Frans Hinskens & Marc van Oostendorp

De palatalisering en velarisering van coronale nasaal-plosief clusters in GTR. Talige, dialectgeografische en onderzoekerseffecten

Abstract

In the late eighties and early nineties of the last century, extensive fieldwork yielded data regarding phonological and morphological variation in 613 local dialects of Dutch as spoken in the Netherlands, the northern, Dutch speaking, part of Belgium and the northwest of France. The project, which was supervised by Ton Goeman, Johan Taeldeman, and eventually Piet van Reenen, resulted in a huge database on the basis of which the dialects can be systematically compared. So far the project resulted in the three-volume *Phonological Atlas of the Dutch Dialects* (FAND) and the first volume of the *Morphological Atlas of the Dutch Dialects* (MAND).

In an earlier publication (Hinskens & Van Oostendorp 2004) we introduced a larger study into the phonological (and partly morphologically conditioned) cross-dialectal and intra-dialectal variation in etymological /nd/ and /nt/ clusters in items such as *hand*, *handel*, *kant* and *kantel*, 'hand', 'trade', 'lace', 'cant; tilt', based on the Goeman-Taeldeman-Van Reenen data among other things.

In this contribution we present a quantitative investigation of an aspect of the reliability of the Goeman-Taeldeman-Van Reenen data for the palatalisation and velarisation of Dutch /nd/ and /nt/ clusters. To this end, we performed a series of statistical analyses in order to (1) detect to which extent the various investigators involved (fieldworkers, transcribers, transcription correctors) affected the data, and (2) determine the weight of the effects excerted on the variability in the data by the investigators, dialect-geography and several internal factors. We will argue how the findings of this part of our study are relevant to both historical phonological and methodological issues.

1. Inleiding. Palatalisering en velarisering in coronale nasaal-plosief clusters[1]

Een onderwerp van veel diachronisch en dialectfonologisch onderzoek is de historische modificatie van fonologische segmenten in verschillende contexten. Zo onderzoeken Coetzee en Van Reenen (1995) de historische nasalering in het Afrikaans. De meerderheid van de commentaren in de drie delen van de FAND bespreken de diverse typen bronnen van segmentmodificatie. Hinskens & Van Oostendorp (2004) richtten zich in het bijzonder op palatalisering en velarisering in etymologische coronale nasaal-plosief clusters in Nederlandse dialecten (in items als *hand*, *handel*, *kant* en *kantel*). Voorbeelden zijn

(1) hond > hoɲt,[2] hoŋk, hoɲt, hoɲc[3]
hand > haɲt, haŋk, haɲt, haɲc

Hinskens & Van Oostendorp (2004) bespreken meerdere fonologische factoren die mogelijk van invloed kunnen zijn op deze processen, zoals de ronding en de hoogte van de voorafgaande klinker en de theoretische implicaties die eventuele empirische bevindingen zouden kunnen hebben.

Grootschalig microvariationistisch werk ontbreekt echter op dit moment.[4] We vragen ons in dit artikel af of dergelijk werk gebaseerd zou kunnen zijn op de Goeman-Taeldeman-Van Reenen (GTR) database, waarin we veel variatie in de realisatie van deze clusters terugvinden. In deze bijdrage gaan we in op de respectieve invloed van verschillende bronnen van deze variatie in de GTR-database, te weten taalinterne en dialectgeografische factoren alsmede de onderzoekers zelf. In sectie 2 presenteren we de onderzoeksvragen, in sectie 3 de methode en in sectie 4 de bevindingen. We sluiten af met een discussie van zowel fonologische als methodologische aspecten (sectie 5).

---

(1) Dank aan de gastredacteurs en aan de anonieme reviewer voor het waardevolle commentaar. Voor tekortkomingen zijn alleen de auteurs aansprakelijk.

(2) Wieringen (Daan 1939).

(3) Maasbracht (Hermans 1994).

(4) Macrovariatie in segmentmodificatie is aan de orde in Hinskens & Van de Weijer (2003), een studie die gebaseerd is op Maddiesons (1984) quotasteekproef van 317 natuurlijke talen uit de diverse taalfamilies.

2. De onderzoeksvragen

De dialecten van het Nederlands verschillen onder meer van elkaar in de wijze waarop fonologische segmenten soms (al dan niet geconditioneerd) gemodificeerd worden. Zo staan de Noordoostelijke dialecten bekend om hun aspiratie van stemloze plosieven (in de volksmond wel aangeduid als de 'natte t') aan het begin van een voet. Veel van dit type feiten heeft uiteraard zijn neerslag gevonden in het materiaal van het GTR-project.

Eerder onderzoek (Van Oostendorp, te verschijnen) heeft duidelijk gemaakt dat er zich in het GTR-materiaal ook 'transcribentenlandschappen' aftekenen, dat wil zeggen dat er verschillen tussen dialecten zijn die hoogstwaarschijnlijk moeten worden toegeschreven aan het feit dat er verschillende transcribenten aan het werk zijn geweest bij de notatie van de dialectfeiten. Van Oostendorp baseerde die bevinding op de hoeveelheden verschillende transcriptiesymbolen die door de diverse GTR-transcribenten gehanteerd zijn. De Vlaamse transcribenten gebruikten er aanzienlijk minder; hun transcripties zijn betrekkelijk fonemisch, die van de Nederlandse transcribenten zijn daarmee vergeleken 'narrow' fonetisch. Een en ander werpt een schaduw over de bevinding uit het tweede deel van Van Oostendorps studie, namelijk het feit dat, gezien de systematiek in de dentale modificatie van de uitspraak van de coronalen in de GTR-transcripties, er in dit opzicht sprake is van een rationele 'kenmerkeconomie', dat wil zeggen dat dialecten een significante neiging vertonen om 'dentalisering', àls het voorkomt, systematisch in te zetten.

