

Ann Marynissen

Een vergelijking van diverse telmethodes in het historisch-linguïstisch onderzoek

Tot voor een aantal jaren was het beginpunt van een onderzoek naar een stadium uit de voorgeschiedenis van het Nederlands vaak een moeizame zoektocht naar feitenmateriaal in uitgegeven bronnen en archieven. Voor één periode uit de taalgeschiedenis is het tijdrovende werk om eigenhandig een representatief corpus samen te stellen verleden tijd: sinds de publikatie van de ambtelijke teksten uit het Corpus van Middel nederlandse teksten tot en met het jaar 1300 door dr. M. Gysseling in 1977 is de 13de eeuw uitzonderlijk goed gedocumenteerd: men beschikt al snel over honderden vindplaatsen van een grammaticaal verschijnsel. Sinds deze bron op computer is overgebracht, is het bovendien niet langer een ondoenbare opdracht om bij het onderzoek duizenden attestaties te analyseren.

Bij die luxe aan materiaal rijzen andere vragen die eerder al aanleiding hebben gegeven tot discussie¹: hoe verantwoordt men statistisch de ongelijke verdeling van het materiaal over de scribenten en over de documenten, m.a.w. welke telmethode is de meest geschikte manier om de frequentieverhoudingen tussen de diverse attestaties te berekenen, zodat de percentages de grammaticale variatie zo adequaat mogelijk weerspiegelen? Op deze vraag heb ik een antwoord geformuleerd in mijn proefschrift, waarin ik de flexie van het substantief in het 13de-eeuwse ambtelijke Middelnederlands in taalgeografisch perspectief heb onderzocht.² In deze bijdrage wil ik mijn standpunt in dit methodologische debat toelichten aan de hand van een concrete vergelijking van de verschillende telmethodes.

De bedoeling van mijn doctoraal onderzoek was een synchrone, structurele beschrijving te maken van het flexiesysteem van het substantief dat in

¹ Nl. tussen Berteloot en Goossens enerzijds en Van Reenen en Berteloot anderzijds (zie Goossens 1979, Berteloot 1984a en 1984b en Van Reenen 1986-1987).

² A. Marynissen, *De flexie van het substantief in het 13de-eeuwse ambtelijke Middelnederlands. Een taalgeografische studie*. Deel I: Tekst, deel II: Tabellen. Leuven, 1993. Dit proefschrift verschijnt in februari 1996 bij het Belgisch Interuniversitair Centrum voor Neerlandistiek, als nr. 2 in de reeks "Studies op het gebied van de Nederlandse taalkunde". Leuven, Uitgeverij Peeters.

de 13de eeuw in de diverse regio's van het Nederlandse taalgebied in gebruik was. Per historische flexieklasse heb ik volgens bepaalde criteria een aantal substantieven uit het Corpus Gysseling geselecteerd³ en van elke attestatie de naamval en het getal bepaald. De schrijfcentra waarin de documenten zijn ontstaan, heb ik gegroepeerd in 7 representatieve dialectgebieden: het Vlaams-Zeeuws, het Hollands, het Brabants, het Utrechts, het Limburgs-Zuidnederrijns, het Noordnederrijns en het IJssellands.⁴ Met de toewijzing van een dialectgebied aan iedere vindplaats is de factor regionale variatie in het onderzoek gevaloriseerd.

De gegevens werden omgezet in kruistabellen waarin per regio af te lezen is welke uitgangen een substantief uit een bepaalde flexieklasse aanneemt als uitdrukking van een bepaalde naamval en getal (zie de voorbeelden als bijlage). De frequentieverhoudingen tussen de varianten nemen in dit type onderzoek een belangrijke plaats in. Er is immers lang niet altijd een een-op-een-verhouding tussen de vorm en de functie van een substantief in de zin: een substantief met een bepaald genus en getal kan in een naamval verschillende vormen aannemen. De variatie die in het buigingssysteem aanwezig is, kan worden afgelezen van de frequentieschommelingen tussen de vormvarianten. Het is dan ook van belang om een telmethode te gebruiken die deze variatie zo nauwkeurig mogelijk weergeeft.

Bij het afwegen van de verschillende telmethodes wil ik er vooraf op wijzen dat de arbitrariteit, die op velerlei gebied in het materiaal van het CG vervat zit, in laatste instantie door de geschiedenis gegeven is en dus nooit helemaal kan worden weggewerkt. Een gelijkmatige verdeling van de vindplaatsen over de onderzochte parameters kan nooit worden bereikt en is wellicht zelfs niet eens wenselijk.⁵

³ De criteria voor de selectie van de substantieven worden uiteengezet in de inleiding van mijn werk. Globaal beschouwd heb ik uit het bestand de lemmata met het grootste aantal attestaties opgevraagd die bovendien evenwichtig gespreid zijn over de regio's, de flexieklassen, de casus en de numeri. De kwantitatieve benedengrens voor selectie, aanvankelijk bepaald op 500 attestaties per substantief, heb ik soepel geïnterpreteerd. In de loop van mijn onderzoek heb ik meer dan 120000 vindplaatsen ontleed.

