

Geslacht en Congruentie in het Moderne Nederlands¹

Mathieu Klomp (0469254)

Masterscriptie voor Taal & Ontwikkeling, Universiteit Utrecht

Begeleid door J.M. van Koppen en F.A.C. Drijkoningen

Augustus 2009

Inleiding

Het Nederlands heeft een verval van woordgeslacht ondergaan dat ertoe geleid heeft dat de taal nu twee geslachten kent, terwijl de Germaanse talen zich traditioneel van drie geslachten bedienden. In het Duits bijvoorbeeld zijn deze drie woordgeslachten tot op heden bewaard gebleven. In het Nederlands zijn het mannelijk en vrouwelijk geslacht samengegaan in één nieuw geslacht wat ik hier als “zijdig” zal aanduiden. Woorden met dit geslacht worden in de volksmond “*de*-woorden” genoemd, daar *de* het bepaald lidwoord is voor woorden met dit geslacht. Hier tegenover staan de onzijdige woorden, de “*het*-woorden”.

De Scandinavische talen van het vasteland hebben een vergelijkbare ontwikkeling naar twee woordgeslachten ondergaan, de ontwikkeling is in deze talen echter samengegaan met het ontstaan van nieuwe voornaamwoorden met overeenkomend geslacht om naar antecedenten te verwijzen. De twee aanwijzend voornaamwoorden voor het enkelvoud maken enkel onderscheid tussen mannelijk/vrouwelijk enerzijds en onzijdig anderzijds, precies zoals dat gebeurt bij het Nederlandse *die* en *dat*. De Scandinavische voornaamwoorden hebben zich ontwikkeld tot persoonlijk voornaamwoorden waardoor er een uitgebreider paradigma is ontstaan dat onderscheid maakt tussen biologisch mannelijk (Ndl.: *hij*) en vrouwelijk (Ndl.: *zij*), en daarnaast tussen grammaticaal zijdig (Ndl.: *die*) en onzijdig (Ndl.: *dat*). De mannelijke en vrouwelijke voornaamwoorden worden dus niet langer in verband met grammaticaal geslacht toegepast.

In het Nederlands heeft zich tot op heden niet in deze richting ontwikkeld. Het Nederlands kent nog altijd drie genera op voornaamwoorden: mannelijk, vrouwelijk en onzijdig. Dit levert een interessant congruentiedilemma op: er is geen één-op-éénrelatie meer tussen genera op voornaamwoorden en lexicale generea. Anders gezegd, het Nederlands beschikt over drie genera op voornaamwoorden om naar twee genera op zelfstandig naamwoorden te verwijzen. De manier waarop dit probleem opgelost wordt is veel minder voor de hand liggend dan de Scandinavische oplossing. In de recente literatuur wordt geconstateerd dat vrijwel uitsluitend mannelijke en onzijdige voornaamwoorden gebruikt worden, waarbij de keuze niet langer gestuurd wordt door het woordgeslacht maar door een nieuw semantisch onderscheid. Dit is een onderscheid op basis van telbaarheid of individuatie (Audring, 2006a & 2006b; De Vogelaer 2006). Grofweg komt het erop neer dat het mannelijke voornaamwoord wordt gebruikt om naar telbare of hoog-geïndividueerde nomina te verwijzen, het onzijdige voornaamwoord om naar niet-telbare of laag-geïndividueerde nomina te verwijzen.

Ik zal in sectie 1 dit nieuwe semantische systeem nauwkeurig beschrijven. Sectie 2 is gewijd aan de verhouding tussen het oude en het nieuwe systeem in de taal. Met deze theoretische

¹ Ik ben veel dank verschuldigd aan Marjo van Koppen voor de begeleiding, het geven van commentaar, terugkoppeling en literatuur; Jenny Audring en Johan Rooryck voor hun hulp en literatuur; Arnout Koornneef voor het opzetten van het concept voor het experiment dat dit onderzoek bevat; Jesper Tijmstra voor het assisteren en adviseren bij de statistische analyse van de resultaten van het experiment; alle taalkundig geïnteresseerde moedertaalsprekers van verschillende talen waarmee ik onderdelen uit deze scriptie heb bediscussieerd; en alle proefpersonen die hebben meegedaan aan het onderzoek.

achtergrond zal ik vervolgens in sectie 3 doormiddel van een experiment onderzoeken in welke facetten van de taal nieuwe systeem is doorgedrongen, en waar we nog steeds het oude systeem vinden. Sprekers geven namelijk (nog) niet in alle contexten de voorkeur aan het nieuwe systeem, ondanks het feit dat het oude systeem in sommige contexten defectief is. Ook is het interessant om te inventariseren of het nieuwe systeem toepassing geniet in contexten waar het oude systeem niet defectief is.

1 Het Semantische Systeem

1.1 De basis

In het traditionele lexicale systeem is de voornaamwoordkeuze gebaseerd op het woordgeslacht van het naamwoord. In (1a) een illustratie van het gebruik van het zijdige *deze* tegenover het onzijdige *dit* in het traditionele Nederlands – de taalvorm waarvan we ons tegenwoordig nog bedienen in de formele schrijftaal.

In de moderne spreektaal zien we echter ook een ander systeem. Dit moderne systeem lijkt de keuze voor een persoonlijk voornaamwoord te berusten op telbaarheid of mate van individuatie van de referent. Naar telbare of hooggeïndividueerde referenten wordt verwezen met mannelijke voornaamwoorden terwijl ontelbare of laaggeïndividueerde referenten, zoals stoffen, samengaan met onzijdige voornaamwoorden (Audring, 2006; De Vogelaer 2006). In (1b) een illustratie van het gebruik van het mannelijke *hij* tegenover het onzijdige *het*.

- (1) a. **Deze/dit** is niet te koop.

(schrijftaal)

| Lexicaal zijdig (Mannelijk/Vrouwelijk) → Deze | Lexicaal onzijdig → Dit |
|--|----------------------------|
| De tafel | Het geld |
| De lucht | Het bureau |
| De liefde | Het gras |
| De lach | Het idee |

- (1) b. **Hij/het** is niet te koop.

(spreektaal)

| Telbaar / hooggeïndividueerd → Hij | Niet telbaar / laaggeïndividueerd → Het |
|---------------------------------------|--|
| De tafel | Het geld |
| Het bureau | De lucht |
| Het idee | Het gras |
| De lach | De liefde |

In tabel (1a) is het woordgeslacht bepalend voor de voornaamwoordkeuze. In de tabel in (1b) zien we geen correlatie met woordgeslacht. Ook vinden we in beide categorieën zowel

concrete (*de tafel, het geld*) als abstracte referenten (*de lach, de liefde*). Het verschil is enkel dat we in de categorie voor mannelijke voornaamwoorden referenten vinden die telbaar zijn (of deze nu abstract zijn of niet), terwijl de categorie voor onzijdige voornaamwoorden gevuld wordt door referenten die niet-telbaar zijn. Dit kan zijn omdat het om een stof gaat, of omdat de referent een bepaalde eigenschap van “vaagheid” heeft waardoor het niet op te vatten is als een telbare entiteit. Deze omschrijving is nog niet precies genoeg. In deze sectie poog ik tot een heldere definitie te komen van de factoren waar het semantische systeem op berust.

1.2 Het systeem volgens Audring (2006, 2009)

Audring (2006a & 2006b) beargumenteert dat de factor die het onderscheid maakt tussen de woorden in de linkerkolom en de rechterkolom in (1) “graad van individuatie” is. Indivduatie is een overkoepelend begrip waar zowel bezielheid, telbaarheid als specificiteit deelaspecten van zijn. Als een lexicaal item beziel, telbaar en/of specifiek is, zou dit leiden tot de keuze voor een mannelijk voornaamwoord. De tegenovergestelde specificaties, onbeziel, ontelbaar en/of niet-specifiek, leiden tot een onzijdig voornaamwoord. Audring hanteert een hiërarchie met verschillende gradaties:

| | | | | | |
|-----|--------------------------------|-----------------|--------------------------------|-----------------|--|
| (2) | Human > | Other animate > | Bounded object / Abstract > | Specific mass > | Unspecific mass, Unbounded abstract |
| | <i>feminine/ masculine</i> | | <i>masculine</i> | | <i>Neuter</i> |

De hiërarchie loopt van hoge mate van individuatie aan de linkerkant naar lage mate van individuatie aan de rechterkant. We zullen de weergave preciezer analyseren. De termen “human” en “animate” hebben betrekking op bezielheid (afnemend naar rechts), “bounded”, “specific”, “unbounded”, “unspecific” hebben betrekking op specificiteit (afnemend naar rechts), “object” en “mass” hebben betrekking op telbaarheid (afnemend naar rechts), “object” en “abstract” hebben tenslotte betrekking op niveau van abstractie. Abstractie is de enige genoemde factor die als irrelevant beschouwd wordt, en dus ook geen heterogene verdeling over de categorieën van de hiërarchie laat zien. De hiërarchie verloopt dus van [+beziel][+telbaar][+specifiek] naar [-beziel][-telbaar][-specifiek], waarbij de afname van de drie parameters niet synchroon loopt. De corresponderende voornaamwoordkeuze verloopt, volgens Audring, van 100% kans op mannelijk (links) naar 100% kans op onzijdig (rechts). Daarbij moet de kanttekening worden geplaatst dat bij de hoogste graad van individuatie, waar we [+beziel] vinden (hieronder vallen mensen, sommige dieren en andere referenten met mensachtige trekken) er congruentie met de biologische sekse van het antecedent beschikbaar komt. Hierdoor vinden we bij de hoogst-geïndividueerde referenten naast mannelijke ook vrouwelijke voornaamwoorden.

1.3 Aanmerkingen op Audring (2006a)

1.3.1 Het gebruik van een graduele hiërarchie

Voordat ik mijn eigen analyse geef, zal ik ingaan op verschillende aspecten van de hierboven beschreven analyse. Audring (2006a) heeft in het Corpus Gesproken Nederlands een set

antecedent-voornaamwoordcombinaties bekeken waarvoor het traditionele lexicale systeem en het nieuwe semantische systeem verschillende geslachten vereisen. Naar *het boek* zou bijvoorbeeld volgens het lexicale systeem verwezen worden met *het* omdat *boek* een onzijdig woord is, maar volgens het semantische systeem met *hij*, aangezien het in de terminologie van Audring (2006a) een vrij hoog-geïndividueerde referent betreft (niet beziel, maar wel telbaar, en wel specifiek). Er wordt gekeken hoe vaak de keuze gebaseerd is op het semantische systeem, en hoe vaak deze is gebaseerd op het lexicale systeem. Vervolgens geeft Audring percentages van het aandeel verwijzing volgens het nieuwe, semantische systeem. Als we deze data willen interpreteren in relatie tot een graduele schaal zoals in (2) blijkt dit echter problematisch. Immers, vrijwel alle antecedent-voornaamwoordcombinaties, waaronder *het boek-hij/het*, bevinden zich niet in een absoluut uiterste van de hiërarchie, maar ergens in het midden van de schaal. De hiërarchie vertelt ons hooguit wat de kans is dat een voornaamwoordgeslacht in de praktijk voorkomt, maar wijst geen specifiek voornaamwoordgeslacht aan dat we onder alle omstandigheden dienen aan te treffen om te kunnen bevestigen dat een spreker het semantische systeem heeft toegepast. Voor ieder item in de set is in het onderzoek van Audring (2006a) *ad hoc* ingeschat of het semantische systeem een mannelijk of een onzijdig voornaamwoord zou voorschrijven, om vervolgens te beoordelen of de spreker zich van dit systeem bediend heeft. Het semantisch systeem wordt in het onderzoek dus gebruikt om een binair onderscheid te maken, terwijl de schaal een gradueel karakter beschrijft. De indruk wordt gewekt dat het gekozen omslagpunt om te gaan van een mannelijk naar een onzijdig voornaamwoord ligt tussen *bounded object* / *abstract* (de laatste categorie waar de verwachting zijdig is), en *specific mass* (de eerste categorie waar de verwachting onzijdig is). Voor het vaststellen van een percentage is dus een binair onderscheid nodig en dat is onverenigbaar met een hiërarchie met een gradueel karakter.

Het hoofddoel van deze sectie is om een model voor het semantische systeem te formuleren. Een graduele schaal helpt ons hier niet veel verder: het betreft een kanselement, dat we niet verder kunnen analyseren. Wanneer een antecedent zich bijvoorbeeld in de buurt van het linker uiterste van de schaal bevindt zal de kans op een mannelijk voornaamwoord groot zijn, maar het voorkomen van een onzijdig voornaamwoord wordt niet uitgesloten. Wanneer we bij dit antecedent dan toch een onzijdig voornaamwoord aantreffen, kunnen we de reden achter die keuze niet verder doorgronden dan constateren dat er een bepaalde kans op is. We zijn niet in staat aan te geven wat nu werkelijk de logica is geweest achter de voornaamwoordkeuze. Door het kanselement is de juistheid nooit te toetsen, en dus kan het niet dienst doen als model. De voorwaarde voor een model is immers dat het getoetst of gefalsificeerd kan worden. Nu is het wel zo dat de hiërarchie een nuttig hulpmiddel is geweest om een reeks observaties overzichtelijk te presenteren. Dit is ook wat Audring (2006a, b) beoogt: het presenteren van observaties, niet het formuleren van een model waarmee we (on)grammaticaliteit kunnen voorspellen. Daarom is dit geen punt van kritiek, enkel een constatering dat we de weergave niet als model kunnen toepassen.

1.3.2 De uitvoering van de graduele hiërarchie

Toch zijn er, ook wanneer we de vraag over de toepasbaarheid van een hiërarchie even buiten beschouwing laten, enkele problemen met de gekozen uitvoering. Individuatie is zoals genoemd een begrip waar zowel beziel, telbaarheid als specificiteit binnen vallen. Echter, omdat deze parameters niet onderling afhankelijk zijn zou dat tot een multidimensionale schaal leiden met één begrip langs iedere as. In (2) worden deze factoren echter enigszins stipulatief naast elkaar gezet langs één as. Dit is alleen gerechtvaardigd wanneer de drie begrippen een logische overgang vormen. Dat is hier niet het geval. Zo staat *Specific mass* bijvoorbeeld voor [+specifiek][−telbaar] –

maar als we alle logisch mogelijke configuraties af gaan vinden we bijvoorbeeld ook iets wat we *unspecific object* zouden kunnen noemen, [-specifiek][+telbaar]. Vervolgens is het probleem dat de volgorde van deze twee in een eindimensionale schaal niet logisch volgt, aangezien beide parameterinvullingen in een eindimensionale schaal even hoog scoren met één [+] en één [-].

Overigens betreft Audring (2006a) naast de drie genoemde parameters ook familieverwantschap als invloed op de graad van individuatie. Het idee is dan ook niet zo zeer dat er een helder afgebakend aantal binaire parameters is, maar dat individuatie een fenomeen is dat op psychologisch vlak door een verscheidenheid aan afwegingen invulling kan krijgen. Maar deze vrijheid maakt het logisch ordenen van de data alleen maar moeilijker.

1.3.3 De gekozen factoren

1.3.3.1 De invloed van specificiteit

Naast het algemene probleem van gemende invloed van factoren is ook de vermeende invloed van één van die genoemde factoren, specificiteit, discutabel. We zullen de argumenten die Audring (2006a) noemt voor de relevantie van specificiteit nader bekijken. Beschouw het volgende voorbeeld uit het Corpus Gesproken Nederlands:

- (3) hij had een **werkvergunning** (z).
 want de afgelopen twee jaar of zo had-ie **dat** (n) niet gehad. (...)
 nou had-ie geld betaald en nou had-ie '**m**' (m) wel.

(uit het Corpus Gesproken Nederlands, uit Audring (2006a))

Hier is te zien hoe achtereenvolgens verwezen wordt met een onzijdig voornaamwoord en een mannelijk voornaamwoord. De referent is telbaar, en dus is enkel het mannelijke voornaamwoord verwacht. Audrings uitleg voor het voorkomen van zowel een mannelijk als een onzijdig voornaamwoord is dat het onzijdige "dat" verwijst naar een niet-specifieke werkvergunning, welke werkvergunning dan ook, en "m" vervolgens naar de werkvergunning waar de persoon in kwestie daadwerkelijk over beschikt. Die redenering wordt ook ondersteund met negatieve evidentie:

- (4) a. Hij heeft nu een partner (z). Dat (n) had-ie eerder niet gehad.
 b. Hij heeft zijn partner (z) meegenomen. *Dat (n) heeft-ie in Engeland ontmoet.

