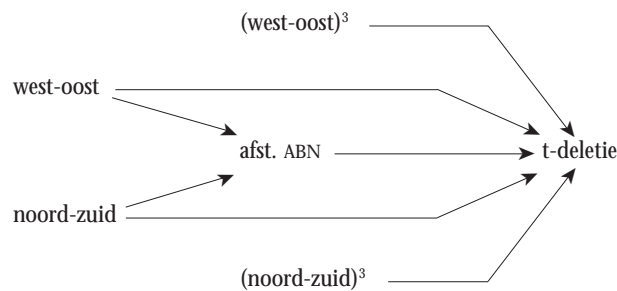
Kiest men een multivariaat model als multiële regressie -in plaats van een serie van bivariate afzonderlijke verbanden- waarin de factoren ten opzichte van elkaar gespecificeerd zijn, dan voorkomt men dit probleem en is men in staat de relatieve sterkte van de invloed van de afzonderlijke factoren echt te schatten. Zo proberen we na te gaan welke invloeden rechtstreeks van belang zijn, in welke mate, en via welke andere verbanden. Met andere woorden, we proberen een causaal model op te stellen.

We specificeren dat “afstand tot de standaardtaal” ten dele bepaald wordt door geografische positie. De gepercipiëerde afstand tot de standaardtaal wordt beïnvloed door enerzijds de afstand tot het standaardtaalcentrum, anderzijds ook indirect door geografische positie. De twee variabelen die de twee dimensies van geografische positie vertegenwoordigen - te weten noord-zuid en west-oost - bepalen, samen met de werkelijke en de gepercipiëerde, subjectieve afstand, het gedrag met betrekking tot t-deletie. Perceptieve afstand is natuurlijk niet bepalend voor de afstand tot het standaardtaalcentrum van het ABN; de werkelijke afstand wordt door de perceptie niet korter, maar omgekeerd kan de afstand tot het standaardtaalcentrum wel de gepercipiëerde afstand bepalen. Op het taalgedrag werken ook een aantal niet-geografische factoren: leeftijd van de informant, sociaal-economische positie (mede door het geslacht bepaald), de sekse van de veldwerker en wat voortvloeit uit de interactie van sekse van informant en veldwerker. Het grote model ziet er dan als volgt uit, waarbij de pijlen, in Fig. 3a, de richting van de beïnvloeding, de causale verbanden, aangeven.

Figuur 3a: Uitgebreid model



Figuur 3b: Uiteindelijk model



Voordat er nog enige aspecten aan de orde komen die met de aard van de afzonderlijke variabelen te maken hebben, moet nu eerst een problematisch punt aangesneden worden. Uit de bespreking van de factor “afstand tot de standaardtaal” en de factor “gepercipiëerde afstand” bleek dat beide wel nauw met elkaar samenhangen, maar toch onderscheiden worden. Omdat beide gemeten zijn ten opzichte van een zelfde centrum zouden ze zelfs zo sterk met elkaar kunnen samenhangen dat de één volstrekt voorspelbaar zou kunnen zijn uit de ander. In statistische zin betekent dit dat zich het probleem van de multicollineariteit voordoet: twee variabelen hangen zeer sterk met elkaar samen, ze correleren te sterk. Dit levert problemen op. De schatting van de gewichten van de betrokken factoren blijft dan wel juist (unbiased) maar de strooiing er om heen wordt groot. Bij perfecte correlatie kan dan zelfs het model niet meer berekend worden. Dit laatste geval doet zich hier gelukkig niet voor, alle variabelen hebben een tolerantie²⁶ die voldoende boven de 0.01 ligt. Maar toch kan een zeer sterke samenhang tussen deze variabelen de schattingen van de overige variabelen vertekenen. Hoewel de correlatie hoog is, wordt de tolerantie van andere variabelen die al in de vergelijking zijn opgenomen daardoor niet onacceptabel klein (vgl. voor deze eis Norusis 1985,55).

In bovenstaand model vormt ieder van de deelsystemen, bestaande uit een afhankelijke variabele en de daar direct op werkende onafhankelijke variabele(n),²⁷ een afzonderlijke vergelijking waarvan de gestandaardiseerde gewichten op iedere variabele geschat kunnen worden en getoetst.

