

Animacy in verschillende teksttypes*

Bob van Tiel & Monique J.A. Lamers

Radboud Universiteit Nijmegen

0. Abstract

Although it is well known that in canonical transitive sentences the subject is animate and the object is lower in animacy than the subject, corpus research does not always corroborate these findings. Probably the most influential corpus study that investigated the use of animacy in transitive sentences is the study of Dahl and Fraurud (Dahl and Fraurud, 1996) on Swedish. They report a preference for constructions with animate subjects and inanimate objects (47,7%). This preference follows the markedness principles of Comrie with a subject that is higher in animacy than the object. Remarkably though, the percentage of constructions that violated this markedness principle was higher than the percentage of unmarked structures. The question arises what causes this pattern. Therefore we replicated the study of Dahl en Fraurud. In our comparative corpus study we used four different corpora that were either newspaper articles or literature texts intended for adults or children. This made it possible to investigate how the use of animacy in transitive relations is affected by genre or target group. Furthermore, the data were analyzed with respect to different animacy classifications: (i) a lexical analysis, in which only lexical items directly representing humans were qualified as animate; (ii) a referential analysis, in which lexical items that represented (a group of) human beings were also qualified as animate (e.g. committee).

Results showed an animate subject-inanimate object preference, which was independent of the animacy classification. The lexical analysis revealed an influence of genre (literature vs. newspapers) as well as target group (adults vs. children). In contrast to the findings of Dahl and Fraurud, who used a corpus with mainly textbook chapters, newspaper and scientific articles, our literature corpora showed higher occurrences of sentences with two animate NPs than with two inanimate NPs. For newspapers, however, a similar pattern was found as in Dahl and Fraurud's study with more sentences containing two inanimate NPs than two animate NPs. For the referent analysis, however, the pattern for the newspaper corpora was reversed: both genres showed a higher incidence of sentences with two animate NPs. With the animacy classification based on the referent instead of the lexical item similar patterns for all corpora were observed. Differences between genre and target group were leveled out and could no longer be observed. Since the pattern revealed high incidences of unmarked structures as well as relative high incidences for sentences with two animate arguments we concluded that the universal preference for animate subjects is confirmed. Furthermore, we argue that the referent analysis is the analysis that best reflects the processing of the reader or listener.

* Dit onderzoek is uitgevoerd in het kader van het "Animacy" project ondersteund door NWO. Dit artikel is gepresenteerd op de TaBu-dag 2007. Wij bedanken het publiek voor hun opmerkingen en suggesties. Ook bedanken wij Helen de Hoop en de overige leden van *Optimal Communication* (Radboud Universiteit Nijmegen) voor hun opmerkingen en suggesties op een eerdere versie van dit artikel.

1. Inleiding

Dat de zin ‘ik ontmoetten hun gisteren’ onwelgevormd is, behoeft geen betoog. Maar hoe zit het met een zin als ‘ik ontmoette de spijker gisteren’? Hij voldoet aan alle grammaticale regels, maar is onmogelijk zinvol te interpreteren zonder te vluchten in allerlei metaforische en poëtische verklaringen. De onwelgevormdheid van deze zin heeft te maken met de animacy van het object en de selectieve restricties van het werkwoord. We noemen een nominale constituent (NP) *animate* als deze levend is of menselijke eigenschappen heeft. Afhankelijk van het werkwoord kan het zijn dat een NP levend of juist niet levend moet zijn. Zo vereist het werkwoord *ontmoeten* dat zowel subject als object menselijk of levend is. Het werkwoord legt als het ware beperkingen op aan de argumenten. Men spreekt daarom van selectieve restricties. Dit samenspel tussen animacy en selectieve restricties speelt een belangrijke rol in de interpretatie van zinnen: spijkers ontmoet je niet, net zoals je geen gordijnen uitlacht of bejaarden afprijst, terwijl bejaarden natuurlijk wel spijkers of gordijnen kunnen afprijzen. Een zin die zich niet houdt aan de selectieve restricties van het werkwoord, is echter niet fout zoals de zin ‘ik ontmoetten hun gisteren’ dat is. Zelfs voor de zin ‘ik ontmoette de spijker gisteren’ is een context te bedenken die de zin begrijpelijk maakt. Misschien gebruik je ‘de spijker’ om een uitzonderlijk dun iemand aan te duiden. Wanneer een woord dat een niet-levende entiteit representeert toch refereert naar een persoon, is er sprake van betekenisoverdracht, zoals in het welbekende voorbeeld “Het broodje ham wil afrekenen.” (afgeleid van Nunberg, 1995). In dit voorbeeld is het niet het broodje ham dat wil betalen, maar de persoon die het besteld heeft. Hoewel een broodje ham geen menselijke eigenschappen heeft, die nodig zijn om de door het werkwoord uitgedrukte handeling uit te voeren, geeft deze zin geen interpretatieproblemen. Zelfs niet wanneer men niet in een restaurant is. De regels rond animacy zijn dus zeker niet zo absoluut en onoverkomelijk als grammaticale regels. Toch gebruiken we liever geen zinnen met betalende boterhammen en sprekende spijkers. In dit artikel zullen we dieper ingaan op de animacy van subject en object in transitieve zinnen.

Net als Dahl en Fraurud (1996) hebben wij een corpusstudie uitgevoerd waarbij we hebben gekeken naar de combinatie van de animacy van subject en object in transitieve zinnen. In het corpus gebruikt door Dahl en Fraurud (1996) kwamen met 47,7 % zinnen met een menselijk onderwerp en een niet-menselijk lijdend voorwerp het meest voor. Zinnen met twee niet-menselijke argumenten werden met 40,8% ook regelmatig aangetroffen. In 8,9% van de gevallen was zowel subject als object menselijk. De combinatie niet-menselijk subject en menselijk object kwam met 2,6% maar zeer zelden voor.