⁴ Uit het Noordnederrijns en het IJssellandse gebied zijn haast geen documenten voorhanden, zodat deze twee dialectgebieden noodgedwongen ondervetegenwoordigd blijven in de studie.

⁵ Met deze eis zou men de taal wel een bijzonder strak keurslijf opleggen. Concreet voor mijn onderzoek kan men bijv. niet verwachten dat ieder substantief ongeveer even vaak in de vier naamvallen en de twee numeri vertegenwoordigd zou zijn. De verdeling van de taalvormen over de naamvallen is niet louter een morfologische aangelegenheid: syntactische en semantische criteria bepalen mee hoe vaak een substantief in welke naamval en in welk getal voorkomt.

Er zijn vier mogelijke manieren om de taalvormen te tellen. Men kan tellen:

1. alle vindplaatsen uit alle documenten.
2. alle vindplaatsen uit alle originele documenten, maar niet de attestaties uit afschriften, vertalingen, vidimi en falsificaties.
3. 1 vorm per (origineel) document.
4. 1 vorm per hand (uit de originelen).

Bij methode 1 en 2 staat de individuele taalvorm centraal: het feit dat een taalvorm geattesteerd is, heeft op zich betekenis. Door alle beschikbare vindplaatsen in aanmerking te nemen voor de berekening van de frequenties krijgt iedere taaluiting statistisch evenveel betekenis.

Wanneer men de voorkeur geeft aan methode 3 of 4, stelt men zich principieel op een ander standpunt. Bij telwijze 4 staat de schrijvershand centraal: iedere scribent heeft statistisch evenveel betekenis, ongeacht het feit of hij veel dan wel weinig heeft geschreven. Men gaat uit van de idee dat iedere scribent een andere taalperformantie realiseert. Om te vermijden dat het taaleigen van een bepaalde schrijver het totale resultaat te veel zou beïnvloeden, tracht men het willekeurige aantal gegevens per document en het toevallige aantal documenten per hand te neutraliseren door slechts één vorm per hand over te houden (Berteloot 1984a, 1984b), eventueel proportioneel verdeeld volgens de verhouding tussen de absolute aantallen voorkomens van de varianten (Van Reenen 1986-1987). Methode 3 sluit enigszins aan bij de volgende, met dit verschil dat niet de schrijver, maar het document het middelpunt vormt. Deze werkwijze kan worden gebruikt als het tekstmateriaal niet in handen is onderscheiden.⁶ Telwijze 3 valt trouwens exact samen met methode 4 als ieder document door een andere hand is geschreven.⁷

Goossens (1979) past de eerste telmethode toe, het tellen van alle beschikbare vindplaatsen. De kaarten in de atlas van Déés e.a. (1980) zijn gebaseerd op een telling van 1 vorm per document. Berteloot (1984a, 1984b) en Van Reenen (1986-1987) tellen 1 vorm per hand. Wanneer een scribent

⁶ Dat is met name het geval bij het corpus waarop Demeersseman haar studie over de meervoudsflexie van drie substantieven in het veertiende-eeuwse Middelnederlands heeft gebaseerd. Aangezien de documenten van het Amsterdamse corpus haast allemaal van verschillende handen zijn, komt het tellen van documenten vrijwel op hetzelfde neer als het tellen van handen (Demeersseman 1988: 17).

⁷ Met ons corpus zouden er, indien we het telsysteem "één vorm per document" zouden toepassen, problemen ontstaan voor het Limburgs. De voornaamste bron voor het Limburgs, het goederenregister van Oudenbiezen, wordt immers beschouwd als één document. De keuze van telmethode 3 zou een enorme reductie van het Limburgse materiaal met zich meebrengen.

variatie vertoont, telt Berteloot beide varianten eenmaal, Van Reenen verdeelt de verschillende vormen proportioneel per hand. Mooijaart (1992) gaat uit van het document als analyse-eenheid, maar definieert een document als "bepaalde hoeveelheid tekst, vervaardigd op één tijdstip en plaats en door één hand", waardoor zij in dat geval 1 vorm per document én per hand telt. Omvangrijke documenten, waaraan meerdere scribenten over een langere periode hebben gewerkt, worden in haar onderzoek opgesplitst en als meerdere documenten beschouwd (Mooijaart 1992: 25).

Ik heb geopteerd voor de eerste telwijze, een methode die door een aantal vakgenoten bekritiseerd wordt.⁸ Ter verantwoording van die keuze heb ik de proef op de som genomen door het materiaal aan alle telmethodes te onderwerpen, om via een vergelijking van de resultaten van alle telwijzen exact te kunnen bepalen in welke mate de keuze van de ene of de andere telmethode de cijfers en mutatis mutandis de conclusies die aan die cijfers worden verbonden, beïnvloedt.

