

# 1

## Inleiding

... Das hier besprochene gehört zur quantitativen Seite des sprachlichen Denkens.

... es werde ja das scheinbar psychisch identische bei seinem Austreten in die Außenwelt bald zu einer positiven Größe, bald wieder zu einer Null.

Baudouin de Courtenay, (1929): “Fakultative Sprachlaute”

### 1.1. Probleemstelling: universaliteit, variatie en structuur

Het probleem dat in dit boek wordt behandeld is de deletie van woordfinale -t in de Nederlandse dialecten: er zijn dialecten waar Standaardnederlandse woorden die eindigen op -t zoals *knecht* of *loopt* (3sing pres) stelselmatig gerealiseerd worden zonder -t: bijvoorbeeld *knech* of *löp*. De kwalificatie ‘stelselmatig’ hoeft niet ‘altijd’ te betekenen. Er is met andere woorden variatie en dat vraagt om een kwantitatieve benadering. Het doel van dit onderzoek is de verwerving van een beter inzicht in t-deletie, en de vaststelling van systematiek binnen de variabiliteit.

T-deletie is een vereenvoudiging van de woordvorm en kan gezien worden als een uiting van het universele fenomeen van vermindering van consonantclusters aan het wordeinde -al dan niet in interactie met het begin van het volgende woord (zie hier speciaal hoofdstuk 6, 10 en 11). Als het om een werkwoord met een uitgangst gaat, betekent de vereenvoudiging van de woordvorm tegelijk het verlies van een morfologisch kenmerk.

Er worden in de literatuur een aantal aspecten van universele aard met betrekking tot t-deletie naar voren gebracht. Deze betreffen:

- a) de woordstructuur, dat wil zeggen de bouw van het woord, en niet alleen op het vlak van het vermijden van consonantclusters<sup>1</sup> maar ook de medebepalende rol die de sonoriteit van het voorafgaande consonantisme speelt (bijv. occlusieven minder sonoor dan fricatieven en die weer minder sonoor dan sonoranten als

---

<sup>1</sup> Vergelijk de conditie No Complex Coda uit de Optimality Theory.

*n, r, l*, waarbij *t* bij grotere sonoriteit meer behouden blijft; zie hier hoofdstuk 6, 10 en 12).<sup>2</sup>

- b) de complexiteit van consonantclusters, waarbij bijvoorbeeld articulatorische verschillen deletie kunnen bevorderen (zie hier hoofdstuk 6), maar waar anderzijds gelijksoortige kenmerken in een cluster ook tot dat resultaat kunnen leiden<sup>3</sup> (Guy 1997).
- c) een gepostuleerde natuurlijke tendens tot deletie als gevolg van 'uitspraakgemak' of toegeven aan de inertie van het menselijk spraakorgaan. Over het algemeen ziet men deze tendens duidelijker bij substandaardtaal en dus ook bij de dialecten en dan vooral bij de groep sprekers uit lagere sociale klassen. Hogere sociale klassen proberen daarentegen hun distinctie ook in uitspraak te realiseren en bieden dan weerstand aan natuurlijke processen (Kroch 1978).  
We zullen zien dat sociale variatie in ons materiaal maar heel weinig van de mate van *t*-deletie verklaart. Het is ook tekenend dat in het westen van het land *t*-deletie meer een stedelijk dan een landelijk verschijnsel is (Kloeker 1950) en dat in de steden Haarlem en Amsterdam *t*-deletie vroeger juist beperkt was tot de hogere sociale klassen en niet tot de lagere (Ponelis 1989; 1993), en pas later in de 18e eeuw in de substandaard bij lagere sociale klassen geattesteerd wordt (Raidt 1974).
- d) de gelaagdheid van het lexicon in relatie tot de morfologische klasse van woorden en beider relatie tot de zinsfonologie (vgl. hoofdstuk 11). Baudouin de Courtenay (1929) wees reeds op rol van de graad van morfologisering en de mate van semantisering van een woordfinale *-t* met betrekking tot het tegengaan van *t*-deletie:

... hängt der relative Grad der Widerstandsfähigkeit eines scheinbar identischen Phonems gegen quantitative und sogar qualitative Wandlungen von dem Grade seiner relativen Semantisierung und Morphologisierung ab.

... Es spielt also hier der Stammaslaut *-t*, als morphologisches Gelenk oder morphologische Grenze zwischen zwei Morphemen, eine höchst wichtige Rolle und erinnert ständig an dieselbe.

In hoofdstuk 6 wordt die morfologische rol en semantisering gezien als conceptuele prominentie, al dan niet samengaand met fonische prominentie. In hoofdstuk 7-10 worden de *t*-suffixen geplaatst binnen het totaal van de ontwikkeling van het paradigma waarbinnen ze functioneren. In hoofdstuk 11 wordt die morfologische

<sup>2</sup> In het Chicano English is het omgekeerde het geval: daar is sprake van meer *t*-deletie naar mate de stamconsonant een grotere mate van sonoriteit heeft, zie Santa Ana (1996) die ook een overzicht geeft van zowel universele als taalspecifieke sonoriteitsrangordes.

<sup>3</sup> Een voorbeeld van de werking van de Obligatory Contour Principle uit de generatieve fonologie, waarbij gelijke klankkenmerken in twee adjacenten klanksegmenten in principe uitgesloten worden. Een van de manieren om zo'n constellatie weg te filteren is deletie van een consonant die gelijke specificaties heeft als zijn nabuur in een consonantcluster.

rol geplaatst binnen het kader van een gelaagd model van het lexicon dat deze morfologische tegenwerking pretendeert te verklaren.

Gezien deze aspecten van universele aard zou t-deletie zowel op het vlak van de fonologische als op het vlak van de morfologische conditionering in alle talen en dialecten hetzelfde patroon moeten volgen. Toegespitst geformuleerd moet deze vaststelling tot de hypothese leiden dat t-deletie in de Nederlandse dialecten en in Engels-Amerikaanse dialecten zowel gelijke condities als gelijke effecten van deze condities vertoont.

Dat nu is op een aantal punten niet het geval. We zullen zien dat er in de Nederlandse dialecten in het algemeen minder sprake is van t-deletie bij nomina dan bij verba. Frequenter treedt deletie op in de positie vóór consonant dan in die voor vocaal of in pauza, en na bepaalde occlusieven of fricatieven is er meer t-deletie dan na andere stamconsonanten. Daarbij is articulatorische prominentie in de Betuwse en Hollandse dialecten belangrijker dan fonische prominentie (hoofdstuk 6), terwijl dat in de Engelse dialecten omgekeerd ligt (Neu 1980; Guy 1980).<sup>4</sup>

De situatie wijkt in belangrijke opzichten af van de toestand in het Amerikaanse Engels, niet alleen door de hoge frequentie van t-deletie van bepaalde flectie-morfemen maar ook door het achterblijven ervan bij nomina en verwante woorden (zie hfdst. 11). Deze constellatie levert problemen op voor een bepaalde theoretisch-fonologische interpretatie: interpretatie binnen een model van het lexicon met niveaus (zie hfdst. 11).

Een ander verschil is de grote rol die het stamconsonantisme in de Nederlandse dialecten speelt tegen het kleine gewicht van het volgconsonantisme, terwijl dat in het Amerikaans Engels duidelijk andersom het geval is (Van Hout 1989). In hoofdstuk 11 wordt het relatief belang van stam- en volgconsonantisme, in samenhang met de rol van de (morfologische) woordcategorieën, nader bepaald.

De constatering van de 'eigen' aard van de Nederlandse t-deletieverschijnselen heeft sommigen er toe gebracht om bij de "achtergebleven" woordvormen herinvoering van een -t te postuleren (Jansen (1977), onder invloed van de standaardtaal (Romaine 1986) of onder invloed van regelmatige morfologische processen als suffixaanhechting of onder invloed van frequentie (Knott 1986). We zullen zien dat dit niet het geval is: zie hoofdstuk 5, 10 en 11.

Anderen gaan eveneens uit van invloed van de standaardtaal bij de achteruitgang van t-deletie, maar veronderstellen een bijzondere weerstand tegen herinvoering bij bepaalde flexie categorieën (Schouten 1984, m.b.t. 3sing pres).

Anderzijds is er voor bepaalde dialecten ook een toename van t-deletie geconstateerd (hoofdstuk 5; Schouten 1982, 1984; Hinskens 1992), hetgeen de globale interpretatie van t-deletie als recessief verschijnsel onder invloed van de standaardtaal verder aantast.

---

<sup>4</sup> Hajek (1997) is eveneens kritisch over universaliteitsaanspraken; in zijn geval betreft het verondersteld universele factoren bij nasalering.

Aan woordfinale t-deletie zitten aspecten van de woordfonologie, de zinsfonologie en de morfologie, van frequentie-uitbreiding of -inkrimping en die aspecten kunnen per dialect verschillen.

Niet alleen verschilt de Nederlandse situatie bij t-deletie sterk van die in het Amerikaanse Engels, ook binnen Nederland zijn er grote verschillen en het is dus onterecht om, zoals Van Ginneken (1933) en Ponelis (1989, 1990) doen, de bestaande Nederlandse gebieden met t-deletie te zien als expansie van t-deletie vanuit het Rijnland (Keulen). Die verschillen liggen op het vlak van de conditionerende rol van het consonantisme, van de woordstructuur en de morfologische conditionering (zie hierna sectie 1.2.1 - 1.2.3.<sup>5</sup> Ook een veronderstelde expansie vanuit het Oudfrans naar de Vlaamse dialecten (Weijnen 1991) is niet aannemelijk (zie hoofdstuk 11).

De Noordoostelijke dialecten vertonen juist minder t-deletie bij woorden met een (ad)nominaal karakter, voor een deel omdat in een aantal woorden -t daar niet woordfinaal is, maar gevolgd wordt door woordfinale -e: bijv. predicatief *vaste*. En de deletie in 3sing pres is er wel zeer gemorfologiseerd (zie hoofdstuk 10) omdat een aantal ontwikkelingen tesamen een subklasse van de sterke werkwoorden creëert waarin t-deletie zich voordoet. Het verlies van een morfologisch element -het suffix- dat resulteert uit de vereenvoudiging van de woordvorm is op die manier gecompenseerd. T-deletie is ook regionaal verschillend als we de Hollandse dialecten vergelijken met de dialecten van het Rivierengebied.

Al deze regionale verschillen binnen de Nederlandse dialecten en de verschillen met het Amerikaanse Engels lijken argumenten tegen het universele karakter. Maar dat hoeft niet het geval te zijn. Ten eerste hoeven universele kenmerken niet een alles of niets karakter te hebben, en ten tweede kunnen algemene kenmerken tot heel verschillende uitkomsten leiden als de taalstructurele inkadering verschilt. Dat wordt in dit boek aangetoond voor de Noordoostelijke dialecten, waar de loop van twee processen, Open Syllabe verLenging (OSL) en syncope van de suffixvocaal, tot andere constellaties in de woordstructuur leidt als in andere Nederlandse dialecten; hetgeen gevolgen heeft voor t-deletie.