Alvorens onze eigen vergelijkende corpusstudie te bespreken, zullen we het gevonden patroon in de studie van Dahl en Fraurud bespreken en laten zien wat er zo opvallend aan hun resultaten is. Het Zweedse corpus, gebruikt door Dahl en Fraurud (1996) bestaat hoofdzakelijk uit zakelijk proza, gericht op volwassenen. Wij veronderstellen dan ook dat zowel genre als doelgroep van de teksten bepalend zijn voor het gevonden patroon. Om dit aan te tonen, hebben wij in onze vergelijkende corpusstudie vier tekstsoorten genomen, twee zakelijke, en twee literaire, waarvan telkens één gericht is op volwassenen (‘moeilijk’) en één op kinderen (‘gemakkelijk’). We verwachten dat het aantal zinnen met een menselijk subject en een niet-menselijk object in literaire en voor kinderen geschreven teksten een stuk hoger zal liggen dan in de moeilijke en zakelijke teksten. Daar waar Dahl en Fraurud transitieve zinnen hebben geanalyseerd puur op de animacy van de NP, zonder zich te bekommeren om de animacy van de entiteit waarnaar deze verwijst, hebben wij ook gekeken naar de animacy van de referent. Zo is een woord als ‘bureau’ lexicaal gezien niet-menselijk. Het kan zowel een meubelstuk zijn, als een locatie of een instantie. In de zin ‘het bureau peilde woensdag de stemming onder potentiële VVD-kiezers’, verwijst het echter duidelijk naar menselijke

wezens, bijvoorbeeld naar de telefonistes die voor het betreffende bureau werken. Uit onze vergelijkende corpusanalyse, waarin we zowel kijken naar de animacy van het lexicale item als van de referent, blijkt dat, naast de samenstelling van het corpus, ook de manier van animacyclassificatie van invloed is op de gevonden animacy-patronen in transitieve zinnen.

2. Animacyhiërarchie

Voordat we ons eigen onderzoek presenteren, zullen we eerst de zogeheten animacyhiërarchie en de problemen die hiermee verband houden bespreken. Het verschil tussen levend en niet-levend is namelijk minder eenvoudig vast te stellen dan je in eerste instantie zou denken. Hoewel biologische kenmerken doorgaans een duidelijk onderscheid mogelijk maken tussen mensen, planten, dieren en niet-levende objecten, is het aanzienlijk moeilijker deze categorieën in een (linguïstische) hiërarchie te plaatsen en binnen deze categorieën gradaties aan te brengen. Zo zijn er talen waarin bepaalde dieren op dezelfde wijze gemarkeerd worden als mensen, terwijl andere dieren, samen met de niet-levende nomina geen markering krijgen (Aissen, 2003). Een van de meest uitgebreide studies op het gebied van de animacyhiërarchie is gedaan door Yamamoto (1999). Een van de door hem gebruikte rangschikkingen is die op basis van empathie. Dieren waartoe we ons sterk verbonden voelen, zoals katten of honden, zijn hierin hoger geplaatst dan dieren waarin we onszelf minder snel inleven, zoals ratten of spinnen. Op die manier staan mensachtige robots ook hoger in de hiërarchie dan staafmixers en boormachines.

Een minder subtiele, maar veelgebruikte hiërarchie is een driedeling waarbij mensen en dieren aparte categorieën vormen. Hoog in de hiërarchie staan mensen, gevolgd door dieren, die op hun beurt weer meer animate zijn dan niet-levende zaken (1). Bij een dergelijke indeling worden planten veelal als niet-levend geïnclassificeerd omdat deze niet vrij kunnen bewegen zoals mensen en dieren dat kunnen.

(1) MENSELIJK >> DIERLIJK >> NIET-LEVEND

(2) MENSEN (animate) >> NIET-MENSEN (inanimate)

In dit artikel zullen wij ons beperken tot een relatief eenvoudige indeling. Dit is de indeling die gebruikt is door Dahl en Fraurud (1996) in bovengenoemde corpusstudie. Zij hebben een tweedeling gemaakt tussen mensen en niet-mensen (2). Hierbij hebben ze gekeken naar het lexicale item, en niet naar de referent. Alleen in het geval van een persoonlijk voornaamwoord is gekeken naar de animacy van het antecedent van dit voornaamwoord. Instanties, organisaties en geografische aanduidingen die personen representeren, werden om die reden als niet-levend geïnclassificeerd. Hoewel ze als lexicale items strikt genomen staan voor niet-levende concepten, verwijzen ze vaak naar menselijke wezens. Daarom kunnen we zonder problemen zinnen vormen als (3), waarin een niet-levend onderwerp ergens iets voor voelt. 'De nieuwe EU-landen' verwijst immers naar de regeringsleiders van die landen.

(3) Vooral de nieuwe EU-landen voelen er niets voor hun Oost-Europese bureaus nog langer te laten wachten, zodat uitstel politiek onhaalbaar lijkt.

Omdat de lexicale items dus in feite mensen representeren die betrokken zijn bij de actie uitgedrukt door het werkwoord, hebben Zaenen en collega's (2004) beargumenteerd dat niet zo zeer naar de animacy van het lexicale item moet worden gekeken, maar naar de referent (Øvrelid, 2004). Voor ons was dit aanleiding om in onze vergelijkende corpusstudie twee analyses uit te voeren, namelijk de lexicale analyse waarbij gekeken is naar de animacy van

het lexicale item zelf, en de zogenaamde referentiële analyse waarbij de animacy van de referent bepalend was.