(uit Audring (2006a))

In (4a) zou *een partner* niet-specifiek zijn, waardoor vervolgens verwijzing met *dat* toegestaan is.² In (4b) is *zijn partner* specifiek, en is gebruik van *dat* ongrammaticaal. Audring (2006) merkt op:

Thus, the distal demonstrative dat has the particular function of referring to unspecific or generic referents. If non-specificity is viewed as an instance of low

² Het is overigens niet geheel duidelijk of dit in taalkunde zin daadwerkelijk een niet-specifieke referent is. De niet-specificiteit wordt in deze grammaticale context niet bevestigd.

individuation, the example supports rather than undermines the analysis proposed here. (p. 101)

Ik wil echter bepleiten dat specificiteit ook in deze gevallen geen rol speelt. Het onzijdige voornaamwoord *dat* in (4) lijkt niet te verwijzen naar een niet-specifieke werkvergunning, eerder naar een pragmatische notie, wat ik in brede zin zou formuleren als een staat van zijn; het “te maken hebben met”, in dit geval, de notie *werkvergunning*. Een andere mogelijkheid is dat *dat* totaal niet congrueert, zelfs niet met een referent buiten de werkelijkheid, en bij gebrek aan congruentiekenmerken als onzijdig aan de oppervlakte komt.

Het is ook niet zonder meer het geval dat wanneer we een niet-specifieke referent nemen, we dan ook inderdaad in staat zijn *dat* te gebruiken.

- (5) *Er staat een man (z) in de tuin. Dat (n) lijkt op mijn vader.

De referent in (5) is telbaar, maar niet-specifiek, getuige de existentiële constructie. Echter, het gebruik van een onzijdig voornaamwoord is evenwel uitgesloten. Beschouw ook de volgende ongrammaticale constructie:

- (6) *Ik zal z’n lievelingsbroek (z) wel voor hem strijken, zodat ie dat (n) straks kan aantrekken.

De keuze voor het onzijdige *dat* schendt zowel het lexicale als het semantische systeem: het lexicale systeem schrijft een zijdig voornaamwoord voor omdat *broek* een zijdig woord is, het semantische systeem schrijft ook een zijdig voornaamwoord voor, z’n *lievelingsbroek* is telbaar en specifiek. Als specificiteit een rol zou spelen, zou de zin goed of in ieder geval iets minder slecht worden als de referent niet-specifieke referent gemaakt wordt, zoals in (6’). Daar is naar mijn oordeel echter geen sprake van. Telbaarheid heeft in tegenstelling tot specificiteit, wel invloed – in (6’’) is de referent specifiek en niet-telbaar:

- (6’) *Ik zal wel een broek (z) voor hem strijken, zodat ie dat (n) straks kan aantrekken.³
 (6’’) Ik zal zijn kleding (z) voor hem strijken, zodat ie dat (n) straks kan aantrekken.

In (3) speelde specificiteit schijnbaar een rol, maar er zijn ook voorbeelden van hetzelfde type waar specificiteit geen bevredigende verklaring biedt. Hieronder enkele treffende voorbeelden gegeven door Audring (p.c.):

- (7) a. Die room kan ik wel hebben, dat fiets ik er wel weer van af.
 b. Pannenkoeken, dat is lekker.⁴

³ Om er zeker van te zijn dat de referent naast onbepaald ook daadwerkelijk non-specifiek is betrek ik het bijwoord “wel”. De positie van “wel” vóór het lijdend voorwerp geeft aan dat deze niet uit de VP verplaatst is, een kenmerk van non-specifieke argumenten.

⁴ Er wordt ook gepleit voor de analyse dat hier sprake is van elipsis, waarbij de referent eigenlijk “pannenkoeken eten” is, en dat “dat” dus een logische keuze is omdat het eigenlijk naar een zin verwijst. Deze analyse hang ik niet aan – “pannenkoeken, dat is lekker” vat ik op als analoog aan “een hamer, hebben we dat?” of “een werkvergunning, dat heeft-ie niet”. In deze gevallen kan er geen werkwoord weggelaten zijn, omdat “hebben” hier niet op een zin kan slaan. In het Zweeds is dit nog duidelijker zichtbaar: “Pannkakor är gott(onz.)” (“Pannenkoeken zijn/is lekker(onz.)”) bevat onzijdige inflectie, die niet

c. Een hamer, hebben we dat?

(7b) en (7c) bevatten (wellicht) een niet-specifieke referent, op basis waarvan de onzijdigheid van het voornaamwoord op het eerste gezicht verklaard zou kunnen worden. Toch is er in al deze voorbeelden iets vreemds aan de hand met de manier van verwijzen. Er geen directe lijn te trekken tussen referent en voornaamwoord: het is in (7a) niet *die room* als zodanig die de spreker *erof fietst*, maar het aangekomen gewicht als gevolg van het nuttigen van *die room*. En of pannenkoeken in (7b) nu wel of niet wordt opgevat als specifiek, de zin staat in het enkelvoud, iets wat we zelfs bij niet-specifieke argumenten in het meervoud niet verwachten (vgl. *Er staan mannen in de tuin*). De specificiteitsverklaring voor het onzijdige voornaamwoord blijft in (7c) in eerste instantie wel overeind, maar daar lijkt het eerder dat door een schijnreferent een pragmatische notie wordt opgeroepen (bijvoorbeeld “het over een hamer beschikken”) waarnaar vervolgens verwezen wordt. Ook hier kan een meervoudige naamwoordgroep gekozen worden, en blijkt dat *dat* niet het niet-specifieke karakter van *een hamer* afspiegelt, maar zich niets aantrekt van de semantische eigenschappen van het antecedent.

De reden dat (7b) in tegenstelling tot (7a) uitgesloten is, lijkt eerder dat in (7a) de zin in het geheel vrije ruimte geeft voor de interpretatie van *dat* als verwijzend naar een pragmatische notie van *een partner hebben*, vertegenwoordigd in de vorm van *dat*, terwijl de zin in (7b) heel anders loopt en niet de vrijheid biedt naar zo’n pragmatische notie te verwijzen. Hier wordt congruentie met het talig antecedent *zijn partner* gedwongen en kan *dat* dus niet gebruikt worden omdat het niet de juiste congruentiekenmerken heeft.

In dit deel heb ik beargumenteerd dat specificiteit niet als beregelende factor gezien kan worden. Daarbij zijn we feitelijk twee mogelijke alternatieve verklaringen tegengekomen: *Dat* verwijst naar een pragmatische notie die niet telbaar is, of *dat* is een syntactisch element dat überhaupt geen enkele congruentie laat zien – noch met een talig antecedent, nog met semantische eigenschappen van een referent in de werkelijkheid. In sectie 1.4 bespreek ik een alternatieve kijk en ga ik in op een keuze tussen deze verklaringen. Voor het punt van dit deel is dat niet van belang – in deze subsectie hebben we gezien dat specificiteit geen bevredigende verklaring biedt.

1.3.3.2 De Invloed van Sekse

Naast de invloed van specificiteit heb ik ook een kanttekening te plaatsen bij de invloed van sekse. Uiteraard is de invloed van sekse niet te ontkennen bij de keuze tussen mannelijke en vrouwelijke voornaamwoorden. Wel is de manier waarop deze parameter in (2) is geïncorporeerd discutabel. Audring (2006a) stelt dat er ook voor bezielde referenten (zoals mensen) verschil kan zijn in individuatie. Zo kan op pejoratieve wijze naar een persoon of groep mensen verwezen worden met het onzijdige *dat*, omdat de referent dan meer naar rechts geplaatst wordt op de schaal.⁵ Echter, de

congrueert met “pannkakor”. Wanneer we dit zinsdeel vervangen door “att äta pannkakor” wordt de grammaticaliteit gered, maar klopt de zin semantisch niet, omdat “gott” enkel op het voedsel kan slaan en niet, zoals in het Nederlands, op een handeling, zoals het eten ervan. “Gott” verwijst noch naar “pannkakor”, noch naar een weggelaten bijzin. De optie die overblijft is dat het naar de notie of het idee van pannenkoeken verwijst. Nog duidelijker is het in een analoge zin als “Vreemde culturen, dat is zijn passie”, dat we onmiddellijk kunnen interpreteren zonder dat we kunnen weten, zoals dat wellicht nog kon bij “pannenkoeken (eten)”, welk werkwoord er is weggelaten.

⁵ Een dergelijke keuze voor een onzijdig voornaamwoord wijt ik niet aan een lagere specificiteit of telbaarheid maar aan verwijzing naar een pragmatische notie, waar door de schijnreferent een soort notie of verzameling wordt opgeroepen.

schaal in (2) is gradueel en eendimensionaal. Het doet de voorspelling dat naar een wat lager geïndividueerd vrouwspersoon (in een tussensegment van de schaal) met een mannelijk voornaamwoord verwezen kan worden. Deze voorspelling is onjuist:⁶

- (8) a. Dati doet maar. (vrouwspersoon)_i
 b. Hij*_i doet maar. (vrouwspersoon)_i

Om dergelijke incorrecte voorspellingen te vermijden zijn subgroeperingen in het model nodig. Dan kunnen we de verschillen tussen de referenten uitdrukken met verschillende ϕ -kenmerken.

1.4 Een alternatief

1.4.1 Parameters

In haar meer recente werk, Audring (2009), houdt Audring dezelfde individuatiehiërarchie aan, met het verschil dat specificiteit uit de hiërarchie is verwijderd, en de term “bounded” wordt nu over de hele linie gebruikt:

*From the usage patterns discussed above, we can distil two parameters that influence pronoun choice. The first parameter is **natural gender** with the values [male] and [female]. As in many gender systems, the masculine and feminine pronouns are used for male and female persons and occasionally animals. The second parameter, which governs the usage of pronouns in inanimate reference, appears under different names in the literature. The best terms are **boundedness** or **countability**, and the values are [bounded]/[unbounded] or [count]/[mass]. The split is that between “discrete entities with a well-defined shape and precise limits” on the one hand and “homogenous undifferentiated stuff without any certain shape or precise limits” on the other (...) (Audring, 2009, p. 128)*

Hiermee komt Audring (2009) op dezelfde relevante kenmerken als Klomp (2007). Klomp (2007) zet de observaties van Audring (2006a) om van een schaal naar een keuzemodel. Een gereviseerde versie van dat model is weergegeven in (9).⁷

⁶ Bij verwijzing naar vrouwelijke dieren worden vaak mannelijke voornaamwoorden gebruikt, echter, het gaat hier niet om een pejoratief aspect, maar om een kwalificatie als onbezielde, waarna op basis van telbaarheid alsnog een mannelijk voornaamwoord gekozen wordt. Bij onbetwistbaar bezielde referenten vinden we die keuze namelijk nooit.

⁷ Het model bestaat uit binaire splitsingen die telkens een ϕ -kenmerk invullen. Als we uitgaan van ϕ -kenmerken met één unieke waarde kunnen we alle witte cellen zien als toegewezen kenmerken. Als we uitgaan van binaire kenmerken kunnen we de witte cellen zien als het toekennen van een positieve waarde aan een kenmerk en de grijze cellen als het toekennen van een negatieve waarde. Ik zal voortaan consistent uitgaan van kenmerken met één unieke waarde, omdat dat extra nadruk geeft aan de grijze cellen als ongemarkeerde condities, maar deze keuze is verder niet relevant voor dit onderzoek.

(9)

| | | | | | |
|------------|--------------|-----------|---------------|---|------------|
| REFERENT → | Telbaar → | Bezield → | Vrouwspersoon | → | Vrouwelijk |
| | | | Manspersoon | → | Mannelijk |
| | | Onbezield | | → | Mannelijk |
| | Niet-telbaar | | | → | Onzijdig |

Allereerst kan het kenmerk [telbaar] eventueel toegekend worden, niet-telbaar is hier de uitgangspositie. Dat is zo gekozen omdat we het onzijdig op veel verschillende plaatsen vinden, situaties die met elkaar gemeen hebben dat er geen sprake is van een telbare referent – we vinden het onzijdig bij stofnamen, bijzinnen, pragmatische noties, semantisch lege constituenten en niet-congruerende voornaamwoorden. Alleen wanneer er sprake is van een telbare referent, wordt het onzijdig niet gekozen. In de volgende subsectie kom ik terug op de verschillende gedaanten van het onzijdig. Als een referent het kenmerk [telbaar] heeft kan het kenmerk [bezield] toegekend worden, hier is het ontbreken van bezielde bij een referent de uitgangspositie. De aanname is wederom dat er veel verschillende typen telbare onbezielde referenten zijn, terwijl een referent als bezield wordt beschouwd als het heel specifiek menselijke eigenschappen heeft. Tenslotte kan het kenmerk [vrouwspersoon] toegekend worden, de uitgangspositie is dat een telbare bezielde referent mannelijk is. Dat is in lijn met de aanname dat het mannelijk geslacht ongemarkeerd is, zoals in Harley & Ritter (2002). Dat zien we ook terug in het Nederlands bij het gebruik van mannelijke voornaamwoorden voor hypothetische personen waar geen geslacht is vastgesteld, en in vaste uitdrukkingen als *Zet 'm op*, en tegenwoordig nog steeds in straattaal – *Ik chill 'm*, *Ik boek 'm*.

In feite is (9) een vereenvoudigde weergave van de “fuzzy” praktijk van het voornaamwoordgebruik in het moderne Nederlands, die door Audring (2006, 2009) is gedocumenteerd en descriptief is weergegeven doormiddel van de individuatiehiërarchie. De weergave in (9) ‘pelt’ telkens een residu aan de rechterkant van de hiërarchie af – in de eerste splitsing correspondeert “telbaar” en “niet-telbaar”, met de grens tussen “bounded object / abstract” en “unspecific mass, unbounded object”, vervolgens correspondeert “bezield” met “human”, en “niet bezield” met de overgebleven rechterkant van de hiërarchie. Tenslotte splitst “bezield” in “vrouwspersoon” en “manspersoon”, waarbij dus een voornaamwoord gevonden wordt op basis van de sekse van de referent. Het grijze gebied tussen “bezield” en “onbezield”, dat in de hiërarchie vertegenwoordigd werd door een tussenklasse “other animate”, is weggelaten. Ook de veronderstelde invloed van specificiteit, en daarmee de categorie “specific mass”, is weggelaten.

De kwalificaties geschieden naar conceptueel inzicht van de spreker. Het is een keuze die ik beschouw als voorafgaand aan de aansturing van het taalsysteem. In minimalistische termen is het een proces dat deel uitmaakt van het conceptueel-intensioneel systeem. Deze overwegingen kennen een waarde toe aan de relevante parameters. Dit geschiedt met behulp van de kennis van de wereld waar de spreker over beschikt. Daaronder valt bijvoorbeeld de wetenschap dat een meisje vrouwelijk is, maar ook kennis van fictieve werelden, bijvoorbeeld de kennis dat een sprekend voorwerp of dier dat men doorgaans als onbezield beschouwd in de context van een sprookje toch als bezield gezien wordt. De invulling van de kenmerken is vervolgens voor het

taalsysteem een vaststaand gegeven. Ik ga er van uit dat wanneer een spreker zich van dit systeem bedient het voornaamwoord niet anaforisch verwijst (met andere woorden, niet verwijst naar een talig antecedent in de context), maar naar een entiteit in de werkelijkheid. In plaats van gebruik te maken van de lexicale eigenschappen van een talig antecedent gebruikt het voornaamwoord kenmerken van de referent in de werkelijkheid, nadat die door de spreker conceptueel vastgelegd zijn.

Uiteraard is het model geen oplossing voor alle grijze gebieden die Audring (2006 & 2009) beschrijft. Het punt wat ik probeer te maken is enkel dat de grijze gebieden het product zijn van grilligheid die voortkomt uit (1) het conflict tussen congruentiesystemen waarvan ik zal beargumenteren dat ze op dit moment actief zijn en (2) de classificatie van bepaalde referenten die voorafgaat aan het taalsysteem in nauwe zin, dat afhankelijk is van de individuele keuzes van de spreker en derhalve grillig is. Omdat het model verder geen variatie toelaat voorspelt het dat wanneer we een overtreding van het semantische systeem vinden, we één van deze twee redenen kunnen aanhalen als verklaring. De punten van twijfel zijn ook concreet geformuleerd – een spreker kan twijfelen over de sekse van de referent, en over het al dan niet toekennen van de kenmerken [beziel] en [telbaar]. Ook geeft het model de gemarkeerdheid van de groepen aan. Het onzijdig is het meest ongemarkeerde geslacht. Verderop in dit onderzoek zullen we, met deze definitie van het systeem en de kenmerken, de praktijk van het conflict tussen het lexicale systeem en het semantische systeem onder de loep nemen.