Ik heb de vier telmethodes getoetst aan de attestaties van de substantieven *ridder* en *parochie* (zie de tabellen als bijlage). Voor *ridder* geef ik de resultaten van de nominatief: de cijfers voor de nominatief enkelvoud weerspiegelen de frequentie en de verspreiding van de *-e*-apocope in de 13de eeuw; uit de gegevens voor de nominatief meervoud blijkt waar en hoe vaak het *-s*-meervoud in die periode voorkomt. De wisselende verhouding tussen de uitgangen *-e* en *-en* in de datief sing. van het feminiene *parochie* gaat terug op een geografische tegenstelling binnen het Vlaams, zodat hier een bijkomende opsplitsing werd gemaakt tussen het westelijke en het oostelijke Vlaams.

Een eerste vaststelling is dat de totaalpercentages, berekend volgens de diverse telmethodes, bijzonder dicht bij elkaar liggen. Bij een vergelijking van de totaalcijfers voor de vormvarianten van de nominatief singularis van *ridder* in de diverse tabellen blijkt dat het verschil tussen de cijfers in de uiterst rechtse kolom van de tabel ten hoogste 5.8% bedraagt.⁹ Ook de percentages voor het substantief *parochie* benaderen elkaar zeer sterk. Het grootste verschil tussen de totaalpercentages van de diverse telwijzen bevindt zich tussen telwijze 3 en telwijze 4: het bedraagt slechts 6.9%.

⁸ Door Berteloot 1984a en 1984b en Van Reenen 1986-1987. Het artikel van Berteloot (1984b) is een reactie op Goossens 1979.

⁹ Dit grootste procentuele verschil van 5.8% bevindt zich tussen de tweede en de vierde telwijze: de vorm op *-e* treedt op in 83.5% van de gevallen wanneer men alle vindplaatsen uit de originele documenten meetelt; indien slechts 1 vorm per hand wordt geteld, maakt de vorm *riddere* 77.7% van de gegevens uit.

Het aantal vindplaatsen waarop de totaalpercentages zijn gebaseerd, loopt echter sterk uiteen. Bij een telling van één vorm per hand wordt slechts 37% van alle bestaande attestaties van *ridder* in de nominatief singularis in aanmerking genomen. Wanneer één vorm per document wordt geteld, bedraagt dat percentage 70%; indien enkel de afgeschreven vormen geëlimineerd worden, wordt dat 87%.¹⁰ Ook bij *parochie* wordt er bij telwijze 4 (1 vorm per hand) het minste materiaal gehonoreerd, nl. slechts 22% van alle beschikbare vindplaatsen, tegenover 80% bij telwijze 3 (1 vorm per document) en zelfs 95% bij telwijze 2 (alle vindplaatsen uit de originele documenten).

Niet enkel de totaalpercentages, ook de percentages voor de afzonderlijke dialectgroepen wijken onderling slechts weinig van elkaar af. De uitkomsten voor de vormvarianten in de datief sing. van *parochie* in het westelijke resp. het oostelijke Vlaams verschillen bij telwijze 1 (alle vindplaatsen) versus methode 4 (1 vorm per hand) amper 0.5% resp. 4.2%. In de enkelvoudsflexie van *ridder* (nom.) wijken de percentages voor de uitgangen \emptyset en *-e* in de diverse dialectgebieden meestal minder dan 3% van elkaar af. Enkel in Holland bedraagt het verschil tussen de hoogste en de laagste waarde bij de eerste en de vierde telmethode meer dan 15%. In vergelijking met de telling van het totale aantal vindplaatsen stijgt bij een telling van één vorm per scribent het procentuele aandeel geapocopeerde vormen in het Hollands van 44.7% naar 60%. Voortgaande op deze cijfers zou ik mijn constatering met betrekking tot de spreiding van de *-e*-apocope lichtjes moeten herzien: de *-e*-apocope blijkt in deze regio verder gevorderd te zijn dan ik op basis van telmethode 1 heb geconcludeerd. Het toepassen van een andere telwijze heeft evenwel in het Hollands procentueel een grote reductie van materiaal tot gevolg.¹¹ De *-e*-apocope is in het 13de-eeuwse Hollands m.i. minder algemeen dan de tabel op basis van een telling van één vorm per hand laat uitschijnen.

Het is evident dat een kwantitatieve reductie van het in sommige gebieden al schaarse materiaal de betrouwbaarheid van de percentages vermindert. Drie van de vier bij telmethode 4 resterende Limburgse vormen van *ridder* in de nominatief singularis hebben een *-e-* als auslautklinker; bij de andere vorm is de slotsjwa gedeleerd. Indien één attestatie meer zou worden

¹⁰ Het absolute aantal vindplaatsen van *ridder* in de nom. sing. volgens de vier tellingen bedraagt resp.: alle vindplaatsen: 479x; alle originele vindplaatsen: 419x; 1 vorm per document: 333x, 1 vorm per hand: 179x.