We hebben nu een uitgebreid, maximaal, model gedefinieerd, waarin allerlei mogelijk relevante variabelen in hun onderlinge samenhang zijn gespecificeerd. Onze stelling was evenwel dat voor het stratum van de “Grundmundart” de geografische factoren doorslaggevend zouden zijn. Dat houdt in dat we met een

²⁶ Tolerantie is een maat voor het al dan niet opnemen van een variabele in de vergelijking. Te kleine tolerantie ten opzichte van de overige variabelen leidt tot het niet opnemen in het model. Tolerantie is de maat gevormd door de betreffende variabelen te laten voorspellen uit de andere onafhankelijke variabelen.

²⁷ Afhankelijke variabelen zijn variabelen waarop een pijl gericht staat.

veel zuiniger model toe zouden kunnen, te weten een model waaruit alle onafhankelijke variabelen met een sociaal karakter verwijderd zijn, en waarin uitsluitend variabelen met een uitgesproken geografisch karakter overblijven. Dat betekent dat de variabelen in de rechterbenedenhoek van fig. 3a zullen verdwijnen en dat van de variabelen links van “t-deletie” mogelijk ofwel “perceptieve afst. ABN” of “afst. ABN” zal verdwijnen: de getallen tussen haakjes geven de volgorde van verdwijnen weer. Zie fig. 3b voor dit uiteindelijke model.

4.3. Overige aspecten van de modellering

We behandelen in sectie 4.3.1 de motivering van de transformatie die toegepast wordt in verband met specifieke eigenschappen van de gegevens en in sectie 4.3.2 de motivering van de toevoeging van twee z.g. “constructed variables”²⁸ of toegevoegde variabelen.

4.3.1. *Transformatie van de onafhankelijke variabele t-deletie*

Multipel regressie veronderstelt dat de Afhankelijke variabele een continu karakter heeft. Dat wil zeggen dat de waarden niet discreet zijn, en bij benadering normaal verdeeld.

De afhankelijke variabele t-deletie heeft echter een bijzonder karakter. De scores ervan zijn proporties van t-deletie, en dat is een heel beperkt bereik van 0 tot 1 (of, indien uitgedrukt als percentages, van 0-100). In zo'n geval is het zeker dat de variantie niet constant is. Dit beperkte bereik, samen met het feit dat deze gegevens alle niet-negatieve getallen zijn, maakt dat de oorspronkelijke gegevens niet de vorm van een normale verdeling hebben. Derhalve zijn ook de varianties niet homogeen verdeeld. Daarmee is niet langer voldaan aan een van de voorwaarden voor lineaire regressie.

Om deze reden werden in Goeman (1986) de gegevens van deze variabele getransformeerd door middel van $\ln(p_i / (1-p_i))$, de zogenaamde logit transformatie. De natuurlijke logaritme \ln garandeert daarbij een bij benadering normale verdeling. De ratio $p_i / (1-p_i)$ van de proportie t-deletie, p_i , werd genomen a) omdat een groot deel van de oorspronkelijke gegevens een binair karakter hadden, en b) omdat dan de invloed van de onafhankelijke variabelen op t-deletie het grootst is rond de waarde 0.5. Dat wil zeggen dat er bij dit soort proporties de neiging bestaat ofwel lage waarden in de buurt van 0, of hoge waarden in de buurt van 1 aan te nemen (Pindyck and Rubinfeld 1982², 287-291).

²⁸ Dit betreft voor wat de afhankelijke variabele t-deletie aangaat een andere transformatie van de gegevens dan in Goeman (1986) -hoewel die tot dezelfde klasse van transformaties behoort (Atkinson 1985, 137 e.v.) en voor wat de twee onafhankelijke variabelen “west-oost” en “noord-zuid” aangaat de twee extra “constructed variables” (Atkinson, ibid. 177 e.v.). In tegenstelling tot situaties waar deze toegevoegde variabelen een bizarre structuur hebben blijken ze in ons geval een duidelijk interpreteerbaar facet te vertegenwoordigen en plausibel te zijn.