1.4.2 *Dat* en Niet-Specifieke Referenten

Ik heb reeds beschreven dat er, wanneer we het idee van de invloed van specificiteit hebben verworpen, nog twee mogelijke verklaringen zijn voor het onverwachte voorkomen van *dat* bij zijdige, telbare referenten. Het voorbeeld in (5), hieronder herhaald als (10), kunnen we verklaren de stellen dat er verwezen wordt naar een hogere pragmatische notie, die geen van de in (9) gespecificeerde kenmerken draagt en dus voor het onzijdig kwalificeert. Audring (2009) houdt echter de andere verklaring aan: *Dat* is een vast element dat geen enkele congruentie laat zien. Hoewel Audring (2009) niet werkt met termen als gemarkeerdheid en kenmerken, is deze aanname consistent met onze bevinding hierboven dat het onzijdig het meest ongemarkeerde geslacht is, dat we dus verwachten terug te vinden indien er geen sprake is van congruentie. Het relevante voorbeeld wordt opnieuw aangehaald en voorzien van een verklaring:

- (10) a. Hij heeft nu een partner (z). *Dat* (n) had-ie eerder niet gehad.
 b. Hij heeft zijn partner (z) meegenomen. **Dat* (n) heeft-ie in Engeland ontmoet.
 (uit Audring (2006a))

[(10a)] shows that the function of the dat is independent of the semantics of its antecedent noun. Even for common gender nouns with person reference, where a neuter pronoun is normally impossible (witness [(10b)]), the right (unspecific) context allows the selection of a neuter gender demonstrative. Neuter personal pronouns are not possible in either [(10a)] or b). (Audring, 2009, p. 122)

Nu zou het eleganter zijn als we deze toepassing van *dat* niet als uitzondering moeten beschouwen maar een plaats in het geheel kunnen geven zonder een discrepantie tussen *dat* en andere voornaamwoorden te hoeven stipuleren. In dat opzicht is de alternatieve verklaring dat *dat* in (10a) naar een pragmatische notie verwijst aantrekkelijker. In deze subsectie vergelijk ik deze twee verklaringen voor dit voorkomen van *dat*.

Duinhoven (1988) merkt ook voor het Middelnederlands op dat *dat* toepassing geniet buiten het onzijdig enkelvoud, maar generaliseert anders:

Dat is gereserveerd voor het onzijdig enkelvoud (eerste en vierde naamval). Het is misschien juister te zeggen, dat dat wordt gebruikt wanneer datgene waarnaar wordt verwezen, niet manlijk of vrouwelijk (noch meervoud) is. Dat kan namelijk op een die-woord volgen wanneer aan een collectief, een verzameling gedacht kan worden. (Duinhoven, 1988, p. 33)

Hier wordt uitgegaan van een exclusieve definitie: Wanneer een naamwoord niet aan de gegeven eisen voldoet, kan *dat* toegepast worden. Dat is het geval bij onzijdige naamwoorden, maar ook wanneer een referent, ongeacht woordgeslacht, als een verzameling opgevat kan worden. Ook dan wordt er blijkbaar niet aan de eis voldaan om *die* toe te passen. Hierbij worden de volgende voorbeelden gegeven:

- (11) a. Torment ende dogen, arbeit ende pine [...] dat deeddi willichlike
"Kwelling en afzien, werk en pijn [...] dat onderging hij bewust"
- b. Der kerken scat, twi soudic vercopen dat
"De schat van de kerk, waarom zou ik dat verkopen?"
- c. Ors ende wapine dat was root
"Paard en wapen, dat was rood"

(Middelnederlands, uit Duinhoven [1988], mijn vertalingen)

Duinhoven (1988) geeft aan dat de toepassing in (11c), in combinatie met een naamwoordelijk gezegde, nog steeds gebruikelijk is, één van de voorbeelden hiervan zijn hieronder herhaald als (11c'). Maar naast (11c) zien we de toepassing in (11b) ook veelvuldig in het moderne Nederlands. Immers, het betreft hier een niet-telbare referent, waarnaar met een onzijdig voornaamwoord verwezen wordt – die observatie was de aanleiding voor dit onderzoek. Een voorbeeld daarvan is hieronder herhaald als (11b'). Ook voor de toepassing in (11a) zijn moderne voorbeelden te bedenken, hieronder herhaald als (11a'):

- (11') a. Je tas en je boeken liggen nog op tafel. Kun je dat even opruimen?
- b. De spinazie ligt nog op het aanrecht. Kun jij dat even in de pan gooien?
- c. Deze steen, dat is een diamant.

We zien in (11) dus drie verschillende manieren waarop eenzijdige of meervoudige naamwoordgroep samen kan gaan met *dat*. De vraag is dan uiteraard of en hoe de motivatie van het voorkomen van *dat* verschilt tussen deze gevallen. In die visie van Audring (2009) heeft het voorkomen van *dat* in (11b') een andere achtergrond dan in (11a') en (11c'): In (11b') is het de parameter telbaarheid op basis waarvan congruentie plaatsvindt, in (11a') en (11c') zien we antecedenten die wel telbaar (en zijdig) zijn en dus moet de verklaring zijn dat *dat* in deze gevallen niet gevoelig is voor enig congruentiesysteem. Dit kunnen we kunnen testen door het lexicaal congruerende *die* te forceren. Als het zo is dat het voornaamwoord *dat* geheel niet congrueert, en dus geen toegang heeft tot de kenmerken van het antecedent, zou gebruik van een zijdig voornaamwoord ongrammaticaal zijn. Dat is alleen in (11c'') het geval:

- (11'') a. Je tas en je boeken liggen nog op tafel. Kun je die even opruimen?

- b. De spinazie ligt nog op het aanrecht. Kun jij die even in de pan gooien?
- c. *Deze steen, die is een diamant

Duinhoven geeft aan dat de constructie in (11c'') een vaste uitdrukking is geworden. De constructie komt dus niet middels een creatief vormingsproces tot stand. Duinhoven wijt dat aan het feit dat naamwoordelijk gezegdes syntactisch ambigu zijn – beide betrokken constituenten kunnen als het onderwerp worden opgevat, en dus is niet duidelijk met welke constituent het voornaamwoord en het werkwoord moeten congrueren. Dat is consistent met de bevinding dat (11c'') ongrammaticaal is – het voornaamwoord heeft dus geen toegang tot de kenmerken van een antecedent. De keuze voor het onzijdige geslacht is consistent met de aanname dat het onzijdig het ongemarkeerde residugeslacht is. Voor deze constructie is het standpunt dat *dat* niet congrueert dus goed te verdedigen.⁸ Echter, in (11b'') is er, net als in (11a'') klaarblijkelijk wel toegang tot de congruentiekenmerken van het antecedent. De aanname dat het hier ook een *dat* betreft dat geen toegang heeft tot de kenmerken van het antecedent, is dus onjuist. Als het voornaamwoord anaforsch kan zijn (dat wil zeggen, in staat is te verwijzen naar een antecedent), en het antecedent heeft een meervoudskenmerk, waarom vinden we dan een voornaamwoord dat dit kenmerk niet weerspiegelt? De anaforische verwijzing van een voornaamwoord is optioneel. We zien in (11a') dat het voornaamwoord wel eventueel in staat is te binden, maar het blijkbaar niet doet, echter zonder dat het semantische verband verloren gaat. Wanneer er op deze manier verwezen wordt is er dus geen sprake van anaforische verwijzing maar verwijzing naar een entiteit buiten de zin, waardoor het voornaamwoord congrueert met semantische eigenschappen van de referent in plaats van met die van een talig antecedent. Verzamelingen of stoffen, alsook pragmatische noties hebben het kenmerk [telbaar] niet en gaan samen met een onzijdig voornaamwoord. Zo kan, zoals we hebben gezien, een onzijdig voornaamwoord ook bij bijvoorbeeld *een hamer* en *een werkvergunning* voorkomen. Dit kunnen we verklaren door te stellen dat er verwezen wordt naar een pragmatische notie, zoals het voorhanden hebben van een hamer, of het verkrijgen van een werkvergunning, en niet naar een telbare referent. Door het ontbreken van het kenmerk [telbaar] valt de keuze op het residugeslacht, onzijdig.

Als deze verklaring juist is voorspelt het dat er voor wat het geslacht betreft geen discrepantie is tussen de toepasbaarheid van *dat* en het persoonlijk voornaamwoord van datzelfde geslacht, *het*, zoals dat aangenomen werd in het bovengenoemde citaat uit Audring (2009). Immers, we nemen nu niet aan dat *dat* een uitzonderlijk niet-congruerend voornaamwoord is, maar het net als *het* het semantische systeem volgt. Concreter, er moet een voorbeeld van het type (11a') zijn waar we *dat* door *het* kunnen vervangen zonder dat het resultaat ongrammaticaal is. Dat is inderdaad mogelijk:

(12) Ik laat mijn boeken gewoon op tafel liggen, ik ruim het morgen wel op.

(13) Die pannenkoeken, vond je dat lekker? Ik vond het echt niet binnen te houden.

Deze voorbeelden tonen aan dat de bevindingen niet exclusief voor *dat* gelden. Als we aannemen dat er verwezen wordt naar een pragmatische notie of collectief is het gedrag van *dat*, en in (12) en (13) *het*, in lijn met het model, zonder dat we een uitvlucht nodig hebben naar een aanname dat *dat* niet congrueert, of dat specificiteit een bepalende factor is. We hebben nu argumenten

⁸ Deze verklaring zou ook van toepassing kunnen zijn op voorbeeld (6c), *Pannekoeken, dat is lekker*. Echter, we zien hetzelfde principe in bijvoorbeeld *Pannekoeken, vind je dat lekker?* Daarmee is het idee van een pragmatische notie ook hier nog steeds een relevante verklaring.

tegen beide standpunten behandeld en de voorbeelden voor beide standpunten kunnen we zien als voorbeelden van verwijzing naar pragmatische noties.

We hebben nu het semantische systeem gedefinieerd. In sectie 2 bespreek ik de theorie achter de concurrentie tussen twee congruentiesystemen (in dit geval, het traditionele systeem van woordgeslacht en het moderne semantische systeem dat ik in deze sectie heb besproken). Dat brengt ons tot enkele verwachtingen die samen met het model in (9) het uitgangspunt zullen vormen voor het experiment dat in sectie 3 beschreven wordt. In sectie 4 volgt een bespreking van de resultaten, waarna in sectie 5 de conclusie volgt.

2 De congruentiehiërarchie

Zoals beschreven zijn er in het Nederlands twee systemen die elkaar beconcurreren: het traditionele lexicale systeem waar de voornaamwoordkeuze geschiedt op basis van het woordgeslacht van een talig antecedent, en het semantische systeem van telbaarheid, waarbij het geslacht van het voornaamwoord bepaald wordt door de telbaarheid van de referent.⁹ Het is mogelijk dat dit nieuwe systeem, dat het probleem van het gat in het voornaamwoordparadigma kan oplossen, het lexicale systeem in dat domein op den duur volledig zal vervangen. Daardoor zou er weer een consistente, stabiele taalsituatie ontstaan. Het is ook mogelijk dat het systeem verder om zich heen zal grijpen dan enkel het domein waarvoor het in eerste instantie een oplossing bood. Het zou door kunnen dringen in andere contexten, misschien uiteindelijk zelfs woordgeslachten veranderen, zodat de telbare naamwoorden *de*-woorden worden en niet-telbare naamwoorden *het*-woorden. Maar het is ook mogelijk dat het hier slechts om een tijdelijke trend gaat, die geen blijvende sporen zal achterlaten.

De congruentiehiërarchie van Corbett (1979) heeft betrekking op taalverandering en is een hiërarchie die bestaat uit verschillende taalkundige contexten. De voorspelling is dat wanneer in een taal een nieuw congruentiesysteem zijn intrede doet, deze zich als eerste zal manifesteren in de context aan de uiterst rechtse kant van de hiërarchie. Naarmate de taalverandering verder doorzet zal het nieuwe systeem stapsgewijs tot de links geplaatste contexten doordringen. De congruentiehiërarchie wordt in Corbett (1979) als volgt weergegeven:

- (15) Attributive > Predicate > Relative Pronoun > Personal Pronoun

Hoewel een hiërarchie niet dienst kan doen als model voor het semantische systeem (zoals besproken in sectie 1.3 hierboven), zou het in dit verband een goede weergave kunnen zijn en voorspellingen kunnen doen. Taalverandering is immers een proces dat zich kenmerkt door grijze gebieden en geleidelijke veranderingen die in één domein beginnen en langzaam uitbreiden over de hele taal (Mesthrie, Swann, Deumert & Leap, 2000).

Het semantische systeem lijkt zich inderdaad voornamelijk te manifesteren in het domein van de persoonlijk voornaamwoorden. Ik zal voor iedere categorie in (15) (waar dat mogelijk is) een Nederlands voorbeeld geven waar het lexicale en semantische systeem tegenstrijdige uitkomsten eisen, en markeren met mijn eigen grammaticaliteitsoordelen:

⁹ Ik ga in dit onderzoek uit van een idealisatie: uiteraard zijn deze twee systemen niet de enige twee die we in het moderne Nederlands kunnen aantreffen. Het doel van dit onderzoek is echter niet om een verscheidenheid aan (regionale) trends te beschouwen. Het semantische systeem op basis van telbaarheid is de wijdst verspreide innovatie die in de literatuur het meest opgemerkt wordt, ik richt mij hier enkel op de invloedssfeer van dit nieuwe systeem en van het lexicale systeem.

- (16) a. **Dat**(lex.)/***die**(sem.) boek ligt op tafel (attributief)
 b. Het boek **ligt**(lex.)/**ligt**(sem.) op tafel (predicatief)
 c. Het boek **dat**(lex.)/??**die**(sem.) op tafel ligt (relatief)
 d. Het boek ligt op tafel. Ik heb ?**het**(lex.)/**hem**(sem.) nog niet gelezen (persoonlijk)

Bij de predicatieve context is geen verschil: Het Nederlands kent geen verschil in vervoeging op basis van geslacht. Deze context is dus niet te toetsen. (16a) is de meest strikte context. Het nieuwe semantische systeem zou hier een zijdig aanwijzend voornaamwoord opleveren, aangezien de referent telbaar is. Naar mijn eigen intuïties is het resultaat, *die boek*, behoorlijk ongrammaticaal. In (16c) leidt het semantische systeem naar mijn mening ook tot een ongrammaticale uitkomst, hoewel ik deze iets minder sterk ongrammaticaal vind dan (16a). In (16d) vind ik toepassing van het semantische systeem zelfs tot een betere uitkomst leiden dan het lexicale systeem.

Merk wel op dat het lexicale systeem alléén in context (16d) defectief is. Daarmee hoeft het voorkomen van het semantische systeem in (16d) niet noodzakelijkerwijs te wijten te zijn aan de voorspelling van de congruentiehiërarchie. Het kan even goed toeval zijn dat de defectieve context nu precies aan het uiteinde van de hiërarchie ligt, en er verder geen sprake is van een verschuiving die zich context voor context voltrekt. Het grammaticaliteitsoordeel in (16c) is dus cruciaal: De keuze tussen *die* en *dat* is niet driedig, het correspondeert met de moderne geslachten zijdig en onzijdig. Er is hier dus geen gat dat gerepareerd dient te worden zoals in (16d). Toch oordeel ik dat het gebruik van het semantische systeem in (16c), hoewel het voor mij niet de voorkeur heeft, ook enige aantrekkingskracht heeft. Wanneer dat oordeel inderdaad breed draagvlak heeft, moet er meer aan de hand zijn dan het repareren van één defectief domein – dan moeten we er inderdaad sterk rekening mee houden dat het hier een algehele taalverandering betreft. Als we constateren dat de taal inderdaad ingrijpender verandert dan alleen in (16d), en de verdeling volgt de congruentiehiërarchie, dan zal dat het idee van de congruentiehiërarchie onderschrijven.