In alle animacyhiërarchieën in de taalkundige literatuur zien we de mens bovenaan staan. Dit is waarschijnlijk te verklaren uit ons antropocentrische wereldbeeld (Dahl, 2000). In het algemeen praat men nu eenmaal meer over mensen dan over bijvoorbeeld korenwolven, yucca's, of kopieerapparaten. Mensen nemen een prominente plaats in het dagelijkse taalgebruik in. In de volgende paragraaf gaan we dieper in op prominentie in de animacyhiërarchie en linguïstische prominentie.

3. Universele animacypatronen in transitieve zinnen

Vaak blijft animacy in de grammatica's van talen onzichtbaar. In het Nederlands zien we het alleen in enkele marginale verschijnselen terugkomen. Zo hebben we het over de man *aan wie* ik me erger, maar over de computer *waaraan* ik me erger. In andere talen is animacy veel meer geïntegreerd in het taalsysteem. Vaak gaat het hierbij om naamvalsmarkering van subject of object. Een voorbeeld hiervan is de markering van het object in Malayalam (Asher and Kumari, 1997) zoals weergegeven in (4) en (5).

(4) avan oru pashuvine vaanni
 hij.nom een koe.acc kocht
 'Hij kocht een koe.'

(5) naan teenna vaanni
 ik.nom kokosnoot.nom kocht
 'Ik kocht een kokosnoot.'

Het levende object in (4) krijgt accusatiefmarkering, terwijl het niet-levende object in (5) ongemarkeerd blijft. Uit crosslinguïstisch onderzoek is gebleken dat transitieve zinnen meestal een levend subject hebben en een niet-levend object (Aissen, 2003, Comrie, 1989, Hawkins, 1983, 2004). De vraag rijst hoe deze voorkeur is ontstaan. Waarom lezen we het liefst een zin met een levend subject en een niet-levend object?

Mensen staan bovenaan in de animacyhiërarchie en zijn daarmee bij uitstek geschikt als onderwerp in een zin. De reden hiervoor is dat subjecten bij de meeste transitieve werkwoorden de rol van agens krijgen toegewezen. De agens is degene die de handeling verricht die door het werkwoord uitgedrukt wordt. In canonieke zinnen (maar niet in passieve zinnen) komt de agens overeen met de syntactische functie of formele categorie van subject (Keenan, 1975). Het verrichten van handelingen vereist specifieke eigenschappen. Mensen voldoen doorgaans aan de criteria voor de prototypische agens. Ze hebben een vrije wil, bewustzijn, en bewegingsvrijheid. Daarom mogen we stellen dat mensen de rol van subject het beste kunnen vervullen (Dowty, 1991, Hopper and Thompson, 2002, Primus, 1999). De meeste dingen bezitten deze eigenschappen niet. Ze kunnen zich niet vrijelijk bewegen, hebben geen bewustzijn en beschikken niet over een vrije wil. Hiermee voldoen ze aan de eigenschappen van de prototypische patiëns. Zoals het label van de thematische rol aangeeft, is de patiëns het zinsdeel waarop de handeling in het werkwoord is gericht. Deze semantische rol is in de meeste transitieve zinnen (maar niet alle) gelijk aan de grammaticale functie van object.

Nomina die hoog in de animacyhiërarchie staan zijn dus de beste subjecten. Naamwoorden die lager op de animacyschaal staan zijn uitstekende objecten. Men zou dus van een standaard zinsmodel kunnen spreken met een menselijk subject en een niet-levend

(niet-menselijk) object. Een dergelijke zin laat wat betreft de animacyhiërarchie een ongemarkeerd patroon zien met een subject dat hoger is in animacy dan het object. Zinnen waarin subject en object gelijk zijn in animacy en zinnen met een object dat hoger is in animacy dan het subject worden dan ook wel gemarkeerd genoemd (Comrie, 1989).

Psycholinguïstisch onderzoek heeft laten zien dat er in taalgebruik een sterke voorkeur is voor de ongemarkeerde of standaard-zinsstructuur. Dit is onder andere gebleken uit onderzoek naar taalbegrip van kinderen. Zinnen waarin het standaardmodel precies wordt omgekeerd interpreteren kinderen in veel gevallen verkeerd (Hendriks et al., 2005). Maar ook volwassenen hebben vaak meer moeite met het begrijpen van zinnen die wat betreft de animacy van de argumenten gemarkeerd zijn. Lamers heeft niet alleen laten zien dat deze zinnen moeilijker te begrijpen zijn, maar ook dat de onderliggende neurologische processen anders zijn (Lamers 2005, Lamers, 2001, de Hoop and Lamers, 2006). Vooropgesteld dat het voor taal essentieel is je boodschap zo duidelijk, efficiënt en eenduidig mogelijk over te brengen, zijn zinnen die van het standaardmodel afwijken niet wenselijk. Onze verwachting is dan ook dat deze zinnen in ons taalgebruik weinig zullen voorkomen en dus minder frequent zullen zijn in corpora.