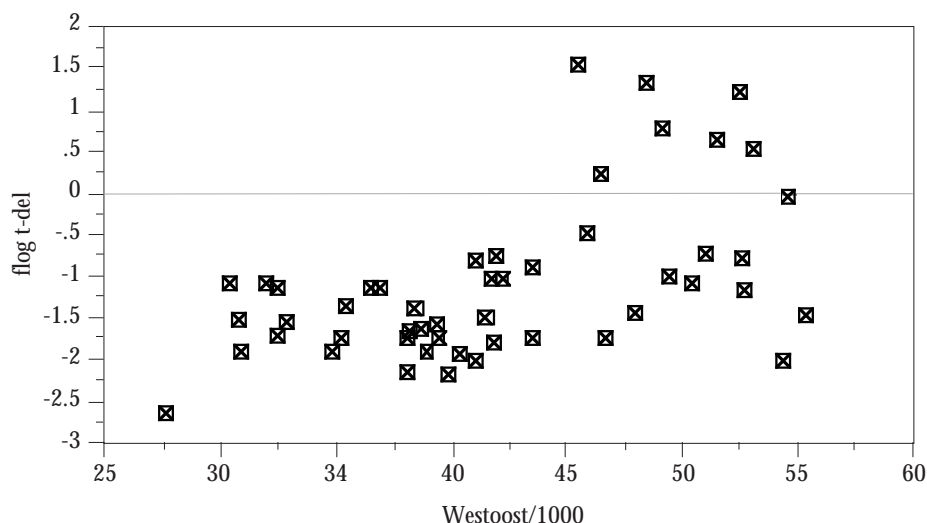
We kunnen een meer algemene oplossing krijgen voor het probleem dat de waarden zich concentreren aan het uiteinde van de schaal. De logit transformatie is een bijzonder geval van een hele familie van transformaties, de zogenaamde “folded power” transformaties (zie Atkinson 1985, 138-144). De “folded power” transformatie bevat een macht tussen 0 en 1; als die 0 benadert krijgt men de logit transformatie, terwijl in de buurt van 1 het resultaat de ongetransformeerde gegevens zijn. Omdat de verdeling van de proporties toch min of meer het gehele interval tussen 0 en 1 bestrijkt lijkt een transformatie die daarvan rekenschap geeft, maar tegelijk van de concentratie van de waarden aan de uiteinden, een goede keus. Derhalve werd de “folded log” gekozen; deze transformatie trekt de waarden -evenals trouwens de logit transformatie- naar het midden van de verdeling zodat deze meer symmetrisch is en de normale verdeling benadert (Mosteller and Tukey 1977). Deze benadering: $(1/2 (\ln p_i)) - (1/2 (\ln (1-p_i)))$ levert een duidelijk lagere standaard error op dan de logit-transformatie. Om de waarde 0, waarvoor \ln niet gedefinieerd is, te vermijden werden de proporties eerst herschaald naar het bereik [0.005-0.995]. De formule daarvoor is $(0.990 * p_i) + 0.005$.

4.3.2. Onafhankelijke variabelen die als toegevoegde variabelen worden beschouwd

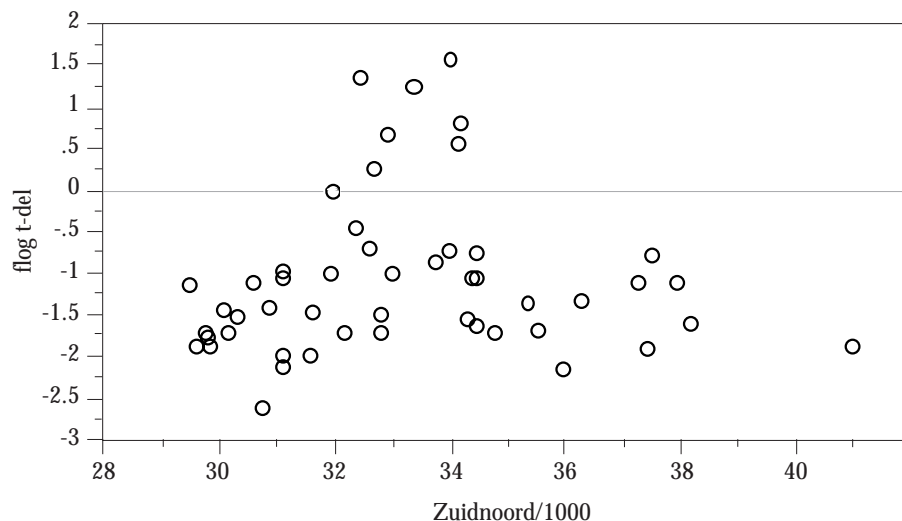
Variabelen die duidelijk geografisch bepaald zijn (zoals bijvoorbeeld wanneer ze uitgedrukt worden in coördinaten) vertonen bijzondere aspecten die te maken hebben met de vorm van het geografische gebied; dat wil zeggen, de vorm van de maximale omvang en van de randen van het onderzoeksgebied.