3 Het onderzoek

3.1 Doel

In sectie 1 is het semantische systeem gedefinieerd doormiddel van het model in (9), met daarbij de kanttekening dat er twee bronnen van ruis zijn: [1] Concurrentie tussen het nieuwe en het oude systeem en [2] twijfel over de toekenning van de relevante ϕ -kenmerken aan de referent. Het eerste hoofddoel van dit onderzoek is om vast te stellen of het semantische systeem boven het lexicale systeem wordt verkozen, en zo ja, hoe sterk de voorkeur is in verschillende taalkundige contexten. Het is dus zaak dat de tweede bron van ruis, de twijfel over de invulling van de kenmerken, tot een minimum gereduceerd wordt in het experiment. Gegeven het bestaan van deze onzuiverheid kunnen we nooit helemaal zeker weten of een antecedent-voornaamwoordcombinatie “conform semantisch systeem” is. Het zou dan ook methodologisch niet helemaal zuiver zijn nu de invloed van dat systeem in getallen uit te gaan drukken. Om die reden formuleer ik de hoofdvraag als volgt:

- I. *Hoe trouw blijft men aan het lexicale systeem in de contexten van (16)?*

Ik zal antecedent-voornaamwoordcombinaties onderzoeken voor de contexten in (16a), (16c) en (16d). Voortaan zal ik naar deze contexten verwijzen met AT (attributief) BT (betrekkelijk) en PS (persoonlijk). Met de congruentiehiërarchie en mijn grammaticaliteitsoordelen in (16) als uitgangspunt is de verwachting dat het semantische systeem de meeste invloed zal hebben in PS, en het minste in AT. Daarvan zouden we een afspiegeling moeten vinden in de trouw aan het lexicale systeem: wanneer het semantische systeem veel aanhang geniet, wijkt men het meest af van het lexicale systeem. De hypothese is dus dat men in AT erg trouw blijft aan het lexicale systeem, en in PS er het sterkst van afgeweken zal worden.

II. Kan de afwijking verklaard worden door concurrentie van het semantische systeem?

Met enige voorzichtigheid is II de vervolgvraag. We kunnen bekijken of er een relatie bestaat tussen telbaarheid en het gekozen woordgeslacht. Omdat het zoals besproken uit principe nooit zeker is hoe een spreker de telbaarheid van een referent daadwerkelijk geïnterpreteerd heeft bij het geven van een grammaticaliteitsoordeel zullen we niet te veel waarde hechten aan de exacte percentages. Maar door testitems te kiezen die duidelijk en ondubbelzinnig wél of niet het kenmerk [telbaar] dragen wordt het probleem tot een minimum gereduceerd en zal deze vraag bij het vinden van overtuigende percentages toch nog positief te beantwoorden zijn. De hypothese is dat wanneer de mate van lexicale congruentie afneemt, de mate van semantische congruentie toeneemt.

Nadat we de invloed van de systemen per context hebben bekeken kunnen we de volgende vraag beantwoorden:

III. Doet de congruentiehiërarchie van Corbett (1979) de juiste voorspellingen voor het antecedent-voornaamwoordrelaties in het hedendaagse Nederlands?

De belangrijkste vragen bij het experiment zijn nu geformuleerd. Om deze vragen te beantwoorden is er, naast AT, BT en PS, nog een vierde context die dezelfde tweestrijd laat zien, en ons dus kan helpen bij het beantwoorden van de vragen:

(16e) Het boek op tafel is **het mijne(lex.)/de mijne(sem.)**

Ik zal met BZ (bezittelijk) aan deze context refereren. Hoe deze context zich exact kwalificeert in termen van de hiërarchie in (15) is niet direct duidelijk. Het vetgedrukte gedeelte is een nominale frase waaruit het naamwoord is weggelaten. Het overgebleven “gat” verwijst terug naar het voorgenoemde naamwoord. Het is interessant om te bekijken hoe het lidwoord zich nu gedraagt als het aankomt op congruentie. Gebaseerd op de nauwe relatie tussen het lidwoord en het weggelaten naamwoord is mogelijk dat het congruentiepatroon gelijkenis vertoont met AT. Aan de andere kant is de relatie tussen het naamwoord en het lege gat vergelijkbaar met de binding tussen naamwoord en persoonlijk voornaamwoord in PS. Mijn hypothese is dat deze constructie vooral op PS zal lijken. De grote afstand tussen antecedent en lidwoord zou er voor kunnen zorgen, net als de grote afstand in PS, dat het conflict met het woordgeslacht minder duidelijk is voor de spreker en er meer vrijheid is om semantisch te verwijzen. Dat is niet noodzakelijk de meest aannemelijke hypothese: Onder de aanname dat het semantische systeem is ontstaan om een probleem op het vlak van de persoonlijk voornaamwoorden te verhelpen en niet verder om zich heen grijpt als taalverandering, kunnen we ook beargumenteren dat dat systeem in dit domein helemaal geen toepassing zal vinden. Naast het verkrijgen van inzicht in de constructie

door het congruentiegedrag te bestuderen is BZ dus ook uitermate geschikt om te inventariseren of het semantische systeem inderdaad verder om zich heen grijpt dan nodig was om het beschreven paradigma-probleem op te lossen. De laatste hoofdvraag luidt:

IV. *Hoe gedraagt de bezittelijke context zich in verhouding tot de andere contexten?*

In Audring (2009), dat nog moest verschijnen ten tijde van dit onderzoek, wordt geconstateerd dat de data uit het Corpus Gesproken Nederlands het patroon volgt dat de congruentiehierarchie voorspelt. Daarnaast lijkt het erop dat naast de grammaticale context ook de woordafstand en de naamval van invloed zijn: Wanneer een voornaamwoord (binnen een gegeven grammaticale context) verder af staat van het antecedent en/of in de eerste naamval staat, is de invloed van het semantische systeem groter. Er zijn een paar kanttekeningen die kunnen worden gemaakt bij het corpusonderzoek naar het CGN waar Audring (2009) over rapporteert. Allereerst onderzoekt neemt zij een zeer breed scala aan semantische groepen en voornaamwoorden. De percentages die aangeven hoe vaak we semantische congruentie aantreffen bevatten daardoor niet alleen de gevallen waar telbaarheid een rol speelt, maar ook gevallen van semantische congruentie die in de schrijftaal ook geaccepteerd zijn, zoals het verwijzen naar *meisje* met *zij*. Beschouw een fictieve dataset in (x)

- (17) a. Daar heb je mijn broertje(o). Hij(m)/*het(o) is erg aardig.
- b. Ik heb een boek(o) gekocht, maar ik heb 'm(m) nog niet gelezen.
- c. Ik heb melk(z) gekocht. Deze(z) staat in de koelkast.

Audring zou voor deze dataset het aandeel semantische congruentie als 2/3 taxeren, want in (17a) en (17b) congrueert het voornaamwoord niet met het lexicale woordgeslacht. In dit onderzoek zijn we echter geïnteresseerd in de invloed van telbaarheid, en niet het type semantische congruentie in (17a). Enkel (17b) bevat namelijk semantische congruentie die niet geaccepteerd is in de schrijftaal, alleen hier bevindt zich dus de taalverandering. Volgens de definities van dit onderzoek is het aandeel innovatieve semantische congruentie 1/2: (17a) wordt buiten beschouwing gelaten omdat zowel spreektaal als schrijftaal hetzelfde voorschrijven, (17b) bevat een keuze voor innovatieve semantische congruentie, (17c) bevat een keuze voor conservatieve congruentie op basis van het woordgeslacht.

Verder beschrijft Audring (2009) een zeer hoge graad van semantische congruentie bij bezittelijke voornaamwoorden, echter, door het ontbreken van een onzijdig bezittelijk voornaamwoord in het Nederlands zijn we niet in staat daaruit af te leiden hoe deze grammaticale context zich verhoudt tot andere contexten. Wanneer we een zin tegenkomen als *Dat bed is zijn pootjes kwijt* zouden we het voorkomen van het zijdige *zijn* bij een lexicaal onzijdig maar semantisch telbaar naamwoord kunnen interpreteren als een voorbeeld voor het semantische systeem, echter, het is niet duidelijk wat de vorm van deze zin zou moeten zijn wanneer de congruentie lexicaal zou zijn, omdat er geen onzijdige tegenhanger van *zijn* bestaat. Als we deze gevallen meetellen rekenen we ten onrechte dat het semantische systeem in deze gevallen een sterke invloed heeft.

Een ander voorbeeld van ruis wordt gevormd door de zelfstandig gebruikte aanwijzend voornaamwoorden: Het is niet altijd duidelijk wanneer er sprake is van een aanwijzende *die*, en wanneer van een zwak mannelijk voornaamwoord *-(d)ie*. Audring (2009) merkt ook op dat de transcripties van het Corpus niet altijd even nauwkeurig zijn. Dit is ook mijn ervaring (Klomp, 2007). Ook wanneer we zeker weten dat we met een aanwijzende *die* te maken hebben is vaak niet duidelijk of de referent een talig antecedent is of een entiteit in de werkelijkheid. Dit laatste

punt wordt ook door Audring (2009) genoemd en de invloed van dit soort factoren wordt met zorg tot een minimum gereduceerd. Maar deze problemen zijn onvermijdelijk en inherent aan corpusonderzoek, zelfs wanneer de data perfect geselecteerd zou zijn: Wanneer we alle data in tact laten is er invloed van een ongewenste factor; Wanneer we een deel van de data elimineren is er een ongewenste vertekening omdat structureel een bepaald type data uit de set is geëlimineerd. Audring (2009) komt tot significante cijfers die de lijn van de congruentiehiërarchie volgen, maar bevestiging door een gecontroleerder onderzoek dat al deze bezwaren ontloopt zou welkom zijn. Het onderzoek dat in deze scriptie wordt beschreven en dat een gecontroleerd experiment bevat, voldoet aan dergelijke eisen. Uiteraard kent het huidige onderzoek weer andere beperkingen en mogelijke bronnen van ruis die ik hieronder zal bespreken. Ik ben daarom van mening dat de twee benaderingen elkaar kunnen ondersteunen. Naast de methode is ook het onderzoeksgebied en het soort bewijs verschillend – het corpusonderzoek analyseert daadwerkelijk geproduceerde voorkomens en levert positief bewijs, het huidige onderzoek legt mogelijke zinnen voor ter interpretatie en levert ook negatief bewijs. Wanneer beide wegen bewandeld worden is er goed zicht op het totaalbeeld.

3.2 Methode

3.2.1 Proefpersonen

Voor dit onderzoek zijn 55 jonge moedertaalsprekers van het Nederlands als proefpersonen gekozen, tussen de leeftijd van 18 en 25. Jonge sprekers hebben de meest recente vorm van de taal als moedertaal verworven. Het is die meest recente vorm waar we nu in geïnteresseerd zijn, omdat het een recente ontwikkeling in de taal betreft.

3.2.2 Stimuli

Er zijn zinnen geconstrueerd voor elk van de in 3.1 beschreven contexten. Het gaat hier telkens om een koppel van een antecedent en een voornaamwoord. De naamwoorden voor het antecedenten zijn van vier verschillende typen: lexicaal zijdig en telbaar, lexicaal onzijdig en telbaar, lexicaal zijdig en niet-telbaar, lexicaal onzijdig en niet-telbaar. Voor ieder type zijn vier eenlettergrepige naamwoorden gekozen:

| | Lexicaal zijdig | Lexicaal onzijdig |
|--------------|-----------------------------------|------------------------------------|
| Telbaar | <i>Bal, Stoel, Fiets, Pen</i> | <i>Bed, Boek, Mes, Huis</i> |
| Niet-telbaar | <i>Verf, Gel, Inkt, Kaas</i> | <i>Zand, Geld, Hout, Vlees</i> |

Alle testitems zijn onbezield. Audring (2006 & 2009) betreft ook bezielde referenten in het bepalen van de invloed van het semantische systeem. In dit onderzoek zijn we niet geïnteresseerd in het aandeel semantische congruentie in zijn geheel. Semantische congruentie op basis van biologisch geslacht vinden we ook in schrifttaal, een register waarin het semantische systeem op basis van telbaarheid niet geaccepteerd is. We onderzoeken hier specifiek congruentie puur op basis van *telbaarheid*, daarin dienen bezielde referenten hier buiten beschouwing gelaten te worden, zoals beschreven in sectie 3.1. Bij andere referenten als dieren is een extra complicatie dat, wanneer er met een mannelijk voornaamwoord verwezen wordt, we niet altijd weten of dat is omdat de referent door die taalgebruiker bezield geacht wordt (bijvoorbeeld een huisdier) of

alleen telbaar (bijvoorbeeld een stuk vee). Alle gekozen naamwoorden in dit onderzoek zijn daarom ondubbelzinnig onbeziel.

Alle naamwoorden zijn in alle vier de te onderzoeken contexten geplaatst, en voor elk van deze combinaties is er een versie met een zijdig voornaamwoord en met een onzijdig voornaamwoord opgesteld. Hieronder een voorbeeld van alle 8 zinnen van één naamwoord.

- (18) a. **Die/dat** bal is van mij. (AT)
 b. Hij heeft een bal op straat gezien **die/dat** hij wil hebben. (BT)
 c. De jongens waren hun bal kwijt. Ze hadden '**m'**'t op het dak gegooit. (PS)
 d. Jouw bal is veel zachter dan **de/het** mijne. (BZ)

Ik zal voortaan de zinnen waarbij met een zijdig voornaamwoord terugverwezen wordt naar het antecedent *de zijdige versie* van de zin noemen, de tegenhanger waar het voornaamwoord onzijdig is noem ik *de onzijdige versie*. We hebben hier 4 items * 4 categorieën * 4 contexten * 2 gekoppelde voornaamwoorden = 128 zinnen. De zinnen zijn in een enquête verdeeld over twee versies van elk 64 van deze zinnen. De twee versies zijn zo ingedeeld dat er nooit in één versie een minimaal paar ontstaat waar enkel het voornaamwoord verschilt (*Die bal is van mij* en *Dat bal is van mij* komen dus niet beide in dezelfde enquête voor). Vanzelfsprekend is de verdeling zo gemaakt, dat beide versies dezelfde stimulifrequentie voor iedere conditie hebben. Ook wordt van ieder naamwoord in twee van de vier contexten de zijdige versie aangeboden, en in de overige twee contexten de onzijdige versie. Zo combineren bijvoorbeeld *bal* en *stoel* in versie 1 in de contexten AT en BT met een zijdig voornaamwoord en in de contexten PS en BZ met een onzijdig voornaamwoord, terwijl versie 2 precies de omgekeerde combinatie bevat. De andere twee woorden uit deze cel, *fiets* en *pen*, krijgen vervolgens in beide versies exact het tegenovergestelde voornaamwoord van *bal* en *stoel*. De zinnen zijn opgesplitst in twee versies omdat we willen vermijden dat proefpersonen hun concentratie of taalgevoel verliezen, wat wellicht kan gebeuren bij het beoordelen van een groot aantal zinnen. Daarnaast is het door de verdeling van de minimale paren minder duidelijk voor de proefpersoon wat de testcondities precies zijn. Op die manier is de proefpersoon gedwongen een spontaan oordeel te geven, in plaats van bewust twee invullingen van een minimaal paar af te gaan wegen. We zijn immers geïnteresseerd in het spontane grammaticaliteitsoordeel, niet in wat de proefpersoon denkt dat juist is. De zinnen zijn op de beschreven manier door de twee versies heen verweven omdat het zo het evenwichtigst gespreid is. Zo voorkomen we dat een inherente discrepantie tussen de twee groepen proefpersonen de resultaten doet vertekenen.

De zinnen zijn per context grotendeels hetzelfde wat betreft de opbouw. Alle zinnen voor AT zijn van het type *Die/dat <naamwoord> is van mij*. In de zinnen voor BT is, zoals in (18b), het antecedent telkens het lijdend voorwerp van de hoofdzin en het betrekkelijk voornaamwoord het lijdend voorwerp van de bijzin. Tussen het antecedent en het voornaamwoord bevinden zich altijd drie woorden; achtereenvolgens een plaatsbepaling en het hoofdwerkwoord van de hoofdzin. De NP in de plaatsbepaling heeft geen lidwoord, zodat de betrekkelijke bijzin nooit opgevat kan worden als betrekking hebbend op deze plaats bepaling (*op straati dieⁱ ..., in bedi dieⁱ ..., enzovoorts*). PS is onderzocht aan de hand van telkens twee zinnen, waarin het antecedent ook weer in de eerste zin het lijdend voorwerp is, en het persoonlijk voornaamwoord dat ernaar verwijst het lijdend voorwerp van de tweede zin is. Tenslotte wordt BZ altijd voorgelegd in een zin van het type *Jouw <naamwoord> is veel <bijvoeglijk naamwoord> dan de/het mijne*.