Eerder genoemd onderzoek van Dahl en Fraurud (1996) bevestigt het hierboven geschetste voorkeurspatroon echter niet. De ongemarkeerde structuur met een subject dat hoger is in animacy dan het object, komt inderdaad in 47,7% van de gevallen voor. Maar daartegenover staat dat 49,7% van de zinnen een subject heeft dat wat animacy betreft gelijk is aan het object. Hiervan vormen de zinnen met twee niet-menselijke NPs met 40,8% het overgrote deel. Zinnen met een subject dat lager is dan het object komen met 2,6% maar heel zelden voor. Met deze resultaten is dus niet onomstotelijk bewezen dat er een sterke voorkeur bestaat voor zinnen van het standaardmodel. Integendeel, meer dan de helft van het corpus bestaat uit zinnen die in meer of mindere mate van dit model afwijken. Hoe is dit te verklaren?

Zoals hierboven al gesteld, vermoeden we dat de corpuskeuze van invloed is geweest op Dahl en Frauruds resultaten. Hun corpus bestond uit zakelijk proza: tekstboeken, brochures, kranten en debatboeken. Deze teksten zijn gericht op volwassenen die, zo mogen we veronderstellen, in staat zijn ongebruikelijke zinnen juist te interpreteren, ondanks dat deze veel meer van het begripsvermogen vergen. We veronderstellen dan ook dat we veel meer ongemarkeerde zinnen zullen aantreffen in teksten die gericht zijn op kinderen. Daarnaast verwachten we dat het genre van invloed is. Zakelijke teksten gaan doorgaans immers over andere onderwerpen dan bijvoorbeeld literaire teksten zoals een roman. Veel lexicale items uit het zakelijke corpus zijn door Dahl en Fraurud als niet-menselijk geclassificeerd, maar verwijzen in feite naar menselijke referenten. Het betreft hier bijvoorbeeld organisaties (NAVO, EU), landen (Amerika, Duitsland) en gemeenschappen (New York, Moskou). In zakelijke teksten komen dergelijke onderwerpen veel voor. Het is dus ook interessant te kijken in hoeverre de resultaten van een corpusstudie naar zakelijke teksten, veranderen als de animacy van de argumenten geclassificeerd wordt op basis van hun referent in plaats van op basis van het lexicale item. De drie hier besproken factoren (de moeilijkheid van de tekst in verband met de doelgroep, het genre of de aard van de onderwerpen van het corpus, en de manier van animacyclassificatie) hebben we in onze vergelijkende corpusstudie onderzocht. De details van deze studie worden in de volgende paragraaf besproken.

4. Vergelijkende corpusstudie

4.1 Zakelijke en niet-zakelijke teksten voor kinderen en volwassenen

Om te achterhalen of het zakelijke karakter van het corpus van Dahl en Fraurud inderdaad de oorzaak is van het hoge aantal gemarkeerde zinnen, hebben we er in dit onderzoek voor

gekozen zowel zakelijke als niet-zakelijke (literaire) teksten te gebruiken. Van beide tekstsoorten hebben we twee typen genomen, één gericht op kinderen en één gericht op volwassenen. Zakelijke teksten voor volwassenen zijn afkomstig uit *de Volkskrant* (van 17 en 18 mei 2006). Voor het zakelijk corpus geschreven voor kinderen is gekozen voor artikelen uit *Kidsweek* (van 12 en 19 mei 2006), een nieuwsblad dat zich profileert als een ‘weekkrant voor jonge mensen vanaf 10 jaar’. Gezien de doelgroep brengt deze krant het nieuws op een toegankelijker wijze dan *de Volkskrant*. *De Ontdekking van de Hemel* van Harry Mulish (1992) vormde het corpus voor de literaire tekstsoort. Criticus Jamie James zei over dit boek: ‘Net als Homerus, Dante en Milton gebruikt Mulisch de kosmos als toneel en zijn thema is de geschiedenis en betekenis van alles’. Iets dat niet snel gezegd zal worden over de literaire teksten die het niet-zakelijk corpus voor kinderen vormen, *Wipneus en Pim helpen Dokter Knippling* (Wijckmade, 1970) en *Wipneus, Pim en de Vliegende Stoel* (Wijckmade, 1977). Het verschil in moeilijkheidsgraad tussen de twee typen literaire werken behoeft geen verdere toelichting. Uit deze vier corpora zijn subjecten en objecten in transitieve zinnen als levend of niet-levend geënclassificeerd.

4.2 Methode en hypothesen

In navolging van het onderzoek van Dahl en Fraurud (1996) is gekozen voor een animacyindeling die is weergegeven in (3) hierboven waarbij mensen hoog in de hiërarchie staan en als levend worden geënclassificeerd en niet-mensen als niet-levend, laag in de hiërarchie¹. Zinnen met een menselijk subject en een niet-menselijk object vormen de ongemarkeerde standaardvorm. Andere combinaties van animacy worden gezien als afwijkend. In tabel I staan van de vier mogelijke categorieën (menselijk – niet-menselijk, menselijk – menselijk, niet-menselijk – niet-menselijk, niet-menselijk – menselijk) voorbeeldzinnen uit de corpora.

Bovenstaande animacyindeling zal gebruikt worden voor twee al eerder genoemde analyses, de lexicale analyse en de referentiële analyse. In de eerste analyse is de animacy van subject en object bepaald op basis van de menselijkheid van het lexicale item ‘zelf’. Deze analyse zal worden aangeduid met de lexicale analyse. In deze methode zijn organisaties, landen en gemeenschappen dus als niet-menselijk geënclassificeerd. Deze analyse komt overeen met die van Dahl en Fraurud. Daarnaast hebben we de zinsdelen ingedeeld op basis van de animacy van hun referenten. In deze referentiële analyse vallen organisaties, landen en gemeenschappen, mits ze naar de personen achter die instanties verwijzen, wél in de categorie ‘menselijk’. Met dit onderscheid is het in principe mogelijk duidelijk te krijgen wat de invloed van animacyclassificatie is.