Universele claims vanuit een theoretische optiek zijn naar voren gebracht op twee gebieden. De eerste is reeds aan de orde geweest en betreft de niveauctuur van het lexicon en de postlexicale processen. De tweede is de claim dat er dan wel

---

<sup>5</sup> Men zou aan Rijnlandse expansie kunnen vasthouden op grond van de veronderstelling dat t-deletie wordt overgenomen, maar dat de condities op Rijnlandse t-deletie geheel worden omgeïnterpreteerd naar de condities die we momenteel in de ontlenende dialecten zien. Dit impliceert dat van bepaalde Rijnlandse factoren het gewicht ongeveer in dezelfde verhouding in Nederlandse dialecten wordt overgenomen, maar dan voor andere factoren. T-deletie is voor het Rijnland variationeel niet onderzocht. Het Limburgse gebied dat door Hinskens (1992) op dit vlak gedetailleerd is geanalyseerd sluit bij het Rijnland aan. Als we uitgaan van de veronderstelling dat dit Limburgse gebied representatief is voor de Rijnlandse constellatie en we vergelijken de situatie in dit specifieke Limburgse gebied met die in de Noordoostelijke dialecten, de Betuwe en Holland dan blijkt uit onze studie (hoofdstuk 10 en 11) dat de verschillen met Limburg niet alleen groot zijn, maar dat er geen zodanige herinterpretatie van de ene groep factoren naar de andere groep aan te wijzen valt.

verschillen tussen dialecten mogen bestaan in algemeen t-deletie niveau, maar dat de waarden voor de afzonderlijke structurele beperkingen - dus het gewicht/het belang van hun invloed - over de dialecten heen constant zijn (Kiparsky 1972, 1988). In hoofdstuk 11 wordt aangetoond dat dat niet het geval is.

In dit boek worden deze onderlinge verschillen duidelijk in het licht gesteld. De concentratie ligt daarbij op het Rivierengebied, Zuid-Holland en op de Noord-oostelijke dialecten.<sup>6</sup> Voor de ligging en de afgrenzing van deze onderzoeksgebieden zie hoofdstuk 12, kaart 1.

T-deletie aan het wordeinde heeft een sterk variabel karakter. Zoals gezegd vereist dit een kwantitatieve benadering. Het gaat daarbij telkens om de systematiek in de structuur van dit taalkundig verschijnsel, in afgrenzing van de onsystematische aspecten. Dat wordt hier zowel in ruimte als in werkelijke tijd onderzocht.

Het blijkt dat een groot deel van de huidige systematiek door het samenlopen van soms toevallige afzonderlijke factoren bepaald is, en dat die systematiek geografisch gezien sterk kan verschillen a) omdat de uitgangssituatie verschilt en b) omdat er per dialectgebied onderscheiden ontwikkelingen hebben plaatsgevonden.

Variabiliteit kan bij een en dezelfde spreker voorkomen. In de linguïstiek wordt het taalgebruik van één informant sinds Bloomfield als idiolect aangemerkt en men neemt aan dat zowel het taalgebruik als het taalsysteem op dit niveau homogeen is. In dit onderzoek is er in principe voor iedere plaats een informant ondervraagd. Dat zou kunnen impliceren dat de variabiliteit op het analyseiniveau van de afzonderlijke informant en in deze ene tijdsdoorsnede relatief homogeen is. Die homogeniteitsassumptie moet evenwel sterk gerelativeerd worden: er is wel degelijk variabiliteit binnen een spreker. De taal van een taalgebruiker is gedurende diens leven niet noodzakelijk homogeen. In de socialisatieperiode van zuigeling tot adolescentie vindt ook de taalverwerving passief en actief plaats, waarbij veranderingen optreden en ook na die periode zijn er leeftijdseffecten, die maken dat variatiepatronen niet altijd constant zijn. Ook binnen eenzelfde tijdsdoorsnede kan er binnenspreker-variantie optreden.

Binnen sprekers bestaat er dus variantie, tussen sprekers bestaan er eveneens verschillen en dus in variatiepatronen. Door de locatie van hun socialisatie en hun domicilie zijn sprekers in de sociale en geografische ruimte gefixeerd.<sup>7</sup> Van dit kenmerk van sprekers maakte de klassieke dialectologie gebruik door een dialectspreker als representant van het dialect van een bepaalde plaats te beschouwen teneinde de variantie binnen en tussen dialecten te kunnen bestuderen. Het hier gerapporteerde onderzoek sluit op die traditie aan maar beperkt zich niet tot het constateren van de overeenkomsten en de verschillen tussen dialecten of groepen van dialecten. Het gaat veeleer om de structuur in die variabiliteit, onder aftrek van de 'toevallige' factoren.

---

<sup>6</sup> De situatie in Limburg komt in dit boek zijdelings aan de orde; ik verwijs naar de dissertatie van Hinskens (1992).

<sup>7</sup> Deze fixatie is relatief vanwege de dagelijkse mobiliteit.

## 1.2. Stand van zaken in de Nederlandse dialecten

De stand van zaken met betrekking tot de studie van t-deletie in de Nederlandse dialecten valt op te delen in drie aspecten: taalstructurele, taalgeografische en diachrone aspecten.

### 1.2.1. Structurele aspecten

Thematisch kunnen er vier groepen onderscheiden worden.

#### a) Stamconsonantisme en volgende omgeving

T-deletie komt overwegend voor na fricatieven, minder na occlusief: Van Ginneken (1933) in het gehele taalgebied; Van Veen (1964) Utrecht; Daan (1965) Amsterdam; Van Hout (1995) Rivierengebied, Noord-Oost Brabant en Noord Limburg. In de twee laatste gebieden gedragen de occlusieven zich mogelijk niet uniform.

Ook is vastgesteld dat t-deletie meer voorkomt na occlusief dan na fricatief en na liquida en nasaal het minst. In de zinsfonologie is het effect van volgende consonant groter dan het effect van een volgende vocaal, in pausa is er de minste deletie: De Vries e.a. (1974) Leiden; Schouten (1982, 1984) Utrecht, Zuiden provincie Utrecht; Goeman en Van Reenen (1985; hfdst. 6), Ottow-Kolman (1989): Zuid-Holland en Rivierengebied; Van Hout (1989) Nijmegen.

Na -s- gedraagt t-deletie zich wat afwijkend: Jansen (1981) Amsterdam; Goeman en Van Reenen (1985) Zuid-Holland en Rivierengebied; Van Hout (1989) Nijmegen.

#### b) Woordstructuur

Er bestaat mogelijk ook een verband met de woordstructuur of de syllabestructuur. In de Ripuarische dialecten zou t-deletie uitsluitend voorkomen in woorden met gerekte klinker. Op deze manier wordt een superlange stamsyllabe (VVCC) gereduceerd tot een van normale lengte (VVC): Frings (1922), Van Ginneken (1933). Deze hypothese van Frings geeft Van Ginneken de mogelijkheid om een relatie met Keulen<sup>8</sup> te postuleren overal waar *gerekte* stamvocaal voorkomt binnen t-deletiegebieden (Kortrijk, Zuidelijk Utrecht en verspreid in de Noordoostelijke dialecten). De mate van t-deletie heeft evenwel geen relatie met de verlenging van een korte stamvocaal, omgekeerd leiden verkorting of verlenging van de stamvocaal ook niet tot meer t-deletie: Goeman (1987a) voor Zuid-Holland en Rivierengebied.<sup>9</sup> Aangetoond is bovendien dat de klinker *verlengingen* in het Rivierengebied hypercorrect zijn en geen natuurlijk fonologisch proces representeren (Goeman en Van Reenen 1986; Goeman 1987a). Hieruit volgt dat een syllabebalans, als gepostuleerd door Van Ginneken, als die al voor het Limburgs geldt, zeker niet overal gelding kan hebben. Van Ginneken verlaat overigens dit idee later weer

<sup>8</sup> Keulen als uitstralingshaard is niet aannemelijk; zie hieronder.

<sup>9</sup> De geografische distributie van verkorting en verlenging is te vinden in Goeman en Van Reenen (1987). Groothuis (1885) observeert eind 19e eeuw na korte stamvocaal verlenging van fricatief bij t-deletie in de westelijke Betuwe. Daar zijn in de huidige dialecten geen voorbeelden meer van gevonden.

(1938), omdat hij de oorzaak in de sonoriteitsopbouw van de syllabe zoekt en niet meer zoals Frings in de more-structuur van de syllabe. Verband met de sonoriteitsopbouw is ook aangetoond door Van Hout (1989a) Nijmegen.<sup>10</sup> We tonen aan dat de woordstructuur, en met name die van de woordstam, in de Noordoostelijke dialecten (hoofdstuk 10) wel een duidelijke rol speelt met betrekking tot t-deletie en dat de sonoriteitsschaal alleen in bepaalde dialecten een rol speelt (Middelwestvlaams en Noordoostelijke dialecten wel, Betuws en Zuid-Hollands niet) en met betrekking tot vocaalsyncope mogelijk alleen in het Middelnederduits van Goor (hoofdstuk 12).

#### c) Morfologische woordklassen

T-deletie komt in de Nederlandse dialecten meer voor in 3sing dan bij part pret en nomina: Van Veen (1964) Utrecht, De Vries e.a. (1974) Leiden, Schouten (1982, 1984), Goeman en Van Reenen (1985, hier hfdst. 6) en Ottow-Kolman (1989), Zuid-Holland en Rivierengebied, Van Vliet (1988), Van Hout (1989) Nijmegen, Scholtmeijer (1993) Zuidelijk Utrecht. Goeman en Van Reenen (1985) tonen de volgende hiërarchie van woordklassen aan: 1) onregelmatige pret, 2) 23sing pres, 3) part pret, nomina en adjectieven. Op een onderdeel is Nijmegen afwijkend. De monomorfematische woorden hebben er een hogere t-deletiefrequentie dan de part pret-vormen: Hinskens en Van Hout (1994). Oostelijk Zuid-Limburg gedraagt zich eveneens anders. T-deletie komt er meer voor bij monomorfematische woorden dan bij flectie-t; 2plur pres vertoont pas recentelijk t-deletie: Hinskens (1992). In de noordoostelijke dialecten komt t-deletie beperkt voor bij nomina en zeer specifiek in 3sing pres, dat wil zeggen, uitsluitend in bepaalde sterke en onregelmatige werkwoorden: Bezoen (1938) Enschede, Entjes (1970) Vriezenveen, Knott (1986) noordoostelijke dialecten. In de Vlaamse dialecten bestaat t-deletie vrijwel uitsluitend bij 23sing en 2plur pres: Goeman (1983 [=1976]; 1976b), Taeldeman (1979 en 1985), De Schutter en Taeldeman (1986). In Kortrijk is t-deletie echter een algemener proces (Vauterin 1961).

#### d) Natuurlijk proces

Gezien het veelvoud aan condities, vooral ook in morfologisch opzicht, zou t-deletie, in plaats van een uniform proces te vormen, uit verschillende processen kunnen bestaan. Zo lijkt bijvoorbeeld het onderscheid woordfonologisch proces versus zinsfonologisch proces zinvol en wordt het bestaan van t-deletie als natuurlijk proces in de Nederlandse dialecten ontkend omdat het proces inmiddels sterk gelexicaliseerd zou zijn (Knott 1986). Maar ook al bestaan deze onderscheidingen, dan nog is er sprake van éénzelfde onderliggend proces voor alle woordvormen (Van Hout 1980) in het Rivierengebied. Dat volgt uit het feit dat alle woordvormen zich op één dimensie -die van een deletie-proces- laten ordenen. Goeman (1988, hier hfdst. 5) betoogt ook dat t-deletie één natuurlijk proces is in het middengebied van het land.