In deze bijdrage onderzoeken wij het GTR-materiaal ter beantwoording van de vragen

1) of er transcribenten- en correctoreneffecten bestaan in verband met palatalisering en velarisering in coronale nasaal-plosief clusters en
2) hoe, in verband met palatalisering en velarisering in coronale nasaal-plosief clusters, het effect van de transcribenten en correctoren zich verhoudt tot enerzijds dialectgeografische en anderzijds de in ons eerdere werk gedetermineerde taalinterne factoren.

We onderzoeken deze vragen aan de hand van de data voor ongelede woorden met een etymologisch /nd/ of /nt/ cluster in de GTR-databank.

3. De werkwijze

In deze sectie bespreken wij achtereenvolgens het materiaal dat de empirische basis vormt van ons onderzoek (par. 3.1) en de diverse stappen in de analyse van dat materiaal (par. 3.2). In sectie 4 gaan wij vervolgens uitgebreid in op de bevindingen.

3.1 Materiaal

De Goeman-Taeldeman-Van Reenen (GTR) databank bevat ongeveer 1,1 miljoen gegevens, dat wil zeggen de antwoorden op 1876 vragen zoals deze in 613 plaatsen in Nederland, Vlaanderen en Frans-Vlaanderen tussen ongeveer 1980 en 1995 zijn verzameld door 45 veldwerkers. De opnames zijn vervolgens getranscribeerd door 29 verschillende onderzoekers - in vrijwel alle gevallen door de veldwerker die het materiaal verzameld had. De transcribenten vertegenwoordigen op het eerste gezicht drie 'scholen', die corresponderen met drie delen van het taalgebied: Vlaanderen, Friesland en Nederland minus Friesland. Een klein deel van de transcripties zijn vervolgens gecorrigeerd door andere transcribenten – voor de Limburgse dialecten bijvoorbeeld omdat informatie over lexicale toon pas later werd toegevoegd.

3.2 Analyses

Ons onderzoek van het GTR-materiaal is kwantitatief van aard en beperkt zich tot de relevante ongelede woorden (zoals *eend*, *hond*, *donder*, *kant* en *winter*). Voor elk van de 613 dialecten zijn gegevens opgenomen voor 28 verschillende woorden. De analyse bestond per dialect per afgevraagd item uit de volgende stappen:

1. identificeren van de dialectvariant;
2. per segment in het cluster bepalen of er sprake is van palatalisering dan wel velarisering;
3. bepalen van een drietal aspecten van de kwaliteit van de voorafgaande vocaal;
4. bepalen van de identiteit van transcribent (in vrijwel alle gevallen identiek aan de veldwerker) en eventuele corrector.[5]

---

[5] Al deze stappen werden met behulp van Python-scripts automatisch uitgevoerd op basis van de KIPA-versie van het GTR-materiaal.

De resultaten van elk van deze stappen werden in gecodeerde vorm in een databestand opgenomen. De *data mining* resulteerde in 17.163 observaties.[6] Op de 17.163 observaties in het databestand werden drie typen van statistische analyse uitgevoerd, namelijk frequentietellingen, kruistabellering (met chi kwadraat-toets en contingentie-coëfficiënt resp. phi) en logistische regressie-analyse. De contingentie-coëfficiënt geeft de sterkte van het verband aan tussen twee nominale variabelen (bijvoorbeeld binaire grootheden van het type: al dan niet palatalisering, maar ook grootheden als sexe, nationaliteit, politieke voorkeur en dergelijke) en ligt tussen 0 (absoluut geen samenhang) en 1 (maximale samenhang). Waar het doel is om een eventuele implicatie vast te stellen, is phi gehanteerd, die eveneens minimaal 0 en maximaal 1 kan bedragen.

De frequentietellingen maken duidelijk hoeveel valide data er zijn, hoeveel daarvan palatalisatie dan wel velarisatie vertonen en welk aandeel van de data getranscribeerd is door elk van de drie 'transcribentenscholen'.

Met behulp van de tweede techniek werd nagegaan of er samenhangen zijn tussen de modificatie van de nasaal en de plosief, hoeveel elke afzonderlijke transcribent getranscribeerd heeft en of er zich ook in dit specifieke deel van het GTR-materiaal transcribenten- en correctoreffecten aftekenen. Zulks ter beantwoording van onderzoeksvraag 1. Eveneens met behulp van kruistabellering werd een eerste inzicht verkregen in de samenhang tussen enerzijds de modificatie van de nasaal en de plosief en anderzijds drie fonologische kenmerken van de voorafgaande klinker.