¹¹ In Holland en ook in Limburg blijft slechts ca. een kwart van het beschikbare materiaal over, elders is dat ca. een derde. Het aantal vindplaatsen dat in aanmerking is genomen voor de telling zakt nl. in Holland van 76 naar 20, in Limburg van 19 naar 4. In Vlaanderen blijven er van de 328 attestaties 121 over, in Brabant 21 van de 39, in Utrecht 11 van de 14.

gevonden, verandert het beeld sterk: in dat geval wordt de slotsjwa in Limburg ofwel haast even vaak wel dan niet gedeleerd (2/5=40%), ofwel heeft in een vijfde van de gevallen *-e*-apocope (1/5=20%) plaatsgehad. Wanneer men daarentegen aan de 19 in totaal overgeleverde Limburgse vormen van *ridder* in de nom. sing. er enkele zou toevoegen, veranderen de verhoudingen niet zo drastisch. De werkwijze waarbij alle beschikbare vindplaatsen worden meegerekend, geeft een nauwkeuriger idee van de mate waarin de *-e*-apocope zich in die periode in Limburg heeft doorgezet dan de werkwijze waarbij slechts een deel van het materiaal in de telling betrokken wordt.

De percentages op basis van een telling waarbij een groot deel van de beschikbare gegevens niet in aanmerking wordt genomen, worden niet enkel minder betrouwbaar doordat de doorgevoerde reductie van het materiaal vaak een gebrek aan gegevens veroorzaakt¹², ze boeten tevens in aan geloofwaardigheid doordat men een groot deel van de regelmatige flexievormen uitschakelt, waardoor de nadruk te sterk komt te liggen op de uitzonderingsgevallen. De observaties die niet worden gehonoreerd wanneer men slechts één vorm per hand of per document telt, zijn immers niet de varianten¹³, maar wel de courante vormen. Het is evident dat de varianten niet uit de selectie mogen verdwijnen: de vormvarianten kunnen immers het begin of het einde van een morfologische verandering aangeven; de frequentieverhoudingen tussen de realisatietypes indiceren in welke richting het flexiesysteem evolueert. Wij hebben bezwaren tegen het feit dat, wanneer slechts één vorm per hand of per document wordt geteld, het overgrote deel van de courante attestaties uit het gegevensbestand verdwijnt. Door alle taalvormen tot één vorm per hand te herleiden, ontstaat er een proportionele wanverhouding tussen de frequente, normaal gebruikte vorm, die de feitelijke Middelnederlandse toestand weergeeft, en de minder gebruikelijke vorm, die wellicht een verandering in het buigingspatroon inluidt of een relictoestand belichaamt.¹⁴

Een fictief voorbeeld, waarbij de attestaties van twee scribenten worden verrekend, kan deze stelling verduidelijken. Van scribent A zijn in totaal

¹² Dées (e.a. 1980: XII) beschouwt een aantal van zes taalvormen als de minimale hoeveelheid om betrouwbare percentages te bekomen. Ook Van Reenen wijst erop dat per plaats of per regio een redelijk aantal getuigenissen beschikbaar moet zijn: "Zes handen is het absolute minimum bij een binaire probleemstelling. Twintig tot dertig handen leiden pas tot echt betrouwbare resultaten per plaats of per streek." (Van Reenen 1986-1987: 134).

¹³ Die worden dubbel geteld door Berteloot 1984a en 1984b, proportioneel verdeeld per scribent door Van Reenen 1986-1987.

¹⁴ Ook Mooijaart wijst erop dat hapaxen niet even zwaar dienen te worden gewaardeerd als de gewone schrijfwijze (Mooijaart 1992: 25).

tien vormen van een bepaald woord overgeleverd: negen ervan hebben vorm X, één vindplaats heeft vorm Y. Scribent B, die minder productief is geweest, heeft slechts twee vormen neergeschreven: eenmaal vorm X, eenmaal vorm Y. Bij een telling volgens het principe “één vorm per hand” worden de negen attestaties X van scribent A slechts als één bewijsplaats gerekend, de acht overige worden niet in aanmerking genomen. Deze schrijver, die vrij consequent dezelfde vorm gebruikt, ziet die vorm slechts eenmaal gehonoreerd. Het is van geen tel hoe vaak hij vorm X heeft opgetekend, waarvan we, aangezien hij hem zo vaak gebruikt, mogen aannemen dat hij de gangbare flexie weerspiegelt.¹⁵ Scribent B, van wie ons slechts twee vormen bekend zijn, ziet beide vormen elk ook eenmaal gevaloriseerd. Door acht regelmatige vormen niet mee te rekenen, krijgen bij het samentellen van de vindplaatsen van beide scribenten de twee uitzonderingen Y te veel gewicht.¹⁶