---

<sup>10</sup> Voor Chicano Engels is die rol ook aangetoond door Santa Ana (1996).

Algemene frequentie-niveaus worden gemodificeerd door de onderscheiden waarden van de afzonderlijke taalinterne, dialectspecifieke constraints (Van Hout 1980). Verderop komt het omgekeerde standpunt aan de orde, namelijk dat de algemene t-deletiefrequentie-niveaus wel variëren, maar dat de waarden voor de afzonderlijke constraints constant zouden zijn (universaliteitsaanspraak): Kiparsky (1972, 1988).

### 1.2.2. *Geografische aspecten*

Voor alle niet-noordoostelijke dialecten wordt aansluiting gepostuleerd met de Ripuarische dialecten, met als uitstralingskern Keulen (Van Ginneken 1933) en datzelfde verband ziet Ponelis (1989, 1990, 1993) in een Middelnederlands continuum van het Zuidoosten (Limburg) -en expanderend vanuit het Rijnland- naar het Zuidwesten: naar Brabant, Vlaanderen, Zeeland en Zuid-Holland. Van Ginneken schetste het taalgeografische beeld voor het gehele taalgebied aan de hand van één, nominaal, voorbeeld: de taalkaart *knecht*. Het Noordoostelijke deletiegebied wil hij ook in dat Ripuarisch-Rijnlandse verband zien, maar hij overweegt een alternatieve aansluitmogelijkheid, namelijk een verband met het Nederduitse t-deletiegebied aan de Elbe-monding. We zagen reeds dat hij (1938) voor het Zuidoosten van het taalgebied terugkomt van het idee van een syllabebalans zodat dit argument voor de Ripuarische connectie is vervallen.

We zullen in hoofdstuk 10 zien dat het geografisch beeld in Duitsland zelf al versnipperd is.

Het platteland van Zuid-Holland vertoont nauwelijks t-deletie, maar de Zuidhollandse steden 's-Gravenhage en Rotterdam wel; minder t-deletie is er in Leiden, nog minder in Gouda maar verder weer algemeen in Utrecht, Woerden, Gorcum, Dordrecht en Culemborg: Kloeke (1950) op grond van gegevens van het eind van de negentiende eeuw uit de enquête van het Aardrijkskundig Genootschap en voor Culemborg op basis van de Belgische enquête Willems. Deletie op het Zuidhollandse platteland wordt wel in de achttiende eeuw in Lekkerkerk aangetroffen (Raidt 1974), maar deze plaats sluit dicht aan op het Utrechtse gebied.

Het Utrechtse t-deletiegebied wordt zelf weer geacht aansluiting te hebben met het (Zuid)oosten en omdat het Zuid-Hollandse platteland nauwelijks deletie vertoont maar de steden wel, is er een Utrechtse expansie gepostuleerd, richting Zuid-Holland (Van Veen 1964, 1966). Binnen Utrecht is er een continuum met groeiende deletie frequentie van West naar Oost geconstateerd (Scholtmeijer 1993). Goeman en Van Reenen (1985, hier hfdst. 6) zien Utrecht en het Rivierengebied als een geheel. Het Rivierengebied is niet homogeen, er kunnen drie groepen van dialecten onderscheiden worden, de middelste groep heeft de laagste deletiefrequentie (Van Hout 1980). Die niet-homogeniteit komt in dit boek ook in hoofdstuk 4 en 5 aan de orde.



### 1.2.3. *Diachrone aspecten*

In de dertiende eeuw zijn er al aanwijzingen voor t-deletie in Utrecht (voornamelijk de stad), Zuid-Holland (de stad Leiden) en in Limburg; in de veertiende eeuw is het verschijnsel geattesteerd in Brabant: Diest, Mechelen en Brussel (Van Loey 1976). Ook in Vlaanderen zijn er attestaties gevonden voor Ieperen, Brugge, Gent en Kortrijk. In deze steden treft Mantou (1972) ook t-deletie aan in *franstalige* oorkonden en rechtsoptekeningen uit de periode 1265-1348; daar komen ook Nederlandstalige plaatsnamen met t-deletie in voor: *Delf*, *Does* en *Deurdrech*.

Vanaf de veertiende eeuw wordt t-deletie ook in de Noordoostelijke dialecten aangetroffen (Kloeke 1956; zie hier hfdst. 10).

In schriftelijke bronnen uit de zestiende, zeventiende en achttiende eeuw, van Hollandse steden, is het verschijnsel eveneens bekend. Het beeld van het eind van de negentiende eeuw vinden we dan nog terug: nauwelijks of niet op het platteland, wel in de steden. Aanvankelijk komt in de zestiende en zeventiende eeuw t-deletie in schriftelijk taalgebruik tot uiting van vooral de sociale bovenlaag, althans in Amsterdam en Haarlem.<sup>11</sup> De volkstaal in de kluchten van Bredero biedt geen voorbeelden van t-deletie op de momenten dat Amsterdams of 'Noord-Hollands' taalgebruik wordt weergegeven. Later, in de achttiende eeuw, lijkt t-deletie zich meer te beperken tot de schriftelijke productie van lagere sociale klassen (Raidt 1974) en in de negentiende eeuw wordt het in Amsterdams dialect aangetroffen (Thijm en van Lennep 1885) waar het in de twintigste eeuw nog weinig frequent is (Daan 1965) maar in de jaren '70 nog steeds voorkomt (Jansen 1981).<sup>12</sup>

Voor Kortrijk is uit de 14e eeuw algemenere t-deletie aangetoond (Debrabandere 1973) dan nu in Vlaanderen voorhanden is (Goeman 1983); zie ook sectie 1.2.1.c. hierboven. In later Middel-Westvlaams zijn er ook attestaties voor Werveke (Vermeersch 1962). In het typische relictgebied Frans-Vlaanderen treft men vaak

<sup>11</sup> Van den Berg (1965): Woerden, Delft, Delfshaven, Rotterdam, Weesp, (alle 16e eeuw), Gouda (16e eeuw); Van Loey (1976): Utrecht (13e eeuw), Leiden (14e eeuw), Dordrecht (15e eeuw); Van Veen: (1964) Utrecht (13e, 14e en 15e eeuw), Crena de Jongh (1959): Delft (17e eeuw), Raidt (1974): Oudewater, Amsterdam (2e helft 17e eeuw), Rotterdam (2e helft 18e eeuw; deze scribent gebruikt ook uitgangsloze 3plur), platteland Holland, Lekkerkerk (3e kwart 18e eeuw; deze scribent gebruikt ook uitgangsloze 3plur pres), Zevenbergen (N.Br. aansluitend bij Holland, midden 18e eeuw; 13plur pres en pret uitgangsloze vormen). Raidt: Amsterdam, weesbriefjes van ongeschoolde schrijvers 1e kwart 18e eeuw. Attestaties uit de sociale bovenlaag zijn beschreven door Snyman (1979) voor Amsterdam (17e eeuw) en door Stoops (1980) voor Haarlem (16e eeuw).

Raidt wijst op het nog ongestandaardiseerd karakter van de Amsterdamse schrijftaal in de tweede helft van de 17e eeuw. In een ongestandaardiseerde situatie hebben lokale uitspraakvarianten de kans om genoteerd te worden. Voor standaardisatie-processen op macro- en microniveau in de zeventiende eeuw zie nu samenvattend Van der Wal (1995). Het microniveau betreft de keuze voor een bepaald dialect als standaard, op het microniveau staan de afzonderlijke sprekers iedere keer weer voor de keuze welke specifieke woordvormvarianten te gebruiken.

<sup>12</sup> Het komt dan vooral voor in meer formeel getint taalgebruik, minder in spontane taal. Jansen verklaart dit uit het feit dat de monitoring van taalgebruik weggedrukt werd ten gunste van de monitoring op het juist beantwoorden van de formele vragen.

de archaische vormen aan, maar t-deletie doet zich daar niet voor. Het feit dat er in Vlaanderen nog een klein groepje dialecten rond Kortrijk, de stad inclusief, algemenere t-deletie kent, tesamen met het feit van vooral de 14e eeuwse algemene t-deletie in Kortrijk,<sup>13</sup> heeft ertoe geleid dat er een samenhang gezien werd met De Oudfranse t-deletie aan het wordeinde (Weijnen 1991). De t-deletie in het Oudfrans lijkt evenwel een ander karakter te hebben en in de direct aansluitende Noordfranse dialecten komt t-deletie juist minder of niet voor, er is dus geen geografisch continuüm tussen Oudfranse en Vlaamse t-deletie. Zie daarvoor verder hoofdstuk 11 in dit boek.

Reeds opgemerkt is dat Weijnen en Ponelis vanwege de geografische aansluiting bij het Rijnland expansie vanuit Keulen aannemen. Weliswaar is t-deletie in het Rijnland nog niet uit het oogpunt van variabiliteit onderzocht, maar het door Hinskens (1992) geanalyseerde gebied maakt deel uit van de Ripuarische dialecten of sluit er nauw bij aan. De constellatie van t-deletie is er totaal anders (zie voor de spreiding hoofdstuk 10, krt. 1-3). Algemene expansie vanuit Keulen is dus niet aannemelijk.

Voor Holland wordt aangenomen dat de oorspronkelijke algemene t-deletie er later onder invloed van de standaardtaal teruggedrongen is (Daan 1965; Raidt 1974; Ponelis 1989; 1990). Dit standpunt wordt ook gedeeld door Romaine (1986) en Knott (1986).

Voor het Hollands-Afrikaans, de overgangsfase van 1739-1851 naar het Afrikaans, heeft Conradie (1981) aangetoond dat er een duidelijke tweedeling is tussen enerzijds monomorfematische woorden en part.pret en anderzijds flectie-t. Flectie-t wordt meer gedeleerd. Pas later treedt er een differentiatie tussen monomorfematische en partpret-vormen op. Het patroon van meer deletie bij flectie-t dan bij partpret geldt er ook bij stammen op sonorante consonant. Ponelis (1989) zet dat deels op rekening van het deflectieproces in het Hollands-Afrikaans.

Een andere kolonistentaal, New Netherland Dutch of New Jersey Dutch vertoont daarentegen geen spoor van t-deletie. Buccini (1996) wijst ter verklaring op het ontbreken van t-deletie in de Amsterdamse en Noord-Hollandse volkstaal en in de Zuid-Hollandse plattelandsdialecten. Het relatieve succes van t-deletie in Afrikaans wijt hij eerder aan imperfect learning dan aan selectie uit Nederlandse dialectkenmerken. Goeman en Van Reenen (1985, hier hoofdstuk 6) vestigen echter voor Afrikaans de aandacht wel op de dialectachtergrond (i.c. die van gouverneur Van Riebeeck c.s).

Niet in real time, maar in apparent time is er sprake van teruggang van t-deletie in Zuid-Utrecht (Schouten 1984) in een aflopend continuüm van de oudste generatie naar de middelste en de jongste generatie. In real time is er eveneens sprake van teruggang, maar anderzijds is er voor bepaalde plaatsen juist een toename van t-deletie geconstateerd. Toename zowel als afname is in real time geconstateerd voor het Rivierengebied (Goeman 1988, hier hoofdstuk 5).