De derde techniek maakt het mogelijk om voor afhankelijke variabelen op nominaal meetniveau te bepalen in hoeverre er sprake is van een significante invloed van één of meerdere onafhankelijke variabelen. De onafhankelijke variabelen kunnen op grond van de analyse-uitkomsten geordend worden naar hun relatieve impact. Deze techniek is in wezen goed vergelijkbaar met de in de kwantitatieve sociolinguïstiek bekende *varbrul* (< 'variable rule') analyse. Vergeleken met de *varbrul*-techniek is het met logistische regressie-analyse betrekkelijk eenvoudig, eventuele beduidende interactie-effecten tussen verschillende onafhankelijke variabelen vast te stellen. Anders dan *varbrul* laat logistische regressie-analyse naast discrete ook continue grootheden toe als on-

---

(6) 28x613=17.164. Voor één dialect (Kloekenummer I069p) bleek echter een relevant gegeven (item 725, het woord *zand*) te ontbreken in de door ons geraadpleegde versie van de database.

afhankelijke variabelen. In concreto hebben wij met deze analyse vastgesteld in hoeverre palatalisering dan wel velarisering van elk van de segmenten in het cluster (de afhankelijke variabelen) samenhangt met de volgende onafhankelijke variabelen:

A - de herkomst van de transcribent en
- een eventuele correctie,

B - het relatieve aantal omringende dialecten die in één of beide segmenten in het cluster palatalisatie dan wel velarisatie vertonen,

C - de hoogte van de voorafgaande vocaal (laag, midden, hoog),
- de voor-/achter-specificatie van de voorafgaande vocaal en
- de +/- rond-specificatie van de voorafgaande vocaal.

De invloed van de omringende dialecten (B) is voor het GTR-materiaal als volgt geoperationaliseerd: er zijn 17 Kloekevakken; sommige daarvan zijn alleen in het noorden bezet (de dialecten in het zuidelijke deel van het taalgebied) en sommige alleen in het zuiden (de dialecten in het noordelijke deel van het taalgebied). De 6 vakken die zowel in het noorden als in het zuiden gevuld zijn (de vakken E, F, G, I, K en L) hebben we in tweeën verdeeld. Van de zo resulterende 23 regio's hebben we vervolgens voor elk relevant item in het GTR-materiaal bepaald hoe vaak de nasaal en de velair in de desbetreffende dialecten gepalataliseerd of gevelariseerd werden. Als bijvoorbeeld in een bepaalde regio 7 van de 30 onderzochte plaatsen de nasaal in het woord *hond* velariseerden, had de variabele voor de /n/ in dit woord in die regio voor velarisering de waarde 7/30 = ,23. Gezien de aard van deze variabele (een continue variabele, meer bepaald een ratio-variabele, die per item per segment in het cluster per modificatietype per regio een andere waarde aanneemt), wordt zijn effect alleen in de regressie-analyse onderzocht.

Met behulp van de regressie-analyse zullen we de vraag beantwoorden of de taalinterne krachten opwegen tegen de buitentalige en in het bijzonder de onderzoekervariabelen. Zijn er wisselwerkingen tussen de herkomst van de transcribent en de dialectgeografische factor? Dit alles in het verlengde van onderzoeksvraag 2.

4. Bevindingen

Voor de nasaal zijn er 14.621 valide data (de overige 2.542 staan in ons bestand te boek als 'missing', aangezien er òf inderdaad geen òf oninterpreteerbare òf anderszins onbruikbare gegevens zijn - dat laatste bijvoorbeeld vanwege

heteroniemen) en voor de plosief 13.681 (3.482 'missing'). In tabel 1 is te lezen hoeveel van de valide data voor beide segmenten palatalisatie resp. velarisatie vertonen.

	palatalisatie	velarisatie
nasaal	4,4 (645)	3,8 (549)
plosief	2,1 (287)	1,2 (160)

Tabel 1. De aandelen palatalisatie en velarisatie in de nasaal en plosief; percentages (tussen haakjes ruwe aantallen)

In de onderzochte coronale nasaal plosief-clusters blijkt palatalisering vaker voor te komen dan velarisering. De nasalen ondergaan de modificatie vaker dan de plosieven; er zijn blijkbaar veel dialecten die in het cluster wel de nasaal maar niet de plosief palataliseren (het omgekeerde komt nauwelijks voor). Relatief gezien is dat verschil voor velarisatie nog groter.

Nadere inspectie leert dat 336 van de 613 dialecten in het relevante deel van de GTR-database palatalisering en/of velarisering van nasaal en/of obstruenten vertonen; daarbij spelen alle 28 relevante woorden in de lijst een rol, dat wil zeggen in elk van deze woorden is in dit opzicht variatie te vinden. Er zijn, met andere woorden, geen absolute woordeffecten.

Kruistabellering maakt duidelijk dat palatalisatie van de nasaal (n=523) amper voorspellingen toelaat over wat er met de plosief gebeurt (eveneens palatalisatie in 248 gevallen, geen palatalisatie in 275 gevallen), terwijl palatalisatie van de plosief (n=284) in niet minder dan 248 gevallen gepaard gaat met palatalisatie van de nasaal. Er is dus sprake van een asymmetrische samenhang (χ^2=5325,086 df=1 p=,000 en phi=,634 p=,000). Een gevelariseerde nasaal (n=325) blijkt in de grote meerderheid van de gevallen (n=235) samen te gaan met een ongevelariseerde plosief, maar omgekeerd gaat een gevelariseerde plosief (n=98) in de grote meerderheid van de gevallen (n=90) samen met een gevelariseerde nasaal, dus ook hier is spake van een asymmetrisch, zij het wat zwakker, verband (χ^2=3299,044 df=1 p=,000 en phi=,499 p=,000).