AT en BZ lenen zich zeer goed voor het opstellen van een basiszin die met alle testitems kan worden ingevuld, zonder dat de zin naar mijn oordeel gefabriceerd of onwelgevoemd voorkomt. De zinnen van BT en PS zijn noodzakelijk langer en hier speelt dat probleem wel – als

we de zin aanhouden van *Hij heeft een bal op straat gezien **die/dat** hij wil hebben* krijgen we bijvoorbeeld *Hij heeft inkt op straat gezien **die/dat** hij wil hebben*. Deze zin is ook goed grammaticaal te beoordelen, maar om te voorkomen dat semantische plausibiliteit gaat interfereren is er toch gekozen voor enige variabiliteit in de precieze opbouw van de zinnen. De zin voor *kaas* is bijvoorbeeld *Mijn moeder heeft kaas op tafel gelegd **die/dat** ik niet lust*. De lengte van de zinnen is dus niet telkens exact hetzelfde, en ook wordt het naamwoord soms voorafgegaan door het onbepaald lidwoord, soms door een bezittelijk voornaamwoord, en soms (bij stofnamen) kaal aangeboden. Toch zie ik dit als een noodzakelijk kwaad dat de welgevormdheid en spontane interpretatie ten goede komt. We zijn geïnteresseerd in de invloed van het gekozen voornaamwoord op de welgevormdheid van de zin, en hoe de proefpersoon de kans inschat dat hij of zij de zin ooit zou horen, puur met betrekking tot het gekozen voornaamwoord. Als de welgevormdheid of semantische plausibiliteit mank gaat op plaatsen in de zin die niet ter zake doen, zou dit het beeld van de oordelen vervormen. Daarom is er gekozen voor een opbouw die steeds in grote lijnen identiek is, maar toch enige variabiliteit toelaat. In de attributieve en bezittelijke context is deze variatie er niet, omdat de zinnen korter en universeeler toepasbaar zijn op verschillende items. Uiteraard bevat geen van de zinnen aanwijzingen voor het lexicale geslacht van het antecedent, zoals de lidwoorden *de* en *het*.

Er zijn in SPSS 17 twee univariate ANOVA-tests uitgevoerd waarin gekeken is naar de hoofdeffecten telbaarheid, zijdigheid en zinstype, en de interactie tussen deze factoren in relatie tot de gegeven cijfers. Deze test is eenmaal uitgevoerd voor de cijfers bij de zinnen die met een onzijdig voornaamwoord werden aangeboden, en eenmaal voor de cijfers bij de zinnen met een zijdig voornaamwoord. De proefpersonen hebben elk 64 zinnen beoordeeld, bestaande uit twee zinnen in 32 minimale condities (telbaarheid * zijdigheid * zinstype). Alvorens de ANOVA-tests uit te voeren zijn de cijfers binnen deze 32 condities gemiddeld.¹⁰ Het door SPSS gegenereerde rapport bij deze test is bijgevoegd in de bijlage.

3.2.3 Taak

De proefpersonen hebben een schriftelijke enquête ingevuld waarin de stimuli in willekeurige volgorde in een lijst zijn voorgelegd. Hun is gevraagd iedere zin te waarderen met een rapportcijfer tussen 1 en 10 waarbij 1 betekent: “Dit is voor volkomen uitgesloten, deze zin bestaat niet” en 10: “Dit klinkt als een gebruikelijke zin, niets mis mee”. Daarbij is ook vermeld dat het niet gaat om wat men op school geleerd heeft, maar om hoe natuurlijk een zin klinkt. Door de opdracht zo te formuleren wordt de invloed van ongewenste factoren zoals schrijftaalgevoel gereduceerd (Featherston, 2007). De proefpersonen mochten ook kommagetallen gebruiken, zodat er geen discreet aantal categorieën was maar een continue schaal waarop de zinnen in ingedeeld werden. Featherston (2007) betoogt overtuigend dat grammaticale welgevormdheid een gradueel begrip is. Daarnaast noemt Featherston (2007) dat zelfs wanneer we een binaire “goed” “fout” optie zouden voorleggen we in het totaalbeeld toch een graduele overgang vinden tussen “goed” en “fout” in de twee meest extreme condities. Gegeven een dergelijke uitkomst zou het aanbieden van een waarderingschaal de resultaten alleen maar optimaliseren, omdat we dan niet eerst de data tot een binair onderscheid hoeven te reduceren om die vervolgens weer gradueel te interpreteren uit de statistische verdeling van de oordelen. Concreet voor dit onderzoek

¹⁰ In zo verre dat dit methodologisch niet zuiver is, is het effect hiervan enkel dat het aantal observaties gehalveerd wordt en de kans op het vinden van significante waarden kleiner wordt, nooit groter. Omdat deze operatie niet tot het vinden van incorrecte significanties kan leiden acht ik dit gerechtvaardigd. Bovendien bleken de gegeven cijfers binnen de minimale condities altijd dicht bij elkaar te liggen. Dit is te controleren in de rauwe data, die bijgeleverd is in de bijlage.

verwachten we in de condities “telbaar, onzijdig” en “niet-telbaar, zijdig” gematigde cijfers te vinden rond het midden van de schaal. Wanneer we een binaire keuze aanbieden zullen we uiteindelijk eenzelfde tussenscore vinden, die zich minder goed laat meten. Bijkomend nadeel daarbij is dat we met een binair onderscheid niet weten of de proefpersonen de grammaticaliteit unaniem als matig beoordelen (waardoor de waardering soms doorslaat naar “goed”, soms naar “fout”), willekeurig hebben gekozen, of sterk van mening verschillen.

3.2.4 Resultaten

3.2.4.1 Significantie

Alle effecten die in de volgende subsectie te zien zijn bleken inderdaad significant. Het significantieniveau bleek ondanks de halvering van het aantal observaties zeer hoog, met in vrijwel alle gevallen een waarde van $p < .001$. Ik zal in de volgende subsectie bij de weergave van de resultaten daarom niet stilstaan bij de individuele significanties – zowel telbaarheid als zijdigheid zijn voor alle contexten significant. Ook verschillen de contexten significant van elkaar als het gaat om de invloed van deze parameters. In de volgende subsectie zullen de resultaten per zinstype besproken worden. De grootte en de richting van deze effecten zal dan goed zichtbaar worden.

3.2.4.2 Weergave van de resultaten

Met de cijfers is een score berekend die aangeeft hoe hoog de zijdige versie scoort. Voor ieder testitem is het gemiddelde cijfer berekend, hier is 1 van af getrokken zodat de waarden tussen 0 en 9 liggen. Vervolgens is het cijfer van de zijdige versie gedeeld door de som van de cijfers van de zijdige en de onzijdige versie. Deze breuk levert een waarde op tussen 0 en 1 en geeft aan hoe groot het aandeel van de score is dat toebedeeld is aan de zijdige versie. Deze waarde noem ik voortaan de *zijdigheidsindex*. Een 0 drukt uit dat de zijdige versie altijd 0 punten heeft gekregen en dat de onzijdige versie alle punten heeft gekregen, een 1 drukt uit dat de onzijdige versie 0 punten en zijdig alle punten krijgt. Mogelijke invloeden op het cijfer door andere factoren dan het voornaamwoord zijn op deze manier tegen elkaar weg gedeeld. Iedere waarde geeft een verhouding aan tussen de twee versies van de zinnen waartussen alleen het voornaamwoordgeslacht verschilt. Of de zin afgezien van het voornaamwoord hoog of laag scoort in verhouding tot de andere zinnen is niet van belang en komt ook niet in deze waarden naar voren. Nu volgen de resultaten van het experiment per grammaticale context.

3.2.4.1 De Attributieve Context (AT)

| <i>Naamwoord is...</i> | Z | O | Z+O |
|------------------------|-----------------|----------|-----------------|
| T | 0,925686 | 0,095907 | 0,510796 |
| N | 0,895597 | 0,06212 | 0,478859 |
| T+N | 0,910641 | 0,079014 | 0,494828 |

Tabel 1. Zijdigheidsindex voor alle testcondities binnen AT. Waarden tot 0,5 geven aan dat verwijzen met een onzijdig voornaamwoord hoger scoort dan met een zijdig voornaamwoord,

waarden boven 0,5 (vetgedrukt) geven aan dat eenzijdig voornaamwoord hoger scoort dan een onzijdig voornaamwoord. Kansniveau ligt dus op 0,5.

In tabel 1 is te zien dat de zijdighe woorden een veel hogere zijdigheidsindex hebben dan de onzijdige woorden. Bij zinnen met zijdige naamwoorden krijgt de combinatie met een zijdig voornaamwoord gemiddeld 91,06% van de punten. Bij onzijdige naamwoorden krijgt de zijdige versie slechts 7,90% van de punten. Dit is conform de verwachting: in AT verwachten we dat de verdeling sterk door het lexicale geslacht bepaald wordt en dat het semantische systeem niet of nauwelijks invloed heeft. Het is volgens de congruentiehiërarchie de context die het minst vatbaar is voor het nieuwe systeem.

De percentages gaan richting perfect 100% zijdig en 0% zijdig, wat zou inhouden dat bij alle zinnen met zijdige naamwoorden de zijdige versie altijd een “10” toebedeeld heeft gekregen en de onzijdige versie een “1”. Bij zinnen met onzijdige naamwoorden zou dan de onzijdige versie altijd een “10” hebben gekregen en de zijdige versie een “1”. Dit is wat we in theorie zouden verwachten als het lexicaal systeem perfecte controle had. Toch wijkt een krappe 10% van de score af van wat het lexicale systeem voorschrijft. We zullen kijken of het hier een afwijking betreft die in de richting van het semantische systeem trekt.

Als het lexicale systeem perfect zou werken zouden we bij zowel de telbare als de niet-telbare condities een zijdigheidsindex van 0,5 zien – er is dan geen enkel verband tussen telbaarheid en geslacht. In de tabel krijgt de zijdige versie bij telbare naamwoorden 51,08% van de punten. Bij niet-telbare naamwoorden is dat 47,89%. Dat houdt dus in dat de zijdige versie bij de telbare naamwoorden iets hoger scoort dan de onzijdige versie, en dat bij de niet-telbare naamwoorden de onzijdige versie hoger scoort dan de zijdige versie. Dat is wat we verwachten te vinden als het semantische systeem van invloed is.

Nu liggen de waarden in de rechter kolom wel zeer dicht bij het kansniveau van 0,5. Ook deze waarden zijn significant, maar het effect is zeer zwak. Het lexicale systeem bepaalt de score bijna volledig, het semantische systeem kleurt het oordeel maar zeer licht en subtiel.

3.2.4.2 De Betrekkelijke Context (BT)

| <i>Naamwoord is...</i> | Z | O | Z+O |
|------------------------|----------------|---------|----------------|
| T | ,836994 | ,340377 | ,588685 |
| N | ,67411 | ,249568 | ,461839 |
| T+N | ,755552 | ,294972 | ,525262 |

Tabel 2. Zijdigheidsindex voor alle testcondities binnen BT. Waarden tot 0,5 geven aan dat verwijzen met een onzijdig voornaamwoord hoger scoort dan met een zijdig voornaamwoord, waarden boven 0,5 (vetgedrukt) geven aan dat een zijdig voornaamwoord hoger scoort dan een onzijdig voornaamwoord. Kansniveau ligt dus op 0,5.

Tabel 2 laat opnieuw zien dat het lexicale systeem een groot deel van de verdeling bepaalt. Zijdige naamwoorden hebben gemiddeld een zijdigheidsindex van 75,56%, onzijdige naamwoorden een index van 29,50%. De percentages liggen wel iets dichter bij kansniveau dan bij AT. Dat houdt dus in dat we in BT een minder sterk effect zien van het lexicale systeem. Dit is goed te rijmen met het idee van de congruentiehiërarchie, waar het verval van het lexicale systeem eerder in BT inzet dan in AT.

Ook hier weer kunnen we bekijken of falen van het lexicale systeem samengaat met toepassen van het semantische systeem. Bij telbare woorden zien we dat met 58,87% de zijdig versie weer een lichte voorkeur heeft dan de onzijdige versie. Bij de niet-telbare woorden is de zijdigheidsindex 46,18%, hetgeen aangeeft dat bij niet-telbare woorden een onzijdig voornaamwoord hoger scoort dan een zijdig voornaamwoord. Dit lijkt sterk op wat we vonden bij AT, alleen is de regelmaat nu wat sterker – de waarden liggen verder van kansniveau af.

We zien bij BT in vergelijking tot AT dat de lexicale regelmaat minder sterk is en de semantische regelmaat sterker. Dat is conform de congruentiehiërarchie.

3.2.4.3 De Persoonlijke Context (PS)

| <i>Naamwoord is...</i> | Z | O | Z+O |
|------------------------|----------------|---------|----------------|
| T | ,746023 | ,468701 | ,607362 |
| N | ,352535 | ,150944 | ,251739 |
| T+N | ,549279 | ,309823 | ,429551 |

Tabel 3. Zijdigheidsindex voor alle testcondities binnen PS. Waarden tot 0,5 geven aan dat verwijzen met een onzijdig voornaamwoord hoger scoort dan met een zijdig voornaamwoord, waarden boven 0,5 (vetgedrukt) geven aan dat een zijdig voornaamwoord hoger scoort dan een onzijdig voornaamwoord. Kansniveau ligt dus op 0,5.

De lexicale regelmaat in tabel 3 is vrij zwak: In zinnen met een zijdig naamwoord is er met een zijdigheidsindex van 54,93% een zeer lichte voorkeur voor een corresponderend zijdig voornaamwoord. Bij de onzijdige naamwoorden lijkt de lexicale congruentie iets sterker; Met een zijdigheidsindex van 30,98% is er een afkeur om met een zijdig voornaamwoord naar een onzijdig naamwoord te verwijzen. De mate waarin het lexicale geslacht bepalend is voor de score van de voornaamwoorden is hier veruit het zwakste van alle contexten: De waarden liggen in deze tabel het dichtst bij het kansniveau van 0,5.

Nu het lexicale systeem zo weinig invloed heeft verwachten we dat dat samengaat met een nog sterkere invloed van het semantische systeem dan we tot nu toe gezien hebben. Dat is inderdaad het geval: Met zijdigheidsindexen van 60,74% voor telbare naamwoorden en 25,17% voor niet-telbare naamwoorden blijkt dat telbaarheid van groter belang is dan woordgeslacht bij het bepalen van de voorkeur voor een zijdig of een onzijdig voornaamwoord.

Deze bevindingen zijn opnieuw in lijn met de voorspellingen van de congruentiehiërarchie. We zien in de rij AT > BT > PS de invloed van woordgeslacht sterk afnemen en parallel daaraan de invloed van telbaarheid toenemen.

3.2.4.4 De Bezittelijke Context (BZ)

| <i>Naamwoord is...</i> | Z | O | Z+O |
|------------------------|-----------------|----------|-----------------|
| T | 0,783293 | 0,437692 | 0,610492 |
| N | 0,676747 | 0,281393 | 0,47907 |
| T+N | 0,73002 | 0,359542 | 0,429551 |

Tabel 4. Zijdigheidsindex voor alle testcondities binnen BZ. Waarden tot 0,5 geven aan dat verwijzen met een onzijdig voornaamwoord hoger scoort dan met een zijdig voornaamwoord, waarden boven 0,5 (vetgedrukt) geven aan dat een zijdig voornaamwoord hoger scoort dan een onzijdig voornaamwoord. Kansniveau ligt dus op 0,5.

De mate waarin in tabel 4 het woordgeslacht tekenend is voor de score van de voornaamwoorden is ongeveer even groot als bij BT. Daarmee plaatst de invloed van het lexicale systeem zich in deze context in de middenmoot: De invloed is groter dan bij PS, maar kleiner dan bij AT.

Ook de invloed van het semantische systeem is vergelijkbaar met die bij BT, en is sterker dan bij AT en zwakker dan bij PS.

4. Discussie

4.1 Interpretatie van de Resultaten

Zoals beschreven kan het vinden van een tussenwaarde zowel duiden op een unaniem oordeel als “matig grammaticaal”, op willekeur, of op grote verschillen tussen de proefpersonen. De scores zijn erg homogeen verdeeld, ook binnen de minimale cellen van vier naamwoorden is geverifieerd of er discrepanties tussen de beoordelingen van de woorden waren, die waren er niet. De minimale cellen lijken dus inderdaad gerechtvaardigde condities, die intern geen verdeeldheid laten zien.

De resultaten komen in grote lijnen overeen met de bevindingen van Audring (2009), en volgen net als de resultaten van Audring (2009) de congruentiehiërarchie van Corbett (1979): De mate van lexicale congruentie is het grootst bij AT, kleiner bij BT en het kleinst bij PS. De cijfers maken het aannemelijk dat deze afname parallel loopt aan een toename van congruentie op basis van telbaarheid.