We verwachten dat we in elk corpus meer gemarkeerde structuren zullen vinden als we de lexicale analyse hanteren. Bij de referentiële analyse zullen veel lexicaal gezien niet-menselijke subjecten toch het label ‘menselijk’ krijgen omdat ze een mens of groepen mensen representeren. Dit zal het aantal afwijkende zinnen verminderen. Bovendien verwachten we dat het verschil tussen deze twee analyses groter zal zijn voor zakelijke teksten dan voor literaire teksten. Zowel in *Wipneus en Pim* als in *De ontdekking van de hemel* staan immers de wedervaringen van mensen centraal. Het gaat om interacties tussen personen of tussen personen en niet-menselijke objecten. Dit semantisch gegeven zal zich ook weerspiegelen in de animacy van onderwerp en lijdend voorwerp. In krantenberichten is het gebruikelijk om met abstracte, niet-menselijke woorden naar concrete, maar weinigzeggende personen te

¹ Deze twee categorieën zullen we in de rest van dit artikel aanduiden als menselijk (M) en niet-menselijk (NM). Wanneer er gesproken wordt over gemarkeerdheid van zinnen worden er structuren bedoeld waarin deze animacyclassificatie is gebruikt, waarbij alleen zinnen met een menselijk subject en een niet-menselijk object als ongemarkeerd worden gezien.

verwijzen. Een voorbeeld hiervan is de krantenkop ‘Iran is geen kleuter, zegt Ahmadinejad tegen Europa’. Wat de president van Iran natuurlijk bedoelt, is dat de machthebbers van zijn land geen kleuters zijn, en hij richt zijn boodschap niet aan het geografisch continent Europa, maar aan de regeringsleiders van haar verschillende landen.

Wat betreft doelgroep of moeilijkheid van de corpora nemen we aan dat schrijvers rekening houden met het taalvermogen van hun doelgroep. Daarom veronderstellen we dat teksten geschreven voor kinderen meer standaardzinnen bevatten dan teksten voor volwassenen.

Uit elk van de vier corpora zijn 200 zinnen met een subject en direct object genomen en geanalyseerd op hun animacy. Alleen actieve zinnen met NPs als subject en object zijn uit de corpora geselecteerd. Passieve zinnen en zinnen met een bijzin als direct object werden dus uitgesloten. Alle te analyseren zinnen bleken hoofdzinnen te zijn. Om de steekproef zo representatief mogelijk te laten zijn, zijn de zinnen uit verschillende delen van de vier corpora gekozen. Hiermee is uitgesloten dat de resultaten te wijten zijn aan toevalligheden in een specifieke sectie van het corpus. Voor het zakelijk corpus gericht op volwassenen kwamen de onderzochte zinnen uit de eerste vijf pagina’s van de Volkskranten van 17 en 18 mei 2006². Voor het zakelijk corpus voor kinderen zijn verschillende artikelen uit de Kidsweeks van 12 en 19 mei 2006 gekozen. De zinnen uit *De ontdekking van de hemel*³ zijn afkomstig uit acht hoofdstukken en de zinnen voor de literaire tekst voor kinderen komen uit de eerste pagina’s van *Wipneus en Pim helpen Dokter Knippeling* en *Wipneus, Pim en de Vliegende Stoel*.

Tabel I: De vier mogelijke zinscategorieën op basis van de animacy van subject en direct object, elk geïllustreerd met een voorbeeldzin uit één van de corpora (M=menselijk; NM =niet-menselijk).

| | | Object | |
|---------|----|--|--|
| | | M | NM |
| Subject | M | Nog steeds hadden Wipneus en Pim de heks Kroepoek niet gevonden. (uit: Wipneus en Pim) | Iemand heeft die ellendige broodrooster natuurlijk weer gebruikt. (uit: De ontdekking van de hemel) |
| | NM | Dat warme, diep-vaderlijke element in mij moet toch ook jou vaak zijn opgevallen. (uit: De ontdekking van de hemel) | Bedrijven nemen steeds meer taken van de scholen over. (uit: de Volkskrant) |

4.3 Resultaten

Allereerst zullen de resultaten van de lexicale analyse worden gepresenteerd en besproken. Vervolgens komen de resultaten van de referentiële analyse aan bod. Om de data uit de vier corpora te kunnen vergelijken met de data van Dahl en Fraurud, zijn de voorkomens zowel in