Van de onafhankelijke variabelen nemen we allereerst de onderzoekers onder de loep, om te beginnen de transcribenten. Los van de vraag of de 17.163 relevante gegevens (voor de 28 woorden voor elk van de 613 onderzochte dialecten) voor de doeleinden van ons onderzoek bruikbaar zijn of niet,

8,5% (n=1.456) hiervan is getranscribeerd door Friese, 30,8% (n=5.292) door Vlaamse en 60,7 (n=10.415) door niet-Friese Nederlandse transcribenten. Er waren in totaal 2 Friese, 2 Vlaamse en niet minder dan 25 niet-Friese Nederlandse transcribenten. Van de Nederlandse (in engere zin) transcribenten hebben er 2 zeer veel en 2 veel van het relevante materiaal getranscribeerd.

Bij de transcriptie van een modificatie van één van de beide segmenten in het cluster blijken de Nederlanders in alle vier de gevallen voorop te lopen, in drie van de vier gevallen gevolgd door de Vlamingen, met de Friezen op de laatste plaats (de contingentie-coëfficiënt bedraagt voor palatalisatie van de nasaal ,086 p=,000, voor velarisatie van de nasaal ,051 p=,000, voor palatalisatie van de plosief ,039 p=,000). Bij de velarisatie van de plosief verdringen de Friese transcribenten hun Vlaamse collega's, al is het verschil tussen beide groepen betrekkelijk gering; de contingentie-coëfficiënt bedraagt hier ,131 (p=,000). Gezien de contingentie-coëfficiënten is in alle gevallen de samenhang weliswaar beduidend maar zeer zwak. Dit laatste kan mede een gevolg zijn van het feit dat 277 van de in deze fase van de analyses onderzochte 613 dialecten geen variatie vertonen in verband met palatalisatie of velarisatie in de onderzochte clusters.

Van de transcripties van de 14.621 valide data voor de nasaal blijken er 2.079 (14,2%) gecorrigeerd te zijn; van die van de 13.681 valide data voor de plosief zijn er 1.789 (13,1%) gecorrigeerd. Er waren 6 correctors, 1 Fries en 5 niet-Friese Nederlanders. In alle gevallen is de proportie modificatie van één van de beide segmenten in het cluster aanzienlijk hoger in het gecorrigeerde dan in het niet-gecorrigeerde materiaal - zoals blijkt uit tabel 2.

	niet-gecorrigeerd	gecorrigeerd
palatalisatie nasaal	3,3	11,1
velarisatie nasaal	3,4	6,0
palatalisatie plosief	1,1	8,8
velarisatie plosief	,8	3,4

Tabel 2. Het aandeel modificatie van elk van beide segmenten in de niet-gecorrigeerde en gecorrigeerde transcripties (in percenten)

Nadere inspectie van de analyseresultaten maakt duidelijk dat in alle vier de gevallen één en dezelfde corrector eindverantwoordelijk is voor de grote meerderheid van de modificaties.

Tot zover de onderzoekersvariabelen. Richten wij ons nu op de taalinterne variabelen, meer bepaald drie dimensies van de voorafgaande vocaal, te weten de hoogte, de voor- dan wel achterspecificatie en de specificatie gerond dan wel ongerond.(7) In tabel 3 zijn de voornaamste uitkomsten samengevat.

	hoogte	voor/achter	+/-rond
palatalisatie nasaal	χ^2=10,295 df=2 p=,006	χ^2=14,690 df=1 p=,000	χ^2=,022 df=1 NS
	conting ,036 p=,006	conting ,043 p=,000	conting ,002 NS
	H<L<M	voor<achter	n.v.t.
velarisatie nasaal	χ^2=15,909 df=2 p=,000	χ^2=17,437 df=1 p=,000	χ^2=67,508 df=1 p=,000
	conting ,045 p=,000	conting ,047 p=,000	conting ,092 p=,000
	H=L<M	voor<achter	ongerond<gerond
palatalisatie plosief	χ^2=20,725 df=2 p=,000	χ^2=16,595 df=1 p=,000	χ^2=4,712 df=1 p=,030
	conting ,054 p=,000	conting ,048 p=,000	conting ,026 p=,030
	L<H<M	voor<achter	gerond<ongerond
velarisatie plosief	χ^2=27,568 df=2 p=,000	χ^2=,672 df=1 NS	χ^2=13,700 df=1 p=,000
	conting ,062 p=,000	conting ,010 NS	conting ,044 p=,000
	L<H<M	n.v.t	ongerond<gerond

Tabel 3. Modificatie van elk van beide segmenten in samenhang met drie aspecten van de kwaliteit van de voorafgaande klinker (NS= niet significant)

De palatalisatie van de nasaal blijkt vooral gevoelig voor de dimensie voor/achter van de voorafgaande vocaal maar in het geheel niet voor de dimensie +/-rond, die van de plosief is vooral gevoelig voor de dimensie hoogte; *alle* samenhangen zijn echter zeer zwak. De velarisatie van de nasaal blijkt vooral gevoelig voor de dimensie +/-rond (met iets meer dan 9% de sterkste samenhang in de gehele tabel), terwijl die van de plosief vooral gevoelig is voor de dimensie hoogte maar hoegenaamd niet voor de dimensie voor/achter. Opmerkelijkerwijs vertoont zowel de palatalisatie als de velarisatie van de plosief de sterkste samenhang met de hoogte van de voorafgaande vocaal.