Wanneer men voor ogen houdt dat bij een telling van één vorm per hand de minder gebruikelijke vormen verhoudingsgewijs te veel benadrukt worden, kan uit een vergelijking van de percentages van de verschillende tellingen worden afgeleid welke van de vormvarianten in het offensief resp. defensief zijn. Zo moet het *-s*-meervoud bij het substantief *ridder* (nom. plur.) in de kustgebieden in de 13de eeuw bijzonder gebruikelijk zijn geweest: bij een telling van het absolute aantal vindplaatsen van de nom. plur. van *ridder*, behaalt de *-s*-pluralis in het Vlaams en het Hollands hogere scores dan wanneer slechts één vorm per hand wordt meegerekend.¹⁷ De *-e*-apocope is in het Hollands in de 13de eeuw daarentegen een vrij nieuw verschijnsel: de cijfers voor de *-e*-loze nom. sing. van *ridder* dalen immers bij een telling van alle vindplaatsen¹⁸, die betrouwbaarder is wat de verhouding tussen de vormvarianten betreft. Anderzijds bevestigt het hogere percentage voor de uitgangsloze nom. sing. *ridder* op basis van alle vindplaatsen in het Utrechts¹⁹ dat deze regio met betrekking tot de apocoperingstendens het voortouw genomen heeft.

¹⁵ Cfr. Goossens 1974: 23: “Het is aan te nemen dat “Normalschreibungen” een veel hogere frequentie hebben dan de “Sonderschreibungen”.”

¹⁶ Volgens de telling van het absolute aantal vindplaatsen: vorm Y: 2/12=17%. Volgens de telling van één vorm per hand: vorm Y: 2/4=50%.

¹⁷ Bij een telling van alle vindplaatsen heeft 97.6% van de Vlaamse en 69.8% van de Hollandse attestaties de meervoudsuitgang *-s*. Bij de telwijze van één vorm per hand, dalen deze cijfers tot 94.7% in het Vlaams resp. 56.3% in het Hollands.

¹⁸ Nl. van 60% naar 44.7% (Hollands).

¹⁹ Nl. 78.6% in plaats van 72.7%.

In de datief sing. van *parochie* komt het effect van het tellen van 1 vorm per document of per hand, vergeleken met het tellen van alle vindplaatsen, het best tot uiting in het Brabants. Het Brabants behoort tot het homogene oostelijke gebied waarin de datief sing. van de feminina op *-e* geflecteerd wordt met de uitgang *-en*. Volgens telmethode 1 eindigt niet minder dan 96.7% van de Brabantse vindplaatsen van de datief sing. op *-en*. Bij het niet-valoriseren van bepaalde attestaties bij telmethode 3 en 4 verdwijnen een aantal regelmatige vormen uit de selectie. De ene attestatie op *-e* krijgt daarentegen te veel belang: de percentages voor de variant op *-en* dalen tot 88.9% (telmethode 3) resp. 87.5% (telmethode 4).²⁰

P. van Reenen heeft Berteloots methode waarbij één vorm per hand wordt geteld en dus niet de taalvorm, maar de scribent centraal staat, gesofisticeerd: wanneer een hand intern variatie vertoont, wordt het getal van één hand proportioneel verdeeld volgens de verhouding tussen de absolute aantallen voorkomens van de varianten²¹: als een schrijvershand bijv. op een totaal van drie attestaties tweemaal de ene en eenmaal de andere vorm gebruikt, wordt deze verhouding gekwantificeerd als 0.66 voor de ene versus 0.33 voor de andere vorm. Alle frequentieverhoudingen worden op die manier herleid tot één, waardoor de toevalligheid van het absolute aantal vindplaatsen per schrijvershand wordt geneutraliseerd.²²

Met deze methode voorkomt men echter evenmin dat het taalgebruik van één bepaalde scribent het resultaat te veel beïnvloedt. Het taalgebruik van de schrijvershand die weinig heeft geschreven en van de scribent die veel heeft geschreven, weegt even zwaar door in het totaal. Bij het samentellen van de resultaten in het voorbeeld van beide scribenten A en B volgens de methode van Van Reenen wordt de ongelijke nauwkeurigheid van beide getallen, waarvan het ene op een groot absoluut aantal gegevens, het andere op een klein absoluut aantal taalvormen is gebaseerd, uit het oog verloren.²³ Er zijn nogal wat scribenten die niet veel hebben geschreven²⁴, zodat het gevaar

²⁰ Het percentage van 96.7% vormen op *-en* in het Brabants bij telmethode 1 is gebaseerd op 30 vindplaatsen: 29x *-en*, 1x *-e*. Bij telmethode 3 resp. 4 worden slechts 9 resp. 8 attestaties in de telling betrokken. Het percentage voor de ene afwijkende vorm op *-e* stijgt daardoor tot 11.1% resp. 12.5%.

²¹ Voorgestelde werkwijze door Van Reenen 1986-1987, in reactie op Berteloot 1984a en 1984b.

²² Berteloot 1984b: 36; Van Reenen 1986-1987: 144.

²³ De telling volgens Van Reenens methode: vorm Y: $(0.1+0.5)/2=30\%$ (vgl. de resultaten van de andere tellingen in voetnoot 16).