<sup>13</sup> Vg. ook de door Mantou gesignaleerde gevallen te Ieperen, Brugge, Gent en Kortrijk.

Toename in apparent time is aangetroffen in Limburgse dialecten (Hinskens 1992) terwijl te Nijmegen een afname van deletie bij partpret is aangetroffen (Hinskens en Van Hout 1994).

Vanwege de geografische verschillen, die ook diachroon een rol spelen, worden in dit boek de verschillen bij t-deletie zowel in geografisch opzicht apart als in relatie met de intertaalige verschillen nader geanalyseerd.

### 1.3. De data

De meeste gegevens voor deze studie zijn afkomstig uit het z.g. Goeman-Taeldeman Project: het project "Fonologie en Morfologie van de Nederlandse dialecten op basis van Veldwerk". De opzet van het project is nader beschreven in Goeman en Taeldeman (1996). Dit project beoogt een representatieve uitsnede van de fonologische en de morfologische verschijnselen in de Nederlandse dialecten te verzamelen en gemakkelijk toegankelijk te maken in een database. De vragenlijst werd ontwikkeld door een commissie van Belgische en Nederlandse dialectologen die ook de selectie maakte van de plaatsen die in het project opgenomen moesten worden.<sup>14</sup> De kenmerken van die plaatsen en de sociale kenmerken van de informanten komen aan de orde in hoofdstuk 2. Van de oorspronkelijke voor Nederland geplande plaatsen zijn er nu 363 getranscribeerd in een enge transcriptie en in de database opgenomen. De manuscript-transcripties zijn in het systeem van de International Phonetic Association (IPA). Voor de invoer in de computer is een aangepaste versie van het IPA ontworpen: keyboard-IPA; die maakt uitsluitend gebruik van de normale, alfanumerieke ASCII-codes teneinde uitwisselbaarheid tussen verschillende besturingsplatforms (Unix, Mac-OS en DOS/Windows) te garanderen. Er wordt momenteel gewerkt aan een conversie van het in het project gebruikte keyboard-IPA naar gewoon IPA.

Het onderzoek waarover in de hiernavolgende hoofdstukken wordt gerapporteerd vond plaats in verschillende fasen van het project, vandaar dat niet in alle hoofdstukken dezelfde aantallen plaatsen voorhanden waren. Een overzicht biedt hoofdstuk 12: voor de onderwerpen (16) en voor de betreffende data (17a-c). Zo was er bijvoorbeeld voor de noordoostelijke dialecten een wat kleiner aantal plaatsen beschikbaar in hoofdstuk 7, 8 en 9 (situatie 1992) dan in hoofdstuk 10 en 11 (situatie 1994). Zo zijn ook de overzichten van hoofdstuk 2 gebaseerd op de situatie van 1992.

## 1.4. Taalvariatie en kwantitatieve modellering

### 1.4.1. *T-deletie en taaltheorie*

In een pre-structuralistisch kader is t-deletie door Baudouin de Courtenay (1929) aan de orde gesteld, zoals hiervoor aangegeven. Hij gaat ervan uit dat een klank

---

<sup>14</sup> Zie voor nadere gegevens en voor de lijst van items uit de vragenlijst Goeman (1984), een monografie waarvoor mede gegevens uit het project voor het eerst integraal konden worden

die drager is van morfologische of semantische informatie minder aan deletie onderhevig is dan een klank die zulke informatie niet draagt. In een structuralistisch kader wordt een gelijksoortige verklaring door Twaddell (1936) gegeven: t-deletie kan optreden als elders in de pret-vorm een klank verschilt van de pres-vorm: *keep* - *kep+ø*, maar blijft achterwege als de twee vormen zouden samenvallen: *heap* - *heap(ed)*.<sup>15</sup> Beide verklaringen kunnen gezien worden als het behoud van een functioneel onderscheid (Kiparsky 1972). Daarna is t-deletie sinds het begin van de kwantitatieve sociolinguïstiek permanent op de onderzoeksagenda geweest.<sup>16</sup>

Labov et al. (1968) verklaren t-deletie in termen van fonetische omgeving van de slot-t, de consonant die eraan voorafgaat, wat er op het desbetreffende woord volgt en de morfologische categorie van het woord, de generatieve fonologische regel wordt voorzien van frequentiegewichten, waarbij ook de deelregels (de condities) van een regel gewichten toegekend krijgen.<sup>17</sup>

Vanaf het begin wordt daar ook op gereageerd door theoretisch taalkundigen. Over het algemeen wijzen die de verbinding van taalstructuur en frequenties af; zo bijvoorbeeld Kiparsky (1972). Frequenties horen in die optiek niet tot de grammatica, de mentale toerusting van de spreker, iets wat nog wel door Baudouin de Courtenay werd aangenomen blijkens het motto van dit hoofdstuk. Kiparsky stelt dat de frequentieverhoudingen geheel of gedeeltelijk voorspelbaar zijn uit universele condities van functionele aard die voortvloeien uit de 'speech performance'. Die condities betreffen enerzijds de fonotaxis: die oefent een positieve invloed op deletie uit, en anderzijds de behoefte aan onderscheid: die oefent een negatieve invloed uit, wat tot behoud van -t leidt.

Guy (1991a) ent zijn model op het model van de lexicale fonologie, dat gaat uit van een lexicon waarin een aantal niveaus te onderscheiden valt, te weten het niveau van monomorfematische woorden, vervolgens dat van woorden met een onregelmatige morfologie, en dan op het niveau van woorden met een regelmatige morfologie. Op ieder niveau kunnen dezelfde regels opnieuw werken en dan komt vervolgens na de lexicale fonologie, dus buiten het lexicon, de postlexicale fonologie (woordgroep- en zinsfonologie) waar nog eens dezelfde regels toegepast kunnen worden; die niveaus vindt Guy terug in de frequentie-verhoudingen bij t-deletie.

Sinds het begin van de jaren '90 convergeren beide onderzoekstradities enigszins en dat komt door de ontwikkeling van Optimality Theory. Optimality Theory gaat uit van *universele* condities, maar daarnaast van *taalspecifieke* rangordening van die condities en van die rangordes kunnen er verscheidene zijn. De theorie

---

gebruikt dankzij de gekozen vorm van computerisering. Zie verder ook Veenhof-Haan en Van Vliet (1989) voor een latere fase van het project en Goeman en Taldeman (1996).

<sup>15</sup> Zo'n functionalistische verklaring is door Bezoen (1938) en Entjes (1970) ook voor Twentse t-deletieverschijnselen gegeven. We komen daar in hoofdstuk 10 nog op terug.

<sup>16</sup> Zie voor een overzicht van de Amerikaanse onderzoeksthema's Labov (1997).

<sup>17</sup> Labov, Cohen, Robins en Lewis (1968), Wolfram (1969), Fasold (1972), Guy (1980), Neu (1980), Guy (1991, 1991a, 1994, 1997, 1997a) Guy en Boberg (1997), Santa Ana (1992, 1996) en Labov (1975, 1994, 1997).

kan op die manier variatie tussen talen en dialecten verklaren, variatie die door Antilla (1997) in frequentie-termen beschreven wordt: er zijn dan binnen een grammatica condities die geordend zijn met andere condities (totale ordening), daarnaast zijn er condities of groepen van condities die ongeordend zijn ten opzichte van een bepaalde condities (partiële rangorde) en dat leidt tot variatie. Er zijn dan namelijk verschillende rangordes mogelijk. Als een woordvorm voorspeld wordt door de ene rangorde en ook nog door andere rangordes, dan is de voorspelde frequentie:

$$p(a) = n / t$$

waarbij  $p$  de waarschijnlijkheid van de vorm  $a$  is, die voortvloeit uit  $n$  verschillende rangordes ten opzichte van  $t$  het totaal aantal rangordes dat mogelijk is.

Een andere mogelijkheid tot frequentievariatie is dat een bepaalde universele conditie niet alleen ongeordend is ten opzichte van een blok andere condities, maar binnen dat blok op elke willekeurige plaats kan voorkomen (z.g. 'floating constraints': Nagy en Reynolds (1997).

Kiparsky (1994) blijft evenwel bij zijn vroegere standpunt dat frequentie niet tot de grammatica behoort<sup>18</sup> en herformuleert Guy's model uit de Lexicale Fonologie in termen van de Optimality Theory (zie hoofdstuk 11).

Binnen een generatief kader met cyclische en postcyclische fonologische regels plaatsen Hinskens en Van Hout (1994) hun bevindingen voor de dialecten van Nijmegen en Rimburch. Monomorfematische woorden deleren significant meer  $t$  dan partpret: een deel van de regels is postcyclisch, maar komt nog voor de zogenaamde postlexicale regels van de zinsfonologie. Zij laten evenwel de status van flexie-morfeem- $t$ , die vaker gedeleerd wordt dan partpret- $t$ ,<sup>19</sup> buiten beschouwing.

In dit boek ligt de nadruk minder op regels en meer op condities,<sup>20</sup> we komen daar aan het eind van sectie 1.4.2 nog op terug.

Beneden het niveau van het klanksegment, dat van de 'features' of klanksegmenten worden de deletieregels ook in termen van features als 'coronaal' en 'sonorant' geformuleerd, al dan niet verbijzonderd naar de overeenkomsten en de verschillen in klankkenmerken die een cononantcluster op  $-t$  uitmaken. Goeman en Van Reenen (1985), hier hoofdstuk 6, bepalen die overeenkomsten en verschillen voor Zuid-Holland en het Rivierengebied zowel op auditorisch als articulatorisch vlak en ze vatten ze samen onder het begrip 'prominentie' en 'complexiteit'. Al vroeg generaliseerde Van Hout (1989a) de segmentele aanpak in de richting van

<sup>18</sup> "...variation comes from competition of grammatical systems (in the individual or in the community), not from a probabilistic component in the rules of the language". Bezwaren tegen frequenties zijn eerder ook te vinden bij Bickerton (1971) die een implicationeel schaalmodel preferereert boven het variabele regelmodel.

<sup>19</sup> Het verschil flexie- $t$  en partpret- $t$  reflecteert mogelijk het verschil tussen flexie en derivatie. Zie hier hoofdstuk 11 en 12.

<sup>20</sup> De hoofdstukken 4-6 zijn eerder ontstaan dan de andere en daardoor wordt er vaker de regelterminologie aangetroffen.

de sonoriteitsopbouw van de syllabe. Later zijn er ook formuleringen in een autosegmenteel kader te vinden (Hinskens 1992 en Guy 1997).

Kenmerkovereenkomst en verschil worden door Guy en Boberg (1997) samengevat onder het hoofd van het Obligatory Contour Principle waarbij gelijke specificaties voor kenmerken binnen een consonantcluster t-deletie in meerdere of mindere mate bevorderen: zo wordt verantwoord dat die deletie bij voorkeur na niet-sonoranten voorkomt en meer specifiek na coronale consonanten. De variabiliteit in de aangetroffen specifieke OCP-effecten kunnen niet in het kader van de Optimality Theory verantwoord worden (Guy 1997), ook niet in de Antilla-versie of de Nagy-Reynolds-versie omdat de meervoudige schending voor hetzelfde kenmerk -die tot een hogere deletiefrequentie leidt- niet gemodelleerd kan worden.