Wanneer we de bevindingen bekijken vanuit het perspectief van de dimensies, dan blijkt hoogte het meest van invloed op de velarisatie van de plosief; de dimensie voor/achter heeft de sterkste impact op de palatalisatie van de plosief en +/- rond op de velarisatie van de nasaal. De palatalisatie van de nasaal is relatief immuun voor deze drie factoren in verband met de kwaliteit van de voorafgaande vocaal.

Dat voor wat betreft de uitkomsten van de univariate (frequentietellingen) en bivariate (kruistabellen) analyses. In de logistische regressie-analyses zijn de onafhankelijke variabelen van alle drie de types ingebracht, dus behalve de beide onderzoekersvariabelen en de drie taalinterne variabelen ook de dialectgeografische variabele, die per item per segment in het cluster aangeeft hoe groot de proportie dialecten in de regio is waar het betreffende modificatietype optreedt.

Wij beperken ons tot een klein deel van de uitkomsten van de logistische regressieanalyse,[8] te weten per segment per modificatietype de signifi-

---

(7) Voor deze en alle volgende analyses zijn de gegevens voor de dialecten waarinde onderzochte palatalisatie en velarisatie geen rol speelt verwijderd, om te voorkomen dat zij de sterkte van de verbanden zouden drukken.

(8) Voor de liefhebber: de gehanteerde methode was *Backward Wald*. De waarschijnlijkheidswaarden voor opname resp. verwijdering van predictoren waren ,05 resp. ,10 (default) en van de onafhankelijke variabelen is 'dialectgeografie' de enige continue. De overige variabelen zijn allemaal categoriaal en voor de analyses behandeld als *indicator* ('al dan niet correctie'; voor- dan wel achtervocaal; geronde dan wel ongeronde vocaal) resp. als *deviation* (de driewaardige variabelen 'herkomst transcribenten' en vocaalhoogte). In dat laatste geval werden alle waarden vergeleken met het *overall* effect, dat wil zeggen het gemiddelde effect van alle waarden van de betreffende variabele. Zie ook Norušis (1997: 4, 49-52).

cante voorspellers. Hierbij staat *criterium* voor de afhankelijke variabele, meer in het bijzonder palatalisatie dan wel velarisatie van nasaal dan wel plosief; *predictor* staat voor een bepaalde waarde van een onafhankelijke variabele of de interactie tussen (bepaalde waarden van) twee onafhankelijke variabelen. In dat laatste opzicht werd in alle vier de gevallen uitsluitend het effect van de (tweeweg oftewel eerstegraads) interactie tussen 'dialectgeografie' en 'herkomst transcribenten' onderzocht; op die manier zijn de invloeden van enerzijds de omringende dialecten en anderzijds veldwerkers / transcribenten ontward.[9] *B* staat voor de waarde van de regressiecoëfficiënt (het relatieve gewicht) en *signif* voor de waarschijnlijkheid dat het effect niet aan de predictor in kwestie maar aan iets anders (bijvoorbeeld het toeval) moet worden toegeschreven. Per criterium worden de predictoren gepresenteerd in volgorde van afnemend gewicht. Voor een eerste inzicht in het belang van de uitkomsten zijn alleen *Nagelkerke* R^2 , *-2 Log likelihood* en de *model* χ^2 nog van belang (hoewel één en ander onze onderzoeksvragen overstijgt). De *Nagelkerke* R^2 geeft de hoeveelheid met de uiteindelijk overblijvende onafhankelijke variabelen / predictoren verklaarde variantie; de waarde van deze coëfficiënt kan liggen tussen 0 (geen enkele verklarende waarde) en 1,00 (alle variantie restloos verklaard door de voorspellers). *-2 Log likelihood* is een maat voor de waarschijnlijkheid van de geobserveerde data gegeven de geschatte effecten (dat wil zeggen de gewichten) van de predictoren. Hoe groter deze waarde, hoe geringer de 'fit' tussen de data en het berekende model (bij gelijkblijvende aantallen observaties). De *model* χ^2 beantwoordt de vraag of er significante verschillen zijn tussen enerzijds het berekende model en anderzijds een model met alleen de constante.

De tabellen 4 tot en met 7 bevatten de voornaamste uitkomsten voor elk van beide modificatietypen voor elk van beide segmenten in het cluster.

---

(9) Voor de data van de 336 dialecten waarin palatalisatie en velarisatie variëren zijn de transcripties van alle Vlaamse (in ruime zin) dialecten gemaakt door Vlaamse transcribenten en alle transcripties van Nederlandse (in engere zin) dialecten door Nederlandse transcribenten. De transcripties van 2 van de in totaal 41 voor deze studie relevante Friese dialecten zijn eveneens gemaakt door Nederlandse, de overige door Friese transcribenten.

predictor		B	signif
dialectgeogr x Fr transcr		-24,974	,000
dialectgeogr x NL transcr		14,704	,000
geen correctie		-,626	,000
voorV		,561	,000
geronde V		,270	,011
-2 Log likelihood		3641,698	
Nagelkerke R^2		,226	
Model	χ^2	808,036	
	df	7	
	signif	,000	

Table 4. Logistische regressie. Criterium: palatalisatie van de nasaal

predictor	B	signif
dialectgeogr x Fr transcr	60,013	,010
dialectgeogr	41,942	,000
dialectgeogr x NL transcr	-25,282	,032
constante	-2,984	,000
Fr transcr	-2,336	,000
NL transcr	1,180	,000
hoge V	,503	,000
geen correctie	-,425	,000
lage V	,199	,027
-2 Log likelihood	3533,533	
Nagelkerke R^2	,132	
Model χ^2	420,229	
df	9	
signif	,000	