²⁴ Cfr. Van Reenen 1986-1987: 142: “De gegevens moeten in voldoende aantallen beschikbaar zijn. Helaas zijn ze dat vaak niet. Slechts in vier plaatsen zijn documenten van

bestaat dat bij het samentellen van de scores van alle scribenten de bij deze manier van tellen ingebouwde onnauwkeurighedsfactor het eindresultaat vertekent.

Mijn motivatie om ook de attestaties uit afschriften, vidimi, vertalingen en falsificaties in de telling op te nemen en zodoende werkwijze 2 terzijde te schuiven ten voordele van methode 1²⁵ vloeit voort uit de centrale positie die de individuele taaluiting inneemt: iedere taalvorm, dus ook een onoorspronkelijke attestatie, is op zich betekenisvol en dient statistisch op identieke wijze te worden gevaloriseerd. Het uitsluiten van vindplaatsen uit afschriften e.d. brengt een reductie van het soms karige beschikbare materiaal teweeg. Vooral de conclusies met betrekking tot het Hollands, dat met zijn vele afschriften uit het cartularium van de graven van Holland de meeste vindplaatsen zou verliezen, boeten in aan betrouwbaarheid indien enkel de vindplaatsen uit de originele documenten worden betrokken in het onderzoek. Het is m.i. dan ook niet toevallig dat het cijfersverschil tussen de telling volgens methode 1 en volgens methode 2 in de nom. sing. en de nom. plur. van *ridder* in het Hollands groter is dan elders.²⁶

Het feit dat de twee telmethodes waarbij wordt uitgegaan van de afzonderlijke taaluiting en niet van de scribent of het document, doorgaans minimale procentuele verschillen opleveren, is niet onlogisch: men mag veronderstellen dat de frequentieverhoudingen tussen de vormvarianten binnen een afschrift en binnen het originele document grotendeels dezelfde zijn. Wanneer dit niet zo is, is er variatie in het geding. Het feit dat een afgeschreven vorm afwijkt van de originele taalvorm, toont aan dat de kopiist de aangeboden originele taalvorm zelfstandig heeft geïnterpreteerd. Een afwijkende vorm in het afschrift is een interessant geval: hij zou op een wijziging in het buigingspatroon kunnen wijzen en dient in ieder geval te worden meegerekend. Ook wanneer het origineel van een tekst niet is overgeleverd,

meer dan 20 handen beschikbaar [...], slechts in 14 plaatsen 6 tot 20 handen. Verreweg de meeste plaatsen zijn maar door een, of twee tot vijf, handen vertegenwoordigd.”

²⁵ Doorgaans wordt het gebruik van afschriften van oorkonden voor taalgeografisch onderzoek afgeraden, zie Berteloot 1984b: 29; Moors 1952: 4-5.

²⁶ In het Hollands neemt, wanneer men de vormen uit afschriften elimineert, het aantal geapocopeerde vormen in de nom. sing. van *ridder* toe met 2.5% (van 44.7% naar 47.2%). In het Vlaams stijgt het overeenstemmende cijfer slechts met drie tiende procent (van 8.5% naar 8.8%). Ook in de nom. plur. worden in het Hollands de tegenstellingen lichtjes uitvergroot door de reductie van materiaal: het aantal -s-meervouden vermindert volgens telmethode 2 met 5.5% (van 69.8% naar 64.3%). In het Vlaams bedraagt het verschil slechts 0.8% (van 97.5% naar 94.7%). In de overige dialectgebieden komt het substantief *ridder* niet in niet-originele documenten voor, zodat telmethode 2 hier niet van toepassing is.

krijgen de vindplaatsen uit de wel bewaarde kopie in onze benadering bijzonder belang.²⁷

Conclusie

Alle besproken telmethodes, zowel het zuiver statistisch tellen van alle vindplaatsen, al dan niet uit de originele documenten, als de selectie van één vorm per document of per hand, leiden naar de waarheid. De frequentieverhoudingen tussen de vormvarianten bij de diverse telmethodes, toegepast op twee concrete gevallen, verschillen slechts miniem, meestal minder dan 10%. Alle studies waarin één van die telmethodes is gebruikt, leveren dan ook statistisch betrouwbare resultaten op. Elke methode heeft zijn eigen verdiensten.

Om drie redenen krijgt de methode waarbij alle vindplaatsen uit alle documenten worden geteld, niettemin mijn voorkeur:

1. Bij tellingen waarbij niet van het absolute aantal vindplaatsen wordt uitgegaan, wordt weliswaar de willekeur met betrekking tot het aantal vormen per hand en per document uitgeschakeld, maar aan de uitzonderingsgevallen wordt overdreven veel gewicht toegekend. Doordat de vormen die door een scribent herhaaldelijk zijn neergeschreven en dan ook de gangbare flexie weergeven, in de telling herleid worden tot één, krijgen de afwijkingen, die minder vaak zijn geattesteerd, meer belang dan ze in werkelijkheid zullen hebben gehad. Wanneer het absolute aantal vindplaatsen wordt geteld, neemt het procentueel aandeel van afwijkende ten opzichte van regelmatige vormen de juiste proporties aan.
2. De omvangrijkste materiaalverzameling biedt de beste garantie om tot statistisch nauwkeurige resultaten te komen. Elke reductie van het beschikbare materiaal vermindert de betrouwbaarheid van de resultaten of maakt het zelfs onmogelijk om conclusies te trekken. Bij de methode om slechts één vorm per hand of per document te tellen wordt slechts een fractie van het beschikbare materiaal werkelijk gebruikt.²⁸ Als de taal-

²⁷ Cfr. Moors 1952: 5: “Op historisch gebied zijn ze [= de charterboeken] wel interessant, omdat ze veel documenten bevatten waarvan het origineel verloren is [...]”.