De resultaten voor BZ zijn echter onverwacht: De hypothese was dat deze context zich zou gedragen als PS, daarbij is genoemd dat een alternatieve hypothese is dat het zich gedraagt als AT, maar geen van beide blijkt het geval. De resultaten van BZ laten zich nog het best vergelijken met die van BT. BT heeft een conditieverschil van 0,126846 op basis van telbaarheid en een conditieverschil van 0,505984 op basis van woordgeslacht (met andere woorden, de invloed van woordgeslacht is veel groter); Voor BZ ligt dat op 0,131422 door telbaarheid tegenover 0,370478 door woordgeslacht. Samen maken de factoren echter slechts 0,5019 aan verschil uit (tegenover 1 wanneer één factor de items perfect zou indelen), hetgeen aangeeft dat noch telbaarheid, noch zijdigheid veel invloed heeft op de voorkeur. Als we bedenken dat we hier spreken over zinnen met *de/het mijne*, is dat niet vreemd – er zijn verschillende alternatieven voor deze constructie zoals *mijnes* of *van mij*, terwijl de constructies PS, BT en AT dergelijke concurrentie niet hebben en veel vaker gebruikt worden. Het is goed mogelijk dat de proefpersonen zich niet dagelijks van deze constructie bedienen, en er daarom minder gevoel over hebben. Het kan ook zijn dat de

constructie zich in principe inderdaad zou spiegelen aan PS, maar – doordat het niet veel gebruikt wordt door de proefpersonen – wat meer als schrijftaal overkomt. Wanneer een zin als schrijftaal overkomt verwachten we meer invloed van het lexicale systeem, omdat schrijftaal traditioneler is. De gematigde cijfers van BZ zouden dan een compromis zijn tussen gelijkenis aan PS enerzijds, en het schrijftaalgevoel dat het lexicale systeem promoot anderzijds.

Eerder heb ik vermeld dat het semantische systeem bij PS een noodzaak is, omdat daar sprake is van een driedig onderscheid dat niet overeenkomt met het tweeledige onderscheid van het woordgeslacht. Daarmee verwachten we direct al een grotere mate van semantische congruentie in deze context, los van de voorspellingen van de congruentiehiërarchie. De resultaten volgen de voorspelling van de congruentiehiërarchie, maar het zou niet juist zijn daar ook uit te concluderen dat het vanwege de congruentiehiërarchie is dat de resultaten zo zijn. De resultaten van AT tegenover BT zijn wel een argument voor de congruentiehiërarchie, de resultaten van BT tegenover PS mogen niet direct zo gezien worden. Hier blijkt BZ, zoals eerder beschreven, inderdaad nuttig: BZ laat semantische congruentie zien. We zien dus in AT dat de semantiek minder meespeelt dan bij BT en BZ. Ten eerste steunt die bevinding het idee van de congruentiehiërarchie, ten tweede bewijst het dat het semantische systeem ook toepassing geniet buiten PS, waarvoor het een noodzaak was.

Voorts dien ik over PS te vermelden, dat ik de de-woorden willekeurig heb gekozen. Daarmee ben ik er van uit ben gegaan dat het aan de proefpersonen niet bekend is wat het traditionele woordgeslacht – mannelijk of vrouwelijk – van de *de*-woorden is. Wanneer bijvoorbeeld een woord vrouwelijk en niet-telbaar is en we vinden een hoge score voor het, hoeft dat geen argument voor het semantische systeem te zijn, omdat *hij* lexicaal gezien evenmin het juiste voornaamwoord is als *het*. Nu bleken de meeste de-woorden (alle behalve *pen* en *verf*) volgens het woordenboek inderdaad mannelijk. Dit is echter een punt waarop woordenboeken niet altijd betrouwbaar zijn en waar we veel dialectvariatie vinden, dus zelfs wanneer we enkel volgens het woordenboek mannelijke woorden selecteren blijft het probleem bestaan. Maar aangezien het merendeel van de proefpersonen uit het midden en noorden van het land afkomstig was ben ik van mening dat het gerechtvaardigd is aan te nemen dat de sprekers niet op de hoogte zijn van een traditioneel mannelijk of vrouwelijk woordgeslacht. Slechts een enkele proefpersoon zal over deze kennis beschikken, en ook dan zullen de oordelen waarschijnlijk beïnvloed worden door de heersende norm. Dit punt kan overigens nooit als argument tegen het semantische systeem gebruikt worden, het is enkel een bron van ruis die het vinden van significante uitkomsten zou kunnen bemoeilijken. Toch bleken de uitkomsten zeer significant, en heeft de keuze van de woorden dus niet tot problemen geleid.

4.2 Algemene Discussie

We hebben nu het semantische systeem nauwkeuriger gedefinieerd en vervolgens bekeken wanneer het systeem wordt toegepast in de huidige taalsituatie. De gevonden verdeling volgt de voorspelling van congruentiehiërarchie, en daarmee spiegelt het voornaamwoordgebruik in het Nederlands zich aan de gevallen van taalverandering die Corbett (1979 & 2006) bespreekt. Als we aannemen dat we hier inderdaad met een dergelijke taalverandering te maken hebben, dan roept dat onmiddellijk een aantal vragen op.

Allereerst dringt zich de vraag op waar de invloed van semantiek zijn oorsprong vindt. Ik zal in sectie 4.2.1 naar een antwoord op deze vraag zoeken.

Als we specifiek naar het Nederlandse voorbeeld kijken is het niet direct duidelijk waarom nu het verwijzingspatroon in zijn geheel aangepast moet worden wanneer slechts een deel van het systeem verandert – het onzijdige geslacht blijft zowel in het nominale als pronominale domein

bestaan, toch zien we toepassing van onzijdige voornaamwoorden bij zijdige referenten, en zijdige voornaamwoorden bij onzijdige referenten. Op het eerste gezicht lijkt het eenvoudiger naar alle zijdige antecedenten met mannelijke voornaamwoorden te gaan verwijzen, en de verwijzing naar onzijdige antecedenten, met onzijdige voornaamwoorden, ongewijzigd te laten. Ook gegeven een totale verandering is het curieus dat het precies telbaarheid is waar de innovatie op is gebaseerd, en niet een andere factor. Ik zal in sectie 4.2.2 een antwoord zoeken op de vraag waar de invloed van telbaarheid vandaan komt.

Tenslotte kunnen we ons afvragen hoe we een dergelijke verschuiving kunnen zien als een taalverandering die cyclisch of omkeerbaar is, en niet als een eenmalige aanpassing in de geschiedenis van een taal die de taal “verbetert.” Talen worden in het paradigma van de taalwetenschap nooit als “verouderd” gezien maar als gelijkwaardige taalvarianten die op ieder punt van de geschiedenis opnieuw voor zouden kunnen komen. De vraag is nu hoe een dergelijke congruentieverandering op die manier gezien kan worden. Er wordt afstand gedaan van een lexicaal systeem en men gaat over op een semantisch systeem. Waar kwam het oude systeem vandaan, en hoe kunnen we de introductie van een semantisch systeem als cyclisch of omkeerbaar kunnen zien? In sectie 4.2.3 zal ik hier op in gaan.

4.2.1 De Oorsprong van Semantische Congruentie

Zowel in het Nederlands als in de voorbeelden van Corbett (1979 & 2006) zien we dat een oud lexicaal systeem van woordgeslachten plaats maakt voor een nieuwe regelmaat. De nieuw geschapen regelmaat is ingegeven door de semantiek.¹¹ We zullen nu de aandacht vestigen op de oorsprong van deze invloed van de semantiek.

Zoals al eerder genoemd kende in het Nederlands ook het traditionele systeem mogelijkheden te congrueren met de semantische eigenschappen van de referent. (19a) is een voorbeeld van lexicale congruentie in de schrijftaal. In deze behoudende taalvorm is de schending van het lexicale systeem in (19b) toegestaan, terwijl het semantische systeem op basis van telbaarheid niet geaccepteerd is:

- (19a) a. Een kind(o)_i heeft een veilige omgeving nodig. Zo kan het(o)_i zich het best ontwikkelen.
 b. Daar loopt het(o) meisje(o)_i dat(o)_i/*die(z) ik gisteren heb gesproken. Zij(v)/*Het(o) is erg aardig.
- (Nederlandse schrijftaal)

Deze breuk met lexicale congruentie vinden we bovendien terug in andere Germaanse taalvormen waar geen sprake is van een congruentieprobleem. Het Duits kent zowel drie nominale als drie pronominale geslachten, toch vinden we ook hier invloed van semantische eigenschappen van de referent (20). Het Zweeds kent twee nominale geslachten en corresponderende pronominale geslachten, maar ook hier is in de spreektaal een opening voor congruentie op basis van semantische eigenschappen van de referent in de werkelijkheid (21).

- (20) Da ist ja *das(o)/die(v) Fräulein(o) Schmidt. *Es(o)/Sie(v) hat mir die Kundendatei ausgedruckt.

¹¹ Ook vormelijke aspecten van de woorden kunnen van invloed zijn, maar volgens Corbett (2006) wordt een congruentiesysteem altijd tenminste deels door de semantiek bepaald.

- “Daar heb je het/de juffrouw(o) Schmidt. Het/Zij heeft de klantgegevens voor me uitgeprint.”
(Duits)
- (21) Han har två barn(o). Ett(o)/En(z) af dem gillar jag verkligen.
“Hij heeft twee kinderen(o). Eén(o)/Eén(z) van hen vind ik erg aardig.”
(Zweeds)
- (22) Den ena sidan (z) av byggnaden (z) var bränd(z)/*bränt (o), resten svartbränd(z)/svartbränt(o) men fortfarande igenkännlig(z)/igenkännligt(o).
De ene kant (z) van het gebouw (z) was verbrand(z/o), de rest (z) svartgeblakerd(z/o) maar nog steeds herkenbaar(z/o)
(Zweeds)

In de voorbeelden hier boven zien we in (20) dat lexicale congruentie is uitgesloten, in (21) en (22) zien we dat hoewel lexicale congruentie als juist beschouwd wordt, semantische congruentie mogelijk is of soms de voorkeur heeft – in het geval van (21) is de referent bezielde en daardoor semantisch zijdig, in (22) roept *resten* (*de rest*) een vage pragmatische referent op waarnaar het beste met een onzijdig voornaamwoord verwezen kan worden, vermoedelijk omdat het onzijdige geslacht ongemarkeerd is en voorkomt bij een gebrek aan kenmerken.

In deze gevallen waar het geslacht niet congrueert met naamwoord is sprake van exoforische verwijzing (Audring, 2009).¹² Exoforische verwijzing is verwijzing waarbij de verwijzende expressie niet gekoppeld is aan een in de context aanwijsbaar antecedent maar naar een referent in de werkelijkheid verwijst. *Zij* in (19b), *die* en *sie* in (20) en *en* in (21) verwijzen naar de persoon in de werkelijkheid en functioneren niet als anafoor die naar een talig antecedent verwijst zoals *het* in (19a). Dit wordt nog duidelijker in (23) en (24):

- (23) Als ze een bloemetje(ev) koopt dan zet ze ze(mv) in de WC want dan blijven ze(mv) langer goed.
(uit Audring, 2009)
- (24) Je kunt het nergens voor gebruiken, want daar is het te klein voor.

In (23) verwijst ze niet direct naar het enkelvoudige *een bloemetje* maar naar de referent die door de referentiële expressie *een bloemetje* opgeroepen wordt – namelijk de bloemen. In (24) verwijst daar niet naar *nergens* (immers, dan zou de betekenis *Het is nergens te klein voor* zijn), maar naar een abstracte groep referenten die door *nergens* opgeroepen wordt.

Exoforische verwijzing is een wijd verspreid fenomeen dat we vinden indien er geen talig antecedent is. Dat kan zijn omdat er geen kandidaat aanwezig is (bijvoorbeeld wanneer een zin met een voornaamwoord in isolatie wordt geuit, zoals bij deiktische verwijzing) of wanneer er aanleiding gezien wordt om af te zien van congruentie met een mogelijk antecedent, zoals in (23). Voor talen met een sterk lexicaal systeem geldt doorgaans dat de kans op deze exoforische verwijzing toeneemt als er zich veel materiaal tussen het antecedent en het voornaamwoord bevindt (verschillende voorbeelden in Corbett [2006] en voor het Nederlands, Audring [2009]). Als er zo veel materiaal tussen staat dat het punt is bereikt dat het woordgeslacht van het

¹² Wanneer een voornaamwoord verwijst naar een entiteit die zich niet in de context bevindt kunnen we dit benoemen als deïxis, echter, een kenmerk van deïxis is doorgaans dat er een nieuwe referent geïntroduceerd wordt. Dat is bij de besproken voorbeelden niet het geval. Anaforische en deiktische verwijzing zijn in de literatuur niet eenduidig gedefinieerd en vormen geen duidelijke complementaire distributie. Voor een discussie over anaforische, deiktische en exoforische verwijzing, zie Audring (2009).

antecedent simpelweg uit het werkgeheugen is verdwenen, kan men enkel nog exoforisch verwijzen of het naamwoord herhalen. Het Spaans heeft een zeer strikt lexicaal systeem waar exoforische verwijzing minimaal is. Wanneer de afstand erg groot wordt, kan een spreker zich van de strategie bedienen om het naamwoord te herhalen, om zo problemen te vermijden. Dit gebeurt bijvoorbeeld in het Spaans:

- (25) a. Compré ésta revista(v)_i. La(v)_i/*Lo(m)_i leí y me gusta.
 “Ik heb dit blad gekocht. Ik heb het gelezen en het beviel me.”
 b. Compré ésta revista. (...) Leí la revista y me gusta.
 “Ik heb dit blad gekocht. (...) Ik heb het blad gelezen en het beviel me.”
 (Spaans)

Ik ga ervan uit dat deze strategie voortkomt uit een universele menselijke beperking is die in alle talen speelt. Ik baseer dit op het feit dat de grootte van het werkgeheugen noodzakelijk een limiet kent. Daarmee heeft de macht van het lexicale systeem altijd een grens, zelfs in een zeer strikte taal als het Spaans. Ik zal geen stipulaties maken over de maximale hoeveelheid materiaal die zich kan bevinden tussen antecedent en voornaamwoord - dat is hier ook niet van belang. Het punt dat ik wil maken is dat de grens er altijd is, hoe ver deze zich ook aan het uiterste van de congruentiehiërarchie bevindt. Voorbij die grens vinden we exoforische verwijzing waarbij kenmerken gebruikt worden anders dan de lexicale kenmerken van het talig antecedent.

Ook in Corbett (2006) worden voorbeelden aangehaald waar anaforische verwijzing (waar we congruentie op basis van woordgeslacht van het talig antecedent vinden) plaats maakt voor exoforische verwijzing (waarbij we congruentie vinden op basis van de semantische eigenschappen van de referent in de werkelijkheid). Een voorbeeld daarvan is het Russisch, waar *vrač* (“dokter”) traditioneel een mannelijk naamwoord is, maar waar de congruentie volgens Corbett (2006) steeds vaker berust op het biologische geslacht van de referent in kwestie, analoog aan de voorbeelden in (19b), (20) en (21).

Exoforische verwijzing is dus een verwijzingsmethode die niet nieuw geïntroduceerd wordt in het uitzonderlijke geval dat een taal in zijn ontwikkeling een congruentiedefect oploopt. Het is continu aanwezig en wint terrein waar anaforische verwijzing faalt. Falen kan veroorzaakt worden door een te grote afstand tussen antecedent en anafoor, maar ook door inherente defecten in het congruentiesysteem waar de anaforische verwijzing gebruik van maakt, zoals semantische dubbelzinnigheden (zoals bij het Russische *vrač*) of gaten in het voornaamwoordparadigma (zoals het ontbrekende zijdige voornaamwoord in het moderne Nederlands) waardoor de aantrekkelijkheid van anaforische verwijzing afneemt. In het Duits en in de Scandinavische talen vinden we de exoforische verwijzing in veel mindere mate, het lexicale systeem kent daar ook geen defecten zoals het Nederlandse systeem en dus is anaforische verwijzing aantrekkelijker dan het in het Nederlands is.¹³ Ook Audring (2006a) redeneert dat het gat in het Nederlandse paradigma het lexicale systeem ondermijnt en zo aanleiding geeft tot exoforische verwijzing. Dit onderzoek alsmede Audring (2009) tonen aan dat de invloedssfeer van het lexicale systeem in het Nederlands krimpt in de richting die de congruentiehiërarchie voorspelt, waarmee het Nederlands zich op dezelfde manier gedraagt als andere talen waar nieuwe congruentiesystemen oprukken.