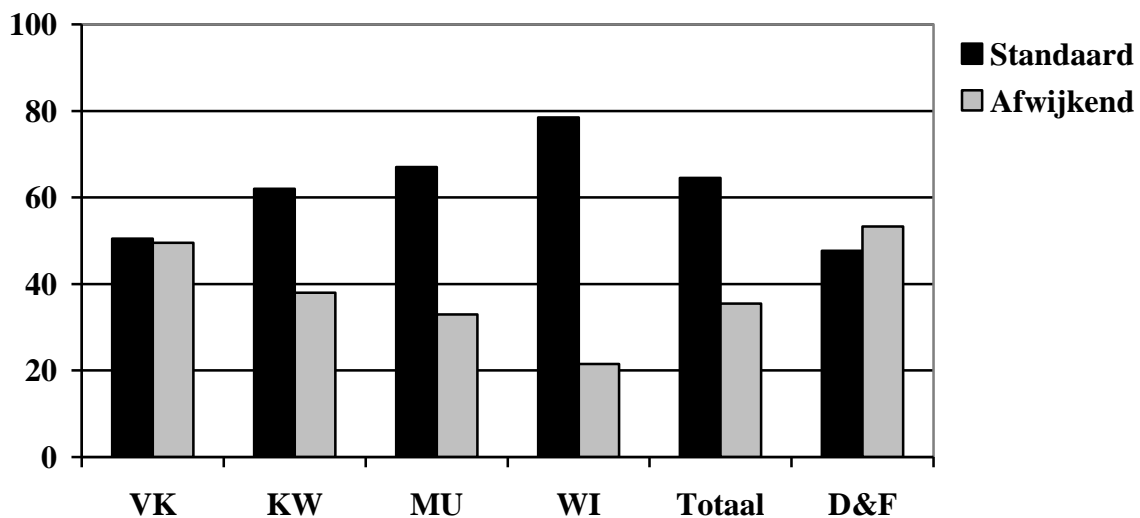
² Het betreft hier de volgende artikelen. Uit de Volkskrant van 17 mei: de hele voorpagina met uitzondering van de CaMu, het artikel ‘‘Verbijsterende’ handelswijze van partijgenoot Verdonk’ van pagina 2, van pagina 4 de artikelen ‘Gekwalificeerde allochtoon maakt bij politie weinig kans’, ‘Er knapte iets in het hoofd van Wilhelm S.’ en ‘Het schimmige karakter van Milli Görüs in Duitsland’, en van pagina 5 ‘Extra druk op Roemenië en Bulgarije’, ‘Uitstel, maar dat betekent helemaal geen afstel’, ‘President Bush zoekt naar ‘de rationele middenweg’’, ‘Einde dreigt al weer voor Linkspartei’ en ‘Britten verdeeld over Britishness als schoolvak’. Uit de Volkskrant van 18 mei: de hele voorpagina met uitzondering van de CaMu, van pagina 2 de artikelen ‘Onderzoek optreden politie Den Haag’, ‘Rekenkamer kraakt plan schooluitval’, Groningse scholen gaan sms’en en msn’en te lijf’, Beleid afrekenen is een kunst’ en ‘Zalm: zicht op overschot begroting’, van pagina 3 de artikelen ‘Deel Rutte-aanhang haakt af bij Verdonk’ en ‘Recht-door-zee Rita moet buigen’, de artikelen ‘Rechter in Turkije vermoord’ en ‘Nu al pessimisme over VN in de rol van redder van Darfur’ van pagina 4 en tenslotte van pagina 5 het artikel ‘Prodi’s kabinet is even groot als broos’

³ Het corpus bestond uit de proloog en de hoofdstukken 1, 2, 20, 21, 22, 34 en 44

aantallen als in percentages weergegeven. Van beide analyses zal eerst een totaaloverzicht van de verschillende zinsstructuren worden gegeven. Vervolgens wordt ingegaan op de details. Uiteindelijk zullen de verschillende analyses met elkaar worden vergeleken.

4.3.1 Lexicale analyse

In figuur 1 en tabel II is een overzicht gegeven van de resultaten van de lexicale analyse voor de verschillende corpora. Tabel II geeft een overzicht van standaardvorm (M/NM voor subject/object) en afwijkende vormen (NM/NM, M/M en NM/M). In figuur 2 en tabel III zijn percentages en aantallen voor alle vier de animacycombinaties apart weergegeven.



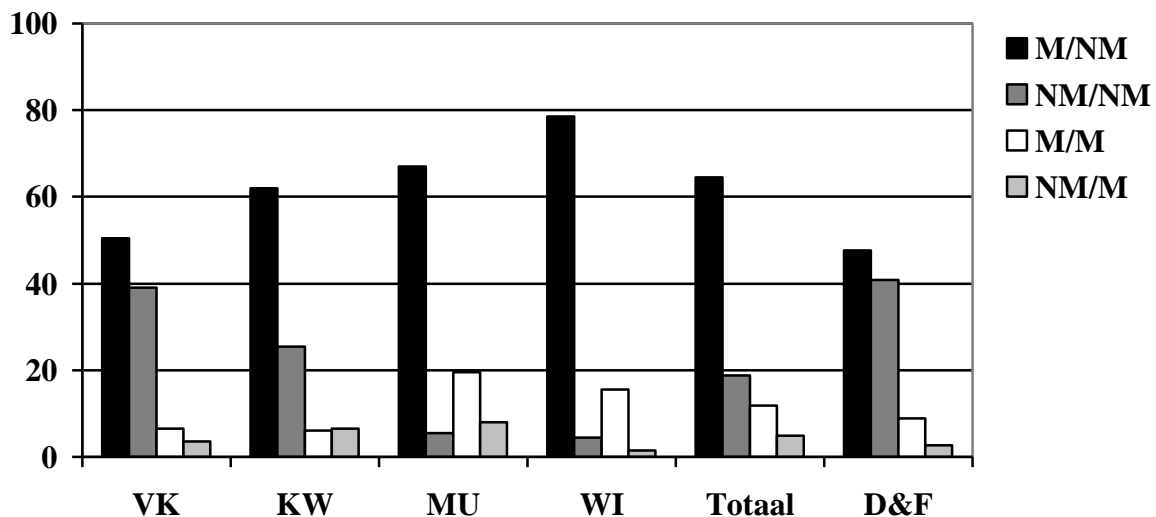
Figuur 1. Staafdiagram van voorkomens volgens de lexicale analyse van transitieve zinnen met de standaard vorm (M/NM) en afwijkende vormen samengenomen (M/M, NM/NM, NM/M) in de verschillende corpora (hier en in alle volgende grafieken en tabellen: VK = Volkskrant, KW = Kidsweek, MU = De ontdekking van de Hemel, WI = Wipneus en Pim, D&F = Dahl en Fraurud).