Een afschrift waarvan het origineel niet voorhanden is, is bijv. het Brugse document 724.

²⁸ Wanneer men bij deze methode ook effectief minder vindplaatsen zou moeten ontleden, zou men sneller tot linguïstische resultaten kunnen komen. Het tijdbesparende effect dat deze werkwijze zou kunnen hebben, is evenwel slechts schijn: om één vorm per hand over te houden, dienen even goed alle gegevens bekeken te worden, waarna een groot deel ervan niet in aanmerking wordt genomen voor de telling.

vormen voldoende vaak zijn opgetekend, bijv. bij *ridder*, *parochie*, beïnvloedt het geringere aantal gegevens de resultaten niet merkbaar. Bij nomina die minder frequent zijn overgeleverd, worden de resultaten daarentegen snel minder betrouwbaar. Bij een klein aantal vindplaatsen kan een miniem verschil in de hoeveelheid attestaties de frequentieverhoudingen aanzienlijk wijzigen.

3. Het tellen van alle vindplaatsen is statistisch veruit de eenvoudigste werkwijze. Ook om die reden geniet deze methode mijn voorkeur.

Bibliografie

- BERTELOOT, A.
1984a *Bijdrage tot een klankatlas van het dertiende-eeuwse Middelnederlands*. Deel I: Tekst, Deel II: Platen. Gent.
- BERTELOOT, A.
1984b "Overwegingen bij de "lieden/luden"-kaart." *TNTL* 100, 29-45.
- DÉES A. & P. TH. VAN REENEN & J. DE VRIES
1980 *Atlas des formes et des constructions des chartes françaises du 13e siècle*. Tübingen.
- DEMEERSSEMAN, H.
1988 *De meervoudsflexie van drie substantieven in veertiende-eeuwse Middelnederlands. Een dialectgeografische studie op het corpus van veertiende-eeuwse ambtelijke teksten*. Onuitgegeven doctoraalscriptie VU, Amsterdam, 1988.
- GOOSSENS, J.
1974 "Historische en moderne taalgeografie". In: A. Van Loey & J. Goossens: *Historische dialectologie*. Amsterdam, 14-33.
- GOOSSENS, J.
1979 "De ambtelijke teksten van het Corpus Gysseling." *TNTL* 95, 247-261.
- MARYNISSSEN, A.
1996 *De flexie van het substantief in het 13de-eeuwse ambtelijke Middelnederlands. Een taalgeografische studie*. "Studies op het gebied van de Nederlandse Taalkunde" nr 2, BICN. Leuven.
- MOOJAART, M.A.
1992 *Atlas van Vroegmiddelnederlandse Taalvarianten*. Utrecht.
- MOORS, J.
1952 *De oorkondentaal in Belgisch-Limburg van circa 1350-1400*. Brussel.
- REENEN, P. VAN
1986-87 "De lange weg naar een betrouwbare en systematische beschrijving van het Middelnederlands." *Spektator* 16, 131-148.

Bijlage

Telling 1: alle vindplaatsen uit alle documenten

ridder

nominatief singularis: *Simpl.*: ridder

Regio:	Vla.	Holl.	Brab.	Utr.	Limb.	B.-V.	H.-B.	
-Ø:	28 8.5%	34 44.7%	1 2.6%	11 78.6%	4 21.1%	0 0.0%	1 100.0%	79 16.5%
-e:	290 88.4%	42 55.3%	38 97.4%	3 21.4%	15 78.9%	2 00.0%	0 0.0%	390 81.4%
-a:	10 3.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10 2.1%
Tot:	328	76	39	14	19	2	1	479

nominatief pluralis: *Simpl.*: ridder

Regio:	Vla.	Holl.	Brab.	Utr.	Limb.	
-Ø:	0 0.0%	1 2.3%	0 0.0%	0 0.0%	2 28.6%	3 1.1%
-e:	0 0.0%	6 14.0%	0 0.0%	1 100.0%	5 71.4%	12 4.4%
-s:	206 97.6%	30 69.8%	2 22.2%	0 0.0%	0 0.0%	238 87.8%
-en:	5 2.4%	6 14.0%	7 77.8%	0 0.0%	0 0.0%	18 6.6%
Tot:	211	43	9	1	7	271

parochie

datief singularis: *Simpl.*: parochie

Regio:	W.Vl.	O.Vl.	Holl.	Brab.	
-e:	290 78.0%	3 15.8%	0 0.0%	1 3.3%	294 69.7%
-en:	82 22.0%	16 84.2%	1 100.0%	29 96.7%	128 30.3%
Tot:	372	19	1	30	422