Table 5. Logistische regressie. Criterium: velarisatie van de nasaal

predictor		B	signif
dialectgeogr		27,521	,000
constante		-3,530	,000
geen correctie		-1,501	,000
hoge V		,634	,000
lageV		-,485	,000
-2 Log likelihood		1642,910	
Nagelkerke R^2		,354	
Model	χ^2	762,745	
	df	5	
	signif	,000	

Table 6. Logistische regressie. Criterium: palatalisatie van de plosief

predictor		B	signif
dialectgeogr		53,856	,000
dialectgeogr x NL transcr		49,837	,000
dialectgeogr x Fr transcr		-48,568	,000
constante		-3,242	,000
Fr transcr		1,927	,000
NL transcr		-1,277	,000
ongeronde V		-1,214	,000
hoge V		,476	,000
geen correctie		-,397	,037
-2 Log likelihood		1242,235	
Nagelkerke R^2		,201	
Model	χ^2	281,237	
	df	9	
	signif	,000	

Table 7. Logistische regressie. Criterium: velarisatie van de plosief

Wat bij vergelijking van de analyseresultaten voor de vier gevallen allereerst opvalt is dat in alle gevallen de verklaarde variantie gering is (wat misschien nog onderstreept wordt door het feit dat in drie van de vier gevallen de constante in de top vier prijkt van predictoren in de regressievergelijking). Het grootste deel van de variantie in de vier gevallen van modificatie blijft met de gehanteerde predictoren / onafhankelijke variabelen onverklaard. Dat zou kunnen betekenen dat we de belangrijkste voorspellers niet te pakken hebben. Dit klinkt dramatisch, maar we moeten bedenken (a) dat het aantal onderzochte onafhankelijke variabelen klein is, (b) dat we van deze onafhankelijke variabelen slechts één enkele interactie onderzocht hebben en, belangrijker, (c) dat de palatalisatie en velarisatie van coronale nasaal-plosief clusters of delen daarvan in de meeste dialecten waarschijnlijk historische processen zijn en dat de verklarende factoren in de loop van de tijd door bijvoorbeeld latere vocaalverschuivingen opaak geworden kunnen zijn. Met andere woorden, in de meeste dialecten zijn deze processen vermoedelijk niet meer productief en daarmee òf lexicale regels òf zelfs volledig gelexicaliseerd en daardoor onvoorspelbaar.

Gezien de waarden van Nagelkerke R^2 laat de variantie in de modificatie van de plosieven (35% en 20%) zich, op basis van de onderzochte onafhankelijke variabelen, beter voorspellen dan die in de modificatie van de nasalen (23% en 13%).[10] Tegelijkertijd laat palatalisatie (23% en 35%) zich, op basis van de onderzochte onafhankelijke variabelen, beter en met minder predictoren voorspellen dan velarisatie (13% en 20%). Ook worden er van de gehanteerde onafhankelijke variabelen voor palatalisatie (van nasalen en plosieven) minder predictoren geselecteerd dan voor velarisatie (van nasalen en plosieven).

In alle gevallen blijkt er een negatief effect uit te gaan van het ontbreken van een correctie van de transcripties (vgl. Tabel 2 hiervóór). Voor de modificatie (zowel palatalisatie als velarisatie) van de plosieven blijkt de factor dialectgeografie de voornaamste voorspeller,[11] voor die van de nasalen de wis-

---

(10) Hetzelfde beeld lijkt naar voren te komen bij vergelijking van de waarden van de -2 Log likelihood, al zijn er in het relevante deel van de data aanzienlijk minder gevallen van modificatie van plosieven dan van nasalen.

(11) Een manifestatie van het algemene inzicht dat nabijheid vaak leidt vaak tot gelijkenis. De samenhang tussen enerzijds taalvariatie en -verandering en anderzijds geografische, temporele en sociale nabijheid kan worden samengevat onder het kopje *proximity principle* (Hinskens, Van Hout & Wetzels 1997: 22). Men vergelijke ook Chambers' (1995: 65-66) 'Dialect laws of mobility en isolation' en de inzichten die verwoord zijn in Weinreich 1953: viii.