Tabel II. Overzicht van voorkomens van transitieve zinnen volgens de lexicale analyse met de standaardvorm (M/NM) en de afwijkende vormen (M/M, NM/NM, NM/M) samengenomen in de verschillende corpora. (hier en in alle volgende grafieken en tabellen: 'Abs.' zijn voorkomens in absolute aantallen; 'Rel.' zijn voorkomens in percentages).

| | Lexicale analyse | | | |
|----------------|------------------|-------------|------------|-------------|
| | Standaardmodel | | Afwijkend | |
| | Abs. | Rel. | Abs. | Rel. |
| Volkskrant | 101 | 50,5 | 99 | 49,5 |
| Kidsweek | 124 | 62 | 76 | 38 |
| Mulisch | 134 | 67 | 66 | 33 |
| Wipneus | 157 | 78,5 | 43 | 21,5 |
| Totaal | 516 | 64,5 | 284 | 35,5 |
| Dahl & Fraurud | -- | 47,7 | -- | 53,3 |

Waar in de data uit het Zweedse corpus van Dahl en Fraurud zinnen met de afwijkende vorm in de meerderheid zijn, treffen we in onze corpora meer ongemarkeerde zinnen aan met een menselijk subject en een niet-menselijk object (M/NM). In tabel II is te zien dat *de Volkskrant*, de zakelijke tekst geschreven voor volwassenen, hetzelfde patroon laat zien als

het corpus van Dahl en Fraurud, dat, zoals eerder gezegd, ook bestond uit moeilijke, zakelijke teksten. Dit in tegenstelling tot *Kidsweek*, waar duidelijk sprake is van een voorkeur voor ongemarkeerde zinnen met de standaard vorm. Ook in *De ontdekking van de hemel* is het aantal zinnen met de standaardvorm hoog. De literaire tekst voor kinderen spant de kroon met slechts 21,5% afwijkende zinnen. Het gemiddelde van alle Nederlandse corpora geraadpleegd in deze studie laat eveneens een voorkeur zien voor de standaardvorm, hetgeen dus afwijkt van de bevindingen van Dahl en Fraurud.



Figuur 2. Staafdiagram van voorkomens van transitieve zinnen volgens de lexicale analyse voor elk van de vier animacy combinaties (M/NM, NM/NM, M/M, NM/M) in de verschillende corpora.

Tabel III. Overzicht van voorkomens van transitieve zinnen volgens de lexicale analyse voor elk van de vier animacy combinaties (M/NM, NM/NM, M/M, NM/M) in de verschillende corpora.

| | Lexicale analyse | | | | | | | |
|------------|------------------|------|-------|-------|------|-------|------|------|
| | M/NM | | NM/NM | | M/M | | NM/M | |
| | Abs. | Rel. | Abs. | Rel. | Abs. | Rel. | Abs. | Rel. |
| Volkskrant | 101 | 50,5 | 79 | 39 | 13 | 6,5 | 7 | 3,5 |
| Kidsweek | 124 | 62 | 51 | 25,5 | 12 | 6 | 13 | 6,5 |
| Mulisch | 134 | 67 | 11 | 5,5 | 39 | 19,5 | 16 | 8 |
| Wipneus | 157 | 78,5 | 9 | 4,5 | 31 | 15,5 | 3 | 1,5 |
| Totaal | 516 | 64,5 | 150 | 18,75 | 95 | 11,85 | 39 | 4,9 |
| D&F | -- | 47,7 | -- | 40,8 | -- | 8,9 | -- | 2,6 |

Uit tabel III blijkt dat zinnen waarin het object in animacy hoger is dan het subject, maar zeer zelden voorkomen. Zinnen met twee menselijke argumenten komen vooral voor in de twee literaire corpora, terwijl het aantal zinnen met twee niet-menselijke argumenten zeldzaam zijn. Deze komen juist vaker voor in beide zakelijke corpora.

Uit een multinominale logistische regressieanalyse blijken Doelgroep, Genre en de interactie tussen Doelgroep en Genre van invloed te zijn op de voorkomens van de verschillende structuren (tabel IV).

Tabel IV. Overzicht van resultaten van multinominale regressieanalyse met Structuur als afhankelijke variabele (M/NM, NM/NM, M/M, NM/M) en Genre (zakelijk, niet-zakelijk) en Doelgroep (kinderen, volwassenen) en de interactie Doelgroep x Genre als afhankelijke variabelen van de lexicale data.

| Structuur ^a | Lexicale analyse | | | |
|------------------------|------------------|--------|-------|----------|
| | B | Wald | Sign. | Exp. (B) |
| NM/NM | | | | |
| Doelgroep | -0,643 | 8,225 | 0,004 | 0,526** |
| Genre | 1,971 | 26,752 | 0,000 | 7,175** |
| Doelgroep x Genre | 0,284 | 0,302 | 0,582 | 1,328 |
| M/M | | | | |
| Doelgroep | -0,285 | 0,456 | 0,499 | 0,752 |
| Genre | -0,713 | 3,911 | 0,048 | 0,490* |
| Doelgroep x Genre | -0,103 | 0,042 | 0,837 | 0,902 |
| NM/M | | | | |
| Doelgroep | 0,414 | 0,720 | 0,396 | 1,513 |
| Genre | 1,702 | 6,823 | 0,009 | 5,487** |
| Doelgroep x Genre | -2,246 | 7,794 | 0,005 | 0,106** |