De verdeling van de scores over de beide geografische dimensies woost en nzuid ziet er schematisch als volgt uit, figuur 4a, 4b:

Figuur 4a: T-deletie van west naar oost



Figuur 4b: T-deletie van zuid naar noord



De vorm van de curve is voor een deel zeker bepaald door de vorm van de geografische ruimte waaruit de steekproef getrokken werd: het gebied heeft min of meer de vorm van een winkelhaak (zie Goeman en Van Reenen (1985, legenda kaarten)).²⁹

Men kan verwachten dat er een trend van hogere orde aanwezig blijft na aftrek van de lineaire combinatie. Op grond van de verdeling van de scores verwachten we dat er een derdemachts-trend aanwezig is. Voor de zekerheid werd eveneens getoetst of er een tweedemachts-trend aanwezig was. Deze lijkt echter minder waarschijnlijk gezien het feit dat de curves niet 1 maar 2 buigingspunten vertonen.

Uit deze toetsing bleek inderdaad dat t-deletie significant afhankelijk was van west-oost, noord-zuid, (west-oost)³ en (noord-zuid)³. De kwadratische factoren bleken inderdaad niet significant. De toevoeging van deze geconstrueerde variabelen en de toetsing van die toevoeging volgt de lijn zoals die door Atkinson (1985, 177) uitgezet is. De toevoeging van de derdemachts variabelen is een alternatief voor een transformatie van de twee oorspronkelijke onafhankelijke variabelen. In dit geval hebben die geconstrueerde variabelen geen bizar karakter. Er bestaat een duidelijke en empirische interpretatie: ze hebben te maken met de spreiding van de gegevenspunten over de geografische ruimte. Er bestaat dus grote kans dat bij een uitgebreidere steekproef met nieuw materiaal uit het defectieve kwadrant van de kaart, deze derdemachtsfactoren niet meer nodig zullen zijn. Temeer daar we weten dat t-deletie ook in het aangrenzende deel van de provincie Utrecht voor-

²⁹ Hier hoofdstuk 6.

komt.³⁰ Mogelijk dat dan ook het multicollineariteitsprobleem (zie einde van paragraaf 3 hiervoor) bij de factoren die met afstand tot de standaardtaal te maken heeft, zich nog minder zal voordoen. Een tweede oorzaak is wellicht gelegen in het feit dat de coördinaten naar verhouding ruw gemeten zijn. In het volgende hoofdstuk zullen we zien dat een preciesere meting van de coördinaten tot een verlaging van de meetfout heeft geleid, hetgeen gevolgen heeft voor de opname van niet-lineaire kwadratische en cubische factoren.

4.4. Toetsing van de modellen

De modellen worden getoetst op de ongestandaardiseerde gewichten (zie appendix 1) maar in de figuren worden de gestandaardiseerde³¹ gewichten weergegeven voor de effecten van de verschillende variabelen. Op die manier ziet men de onderlinge belangrijkheid van de factoren beter omdat de schalen binnen het model verschillend van aard zijn: proporties hebben een bereik van 0 tot 1, maar geografische positie loopt in Nederland van 0 tot in principe een paar honderd kilometers, terwijl leeftijd weer kan variëren van 0 tot boven de honderd. Door standaardisatie worden de schalen waarop de onderscheiden variabelen gemeten zijn onderling vergelijkbaar.

De toetsing van het uitgebreide model geschiedt tragsgewijs: eerst worden de afhankelijke variabele en de daar direct op werkende onafhankelijke variabelen getoetst, en vervolgens de variabelen die daar weer direct op werken.³² De gebruikte methode is die van backward regression; dat wil zeggen dat de direct betrokken onafhankelijke variabelen eerst alle tegelijk ingevoerd en geschat worden, en dat de significanties voor die schattingen berekend worden. Vervolgens worden er stuk voor stuk onafhankelijke variabelen uit het model verwijderd om te kijken of ze wel nodig zijn, dit gaat door totdat er uitsluitend variabelen overblijven die significant zijn.