selwerking tussen dialectgeografie en de herkomst van de transcribent, meer in het bijzonder als het om een Friese transcribent gaat; in dat laatste verband blijkt de richting van het effect op de transcriptie van palatalisatie tegengesteld te zijn van dat op de transcriptie van velarisatie. Die wisselwerking is negatief voor wat betreft palatalisatie van nasalen, dat wil zeggen dat naar gelang de hoeveelheid omringende dialecten met palatalisatie van nasalen toeneemt resp. afneemt, de Friese transcibenten minder resp. meer gepalataliseerde nasalen signaleren; dezelfde wisselwerking is positief voor wat betreft de velarisatie van nasalen, dat wil zeggen dat naar gelang de hoeveelheid omringende dialecten met velarisatie van nasalen toeneemt resp. afneemt, de Friese transcribenten meer resp. minder (zij het veel meer resp. veel minder) gevelariseerde nasalen signaleren - een soort hypercorrectie. Een dergelijk interactie-effect is er ook tussen dialectgeografie en de Nederlandse (*strictu senso*) transcribenten, zij het dat deze voor de beide gevallen de omgekeerde richting heeft, vergeleken met de Friese transcribenten. Voor de velarisatie van nasalen en plosieven kunnen we de effecten van de wisselwerking tussen dialectgeografie en de herkomst van de transcribent vergelijken met de hoofdeffecten van de factor herkomst van de transcribent; we stellen dan vast dat de richting van het effect van Friese resp. Nederlandse transcribenten in interactie met dialectgeografie het omgekeerde is van die van het hoofdeffect. Dat wil zeggen dat de Friese transcribenten vergeleken met alle drie de groepen van transcribenten gezamenlijk in het algemeen minder velarisatie in nasalen signaleren, behalve voor gebieden waar de velarisatie in nasalen al betrekkelijk veel (resp. weinig) voorkomt - daar hebben de Friese transcribenten juist veel meer (resp. veel minder) velarisatie gesigaleerd dan gemiddeld voor de Friese, Nederlandse en Vlaamse transcribenten gezamenlijk; het tegendeel is het geval voor de Nederlandse transcribenten. Het exacte spiegelbeeld van deze hele constellatie tekent zich af bij velarisatie van de plosieven. In het geval van de palatalisatie van plosieven (en alléén daar) zijn er geen transcribenteneffecten, noch als zodanig noch in wisselwerking met de dialectgeografische verbreiding van de modificatie.

Voorts valt op dat de predictoren en hun relatieve gewicht voor velarisatie van de nasaal lijken op die voor velarisatie van de plosief. Met andere woorden, er is een patroon in de variantie van velarisatie. In het patroon van de velarisatie speelt, in tegenstelling tot dat van de palatalisatie, de (academische) herkomst van de transcribenten ook een autonome rol. Van de drie taalinterne onafhankelijke variabelen speelt in het velarisatiepatroon in beide gevallen

hoogte een rol. De dimensie voor/achter speelt uitsluitend bij palatalisatie van de nasaal een significante rol.

In Tabel 8 vatten we de resultaten van de logistische regressie-analyses samen.

	pal nas	vel nas	pal plos	vel plos
A				
# Fr transcr		--		+
NL transcr		+		--
# geen correctie	--	--	--	--
B				
# dialectgeografie		+	+ **Z**	+ **Z**
A x B				
# dial.geogr. x Fr transcr	-- **Z**	+ **Z**		--
dial.geogr. x NL transcr	+	--		+
C				
# hoge Voc		+	+	+
lage Voc		+	--	
# voorVoc	+			
# geronde Voc	+			
ongeronde Voc				--

Tabel 8. Significante positieve (+) en negatieve (--) effecten op de modificatie van elk van beide segmenten (op basis van de logistische regressie-analyses). Per modificatietype per segment is het zwaarste effect aangeduid met een 'Z'

De bevindingen voor de onderzoekerseffecten (herkomst transcribenten, al dan niet correctie) stroken grotendeels met die uit de chi-kwadraatanalyses. Dat geldt ook voor de meerderheid van de effecten van de taalinterne onafhankelijke variabelen:

- in het geval van de palatalisatie van de nasalen blijkt hoogte wel (kruistabellering) of juist niet (logistische regressie), +/- ronding niet (kruistabellering) of juist wel (maar zwak, logistische regressie) van belang. Maar in alle analyses komt de dimensie voor/achter als zwaarste interne voorspeller naar voren;
- in het geval van de velarisatie van de nasalen komen in de kruistabellering +/- ronding als sterkste en hoogte als zwakste significante voor-

spellers naar voren, maar in de logistische regressie blijkt alleen hoogte signifant invloed uit te oefenen;
- in het geval van de palatalisatie van de plosieven blijkt hoogte de zwaarste (kruistabellering) resp. enige (logistische regressie) beduidende taalinterne factor;
- in het geval van de velarisatie van de plosieven, tenslotte, is hoogte blijkens de kruistabellering de zwaarste en +/- ronding de op één na zwaarste factor, terwijl die verhouding in de logistische regressie-analyse omgekeerd blijkt. In geen van beide typen van analyse speelt de factor voor/achter een rol van betekenis.

De (al met al betrekkelijk geringe) verschillen tussen de bevindingen uit de kruistabellering en de logistische regressie worden waarschijnlijk grotendeels veroorzaakt door het feit dat elke afzonderlijke taalinterne factor in de kruistabellering de enige onafhankelijke variabele vormt en in logistische regressie één van de zes onafhankelijke variabelen.

Onze bevindingen betekenen, in het licht van de onderzoeksvragen, het volgende:

1) hoofdeffecten van de transcribenten zijn er voor velarisatie van zowel nasalen als plosieven. Correctie-effecten zijn er in alle gevallen;
2) de sterkste effecten gaan, in verband met palatalisering en velarisering in plosieven, uit van de dialectgeografie (B), resp. van de interactie met de herkomst van de transcribenten (palatalisering en velarisering in nasalen), en de zwakste effecten gaan uit van de onderzochte taalinterne factoren (C). De onderzoekerseffecten (transcribenten en correctoren, A) zitten hier globaal gesproken tussenin, al lijken zij qua gewicht meer op de taalinterne factoren.