In hoofdstuk 11 wordt aangetoond dat Guy's formulering in termen van het model van de lexicale fonologie niet voldoet voor de data uit Nederlandse dialecten. Ook Kiparsky's herformulering in termen van OT voldoet niet.

Boven het niveau van het segment speelt de woord(stam)structuur een rol en dan met name de syllabestructuur en de voetstructuur. De syllabestructuur bepaalt de woordstructuur doordat er duidelijke voorkeuren in syllabebouw bestaan. Dit betreft de minst gemarkeerde vorm (CV), de regeling van toegestane of geprefereerde slotconsonantclusters, de preferentie voor lichte of zware syllaben (CVC(C), CVV(C)) en de rol van de voetstructuur, dat wil zeggen de toegestane typen in de opeenvolging van een zware plus een of meer lichte syllaben. Zo hadden woorden met een flectie-t vroeger twee syllaben die samen een voet vormden: *breket* 3sing pres 'breekt': CCVC+VC. De structuur van suffixen, meestal VC, wijkt duidelijk af van de geprefereerde syllabestructuur CV, ze zijn meestal onbeklemtoond.<sup>21</sup>

De neerslag daarvan is ook te vinden in studies over Open Syllabevelenging (OSL) (Ritt 1994, Kusmenko 1995, Riad 1995) die zich in het woord binnen het domein van de voet. Zo had bijvoorbeeld *breekt* in zijn vroegere vorm *breket* een korte vocaal die door OSL wordt gerek: ...VCVC > ...VVCVC.

Deze studie over t-deletie stelt zich niet ten doel om een zelfstandige bijdrage aan dit theoretisch debat te geven. Er wordt in hoofdstuk 10 wel met woordstructuurtemplatens gewerkt (VC, VVC, VCC) maar die hebben hier een pre-theoretisch karakter. Ze zijn echter heel wel in een moren-theoretisch kader met lichte en zware lettergrepen te interpreteren en onder te brengen in een theorie die gebruik maakt van voetstructuur.

#### 1.4.2. *De modellering van variatie*

Deze studie gaat uit van een variationeel kader. Dat betekent niet dat hier variatie door middel van frequentieverhoudingen in de regelstructuur wordt "ingebouwd". Over het nut van dit inbouwen, waarbij aan taalregels en subregels frequenties "hangen", neem ik een agnostisch standpunt in. De variatielinguïstiek is er mijns

<sup>21</sup> In het Westgermaans dragen ze in 23sing aanvankelijk wel klemtoon (Boutkan 1995).

inziens meer mee gebaat als het normale “handwerk” beoefend wordt: uitgaan van, meestal partiële, theorieën over een verschijnsel en kijken welk deel van de variatie systematisch is en welk deel van de variatie verstoringen zijn om te zien waar er gegeneraliseerd kan worden en of een hypothese weerlegd of, voorlopig, gehandhaafd kan worden. We zijn voornamelijk geïnteresseerd in de taalstructuur en in de geografische structuur, dus in de systematiek in de variatie. De hoofdscheiding die we aanbrengen is dus in de systematiek tegenover andere variatiebronnen die beschouwd kunnen worden als verstoringen.

Geobserveerde waarden (data) = systematische ‘echte’ waarden + verstoringen

De component *systematische waarde* kan bestaan uit een combinatie van meerdere relevante variabelen en hun waarden.

De component *verstoringen* kan toch nog een zekere systematiek vertonen:

- a) verstoring door de verkeerde variabele of de verkeerde proefpersonen voor het gestelde doel die de validiteit van de gegevens beïnvloeden (dit komt aan de orde in hoofdstuk 2);
- b) verstoring door systematische vertekening in een bepaalde richting die de betrouwbaarheid van de gegevens nadelig beïnvloedt (dit komt aan de orde in hoofdstuk 3);
- c) verstoring door niet opgenomen relevante variabelen die alsnog tot meer systematiek zouden leiden (dit is aan de orde in hoofdstuk 3-5);
- d) verstoring door systematische onderlinge samenhang tussen aparte onderzoekseenheden; naburige onderzoekseenheden vertonen dan stelselmatig gelijksoortige scores in positieve of negatieve richting. Dat is bijvoorbeeld het geval in de regionale samenhang tussen gegevens die leidt tot wat we ‘kaartbeeld’ noemen (dit is aan de orde in hoofdstuk 3 en 5, en in de gevolgde procedures voor de automatisch getekende dialectkaarten).

De fouten door *a* en *b*<sup>22</sup> moeten zo klein mogelijk gehouden worden, *c* kan ondervangen worden door de juiste selectiviteit toe te passen met betrekking tot de variabelen; een deel daarvan kan als ze bij statistische toetsing irrelevant blijken verder veronachtzaamd worden, en voor *d* zijn specifieke technieken ontwikkeld (rekenschap geven van autocorrelatie).

---

<sup>22</sup> In de klassieke testtheorie geldt dat de betrouwbaarheid zijn validiteit is ten opzichte van een gelijksoortige test. De validiteit is dan de correlatie van de debetreffende test met de gelijksoortige maar die validiteit kan nooit groter zijn dan de berekende betrouwbaarheid. De klassieke testtheorie gaat het meest op voor experimentele onderzoeksopzetten maar niet voor survey onderzoek of veldonderzoek. De cruciale vooronderstelling in de klassieke testtheorie is dat alle meetfouten toevallig zijn. In survey- en veldonderzoek is deze vooronderstelling niet te handhaven (Zeller en Carmines 1980). Doordat de situatie minder manipuleerbaar is dan in een experiment kunnen er systematische afwijkingen optreden. De verklaarde variantie is in survey-onderzoek dus meestal lager dan in experimenteel onderzoek.

We gebruiken in hoofdstuk 2 en 3 de metafoor van het meten met een thermometer om validiteit en betrouwbaarheid te illustreren: een groot deel van wat ons in dit boek bezighoudt is inderdaad meten: we verbinden abstracte theoretische begrippen als bijvoorbeeld [sonorant] met door ons aangebrachte groeperingen in de gegevens die dan fungeren als empirische indicatoren voor die abstracte begrippen. We proberen door de foutenbron (de niet verklaarde variatie) zo klein mogelijk te houden hun relatieve invloed zo zuiver mogelijk te schatten, zodat we uitspraken over de theoretische begrippen kunnen doen.

Voor de rol van variatie en informatieve meerwaarde van kwantitatieve gegevens verwijs ik naar de inleidingen van hoofdstuk 7 en 8. De kruisklassificatie van deterministische en kwantitatieve onderzoekstradities die in die hoofdstukken opgesteld is, werd later, ook toegepast door Guy (1994).

	Modeleigenschappen	
	kwalitatief	kwantitatief
deterministisch	A	B
probabilistisch	C	D

Deterministische modellen postuleren een ideale wereld waarin alle geconstateerde variatie veroorzaakt wordt door variatie in informanten/proefpersonen en variatie in stimuli/variabelen. Er bestaat geen onsystematische variatie of verstoring.<sup>23</sup> Dat betekent dat iedere schending tot verwerping van het model zou moeten leiden, maar in de praktijk gebeurt dat niet; deze modellen beantwoorden niet de vraag of ze bij de gegevens passen, maar wel of ze die gegevens in ‘voldoende’ mate benaderen. Het implicatieve schaalmodel dat voor creool-continua en voor fonetische en lexicale diffusie aangewend is,<sup>24</sup> hoort tot deze categorie.

Probabilistische modellen gaan ervan uit dat de wereld minder ideaal in elkaar zit; ze geven ook rekenschap van variatie in informanten en variabelen, maar bovendien houden ze rekening met onsystematische en met onverklaarde variatie. Daardoor kunnen ze de vraag beantwoorden of het model bij de data past. Daar bestaan statistische criteria voor de “goodness of fit” tussen model en data.<sup>25</sup>

Kwalitatieve modellen zijn gebaseerd op kwalitatieve data: iets komt voor of niet waarbij de mate van voorkomen er niet toe doet, of iets hoort tot een bepaalde categorie of niet: bijv. *oud* - *niet oud* en dat kan ook multicategoriaal zijn: *jong* - *middelbaar* - *oud*.

<sup>23</sup> Zie Torgerson (1958); er zijn overigens ook probabilistische schalen mogelijk. Een sociolinguïstische deterministische variant is: “vrije variatie bestaat niet, variatie is altijd sociaal bepaald”.

<sup>24</sup> Cfr. Bickerton (1971; 1973; 1975), Bailey (1973), Cheng (1972) en door Chen en Wang (1975) is gebruikt. Het model is voor taalkundige doeleinden het eerst geïntroduceerd door DeCamp (1971b: noot 1). De schaalmethode is door Van Hout (1980) toegepast op t-deletie data uit het Rivierengebied.

<sup>25</sup> Baayen (1998) poneert het IJsbervmoeden: “Deterministische regels en intuïties zijn slechts het topje van een ijsberg van merendeels probabilistische regels en intuïties”.



Kwantitatieve modellen zijn gebaseerd op kwantitatieve data: een verschijnsel heeft een aantal voorkomens, of de categorie heeft een (semi)continue karakter: bijv. leeftijd.

Kwantitatieve gegevens zijn altijd te dichotomiseren en worden dan kwalitatief van aard: leeftijdsindeling in de klassen *jong-oud*; aantal voorkomens van 1-35 > ja (of 1), niet voorkomen > nee (of 0). Dit leidt wel tot informatieverlies. Omgekeerd is onder bepaalde assumpties een kwalitatieve indeling te waarderen als een kwantitatieve. Als een categorie uit deelcategorieën bestaat en er is een stricte rangorde tussen die deelcategorieën en als aan die rangorde een continue schaal ten gronslag kan liggen dan kan de rangorde als een kwantitatieve schaal gezien worden: bijvoorbeeld *leeftijd* is gecodeerd als kind/adolescent - jongvolwassene - middelbaar - oud levert een schaal van 1 - 4; zeker als deze vier klassen een ongeveer gelijk bereik in jaren hebben.

Alle multivariate statistische technieken als Variantieanalyse en Regressie horen thuis in kwadrant D.

In deze taxonomie van modellen kunnen ook een aantal taalkundige theorieën of modellen ondergebracht worden. Voor ieder kwadrant geven we een voorbeeld. Er is al gewezen op het feit dat de implicatieve analyse in kwadrant A thuishoort.<sup>26</sup> De generatieve fonologie uit het pre-Optimality Theory (OT) tijdperk is ook een voorbeeld van A: ze is kwalitatief/categoriaal en deterministisch; maar ook Kiparsky's opvatting van OT hoort daartoe. Het OT-model van Antilla (1995) is een voorbeeld van B (kwantitatief en deterministisch).<sup>27</sup> Het variabele regelmodel dat ontwikkeld is door Labov, en later door Cedergren en Sankoff (1974) verfijnd, is een voorbeeld van C: kwalitatief/categorisch<sup>28</sup> en probabilistisch.