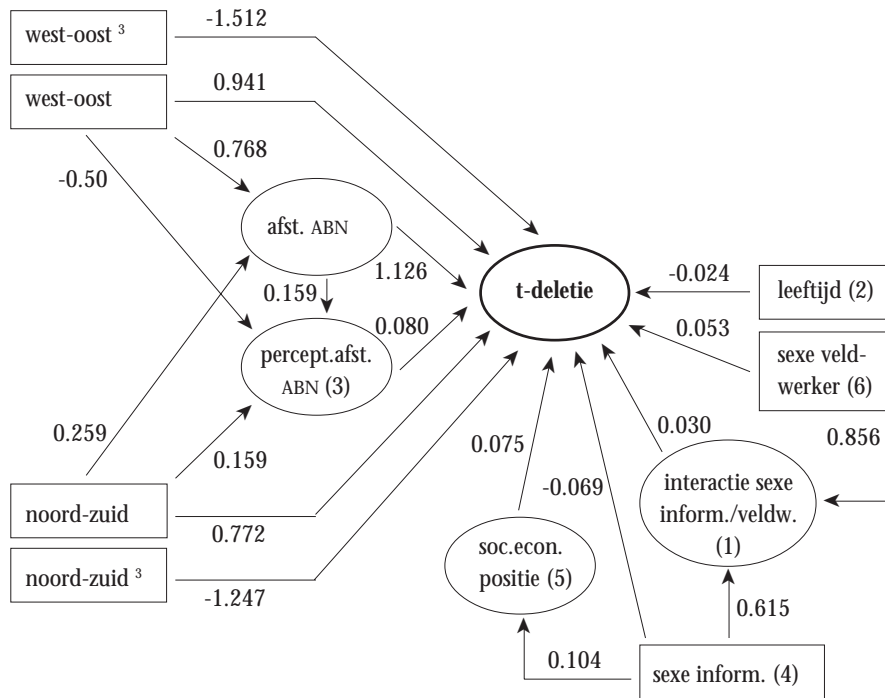
Naarmate er variabelen verwijderd worden neemt de verklaarde variantie enigszins af: de mate van teruggang is een indicatie voor het relatieve (on)belang van de desbetreffende variabelen.

³⁰ Ook in Limburg en het oostelijke gedeelte van Nederland, te weten de Achterhoek en Overijssel, komt t-deletie voor. De intern-talige condities zijn daar evenwel geheel anders dan in het Rivierengebied en Zuid-Holland/Utrecht. In Limburg is het proces grotendeels fonologisch bepaald naast morfologisch, in het Noordoosten niet alleen fonologisch maar daarnaast ook morfologisch. Iets gelijksoortigs doet zich wat betreft de intern-talige condities ook in Vlaanderen voor: t-deletie uitsluitend bij finiete werkwoorden indien voor consonant (Taldeman 1979, 98; Goeman 1983, kaart 5 en 5a).

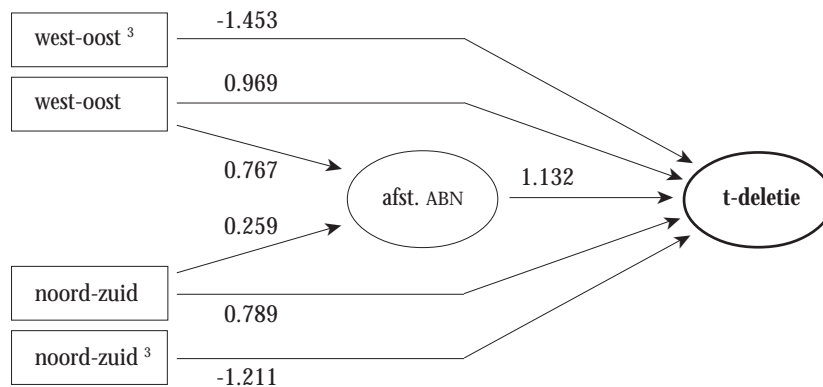
³¹ Dat wil zeggen dat de standaarddeviatie (variantie) van de variabelen gelijk is aan 1.

³² Saris en Stronkhorst (1984, 174) wijzen erop, dat in het geval dat een of meer variabelen compleet afhankelijk zijn van een of meer van de andere variabelen er geen Generalized Least Squares- of Maximum Likelihood-schatting kan plaatsvinden. In dat geval kan nog wel de Unweighted of Ordinary Least Squares-schatting (OLS) gebruikt worden. Omdat de situatie zich hier voordeed is OLS gebruikt. Verder is vanwege de overidentificatie van het model de procedure van Goldberger (1970) gevolgd.