5. Afsluitend

De hier gepresenteerde verkenning van het GTR-materiaal heeft in onze ogen zowel fonologische als methodologische implicaties. Wat betreft de (historische) fonologie zijn deze implicaties mogelijk het geringst. In Hinskens en Van Oostendorp (2004) zetten wij twee benaderingen van de hier behandelde processen tegenover elkaar: een benadering die we toeschrijven aan Kieft (1945) en een andere benadering die we toeschrijven aan Van Ginneken (1935). We

merken in dat eerdere artikel op dat uit de beschouwingen van de laatste volgt dat er een samenhang moet zijn tussen palatalisering en ontronding resp. tussen velarisering en ronding, terwijl Kieft een dergelijke voorspelling niet maakt. Onze uitkomsten voor het moderne dialectmateriaal weerspreken Van Ginneken; voor de gevallen waarin de dimensie +/-rond uit de multivariate analyses als significant naar voren komt, is de richting van het effect tegengesteld aan Van Ginnekens opvatting. Bovendien blijken er in alle gevallen behalve de palatalisatie van de nasalen samenhangen te zijn met klinkerhoogte, iets dat Van Ginneken noch Kieft had voorzien.

Wat betreft de methodologie dient er, naar onze overtuiging, aandacht besteed te worden aan de bevinding dat er inderdaad transcribenten- en correctoreneffecten aanwezig zijn in de aangetroffen variatie. Voor vrijwel ieder denkbaar gebruik van het GTR-materiaal zijn dit soort effecten -hoe bescheiden of voorwaardelijk misschien ook- te zien als ongewenste 'ruis'. Daarom, maar ook ten behoeve van de verdere vergroting van de betrouwbaarheid van de bevindingen, lijkt het raadzaam

a) aan de GTR-database de opnamen toe te voegen
b) aan de GTR-database per dialect per item meerdere transcripties toe te voegen, liefst van vertegenwoordigers van alle drie de 'scholen', in plaats van één enkele (al dan niet gecorrigeerde) transcriptie
c) voor specifieke onderzoeksvragen de GTR-data aan te vullen met andere relevante data, zoals bijv. data uit de SGV enquête, uit de RND, uit de Meertensvragenlijsten, uit monografieën en eventueel zelfs uit nog te verzamelen materiaal. Deze aanvulling zal ook de generaliseerbaarheid van de bevindingen ten goede komen, indachtig Labov's (1972: 102, 118-9) *principle of convergence*: "the value of new data for confirming and interpreting old data is directly proportional to the differences in the method used to gather it".

Voor zijn niet te overschatten aandeel in de transcriptie van het GTR-materiaal en de inrichting van de (in veel opzichten indrukwekkende) database die wij voor het hier beschreven onderzoek gebruikt hebben kan Piet van Reenen niet genoeg geprezen worden. Niet alleen voor de voortgang van de werkzaamheden ten behoeve het tweede deel van de MAND is het te hopen, dat hij -ook na zijn pensionering- nog vele jaren actief kan blijven als onderzoeker.

Bibliografie

Chambers, Jack,
1995, *Sociolinguistic theory*. Oxford etc. (Blackwell)

Coetzee, Anna E. & Reenen, Piet van,
1995, "Die Afrikaanse nasalering en nienasalering se verband met 17de-eeuse Nederlands," *SATT*, 13/2, 1995, 62-73.

Daan, Jo,
1950, *Wieringer land en leven in de taal*. Proefschrift, Universiteit van Amsterdam.

Ginneken, Jacques van,
1935, "De correlatie van harde en weeke medeklinkers in het Oud- en Middelnederlandsch". *Onze Taaltuin*, 3, 10, 302-317.

Goeman, Ton & Johan Taeldeman,
1996, "Fonologie en morfologie van de Nederlandse dialecten. Een nieuwe materiaalverzameling en twee nieuwe atlasprojecten". *Taal en Tongval* 48, 38-59.

Hermans, Ben,
1994, "Palatalisatie in een Limburgs dialect: dissimilatie als gevolg van OCP." In: G. Booij & J. van Marle (eds), *Dialectfonologie*. Amsterdam (Cahiers van het P.J. Meertens-Instituut), 60-82.

Hinskens, Frans & Marc van Oostendorp,
2004, "Coronale nasaal-plosief clusters in de Nederlandse dialecten. Een verkenning", in: J. de Caluwe, G. de Schutter, M. Devos & J. van Keymeulen (red.), *Taeldeman, man van de taal, schatbewaarder van de taal. Liber amicorum voor Johan Taeldeman*. Gent (Academia Press), 519-545.

Hinskens, Frans, Roeland van Hout & Leo Wetzels,
1997, "Balancing data and theory in the study of phonological variation and change", in: F. Hinskens, R. van Hout & L. Wetzels (Hsg.), *Variation, change and phonological theory*. Amsterdam / Philadelphia (Benjamins). 1-33

Hinskens, Frans & Jeroen van de Weijer,
2003, "Patterns of segmental modification in consonant inventories: a cross-linguistic study", *Linguistics* 41, 6, 1041-1084.

Kieft, Antonij,
1945, "Gutturalisering en dentalisering in het Nederlands". *Nieuwe taalgids* 38, 111-113.

Labov, William,
1972, "Some principles of linguistic methodology", *Language in society* 1, 97-120.

Norušis, Marija/SPSS Inc.,
1997, *SPSS Professional Statistics 7.5*. Chicago (SPSS Inc.).

Oostendorp, Marc van,
te verschijnen, "Na de FAND", *Taal en Tongval. Atlassenthemanummer*.

Weinreich, Uriel,
1953, *Languages in contact; findings and problems*. New York (Publications of the Linguistic Circle of New York, 1)