Modellen die een probabilistische implicatieve schaal gebruiken vallen ook in C: ze zijn kwalitatief/categorisch, maar met een probabilistische component<sup>29</sup>

<sup>26</sup> Deze deterministische scalogramanalyse werd speciaal ontwikkeld met het oog op kwalitatieve data door Guttman (1944).

<sup>27</sup> OT is een model met gereduceerde informatie ten opzichte van het oorspronkelijke model waar het uit afgeleid en dat behoort tot de z.g. neurologische netwerken. Het oorspronkelijke model hoort tot kwadrant D. De kwantitatieve informatie uit het oorspronkelijke model is in OT gereduceerd tot informatie over de rangorde van constraints. Het oorspronkelijke model kent daarentegen numerieke constraints, zie Smolensky 1986a,b). Naar hun belangrijkheid hebben constraints een bepaald gewicht. Een belangrijke constraint heeft een zwaar gewicht, minder belangrijke constraints hebben lagere gewichten. De constraints zijn niet 'hard' in de zin dat er aan voldaan moet worden, maar het zijn 'zachte' constraints waaraan in zekere mate, d.w.z. op de best mogelijke wijze voldaan wordt: het doel is om een oplossing te vinden waarin aan de belangrijke constraints zoveel als mogelijk is voldaan (McClelland en Rumelhart 1989: 50). Er bestaan overigens ook deterministische connectionistische modellen of neurale netwerken. Smolensky is ook betrokken bij de ontwikkeling van OT.

<sup>28</sup> Het programma kent voor de afhankelijke variabele twee (binomiaal) of meer (multinomiaal) categorieën. De Goldvarb-versie kent alleen de binomiale mogelijkheid.

<sup>29</sup> Over de verhouding variabele regelmodellen en implicatieve analyse als meetmodel zie Goeman (1983a). Rousseau en Sankoff (1978) hebben het variabele regelmodel met het categoriale implicatieve model verbonden. Dat gebeurt doordat afzonderlijke sprekers aan optimaal

en ook de Correspondentie-Analyse van Benzécri (1973a,b) voor categorische variabelen hoort hiertoe.<sup>30</sup> In de sociolinguïstiek overheerst het variabele regelmodel, maar daarnaast wordt ook het logit-regressiemodel gebruikt, bijvoorbeeld door Van Hout (1989a, b). Voor zowel het variabele regelmodel als voor het logit-model valt echter ook te verdedigen dat ze in kwadrant D horen; de nadruk ligt dan op de onafhankelijke variabelen die als kwantitatief karakter hebben: aantallen voorkomens. Logit-model en variabele regelmodel zijn beide varianten van de klasse van modellen waar de variabele een logistische transformatie heeft ondergaan (een vergelijking van beide vindt plaats in hoofdstuk 11). Voorbeelden van dergelijke logmodellen komen verderop in dit boek voor.

Het variabele regel-model is dus niet de enige manier die men onder D kan tegenkomen. Het meeste van wat in dit boek aan de orde is hoort eveneens onder D thuis, zonder zich aan het variabele regel-model te committeren.

Deze studie naar t-deletie is meer toetsend dan exploratief. Maar met nadruk moet gezegd worden dat de studie geen experimentele opzet heeft. Over het algemeen zal men dus geen specifieke variantie-analyses aantreffen. Wel is er gebruik gemaakt van technieken die toestaan om voor de in het materiaal aanwezige talige variabelen, buitenlinguïstische aspecten als sociale positie, leeftijd en geslacht van de informanten te controleren.

Uitgangspunt zijn duidelijke modellen, opgesteld op basis van inhoudelijke theoretische overwegingen van taalkundige, specifiek dialectologische, aard. Uit die modelmatige aanpak vloeien voorspellingen voort die toetsbaar zijn en uit de meer of mindere frictie met de data kan meer inzicht verkregen worden over de structuur van de taalvariatie.

In Goeman (1983) signaleerde ik dat met de jaren bij de variabele regelmodellen het regelkarakter afnam waardoor de nadruk veel meer kwam te liggen op de conditionerende factoren voor een fonologisch proces. Die conditionerende factoren interpreteerde ik toen in termen van taalinterne contexten, de distributie. De ontwikkelingen van de laatste jaren in de taalkundige theorievorming, met de

---

samengestelde groepen worden worden toegedeeld. Een directe benadering kan m.i. via Correspondentie-Analyse plaats vinden; dit probabilistische model ordent namelijk net als implicatieve analyse sprekers en variabelen tegelijk, maar legt niet de procrustes-restrictie op dat er sprake moet zijn van slechts 1 schaal (dimensie) zoals bij de implicatieve schaal het geval is.

<sup>30</sup> Correspondentie-Analyse zoekt het optimale aantal dimensies waarop individuen en variabelen geordend zijn. Voorbeelden van toepassingen van Benzécri's ideeën in de dialectologie zijn Cichocki (1993) voor een groep franstalige dialecten in Canada (Acadien) en Mooijaart en Van der Heijden (1992) voor de determinatie van de regionale herkomst van onbekende Middelnederlandse handschriften, maar een beperking van dit soort modellen ligt in hun exploratief karakter. Bij Mooijaart en Van der Heijden blijft onduidelijk of de reductie tot twee dimensies (die natuurlijk gekozen zijn vanwege de interpretatie met de twee geografische coördinaten) wel de meest optimale is. De breuk in hoogte van de eigenwaarden van bijvoorbeeld de eerste 7 dimensies zou dan duidelijk bij 2 moeten liggen ('screetest'). Zo niet, dan zijn ook de verkregen schattingen niet optimaal. Bovendien is het dan mogelijk dat niet dimensie 1 en 2 de geografische coördinaten optimaal representeren, maar dat andere dimensies dat beter zouden doen.

grote nadruk op structuurconstraints en op representatieve in plaats van derivatieve modellen,<sup>31</sup> hebben laten zien dat die distributionele contexten in termen van structurele constraints geïnterpreteerd kunnen worden. Dit boek staat dus niet primair in een 'regel'-kader; de meeste nadruk ligt op de relevantie van de condities.

#### 1.4.3. *De geografische modellering van variatie*

Ook bij geografische data kunnen we een onderscheid aanbrengen tussen het systematische en toevallige verstoringen. Hiervoor is reeds gewezen op het feit dat we, als we een kaart zien, onmiddellijk de gebieden waar een verschijnsel voorkomt proberen te herkennen. Deze gebiedsvorming brengt evenwel bepaalde problemen met zich mee voor een statistische evaluatie van geografische gegevens.<sup>32</sup> De extra samenhang binnen bepaalde gebieden in de geografische ruimte betekent dat gegevens voor een bepaalde plaats samenhangen met gegevens uit plaatsen in de buurt: autocorrelatie. Dit probleem komt aan de orde in hoofdstuk 3-5. Er zijn oplossingen voor dit probleem en die zijn toegepast voor de kaarten in hoofdstuk 7, 8, 10 en 11.

De geografische kaarten in deze studie bestaan grofweg uit 5 typen:

- a) eenvoudige distributiekaarten zoals in hoofdstuk 9: ze geven een verschijnsel in zijn categoriale vorm weer: een vormvariant komt voor of komt niet voor. Het kaartbeeld ontstaat doordat gelijksoortige vormvarianten meestal in elkaars nabijheid voorkomen.
- b) absolute frequentiekaarten, zoals de meerderheid van de kaarten in hoofdstuk 6: de relatieve frequentie wordt daar als absoluut getal weergegeven. Het kaartbeeld ontstaat doordat gelijksoortige aantallen zich vaak in elkaars buurt voordoen: de zwartste tekens komen in elkaars buurt voor, de witste tekens onderling ook.
- c) frequentiekaarten waar gecorrigeerd is voor bepaalde effecten, die het effect waar men belangstelling voor heeft zouden kunnen verstoren of beïnvloeden: een voorbeeld is kaart 11 in hoofdstuk 11. Het regionale effect is daar 'opgeschoond' van de effecten van taalstructurele aard. Dit kan gemakkelijk als het aantal regionale gebieden/plaatsen niet al te groot is, ze kunnen dan als waarden van de geografische variabele regio (of plaats) in een loglineair model opgenomen worden.
- d) gegeneraliseerde frequentiekaarten zoals in hoofdstuk 7, 8, en 10: de frequenties zijn daar in frequentieklassen ingedeeld. Zo op het oog lijken die frequentieklassen het karakteristieke van die kaarten te zijn, maar dat is niet zo: de kaarten bevatten zowel geïnterpoleerde als geëxtrapoleerde waarden. Er wor-

<sup>31</sup> Cumulerend in Optimality Theory.

<sup>32</sup> De technische term hiervoor is dat geografische variabelen 'geregionaliseerde' variabelen zijn, dit in tegenstelling tot gewone variabelen die 'random' variabelen zijn.

den voorspellingen gedaan voor plaatsen zonder gegevens, op basis van bekende waarden in de omgeving. Ook de hoofdstukken 4 en 5 bevatten kaarten van de voorspelde waarden waar een klassenindeling overheen is gelegd, maar de schatting van de voorspelde waarde is een polynomische functie (een vergelijking met kwadratische, cubische en eventuele hogere machtsvormen). Polynomische functies zijn vooral geschikt als het aantal geografische datapunten niet al te groot is. Een bijkomend voordeel ervan is dat ze als multiële regressiemodellen geformuleerd kunnen worden, want de statistische eigenschappen van deze modellen zijn zeer goed bekend.

Het specifieke van deze laatste groep kaarten is de conceptie die er aan ten grondslag ligt, namelijk dat de structuur van de gegevens niet gelegen is in de absolute frequenties van de data die een relatief onherkenbaar beeld van afzonderlijke pieken en dalen zou vertonen. Dat is eigenlijk ook al enigszins een probleem bij de kaarten die in de hiervoor genoemde categorie b vallen. De echte structuur van de gegevens komt beter tot uitdrukking als men rekening houdt met het feit dat de data verstoringen, error, bevatten die geen deel uitmaken van de eigenlijke structuur.

Uit de data voor de kaarten uit hoofdstuk 7, 8 en 10 worden verstoringen weggefilterd door een bewerking, waardoor de werkelijke geografische structuur van de data tot zijn recht komt en die functie houdt rekening met het feit dat de plaatsen niet alle even dicht over het onderzoeksgebied gespreid zijn.

Die bewerking bestaat uit 2 gedeelten (zie Wattel 1990; Wattel en Van Reenen 1995).<sup>33</sup> Het eerste deel betreft de invloeden van plaatsen op elkaar waarbij voor elke afzonderlijke plaats er een effect bestaat vanuit die plaatsen die binnen een bepaalde kring met betrekking tot die plaats liggen. Plaatsen die buiten die kring vallen hebben geen invloed meer. Het tweede deel geeft een correctie zodanig dat met toenemende afstand het effect afneemt; op die manier worden kleine effecten van plaatsen dichtbij niet weggedrukt. Duidelijke lokale effecten mogen niet geheel "verdrinken". De zo gevonden "opgeschoonde", gefilterde waarden kunnen in frequentieklassen worden ingedeeld en via een algoritme kan men op een niet-impressionistische manier isoglossen (in feite: isoplethen) construeren tussen de plaatsen.<sup>34</sup>

Overigens kan ook voor categoriale data (wel-niet voorkomend) een isoplethenkaart geschat worden. De frequentieklassen geven dan de kans aan dat een bepaalde variant voorkomt en een lijncontour van de drempelwaarde kan de scheiding tussen wel en niet voorkomen weergeven (Van Reenen en Wattel 1995).