Figuur 5a: Uitgebreid model: schattingen



Figuur 5b: Uiteindelijk model: schattingen



In figuur 5a is met cijfers de volgorde aangegeven waarin de variabelen uit het model verwijderd worden. De verklaarde variantie van het volledige model is 48%. Dit lijkt weinig, maar bij survey-gegevens is het heel gewoon nog lagere verklaarde variantie aan te treffen. In appendix 1 staan de tabellen van de variantie-analyse van het volledige model en de geschiedenis van de backward-regressie, alsmede de significantie van de variabelen berekend met de t-toets.

Het beperkte model, dat het beste en dus het definitieve model is, staat in figuur 5b.

Bij deze modellen geven de gewichten op de pijlen de vermenigvuldigingsfactoren aan waarbij een verschuiving van één standaarddeviatie in de onafhankelijke variabele leidt tot een verschuiving in de orde van dat gewicht op de schaal van de afhankelijke variabele. Dat leidt tevens tot de voorspelde score voor t-deletie onder invloed van die onafhankelijke variabele. Het is de preciese maat waarmee die factor mede de oorzaak is van t-deletie. De cubische factoren (west-oost) en (noord-zuid) werken in omgekeerde richting (negatieve gewichten). Ze vertegenwoordigen een noordwestelijk zwaartepunt tegenover een effect dat van het Noord-Westen naar het Zuid-Oosten groter wordt.

“Afstand tot de standaardtaal” en “geografische positie in termen van coördinaten” blijken de sterkste invloed te hebben. De verwijdering van “gepercipiëerde afstand” blijkt de standaard error sterk terug te brengen bij de andere afzonderlijke variabelen, en wel voornamelijk bij de variabele: “afstand tot de standaardtaal”. Deze vermindering door verwijdering is het gewone en gewenste gevolg bij multicollineariteit. “Objectieve afstand” is dus in staat een groot deel van de taak van “gepercipiëerde afstand” (= subjectieve afstand) over te nemen. Het beperkte model verklaart maar liefst 46% van de variantie.³³ De bijdrage van de factoren die niet meer in het beperkte uiteindelijke model zitten is dus zeer beperkt, het uitgebreide model verklaarde immers maar 2% meer. De gewichten van de uit het model verwijderde factoren waren kleiner dan 0.10 (gestandaardiseerd).

Ondanks dat “gepercipiëerde afstand” niet in het beperkte, uiteindelijke model kan worden opgenomen, correleert het zeer hoog met “afstand tot de standaardtaal”. In zekere zin zou men kunnen zeggen dat “gepercipiëerde afstand” daarom via de variabele “afstand tot de standaardtaal” impliciet in het model verantwoord wordt.

Men kan tegenwerpen dat er bij gebruikmaking van de eerder genoemde andere operationalisatie van het begrip “gepercipiëerde afstand” in termen van richting wel een duidelijk verband geconstateerd zou kunnen zijn. Maar we zagen in sectie 4.1.2.3 dat dat niet aannemelijk is. Wellicht ook is “gepercipiëerde afstand” toch geen lineair gegeven, omdat perceptie van een objectief gegeven voorzover gekoppeld aan de werkelijke, fysieke, stand van zaken, vaak pas een lineair verband te zien geeft na het nemen van logaritmes van beide kanten. De psychofysische wet van Fechner geeft dat verband weer als: $F = bP^x$, waarin F de

³³ De verklarende waarde van dit model is dus sterk verbeterd ten opzichte van het model uit hoofdstuk 3, sectie 3.2.3, dat minder variabiliteit verklaarde: nl. 25%.

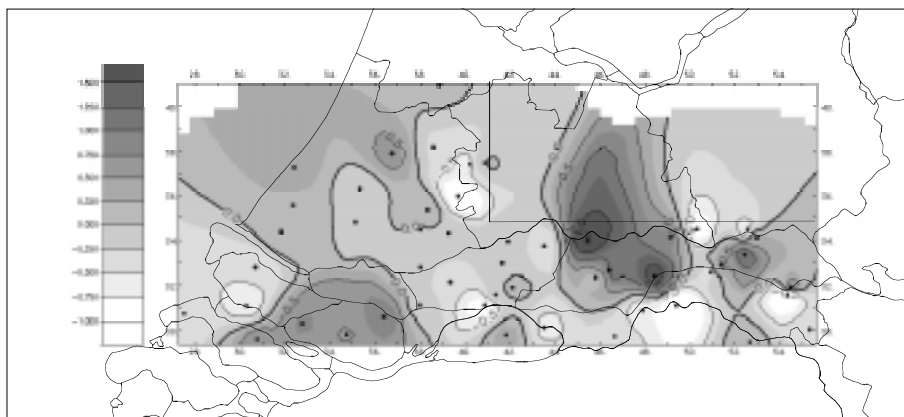
fysieke schaal is, b een constante, P de perceptuele schaal, en x een macht is, die kan verschillen naar gelang de perceptuele modaliteit in kwestie (beoordelen van de luidheid van geluid of beoordelen van de gelijkheid tussen dialecten). Daarop zou de correlatie van $\text{perc}af$ met $wocub$ en $nzcub$ (zie de betreffende tabel 1. hierboven) kunnen wijzen.