Ik heb nu beschreven waar invloed van semantische kenmerken vandaan komt. Ik stel dat de semantische verwijzing ontstaat door exoforische verwijzing, die aantrekkelijker wordt wanneer anaforische verwijzing om welke reden dan ook problematisch wordt. Bij anaforische

¹³ Toch is het niet zo dat er helemaal nooit exoforische verwijzing optreedt wanneer er een potentieel talig antecedent aanwezig is, getuige voorbeelden (19) en (20). Het zou interessant zijn te onderzoeken of er een oorzaak aangewezen kan worden voor het voorkomen van exoforische verwijzing in deze talen.

verwijzing treedt congruentie op met de lexicale kenmerken van het antecedent, bij exoforische verwijzing ontbreekt een talig antecedent en de bijbehorende kenmerken, en worden de semantische kenmerken van de referent in de werkelijkheid gebruikt. Bij exoforische verwijzing naar bezielde referenten zien we doorgaans dat het biologische geslacht dan de vorm van het voornaamwoord bepaalt. Bij onbezielde referenten is niet op voorhand duidelijk welk geslacht we exoforische verwijzing zullen aantreffen. De logische aanname is dat men naar onbezielde referenten verwijst met onzijdige voornaamwoorden, omdat het onzijdige geslacht zoals beschreven de universeel ongemarkeerde uitgangspositie is. Vanuit dat perspectief is het onverwacht dat we in het Nederlands voor een deel van de onbezielde referenten, namelijk telbare onbezielde referenten, toch een ander geslacht aantreffen. Hieraan is de volgende subsectie gewijd.

4.2.2 De Oorsprong van de Invloed van Telbaarheid

We zijn nu aangekomen bij de vraag waarom telbaarheid voor de exoforische verwijzing in het Nederlands relevant is. Ritter & Harley beschrijft een kenmerkengemetrie waarin alle kenmerken gevat zijn die in talen kunnen meespelen bij de congruentie van een voornaamwoord. De weergave heeft een implicationeel karakter: Wanneer een taal bijvoorbeeld over een tweevoud (dualis) beschikt, kent het ook altijd een meervoud (pluralis). Ook woordgeslacht heeft een plaats in het model, al worden de factoren die daarin meespelen niet verder uitgewerkt. We zouden het model zo kunnen uitbreiden dat we het model in (9) met telbaarheid er op een bepaalde manier aan ‘koppelen’. Dat zou ons echter niet veel verder helpen: Men zou alle semantische factoren die we vinden in de talen van de wereld aan het model kunnen toevoegen, en kunnen beweren dat in iedere taal een andere tak actief is, ook wanneer we eigenlijk te maken hebben met dezelfde tak in verschillende talen. Dat brengt ons niet tot diepere inzichten. Wat het model wel oplevert is de voorspelling dat onderscheid op basis van telbaarheid alleen voorkomt als een taal ook onderscheid maakt op basis van bezieldeheid, omdat bezieldeheid in het model in (9) lager geordend is dan telbaarheid. Wanneer het telbaarheidsparameter wegvalt, vallen daarmee ook de parameters die het regeert weg. Dit is een toetsbare voorspelling waarmee we de positie van telbaarheid in het model kunnen bepalen.¹⁴ Echter, de vraag waarom nu precies telbaarheid relevant is in het Nederlands, blijft nog altijd onbeantwoord.

We hebben gezien dat het in het Middelnederlands al voorkwam dat naar stoffen en verzamelingen soms verwezen werd met een onzijdig voornaamwoord (Duinhoven, 1988). Het onzijdige geslacht vinden we ook wanneer het geen nominale constituent betreft waarnaar verwezen wordt, zoals in *Muzikaal, dat ben ik niet en Ik ga verhuizen, wist je dat?* Dat geldt voor zowel het modern Nederlands als voor het Middelnederlands. Zoals besproken kunnen al deze voorbeelden opgevat worden als situaties waarin het voornaamwoord geen kenmerken van een antecedent reflecteert – omdat die er niet zijn, of omdat er geen toegang is. Met de overgang van het driedelige woordgeslacht naar het tweeledige (hetgeen tot het congruentieprobleem leidt) verwachten we zoals in 4.2.1 besproken veel vaker onzijdige voornaamwoorden aan te treffen. Immers, het traditionele systeem defectief is geworden en exoforische verwijzing is aantrekkelijker. Bij exoforische verwijzing naar onbezielde referenten verwachten we het onzijdig te vinden, omdat onbezielde referenten geen semantisch (dat wil zeggen, biologisch) geslacht

¹⁴ Het model zou dus voorspellen dat een taal eerder de parameter telbaarheid zou kennen dan bezieldeheid. Dat is op het eerste gezicht onwaarschijnlijk, omdat we de invloed van bezieldeheid vaak zien in talen (Corbett, 2006), terwijl de invloed van telbaarheid zoals we die in het Nederlands zien vrij uitzonderlijk lijkt. Toch lijkt telbaarheid ook in andere talen als parameter aanwezig, zie Audring (2009) voor een uitgebreide discussie over telbaarheid in de Germaanse talen.

hebben waar de congruentie op gebaseerd kan worden. Het gevolg zou dus zijn dat er een enorm aandeel verwijzing met onzijdige voornaamwoorden ontstaat.

Blijkbaar is van de onbezielde referenten toch een gedeelte mannelijk, namelijk de telbare referenten. Waarom is dat zo? Het is mogelijk dat het hier om een tegenbeweging gaat. Voor er “vage” referenten als verzamelingen en kenmerkloze referenten (denk ook aan bijzinnen en bijvoeglijk naamwoorden) werd al veel langer – in ieder geval in de middeleeuwen al – het onzijdig gebruikt. Door exoforische verwijzing dreigen nog veel meer referenten (namelijk alle onbezielde referenten) in deze restgroep te belanden. Het gebruik van mannelijke voornaamwoorden bij duidelijk telbare referenten kan hierop een tegenbeweging zijn. De spreker is niet langer in staat congruentie op basis van woordgeslacht toe te passen, maar het gebruik van het mannelijk geslacht kan aangeven dat er niet naar een kenmerkloze entiteit verwezen wordt, maar naar een entiteit die wel degelijk kenmerken heeft.¹⁵ De innovatie op basis van telbaarheid is vanuit deze redenering veel minder omvangrijk en ingrijpend dan hij op het eerste gezicht lijkt: De tendens naar kenmerkloze referenten met onzijdig te verwijzen bestond al, en blijft bestaan, maar breidt zich uit naar alles dat niet telbaar is; de tendens om naar referenten met kenmerken te verwijzen met overeenkomend woordgeslacht wordt omgezet in een tendens naar deze referenten te verwijzen met het mannelijk geslacht, omdat het woordgeslacht niet meer beschikbaar is. In termen van kenmerkenspecificatie zoals in Rooryck (1994) kunnen we dat als volgt formuleren: We kunnen [mannelijk] en [vrouwelijk] als lexicale kenmerken beschouwen, onzijdig kunnen we zien als het ontbreken van deze kenmerken. Dat is plausibel gezien het feit dat, zoals we gezien hebben, het onzijdig in uiteenlopende situaties voorkomt waar het niet naar een naamwoord verwijst. Bij het wegvallen van het onderscheid tussen de woordgeslachten mannelijk en vrouwelijk gaat de kenmerkenspecificatie van de naamwoorden op de helling. Bij verwijzing naar vage referenten ontbreekt enig kenmerk voor geslacht zoals dat altijd al was, voor referenten die discreet genoeg zijn (dat wil zeggen, zich genoeg van vage referenten onderscheiden) dient nog steeds gecongrueerd te worden met een kenmerk voor geslacht, maar deze naamwoorden zijn nu ondergespecificeerd voor geslacht, hebben [α geslacht], terwijl de restgroep van vage referenten nog altijd het geslachtskenmerk volledig mist. De keuze voor mannelijke voornaamwoorden boven vrouwelijke voornaamwoorden is logisch uit het model in (9), waar er voor het gebruik van een vrouwelijk voornaamwoord expliciet het kenmerk [vrouwelijk] nodig is. Het mannelijk geslacht is dan de uitgangspositie. Terugkomend op de definities in het model in (9) kan het hier beschreven kenmerk [α geslacht] gezien worden als synoniem voor wat ik eerder [telbaar] heb genoemd.

4.2.2 Congruëntieverandering als Cyclische Taalverandering

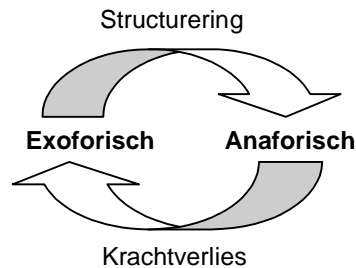
Het laatste punt waar deze discussie zich op zou richten is het vinden van een manier waarop we deze verandering als cyclisch of omkeerbaar kunnen zien, in plaats van als eenmalig en permanent.

Als we in de Nederlandse situatie te maken hebben met een werkelijke taalverandering en het is een taalverandering die zich ongestoord kan voltrekken, is de voorspelling dat ooit het punt bereikt wordt dat het lexicale systeem verdwenen is en er in alle grammaticale contexten het geslacht van het semantisch systeem gebruikt wordt. We zouden dan het volgende vinden:

¹⁵ Of naast de samenhang verzameling-onzijdig ook deze tegenbeweging telbaar-mannelijk reeds bestond in het Middelnederlands, is mij niet bekend. Duinhoven (1988) geeft hiervoor geen aanwijzing. Ook bij het doornemen van Middelnederlandse teksten heb ik hiervan geen voorbeelden kunnen vinden. Ik ga ervan uit dat dit een latere tegenbeweging is. We zien in de resultaten van het onderzoek ook dat de correlatie niet-telbaar-onzijdig sterker is dan telbaar-zijdig.

- (26) a. **Die(z)/*dat(o)** boek ligt op tafel (attributief)
 b. **De(z)** boek **die(z)/*dat(o)** op tafel ligt (relatief)
 c. **De(z)** boek ligt op tafel. Ik heb **hem(z)/*het(o)** nog niet gelezen (persoonlijk)

Dit is het punt dat er niet langer een tweestrijd is tussen twee systemen, en alle contexten weer consistent zijn. Als dit punt bereikt is en de nieuwe verwijzing is ingesleten, kan exoforische verwijzing plaats maken voor het lexicaliseren van het resultaat van die exoforische verwijzing.¹⁶ De exoforische verwijzing resulteert zo in het vastleggen van een nieuw lexicaal geslacht. Dit brengt de taal aan het begin van de cirkel, waar de taal een breed verspreid en inzichtelijk lexicaal systeem heeft. Als door verdere taalontwikkelingen de inzichtelijkheid verdwijnt wordt de invloedssfeer van het lexicale systeem weer kleiner en wordt exoforische verwijzing aantrekkelijker, en zo wordt de cyclus opnieuw op gang gebracht. In het hedendaagse Nederlands zien we telbaarheid als structurerende factor die via de exoforische verwijzing gesystematiseerd wordt, in een andere tijd of taal kan het een andere semantische eigenschap zijn. In de afbeelding hieronder een schematische weergave van deze cyclus.



Afbeelding 1. Schematische weergave van de wisselwerking tussen exoforische en anaforische verwijzing. Exoforische verwijzing schept regelmaat die gelexicaliseerd wordt. Het ontstane systeem verliest kracht door inconsistenties, ondoorzichtige generalisaties of andere oorzaken, waardoor zoals besproken in 4.2.1 exoforische verwijzing meer naar de voorgrond treedt.

Verder onderzoek kan deze visie verder bevestigen of ontcrachten. Het biedt in ieder geval een antwoord op de vraag die we voor deze subsectie stelden - zo is er een cyclus die zich oneindig kan herhalen. De volgorde van de congruentiehiërarchie is in deze visie ook logisch. Immers, hoe losser de grammaticale relatie tussen antecedentkandidaat en voornaamwoord is, hoe groter de kans is dat er niet meer naar een antecedent verwezen wordt en voor exoforische verwijzing wordt gekozen, omdat het werkgeheugen het in de periferie als eerste af laat weten. Audring (2009) bevindt dat er binnen een grammaticale context ook een significant effect is van afstand, waarbij een grote afstand tussen potentieel antecedent en voornaamwoord correleert met een groter aandeel exoforische verwijzing. Het is moeilijk directe evidentie te vinden voor deze visie die ik hier beargumenteer, maar gegeven de premissen dat een taal volgens de congruentiehiërarchie kan veranderen, en dat het proces niet eenmalig is maar zich kan herhalen,

¹⁶ Dit is geen vreemde aanname: Het is bekend dat lexicaliseren van de resultaten van een creatief vormingsproces schering en inslag is. In de morfologie zien we voortdurend dat de mogelijkheid om een affix creatief toe te passen verdwijnt en de resultaten van het vormingsproces gelexicaliseerd worden. Ieder etymologisch woordenboek bevat talloze voorbeelden hiervan. Een voorbeeld is het morfeem *-baar*, dat verwant is aan het Engelse *to bear* (dragen). Het morfeem vinden we nu in een breed scala aan woorden waarvan de betekenis niet langer transparant is - *drinkbaar* betekent dat iets gedronken kan worden, *leefbaar* betekent echter niet dat iets geleefd kan worden. Het is deze asymmetrie waaruit we concluderen dat de woorden en hun specifieke betekenissen gelexicaliseerd moeten zijn. Kennelijk heeft opslag de voorkeur gehad boven actieve computatie met het oorspronkelijk losse woord *baar*.

is een deel van deze redenering noodzakelijk waar – de instroom van semantische kenmerken bij de verwijzing moet kunnen evolueren in een nieuw vastgelegd lexicaal systeem. Daar komt bij dat wanneer deze visie verworpen wordt, het onverklaarbaar is waar de huidige woordgeslachten ooit vandaan zijn gekomen. In het geval van het Nederlands is het denkbaar dat de telbaarheid zich ooit in alle contexten zal doen gelden en zo een nieuw woordgeslacht vastlegt. Wat echter wel problematisch blijft is wat er gebeurt voor woorden waarbij exoforische verwijzing geen eenduidig geslacht naar voren brengt, zoals het woord *dokter* dat zowel naar een mannelijk of een vrouwelijk persoon kan verwijzen.

Ik heb nu beschreven hoe de verandering als cyclisch gezien kan worden. Ik heb nu een sterk geïdealiseerd beeld van een taalverandering geschetst. Maar de voorspelling is zeker niet dat bij een dergelijke verandering in de praktijk een taal ooit volledig consistente punten bereikt. Inconsistentie zien we ook bij de werkwoordsvormen: sterke en zwakke werkwoorden bestaan naast elkaar, en volgen toch elk een logisch patroon. Een soortgelijke verdeling zien we vaak bij woordgeslachten. Corbett (2006) komt met verschillende voorbeeldparadigma's waarin sprake is van een hybride indeling, waarin de woorden op basis van verschillende semantische criteria worden ingedeeld. Een dergelijke hybride indeling kan het gevolg zijn van ontlening of een reeks veranderingsprocessen die het punt van voltooiing niet gehaald hebben. Ik krijg de indruk dat het Nederlands ook zo'n hybride indeling heeft, en dat telbaarheid in het traditionele Nederlands ooit eerder verantwoordelijk is geweest voor een deel van de indeling van de woordgeslachten. *Water*, *vuur*, *geld* en *gras* zijn namelijk onzijdig, terwijl *druppel*, *vlam*, *munten* en *spriet* zijdig zijn. Ook zijn *steen*, *draad* en *haar* als telbare entiteiten zijdig, de stofnamen *steen*, *draad* en *haar* zijn onzijdig. Het is onwaarschijnlijk dat het hier om puur toeval gaat. De zijdigheid van *lucht*, *verf* en *inkt* is niet op grond van telbaarheid te verklaren, het woordgeslacht zal hier op een andere manier tot stand zijn gekomen. Het lexicale systeem zoals we dat kennen is in mijn visie dus niet het product van één perfect doorgevoerde semantische eigenschap die gelexicaliseerd is, maar van een verzameling van verschillende semantische tendensen uit het verleden die niet langer actief zijn en waarvan het resultaat gelexicaliseerd is. Het zou interessant zijn te onderzoeken of zo'n groep woorden die een bepaalde regelmaat volgt ook dezelfde etymologische achtergrond heeft, zoals bij een groep sterke werkwoorden die een bepaald consistent patroon volgen.

6. Conclusie

Er zijn in het Nederlands twee congruentiesystemen actief: Het traditionele lexicale systeem op basis van woordgeslacht en het moderne semantische systeem waarvan we in sectie 1 hebben vastgesteld dat bezielde en telbaarheid de bepalende parameters zijn. Een gedeelte van deze semantische verwijzing bestaat ook in de schrijftaal, waar het semantische systeem op basis van telbaarheid niet gebruikt wordt: Semantische verwijzing door persoonlijk voornaamwoorden bij bezielde referenten (op basis van sekse) is dus geen deel van de innovatie. De innovatie betreft de semantische verwijzing bij onbezielde referenten, en mogelijk ook een tendens semantisch te verwijzen naar bezielde referenten in contexten waar dat voorheen niet gebeurde, zoals bij de betrekkelijk voornaamwoorden (*het meisje die...*) (Audring, 2009).