---

<sup>33</sup> Van Reenen en Wattel (1995) geven een wat andere presentatie van de gevolgde procedures.

<sup>34</sup> De techniek is algemeen en kan ook gebruikt worden voor andere drie-dimensionele problemen, zoals (Wattel 1989): hoe interpoleer ik voor niet opgegraven gedeelten bij een opgraving geschatte waarden vanuit de gevonden hoogtewaarden en locatiewaarden van het gedeelte dat wel kon/mocht worden opgegraven, omdat het terrein gedeeltelijk is verstoord of omdat een gedeelte onopgegraven voor toekomstige generaties moet blijven zitten.

- e) de geschatte waarden -ingedeeld in frequentieklassen- worden verkregen door Kriging. Kriging houdt rekening met het feit dat de geografische variabele niet helemaal een random variabele is, maar geregionaliseerd is, dat wil zeggen dat de afzonderlijke waarden nauw kunnen samenhangen met die van andere plaatsen in de buurt (vg. een voorgaande noot). Kriging is geen extrapolatietechniek, maar een geografische *interpolatietechniek* voor de schatting<sup>35</sup> van plaatswaarden van plaatsen die niet in de steekproef zitten op basis van de wel bekende waarden van plaatsen uit de steekproef. Kriging optimaliseert de uit de data te schatten waarden<sup>36</sup> via een afstandsbepaling die men vooraf door variogramschatting moet verwerven. Deze afstandsbepaling heeft de vorm van een functie die het gedrag van plaatsen beschrijft die elkaars scores negatief of positief beïnvloeden (autocorrelatie). De variogramschatting geeft bovendien informatie over de invloed van scherpe overgangen, breuken, in het geografische patroon en over de invloed van de richting van die breuken.

De grote hoeveelheid informatie over de geografische distributie die bij Kriging verdisconteerd is heeft tot gevolg dat Kriging exacte schattingen geeft: voor de datapunten uit de steekproef wordt als geschatte waarde de werkelijk aangetroffen waarde voor dat punt gegeven, of zeer nauw benaderd. Bij andere methoden, zoals hierboven aan de orde geweest, zullen in het algemeen de voorspelde waarden een grotere afwijking van de datawaarden vertonen. Een nadeel van Kriging is dat op die manier het mogelijke effect van andere verklarende variabelen weliswaar meegenomen is, maar niet meer bestudeerd kan worden.<sup>37</sup> Normaal zou het effect van niet opgenomen variabelen in de error-component tot uitdrukking komen, zoals we in de vorige sectie hebben betoogd, maar die error-component, de foutenbron, is in deze techniek extreem geminimaliseerd. Dat wil zeggen dat

---

<sup>35</sup> Zie voor de procedures Deutsch en Journel (1998); meer inleidend is, met Nederlands illustratiemateriaal, Burrough en McDonnell (1998), waar de relatieve voor- en nadelen van (inverse) afstandsmethoden als bijv. door Wattel gebruikt en geostatistische methoden als Kriging nader worden besproken.

<sup>36</sup> Dus niet via een kleinste kwadratenmethode (OLS = Ordinary Least Squares).

<sup>37</sup> Voor exploratieve voorspellingen is dat juist irrelevant. Kriging is oorspronkelijk ontwikkeld voor de exploratie van bepaalde gesteenten of erts in de mijnbouw (Krige, 1951) en is vooral door Franse geostatistici verder ontwikkeld. Er bestaat in de mijnbouw, gezien de exploratiekosten, behoefte aan een zeer nauwkeurige voorspelling in welk gebied gesteente of erts aangetroffen zal worden met de daarbij behorende waarschijnlijkheidsgrenzen voor de geschatte waarden. De factoren die tot het ontstaan van erts hebben geleid zijn voor exploratieve doeleinden niet direct relevant, het gaat alleen om de preciese localisatie. In tegenstelling tot wat verwacht zou kunnen worden op grond van de naam is Co-Kriging geen methode om covariatie te schatten. Integendeel, er worden uit covariërende variabelen waarden voor de afhankelijke variabele afgeleid voor die plaatsen waar wel data voor de onafhankelijke variabelen voorhanden zijn, maar geen data voor de afhankelijke variabele zelf. Het is dus een vorm van missing data imputation. Een praktisch nadeel van Kriging zou kunnen bestaan uit het feit dat de procedure de data aan de rand van het onderzoeksgebied verbruikt om tot nauwkeurige schattingen aan de periferie te kunnen komen. Dat is ook aan kaart 5 in hoofdstuk 10 te zien: de geschatte waarden lopen niet tot de landsgrenzen.

het effect van alle niet-geografische variabelen grotendeels in de geschatte waarden zitten verwerkt. Om dat te ondervangen zou men eerst een regressieanalyse met relevante niet-geografische variabelen uit moeten voeren en op de residuen, die normaliter als error worden beschouwd, Kriging toepassen. Dit lost het probleem maar ten dele op, want ook deze schattingen zullen niet geheel juist zijn, want de meest preciese schattingen krijgt men als alle relevante variabelen tegelijk in het model zijn opgenomen (vgl. hoofdstuk 4). Een voorbeeld van Kriging is kaart 5 hoofdstuk 10. De schatting is exact, maar omvat tevens het effect van eventuele niet geziene andere variabelen als bijvoorbeeld de invloed van sociale factoren en van specifieke taalkundige categorieën. Zie voor een techniek waar dat wel kan, maar die weer zijn eigen specifieke nadeel heeft, punt c. hiervoor.

Voor de kaarten in hoofdstuk 7, 8 en 10 is op een onderdeel een waarschuwing op zijn plaats.<sup>38</sup> Gebieden die vrijwel leeg zijn op de kaart zullen zeer onzuivere *extrapolaties* vertonen in de voorspelde waarden. Dat geldt temeer als er in werkelijkheid een andere variant voorkomt. Zo moeten de voorspelde waarden voor Friesland niet gebruikt worden. De waarden zijn er de resultante van de waarden in noordelijk Noord-Holland en Groningen, Drente en de Stellingwerven, en een beetje in toom gehouden door de ene plaats in de database op een van de eilanden. In feite doen zich in Friesland geen eenheidspluralia op -n voor in het presens, maar op -e. De argumentering in deze studie hangt daar niet op. Friesland zelf is geen voorwerp van onderzoek. Wel worden te bestemder plaatse facetten van het Fries te berde gebracht, maar die steunen op bestaande literatuur.

#### 1.4.4. *De diachrone modellering van variatie*

Voor de in deze studie geanalyseerde tijdreeksen zijn drie technieken gebruikt. Multipiele regressie, splines en loglineaire analyse.

In hoofdstuk 5 worden twee tijdsdoorsneden vergeleken door multipiele regressiemodellen. Dat gebeurt op twee manieren. Ten eerste worden de data van 1950 als variabele in de multipiele regressie van 1982 meegenomen. Vervolgens worden 3 multipiele regressiemodellen met elkaar vergeleken: die voor 1950, die voor 1982 en die voor de verandering van 1950 naar 1982.

In hoofdstuk 8 wordt modellering door splines gebruikt voor goed gedateerde, telkens apart ingeschreven fragmenten, die niet noodzakelijkerwijs in hun chronologische volgorde in het manuscript staan. De procedure wordt uitgelegd in de appendix bij dat hoofdstuk (Anhang 1). We volstaan hier met een concreet voorbeeld. Zo kan bijvoorbeeld 13 febr. 1434 op bladzijde 12 van een manuscript komen, en 13 mei 1937 op bladzijde 10 als daar nog plaats was. Dit keer wordt de modellering van de tijdsvolgorde op grond van de paginering niet gebruikt om de systematiek te scheiden van de vertekening, maar het gaat er in dit geval om het totale beeld, inclusief de verstoring (in dit geval vertoring van het juiste

---

<sup>38</sup> Zie ook de betreffende noot in de secties over de kartografie in hoofdstuk 7 en 8.

chronologische beeld door bladzijdegebruik), te modelleren. Bij de plaatsing van de ongedateerde fragmenten heeft dezelfde systematiek (aanhouden van de chronologie) mét dezelfde storingsbron (plaatsing op voorhanden lege bladzijde-(gedeelten)) een rol gespeeld als bij de plaatsing van de wel gedateerde fragmenten. Als de chronologische tijdsas tegen de bladzijdevolgorde afgezet wordt ontstaat er zeker geen rechte lijn: de lijn vertoont veel op en neergang op bepaalde trajecten. Liniare interpolatie, door een strakke lijn is dus niet juist, maar splines kunnen de 'brokkeligheid' goed modelleren. De gevonden 'brokkelige' curve wordt gebruikt om de fragmenten zonder datering op de curve te interpoleren op basis van hun bladzijdevoorkomen en zo hun meest waarschijnlijke plaats in de chronologie vast te stellen. De totale reeks wordt vervolgens gebruikt in de analyse.

In hoofdstuk 2 wordt een logistisch loglineair model, dat bekend staat als het variabele regelmodel, gebruikt buiten zijn regelcontext, omdat het een toepassing is van een 'normale' statistische techniek; in hoofdstuk 11 wordt een vergelijking van dit specifieke model met meer algemene modellen gemaakt.

### 1.5. Indeling van deze studie

In verband met het onderscheid dat gemaakt wordt in deze studie tussen systematische patronen en verstoringen in de data gaan we eerst na welke factoren een storende rol zouden kunnen spelen.

Meetfouten vormen een van de storingsbronnen die de werkelijke structuur van de data verhullen. Daarom wordt vooraf de vraag aan de orde gesteld in welke mate eventuele meetfouten het beeld vertekenen. Die vraag wordt gesplitst in de vraag naar de validiteit en vraag naar de betrouwbaarheid van de gegevens.

In hoofdstuk 2 komt de validiteit van de dialectdata aan de orde oftewel de vraag of de data representeren wat we denken dat ze doen: namelijk dialect.

In hoofdstuk 3 komt de betrouwbaarheid van de data aan de orde: hoe nauwkeurig zijn deze data.

In hoofdstuk 4 bekijken we in welke mate de data dialect als geografisch fenomeen representeren en welke rol andere, waaronder sociale factoren, die eigen zijn aan de data, eventueel een rol meespelen. We zullen zien dat de geografische variatie voor t-deletie het meest saillant is, vergeleken met de niet-significante of zwakkere impact van sociale variabelen. Vandaar dat de aandacht in de overige hoofdstukken voornamelijk aan de niet-sociale variabelen gewijd is. Ook in de hoofdstuk 3 wordt de betrouwbaarheidsvraag toegespitst op t-deletie.

In hoofdstuk 5 wordt bekeken of en in hoeverre t-deletie in werkelijke tijd ontwikkelingen vertoont; een van de conclusies is dat de invloed van de standaardtaal het geografisch patroon van meer of minder t-deletie niet kan verklaren.