Om meer zekerheid in deze materie te krijgen is evenwel nader onderzoek nodig. Voorlopig blijft het zo dat van de twee enkel de variabele: "afstand tot de standaardtaal" significant van invloed is.

4.5. De residuen

Het is mogelijk het onverklaarde deel van de variantie, de residuen, te karteren. De positieve residuen geven aan waar het verschijnsel t -deletie in sterkere mate voorkomt dan voorspeld wordt door het model, en de negatieve residuen waar het in mindere mate dan voorspeld voorkomt. Als er nu geografische samenhang, dat is gebiedsvorming, blijkt te zijn bij de positieve residuen, dan wijst dat erop dat t -deletie extra sterk streekgebonden is boven de algemene tendens uit. Dat is te zien op kaart 3 voor het beperkte model. De geografische verdeling wijst niet op het effect van een spreiding via een netwerk van steden. Met name in de gebieden met positieve residuen is het dus de moeite waard de intern-talige condities nader op te sporen, dat wil zeggen voor het Rivierengebied (Betuwe) en Zuid-Holland. Die intern-talige condities komen aan de orde in Goeman en Van Reenen (1985): hier hoofdstuk 6.

Kaart 3: Residuen van het uiteindelijke beperkte model



4.6. Conclusie

Alle resultaten overziende valt te concluderen dat de geopperde veronderstellingen in het algemeen bevestigd worden.

Het eerste doel was om aan te tonen dat het dialect, in de zin van het stratum "Grundmundart", voornamelijk geografisch bepaald is, en dat het verzamelde materiaal in dat opzicht een "ortsgetreue Darstellung" van de dialecten levert. Sociale factoren/variabelen verklaren relatief weinig en zijn niet significant. Dat betekent dat het dialect uitsluitend *relatief* homogeen is maar zeker niet *absoluut*. Er is geen veldwerkers-effect dat significant is.

Het tweede doel van dit hoofdstuk was te laten zien dat "Afstand tot de standaardtaal" effect sorteert, en wel significant. Het maakt dus uit of een dialect zich verder van het centrum van de gesproken standaardtaal bevindt. Dit is in lijn met Kloeke's theorie over de Hollandse expansie. Daarentegen bleek de "gepercipeerde afstand" niet een zelfstandige rol te spelen. Het grootste deel van het mogelijke effect daarvan is al opgenomen in de variabele "afstand tot de standaardtaal". Hierbij past voorlopig enige reserve, maar toch is dit resultaat een indicatie dat Daan (Daan en Blok 1970) gelijk heeft gehad met de subjectieve opvattingen over gelijkheid tussen dialecten rechtstreeks op de kaart af te beelden in een rangorde die de afstand tot de standaardtaal weergeeft.

We zijn op zoek gegaan naar de begrenzing van ideeën over heterogeniteit en subjectieve inschattingen van sprekers in het kader van traditioneel dialectonderzoek. Voor wat er vervolgens aan onverklaarde variabiliteit resteert, blijven er nu nog twee mogelijkheden over die in aanmerking komen voor een systematische verklaring van het gedeelte variabiliteit dat nog onverklaard bleef (54%): specifieke samenhang binnen dialectgebieden en intertaalige condities van fonologische en andere aard. Aan die specifieke samenhang is het volgende hoofdstuk gewijd, de intertaalige condities worden in de daaropvolgende hoofdstukken besproken.