Daarbij dient wel de kanttekening geplaatst te worden dat ook voor het Middelnederlands een tendens is geconstateerd naar niet-telbare referenten met het onzijdig te verwijzen – dat omgekeerd ook geldt dat naar telbare referenten met mannelijke voornaamwoorden werd verwezen, zoals dat in het moderne Nederlands gebeurt, heb ik noch middels de literatuur, noch middels Middelnederlandse data kunnen bevestigen.

In het experiment in 3 hebben we enkel naar onbezielde referenten gekeken, de vragen die we hierbij stelden zullen nu beantwoord worden.

I. Hoe trouw blijft men aan het lexicale systeem in de contexten van (12)?

Het lexicale systeem is het sterkst in AT, dan in BT, dan in BZ, dan in PS. Dit is conform de verwachting – AT beschrijft de meest nauwe syntactische relatie, PS de meest losse.

II. Kan de afwijking verklaard worden door concurrentie van het semantische systeem?

De afname van de invloed van het lexicale systeem van AT naar PS valt inderdaad samen met een toename van de voorspelbaarheid op basis van telbaarheid. Dit geldt voor alle onderzochte contexten.

III. Doet de congruentiehiërarchie van Corbett (1979) de juiste voorspellingen voor het antecedent-voornaamwoordrelaties in het hedendaagse Nederlands?

De resultaten volgen de voorspellingen van de congruentiehiërarchie van Corbett (1979), en zijn daarmee in lijn met de resultaten van Audring (2009) waar dit langs de weg van corpusonderzoek ook bevonden werd. Enige nuancering is daarbij wel op zijn plaats: Omdat de persoonlijk voornaamwoorden drie genera kennen en de antecedenten waarnaar verwezen wordt slechts twee is er hier een probleem voor het traditionele lexicale systeem. De invloed van dat lexicale systeem is als gevolg daarvan noodzakelijk minder sterk dan in de andere contexten. In de andere contexten kennen de voornaamwoorden wel het tweeledige onderscheid dat correspondeert met de woordgeslachten. Echter, ook wanneer we PS buiten beschouwing laten zijn de resultaten in lijn met de congruentiehiërarchie, een constatering die de congruentiehiërarchie steunt en aanwijzing is dat het in het Nederlands een taalverandering betreft die zich laat vergelijken met alle andere voorbeelden van taalverandering die Corbett (2006) aanhaalt.

IV. Hoe gedraagt de bezittelijke context zich in verhouding tot de andere contexten?

BZ plaatst zich in de middenmoot. Het is niet direct duidelijk waarom dit zo is, mogelijk omdat de constructie in principe PS volgt maar voor de proefpersonen een hoger schrijftaalgehalte had, waardoor de invloed van het lexicale systeem toch hoger is geweest. Het is hoe dan ook niet zonder meer duidelijk dat BZ inderdaad analoog is aan PS. Wel laat BZ samen met BT ons zien dat het semantische systeem wel degelijk toepassing vindt buiten PS, het domein waarvoor het noodzakelijk is.

Nu heb ik hier slechts 16 naamwoorden van het Nederlands aan onderzoek onderworpen, in een klein aantal zinsconstructies. In combinatie met de resultaten van Audring (2009) begint een robuuster beeld te ontstaan van de congruentie in het Nederlands, maar meer data op basis van grammaticaliteitsoordelen (over meer dan 16 naamwoorden) zou een goede aanvulling zijn. Naast verder onderzoek aangaande de verdeling van de voornaamwoorden zou het ook zeer interessant zijn als ook het waarom van het semantische systeem, van de invloed van telbaarheid, verder belicht wordt. Bij gebrek aan literatuur hierover heb ik in de discussiesectie geprobeerd een verklaring te formuleren voor de invloed van telbaarheid en heb ik daarbij verwachtingen

uitgesproken. Dat geschetste scenario kan op juistheid getoetst te worden. Verder zou nader onderzoek in het Middelnederlands zou tot interessante inzichten kunnen leiden. Ook zou onderzoek naar het Oudnederlands kunnen bijdragen aan ons begrip van de woordgeslachten die we in het moderne Nederlands kennen. Als het zo is dat woordgeslachten altijd doormiddel van semantiek tot stand komen (Corbett, 2006), dan is de kans groot dat de rol van telbaarheid in het Nederlands niet nieuw maar juist heel oud is.

I. Appendix

I.1 Sterke, Zwakke en Aanwijzende Voornaamwoorden

I.1.1 Invloed van Bezieldeheid

Zoals beschreven vinden we in het Nederlands twee groepen antecedenten waarnaar met mannelijke voornaamwoorden verwezen wordt. Beschouw de weergave van het model in (9), hieronder herhaald als (27):

(27)

| | | | | | |
|------------|--------------|------------|---------------|---|------------|
| REFERENT → | Telbaar → | Bezielde → | Vrouwspersoon | → | Vrouwelijk |
| | | | Manspersoon | → | Mannelijk |
| | | Onbezielde | | → | Mannelijk |
| | Niet-telbaar | | | → | Onzijdig |

De kenmerkenset [telbaar] en [telbaar][bezielde] selecteren beide het mannelijk geslacht. Om de kwestie helder te houden ben ik in het hele onderzoek niet ingegaan op verschillen tussen deze twee groepen. Toch brengt de toekenning van [bezielde] wel degelijk een onderscheid met zich mee dat nader onderzoek verdient. Zoals Audring (2009) ook opmerkt kunnen sterke voornaamwoorden niet gebruikt worden als er naar bezielde referenten verwezen wordt. Ter illustratie:

(28) Wat Jan_i/het huis_j betreft, ik heb ‘m_i/j/hem_i/’_j gezien.

- (29) a. Jan_i lette niet op, waardoor hij_i/-ie_i struikelde.
b. Het mes_i is geslepen, waardoor hij_i/-ie_i extra goed snijdt.

De reden hiervoor is niet duidelijk. Zodra we nadruk willen leggen bij onbezielde referenten vinden we een aanwijzend voornaamwoord. Dat kan logisch twee oorzaken hebben: Het aanwijzend voornaamwoord is feitelijk de sterke vorm in het paradigma van de onbezielde referenten, of de sterke vorm is niet beschikbaar, en de aanwijzende vorm is dan het

voorkeursalternatief. Dit laatste is beter, aangezien we bemerken dat bezielde referenten ook naast de sterke en de zwakke vorm de aanwijzende vorm kunnen krijgen:

(30) Wat Jan_i/het huis_j betreft, die_{i/j}/hem_{i/'j} heb ik gezien.

Kennelijk zijn aanwijzend, sterk en zwak alle drie beschikbaar voor referenten die [bezielt] hebben, en enkel aanwijzend en zwak voor referenten zonder dit kenmerk. Dit kunnen we formaliseren door te stellen dat de sterke vormen gespecificeerd zijn als [bezielt], waardoor ze niet kunnen congrueren met referenten zonder het kenmerk [bezielt]. De aanwijzende en zwakke vormen zijn niet gespecificeerd voor bezieldheid, waardoor ze op dit vlak met elk type referent kunnen congrueren. Dat is echter enkel een herformulering van de observatie en biedt geen diepere verklaring.

Het fundamentele verschil tussen sterke en aanwijzende voornaamwoorden is dat de laatste groep naast een identificerende functie ook een aanwijzende functie kan hebben – dat wil zeggen, met een aanwijzend voornaamwoord kan een specifieke referent uit een verzameling aangeduid worden, met een sterk voornaamwoord kan dat niet (Duinhoven, 1988). Wellicht is het zo dat bij nadruk een niet-aanwijzende vorm alleen beschikbaar is als de referent uniek is, en zijn bezielde referenten uniek. Wanneer er dan nadruk gelegd wordt, is er geen noodzaak de referent aan te wijzen uit een verzameling, en hoeft er geen aanwijzend voornaamwoord gebruikt te worden. Wanneer de referent niet uniek is, dient expliciet gemaakt te worden welke entiteit uit een verzameling bedoeld wordt. Dit idee heeft onderbouwing nodig vanuit verder onderzoek. Hier laat ik het bij de observatie dat sterke voornaamwoorden niet beschikbaar zijn voor onbezielde referenten.

In Craenenbroeck & Van Koppen (2000) wordt aangenomen dat er geen sterk onzijdig voornaamwoord bestaat (**hèt*) en men als gevolg van dit gat in het paradigma terugvalt op het aanwijzende *dat*. In het licht van de observatie hierboven lijkt het mij echter juist om het om te draaien: Omdat sterke voornaamwoorden enkel bij bezielde referenten voorkomen, en deze bezielde referenten semantisch uitsluitend mannelijk of vrouwelijk kunnen zijn, volgt daar logisch uit dat de sterke onzijdige positie in het paradigma leeg blijft.

I.1.2 Invloed van Positie van het Voornaamwoord

Wanneer het voornaamwoord in de eerste naamval staat geldt niet altijd dat een sterk voornaamwoord enkel naar een bezielde referent kan verwijzen:

(32) Wat Jan_i/het bed_j betreft, hij_{i/j} is erg groot.

Ook Audring (2009) maakt deze observatie, hier wordt echter niet verder ingegaan op een verklaring. Ik neem aan dat *ie* een cliticum is (Van Craenenbroeck & Van Koppen, 2000). Als we (29b) met (32) vergelijken zien we een cruciaal verschil: in (29b) is er positie waar een cliticum mogelijk is, in (32) is dat niet het geval. Als we aannemen dat een cliticum een hoofd is en een (sterk) voornaamwoord een frase (Cardinaletti & Starke, 1994), kan in (32) geen cliticum voorkomen omdat er een frase nodig is om het EPP van de CP te vervullen. Immers, alleen een XP kan in [Spec, CP] staan. Uitsluitend in deze gevallen wordt de zin gered door *hij*, die wel een XP structuur heeft en de positie kan opvullen.

Hoewel een sterk voornaamwoord dus in principe niet naar een onbezielde referent kan verwijzen is dat blijkbaar wel mogelijk indien het zwakke voornaamwoord niet beschikbaar is. Omdat het zwakke voornaamwoord alleen in de eerste naamval clitisch is en niet altijd geplaatst

kan worden, vinden we bij onbezielde referenten dus enkel in de eerste naamval een sterk voornaamwoord, en alleen wanneer een cliticum in die positie niet mogelijk was geweest. Dit kunnen we controleren – we verwachten dus bij onbezielde referenten in het meervoud enkel “ze” en nooit “zij” te vinden. Immers *ze* is, in tegenstelling tot *ie*, niet clitisch (Van Craenenbroeck & Van Koppen, 2000) en hoeft dus niet gered te worden door *zij*, zoals *ie* door *hij* gered wordt in (32). Dat is inderdaad het geval:

(33) Wat de glazen_i betreft, *zij_i/ze_i staan op tafel. Ik heb *hun_i/ze_i op tafel gezet.

Het voorkomen van een sterk voornaamwoord bij onbezielde referenten lijkt dus inderdaad gerelateerd aan het niet kunnen voorkomen van een zwak voornaamwoord. Dit is te formaliseren in het paradigma van de Optimaliteitstheorie als een conflict van eisen dat resulteert in een keuze voor *hij* als optimale, doch niet perfecte keuze. De spreker is gedwongen te kiezen om een voornaamwoord te gebruiken dat niet helemaal juist congrueert, of een voornaamwoord te gebruiken dat wel juist congrueert, maar waardoor het EPP geschonden wordt. Het EPP is dan hoger geordend dan “congrueer correct”, waardoor het incorrect congruerende voornaamwoord, in de afwezigheid van betere concurrenten, de verkozen oplossing is. In termen van kenmerken zouden we kunnen stellen dat *hij* niet gespecificeerd is voor bezieldeheid. We kunnen de in de vorige subsectie gepostuleerde aanwezigheid van het kenmerk [bezielde] dan iets verfijnen en stellen dat sterke voornaamwoorden ondergespecificeerd zijn voor bezieldeheid en een invulling voor dit kenmerk uit de context kunnen halen. Dit kunnen we aangeven als [α bezielde] (Rooryck, 1994). Het probleem is echter dat *ie* deze specificatie noodzakelijk ook draagt, daar het ook naar beide type referenten kan verwijzen. Daarmee slagen we er dus niet in de ongrammaticaliteit van *hij* in (29b) te verklaren. We kunnen dit pas verklaren wanneer we aannemen dat, naast deze kenmerkspecificatie, *ie* geen beginwaarde (default) heeft, terwijl *hij* standaard is ingevuld als [bezielde]. Het kenmerk kan aangepast worden, maar wanneer er concurrentie is met een element zonder de invulling heeft het niet-ingevulde element de voorkeur. Immers, daar hoeft enkel een waarde ingevuld te worden, bij *hij* moet eerst de staande waarde verwijderd worden. Het gedrag van *hij* en *ie* is dus een argument voor het idee dat er naast onderspecificatie ook uitgangspunten voor ondergespecificeerde elementen bestaan.

1.2 Conclusie

De hier geschetste relatie tussen sterke voornaamwoorden en de parameter bezieldeheid is theoretisch erg interessant en roept verschillende vragen op. Zo kunnen we afvragen of het bezieldeheid is dat een vereiste is voor het kunnen voorkomen van een sterk voornaamwoord, of dat het de identificerende functie van het aanwijzend voornaamwoord is dat vereist is bij niet-unieke referenten. Het gedrag van *hij*, dat, wanneer genoodzaakt, kan voorkomen bij onbezielde referenten is onverwacht, en biedt vragen voor verder theoretisch onderzoek naar de ϕ -kenmerken, en praktisch onderzoek naar de condities waarin *hij* bij onbezielde referenten voor kan komen. Omdat deze toepassing van *hij* een noodoplossing lijkt te zijn, is het goed mogelijk dat sprekers bij het verwijzen naar onbezielde referenten eerder dan bij bezielde referenten geneigd zijn zinsconstructies te gebruiken die het clitische element *ie* mogelijk maken. Dat zou een interessante hypothese voor verder onderzoek kunnen zijn.

Referenties

- Audring, J. (2006a). 'Pronominal Gender in Spoken Dutch'. Amsterdam: *Journal of Germanic Linguistics* 18.2 (2006):85-116
- Audring, J. (2006b). 'Genusverlies en de betekenis van voornaamwoorden', Hüning, M., Verhagen A., Vogl, U., van der Wouden, T. *Nederlands tussen Duits en Engels*, Leiden: Stichting Neerlandistiek Leiden. pp. 71-88.
- Audring, J. (2009, te verschijnen). 'Reinventing Pronoun Gender'. Dissertation VU University Amsterdam.
- Corbett, G.G. (1979) 'The agreement hierarchy'. In: *Journal of Linguistics*, 15, 203-224.
- Corbett, G.G. (2006). 'Agreement'. Cambridge University Press, Cambridge, UK, 2006.
- Craenenbroeck, J. van & M. van Koppen (2000). 'On the pronominal system of Dutch dialects', Ms, Leiden University.
- Duinhoven, A.M. (1988). 'Middel nederlandse syntaxis: synchroon en diachroon. Deel 1: De naamwoordgroep'. Uitgeverij Martinus Nijhoff.
- Featherston, F.(2007) 'Data in generative grammar: the stick and the carrot', *Theoretical Linguistics*, 33, pp. 269-318
- Harley, H. and E. Ritter. (2002). 'Person and number in pronouns: a feature-geometric analysis'. *Language* 78:482-526.
- Mesthrie, R., Swann, J., Deumert, A. and Leap, W.L. (2000). 'Introducing Sociolinguistics'. Edinburgh University Press.
- Rooryck, J. (1994). 'On two types of underspecification: towards a feature theory shared by syntax and phonology'. *Probus* 6, 207—33
- Vogelaer, G. de (2006). 'Pronominaal genus bij 'Zuid-Nederlandse' taalverwervers. van grammaticaal naar semantisch systeem'. In: Matthias Hüning, et al. (Hrsg.), *Nederlands tussen Duits en Engels*. Handelingen van de workshop op 30 september en 1 oktober 2005 aan de Freie Universität Berlin. Leiden: SNL, 89-102 (SNL-reeks, 15).

Bijlagen

Bijlage 1: De enquête.

De data is verzameld middels 128 zinnen die gelijkmatig over 2 enquêtes verdeeld waren (zoals beschreven in het onderzoek). Iedere uitgedeelde enquête bevatte de stimuli in een willekeurige volgorde, hieronder twee willekeurige exemplaren van beide enquêteversies. (De enquête bevat meer dan 64 zinnen, er zijn namelijk zeven andere zinnen aangeboden van een type dat niet gebruikt is en niet in dit onderzoek is beschreven.)

Bijlage 2: De data.

Alle rauwe data die in het experiment is verzameld.

Bijlage 3: De statistische gegevens.

Een uitdraai van de uitkomsten van de univariate ANOVA-tests die SPSS genereerde met wat toegevoegde toelichting.