Vervolgens komen in hoofdstuk 6 de intertaalige factoren aan de orde die een rol bij t-deletie spelen. Gewezen wordt op de hiërarchie in deletie-frequentie van morfologische woordklassen. Er wordt uitgegaan van een paar generalisaties van de taalinterne constraints: conceptuele en fonische prominentie, conceptuele en articulatorische complexiteit en de rol van woordfrequentie. Mogelijke onder-

linge systematische verbanden komen in tweezijdige, bivariate analyses aan de orde. Tot en met dit hoofdstuk zijn de data allemaal afkomstig uit Zuid-Holland en het Rivierengebied, te weten zuidelijk Utrecht en de Betuwe. De hoge waarden voor t-deletie in het laatstgenoemde gebied zijn taalextern deels te verklaren vanuit de contactsituatie. Het gebied vormt een, inmiddels passieve, contactzone tussen Utrechts en Brabants en tussen Hollands en Gelders. Dit hoofdstuk bevat waarschijnlijk de eerste poging in de literatuur om dialectgeografische variatie via een eenvoudig lineair model te verklaren. Verschillen tussen de geografische resultaten van dit hoofdstuk en van eerdere hoofdstukken zijn toe te schrijven aan het feit dat de geografische coördinaten een oplopende precisie vertonen (oplopend van hoofdstuk 3 en hoofdstuk 6 via hoofdstuk 4 naar hoofdstuk 5 en latere hoofdstukken). Daardoor neemt de error af. Hier is dus sprake van een duidelijke relatieve reductie van de meetfout. Een van de taalinterne bevindingen van hoofdstuk 6 is dat t-deletie in de Nederlandse dialecten het meest frequent bij werkwoordsuffixen voorkomt. Deze situatie wordt ook door andere studies geconstateerd, die voor, gelijktijdig met, of na wat hier hoofdstuk 6 is, zijn verschenen.

Om die reden wordt er in hoofdstuk 7 tot en met 9 speciaal aandacht geschonken aan de inbedding in de werkwoordsmorfologie en dus aan het gehele presensparadigma.

Hoofdstuk 7 geeft de achtergrondkennis voor de rol van t-deletie bij werkwoorden doordat daarin een totaaloverzicht van de suffixen van de presensvervoeging wordt gegeven. In een aantal dialecten vervult het t-morfeem meer functies dan in de standaardtaal. Het t-suffix vervult verscheidene functies: het is niet alleen 3sing markering, maar in sommige dialecten ook 2sing en/of 2plur en in de Noordoostelijke dialecten zelfs het complete meervoud: eenheidspluralis op -t. T-deletie in 123plur is maar zeer beperkt voorhanden en het ontbreken van t-deletie kan fonologisch verklaard worden (zie hoofdstuk 10). Als t-deletie wel optreedt kan dat schijnbaar zijn. De zero-vormen in 123plur kunnen ook het resultaat zijn van suffixconcurrentie tussen -t en -n.

Hoofdstuk 8 werkt daarom een onderdeel van het voorafgaande hoofdstuk over het presensparadigma verder uit: het gaat hier over de ontwikkelingen over een zeer lange tijd bij de eenheidspluralis op -t in het Noordoosten van het taalgebied. Alle personen van de pluralis vertonen hier een t-suffix, maar dat suffix ondervindt concurrentie van andere vormen, met name van die op -n. Hoofdstuk 5 was ook al gewijd aan veranderingen van t-deletie in werkelijke tijd, maar daar betrof het deletie in zijn totaliteit, en de tijdsspanne is aanmerkelijk korter. In een naschrift wordt een vergelijking gemaakt met Middelenegse suffixconcurrentie omdat suffixconcurrentie tot zero-vormen kan leiden die dus niet gezien moeten worden als t-deletie.

Mede als vervolg op een sectie in hoofdstuk 6 waarin is aangetoond dat de hogere frequentie van t-deletie rechtstreeks samenhangt met het feit of er in inversie een nulsuffix alterneert, is hoofdstuk 9 gewijd aan de suffixen van de 2e persoon enkelvoud in inversie: *jij loopt-loop jij*, waarvan aangetoond wordt dat het ontbrekende suffix (zero-suffix) in een groot aantal gevallen niet op t-deletie



terugvoerbaar is: de mogelijke processen worden gedemonstreerd aan het monosyllabische werkwoord *doen*: *jij doet-doe jij*.

Tegen de achtergrond van de drie voorafgaande hoofdstukken komen vervolgens in hoofdstuk 10 de t-loze 3sing vormen bij sterke werkwoorden in de Noordoostelijke dialecten aan de orde in interactie met de syncope van de oorspronkelijke vocaal van het suffix en in interactie met Open Syllabe verLenging.

Verklaard wordt waarom de eenheidspluralis op -t geen deletie vertoont. Voor de deletie van het t-suffix in 3sing worden factoren als wortel- en lettergreepstructuur, sonoriteit van de stamconsonant en verkorting of verlenging van de stamvocaal alsmede woordfrequentie gescreend op hun relatieve invloed, en wordt het optreden in een beperkte klasse van woordvormen verklaard. Het ontstaan van deze constellatie is het gevolg van diachroon aanwijsbare toevalligheden. Hoewel universele natuurlijke ontwikkelingen wel een rol spelen is het uiteindelijke hedendaagse synchrone patroon niet het effect van monocausale, natuurlijke processen. De processen zijn natuurlijk, maar hun gezamenlijk effect niet.

Was hoofdstuk 6 voornamelijk bivariaat van aard, hoofdstuk 11 is evenals het vorige sterk multivariaat van aard. Hoofdstuk 11 neemt onderdelen van zowel hoofdstuk 6 als het voorafgaande hoofdstuk 10 op. De effecten voor taalstructurele beperkingen op t-deletie hebben niet het gepostuleerde karakter van constanten. T-deletie in het Rivierengebied en in Zuid-Holland verschillen in algemene zin van elkaar, maar ook het gewicht van de dooslaggevende condities verschilt in die twee dialectgroepen. De noordoostelijke t-deletie wijkt zeer sterk af. Het proces van t-deletie in het 13e eeuwse Westvlaams verschilt zowel van het contemporaine Westvlaams als van het 13e eeuwse Oudfrans. Tenslotte komt een model van het lexicon uit de Lexicale Fonologie aan de orde met een verschillende status voor wordeinde-t naar niveau en met een gradueel verschillende regelmatigheid. Dit model zou een verklaring kunnen bieden voor de morfologische conditionering van het deletie-proces. De t-deletiedata evenwel van geen der onderzochte dialecten bieden evidentie voor zo'n model.

Hoofdstuk 12 vat de conclusies van de voorafgaande hoofdstukken op zo'n manier samen dat het kan volstaan voor wie verder geen kennis wil nemen van de afzonderlijke hoofdstukken. Het perkt bovendien aspecten voor nader onderzoek af. Leest men daarbij dit inleidende hoofdstuk, dan heeft men ook een beeld van de onderzoeksvragen die in de literatuur aan de orde zijn of geweest zijn.

### 1.6. Verantwoording van eerder gepubliceerd materiaal

Een aantal hoofdstukken of hoofdstukdelen hebben eerder als afzonderlijke publicatie het licht gezien,<sup>39</sup> of zijn sterk bewerkt. Deze worden hierna genoemd.

<sup>39</sup> Ik dank de uitgevers van de tijdschriften *Taal en Tongvalen Jahrbuch des Vereins für niederdeutsche Sprachforschung*, en ook de uitgevers ICG Publications en John Benjamins Publishing Co voor hun toestemming om de betreffende artikelen hier als hoofdstuk op te nemen.

- Hoofdstuk 3 sectie 3.2-3.3 is een bewerking van: Goeman, A. (1986). Alleged unreliability of European continental dialect geography. A statistical appraisal. *Sprachnormen: lösbar und unlösbar Probleme/Kontroversen um die neuere deutsche Sprachgeschichte/Dialektologie und Soziolinguistik: Die Kontroverse um die Mundartforschung [Kontroversen, alte und neue. Akten des VII. Internationalen Germanisten-Kongresses Göttingen 1985, Bd. 4]*, ed. Polenz, P., J. Erben en J. Goossens, 305-318. Tübingen, Max Niemeyer Verlag.
- Hoofdstuk 5 is eerder verschenen als: Goeman, A. (1988): Les modèles spatiaux et temporels dans la dialectologie. Comment séparer les tendances régionales et générales des fluctuations systématiques au niveau local. *Distributions spatiales et temporelles, constellations des manuscrits/Spatial and Temporal Distributions, Manuscript Constellations. Etudes de variation linguistique offertes à Anthonij Dees à l'occasion de son 60ème anniversaire/Studies in language variation offered to Anthonij Dees on the occasion of his 60th birthday*, ed. Reenen, P. en K. van Reenen-Stein, 77-92. Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Hoofdstuk 6 heeft gecirculeerd als: Goeman, A. and P. van Reenen (1985). *Word-final t-deletion in Dutch dialects: the roles of conceptual prominence, articulatory complexity, paradigmatic properties, token frequency and geographical distribution*. [Vrije Universiteit Working Papers in Linguistics 16]. Amsterdam: Vakgroep Algemene Taalwetenschap VU.
- Hoofdstuk 7 is gepubliceerd als: Goeman, A. (1992). De Presensvervoeging in de dialecten van Nederland; ja-nee versus gradueel. *Taal en Tongval* 44, 97-139.
- Hoofdstuk 8 is, afgezien van de Nachtrag, gepubliceerd als: Goeman, T. (1993). Der Einheitsplural im Präsens Indikativ der (Ost) Niederländischen Dialekte. Synchrone und diachrone Variation: rezente Dialekterhebungen und mittelalterliche Urkunden. *Niederdeutsches Jahrbuch, Jahrbuch des Vereins für niederdeutsche Sprachforschung* 116, 134-169.
- Hoofdstuk 9 is gepubliceerd als: Goeman, A. (1992a). Voorbeelden van Paardekooper z'n "vaste aansluiting": werkwoordsuffigering in enclise bij monosyllabische werkwoorden, 2e persoon enkelvoud. *De Binnenbouw van het Nederlands. Een Bundel artikelen voor P.C. Paardekooper*, ed. Bennis, H. en J.W. de Vries, 95-106. Dordrecht: ICG Publications.
- Hoofdstuk 11, sectie 11.4 en 11.5 zijn een nadere uitwerking van een squib die oorspronkelijk gepubliceerd is als: Goeman, A. (1995). Een exponentieel model met niveaus in de Lexicale Fonologie is voor t-deletie in Nederlandse dialecten niet van toepassing. *Tabu* 25, 142-147.

Het betreffende onderdeel van hoofdstuk 3 is sterk bewerkt en dat geldt ook voor het onderdeel van hoofdstuk 11. De Nachtrag van hoofdstuk 8 en de overige hoofdstukken of hoofdstukdelen verschijnen hier voor het eerst.

Voor publicatie in dit boek moesten de kaarten voor hoofdstuk 7 en hoofdstuk 8 opnieuw getekend worden. Inmiddels is de programmatuur veranderd en daardoor zijn de kaarten op onderdelen anders. In de meeste gevallen heeft dat geleid tot meer generalisatie. Enkel in de gevallen dat de maximale frequenties lager dan of gelijk aan 21% waren is de klassenindeling in 7 frequentieklassen gespreid van 0-21. Daar is dus meer detail te zien dan in de oorspronkelijke publicatie.