Telling 2: alle vindplaatsen uit de originele documenten

ridder*nominatief singularis: Simpl.: ridder*

Regio:	Vla.	Holl.	Brab.	Utr.	Limb.	B.-V.	H.-B.	
-Ø:	27	17	1	11	4	0	1	61
	8.8%	47.2%	2.6%	78.6%	21.1%	0.0%	100.0%	14.6%
-e:	273	19	38	3	15	2	0	350
	88.6%	52.8%	97.4%	21.4%	78.9%	100.0%	0.0%	83.5%
-a:	8	0	0	0	0	0	0	8
	2.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.9%
Tot:	308	36	39	14	19	2	1	419

nominatief pluralis: Simpl.: ridder

Regio:	Vla.	Holl.	Brab.	Utr.	Limb.	
-Ø:	0	0	0	0	2	2
	0.0%	0.0%	0.0%	0.0%	28.6%	0.8%
-e:	0	4	0	1	5	10
	0.0%	14.3%	0.0%	100.0%	71.4%	4.0%
-s:	198	18	2	0	0	218
	97.5%	64.3%	22.2%	0.0%	0.0%	87.9%
-en:	5	6	7	0	0	18
	2.5%	21.4%	77.8%	0.0%	0.0%	7.3%
Tot:	203	28	9	1	7	248

parochie*datief singularis: Simpl.: parochie*

Regio:	W.Vl.	O.Vl.	Brab.	
-e:	280	3	1	284
	77.8%	15.8%	4.2%	70.5%
-en:	80	16	23	119
	22.2%	84.2%	95.8%	29.5%
Tot:	360	19	24	403

Telling 3: 1 vorm per origineel document

ridder*nominatief singularis: Simpl.: ridder*

Regio:	Vla.	Holl.	Brab.	Utr.	Limb.	B.-V.	H.-B.	
-Ø:	22	17	1	9	1	0	1	51
	8.4%	60.7%	4.3%	75.0%	16.7%	0.0%	100.0%	15.3%
-e:	232	11	22	3	5	1	0	274
	88.5%	39.3%	95.7%	25.0%	83.3%	100.0%	0.0%	82.3%
-a:	8	0	0	0	0	0	0	8
	3.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.4%
Tot:	262	28	23	12	6	1	1	333

nominatief pluralis: Simpl.: ridder

Regio:	Vla.	Holl.	Brab.	Utr.	Limb.	
-Ø:	0	0	0	0	1	1
	0.0%	0.0%	0.0%	0.0%	33.3%	0.4%
-e:	0	3	0	1	2	6
	0.0%	17.6%	0.0%	100.0%	66.7%	2.7%
-s:	192	10	2	0	0	204
	98.0%	58.8%	33.3%	0.0%	0.0%	91.5%
-en:	4	4	4	0	0	12
	2.0%	23.5%	66.7%	0.0%	0.0%	5.4%
Tot:	196	17	6	1	3	223

parochie*datief singularis: Simpl.: parochie*

Regio:	W.Vl.	O.Vl.	Brab.	
-e:	254	1	1	256
	78.9%	14.3%	11.1%	75.7%
-en:	68	6	8	82
	21.1%	85.7%	88.9%	24.3%
Tot:	322	7	9	338

Telling 4: 1 vorm per hand uit de originelen

ridder*nominatief singularis: Simpl.: ridder*

Regio:	Vla.	Holl.	Brab.	Utr.	Limb.	B.-V.	H.-B.	U.-B.	
-Ø:	10 8.3%	12 60.0%	1 4.8%	8 72.7%	1 25.0%	0 0.0%	1 100.0%	0 -	33 18.4%
-e:	104 86.0%	8 40.0%	20 95.2%	3 27.3%	3 75.0%	1 100.0%	0 0.0%	0 -	139 77.7%
-a:	7 5.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 -	7 3.9%
Tot:	121	20	21	11	4	1	1	0	179

nominatief pluralis: Simpl.: ridder

Regio:	Vla.	Holl.	Brab.	Utr.	Limb.	
-Ø:	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 50.0%	1 1.0%
-e:	0 0.0%	3 18.8%	0 0.0%	1 100.0%	1 50.0%	5 5.0%
-s:	71 94.7%	9 56.3%	2 33.3%	0 0.0%	0 0.0%	82 82.0%
-en:	4 5.3%	4 25.0%	4 66.7%	0 0.0%	0 0.0%	12 12.0%
Tot:	75	16	6	1	2	100

parochie*datief singularis: Simpl.: parochie*

Regio:	W.Vl.	O.Vl.	Brab.	
-e:	62 77.5%	1 20.0%	1 12.5%	64 68.8%
-en:	18 22.5%	4 80.0%	7 87.5%	29 31.2%
Tot:	80	5	8	93