Appendix

Uitgebreide model

Backward: afhankelijke variabele: tdel

Stap 1 t/m 11 in: onafhankelijke variabelen afst,sexinf, lftijd, socec, nzcub, sexveldw., woost, isexen, percaf, nzuid, wocub (toetsing voor de vergelijking in zijn geheel):

R= 0.6903; R²=0.4765; F-toets=3.310, sign. F=0.003 * (*=significant)

Ongestandaardiseerde gewichten en t-toets voor de afzonderlijke variabelen

VARIABELE	B-GEWICHT	t-WAARDE	t-TOETS	(*=SIGNIFICANT)
afst	0.2318	1.879	0.0676	(net niet sign.)
sexinf	-0.15722E-08	-0.299	0.7666	
lftijd	-0.19912E-09	-0.177	0.8602	
socec	0.56573E-09	0.594	0.5559	
nzcub	-0.00985	-2.922	0.0057	*
sexveldw.	0.99943E-08	0.183	0.8555	
woost	0.27997	2.365	0.0229	*
isexen	0.60060E-16	0.094	0.9258	
percaf	0.01076	0.216	0.8299	
nzuid	0.50672	1.898	0.0649	(net niet sign.)
wocub	-0.00224	-2.209	0.0329	*
(constant)	-4.50582	-3.105	0.0035	*

Stap 12 t/m 17 verwijdert in deze volgorde: isexen, lftijd, percaf, sexinf, socec, sexveldw (toetsing voor de vergelijking in zijn geheel):

R= 0.6784; R²= 0.4602; F-toets= 7.845; sign. F = 0.000 *

Dit levert het uiteindelijke, beperkte, model. De t-toets waarden van de afzonderlijke variabelen van dit beperkte model volgen aan het eind van deze appendix.

Indirecte invloeden in het uitgebreide model

Afhankelijke variabele afst; onafhankelijke variabelen nzuid,woost(toetsing voor de vergelijking in zijn geheel):

R=0.85598; R²=0.7327; F-toets= 67.1563; sign. F=0.0000 *

Ongestandaardiseerde gewichten en t-toets voor de afzonderlijke variabelen

VARIABELE	B-GEWICHT	t-WAARDE	t-TOETS	(*=SIGNIFICANT)
nzuid	0.82540	3.439	0.0012	*
woost	1.10884	10.196	0.0000	*
(constant)	-6.58724	-4.646	0.0000	*

Afhankelijke variabele socec; onafhankelijke variabele sexinf(toetsing voor de vergelijking in zijn geheel):

R=0.10426; R²= 0.01087; F-toets= 0.54951; sign. F=0.4620; F=niet significant

Ongestandaardiseerde gewichten en t-toets voor de afzonderlijke variabelen

VARIABELE	B-GEWICHT	t-WAARDE	t-TOETS	(*=SIGNIFICANT)
sexinf	0.31	0.741	0.4620	
(constant)	0.3E+09	8.179	0.0000	* (=vanwege 0-1-karakter van sexinf de waarde voor 'vrouw')

Afhankelijke variabele percaf; onafhankelijke variabelen: nzuid, woost, afst(toetsing voor de vergelijking in zijn geheel):

R=0.93370; R²=0.87179; F-toets=108.799; sign. F=0.0000 *

Ongestandaardiseerde gewichten en t-toets voor de afzonderlijke variabelen

VARIABELE	B-GEWICHT	t-WAARDE	t-TOETS	(*=SIGNIFICANT)
nzuid	0.77250	2.707	0.0094	*
woost	-0.10966	-0.535	0.5952	
afst	1.36742	8.971	0.0000	*
(constant)	-1.59564	-0.879	0.3839	

Afhankelijke variabele interactie sexe; onafhankelijke variabele sexinf, sexveldw (toetsing voor de vergelijking in zijn geheel):

R=0.91882; R²=0.84423; F-toets=132.7797; sign. F=0.0000 *

Ongestandaardiseerde gewichten en t-toets voor de afzonderlijke variabelen

VARIABELE	B-GEWICHT	t-WAARDE	t-TOETS	(*=SIGNIFICANT)
sexinf	.69444E+07	10.553	0.000	*
sexveldw	.81481E+08	14.683	0.000	*
(constant)	-0.61111E+15	-8.775	0.000	*

Uiteindelijke, beperkte, model. Ongestandaardiseerde gewichten en t-toets voor de afzonderlijke variabelen

VARIABELE	B-GEWICHT	t-WAARDE	t-TOETS	(*=SIGNIFICANT)
afst	0.23310	2.854	0.0065	*
nzcub	-0.00957	-3.167	0.0027	*
woost	0.28814	2.761	0.0082	*
nzuid	0.51710	2.157	0.0363	*
wocub	-0.00215	-2.457	0.0178	*
(constant)	-4.61105	-4.854	0.0000	*