^ade referentiecategorie is de M/NM structuur

Nagelkerke R²: 0,19

* p<0,05; ** p<0,01

Om specifieker te achterhalen wat de invloed is van doelgroep en genre hebben we de voorkomens van de verschillende animacycombinaties gerangschikt op voorkeur op basis van de voorkomens (tabel V en VI). Zoals eerder aangegeven is er een voorkeur voor niet-gemarkeerde combinaties met een menselijk subject en een niet-menselijk object. Van alle combinaties komt deze in alle corpora het meest voor. Opvallend is dat voor beide zakelijke corpora dezelfde rangschikking gevonden is als in de studie van Dahl en Fraurud, met een voorkeur voor structuren met twee niet-menselijke boven twee menselijke argumenten (NM/NM > M/M) (Tabel V). Uit chi-kwadraat analyses⁴ voor elk corpus afzonderlijk blijkt in *de Volkskrant* de animacy van subject en object onafhankelijk te zijn van elkaar ($X^2(1) = 0,580$; $p=0,446$). Echter, voor de *Kidsweek* was dit niet het geval ($X^2(1) = 5,252$; $p<0,022$). Zoals af te lezen is in tabel 5 zijn er in dit corpus relatief veel zinnen gevonden waarbij zowel het subject als het object afwijkt van de ongemarkeerde structuur. Dit aantal is vrijwel gelijk aan de zinnen met de animacycombinatie van een levend subject, maar ook een levend (en dus gemarkeerd) object (13 NM/M vs. 12 M/M).

De twee niet-zakelijke corpora laten een andere rangschikking zien met een hoger aantal zinnen met twee menselijke argumenten dan niet-menselijke argumenten (31 M/M vs. 9 NM/NM) (tabel VI). De structuur waarbij zowel het subject als het object gemarkeerd zijn (NM/M), komt met 3 voorkomens het minst voor in het niet-zakelijke corpus voor kinderen (Wipneus en Pim). Uit de chi-kwadraat toets van dit corpus blijkt dat de animacy van subject en object onafhankelijk van elkaar zijn ($X^2(1) = 0,579$; $p<0,433$). In het niet-zakelijke corpus voor volwassenen komen structuren met een niet-menselijk subject en een menselijk object vaker voor dan structuren waarin alleen het subject gemarkeerd is (16 NM/M vs 11 NM/NM). Uit de chi-kwadraattoets van dit corpus blijkt dan ook dat de animacy van subject en object afhankelijk van elkaar zijn ($X^2(1) = 15,791$; $p<0,001$).

⁴ Indien in een van de cellen een waarde lager was dan de laagste verwachte waarde is de significantie volgens de Fischer's exact test aangegeven.

Tabel V. Rangschikking van structuren in zakelijke teksten op basis van de voorkomens van verschillende animacy combinaties in transitieve zinnen in de corpora volgens de lexicale analyse. Voorkomens zijn in procenten weergegeven (voor absolute aantallen zie tabel 3).

| | Zakelijke teksten | | | |
|----------------|-------------------|-------|-----|------|
| | M/NM | NM/NM | M/M | NM/M |
| Volkscrant | 50,5 | 39 | 6,5 | 3,5 |
| Kidsweek | 62 | 25,5 | 6 | 6,5 |
| Dahl & Fraurud | 47,7 | 40,8 | 8,9 | 2,6 |
| Hiërarchie | M/NM | NM/NM | M/M | NM/M |

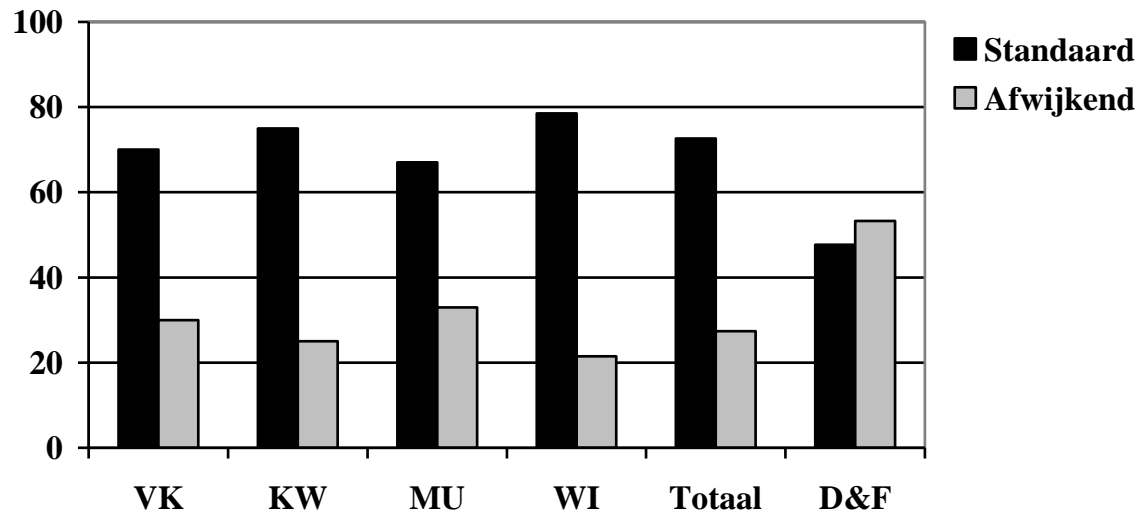
Tabel VI. Rangschikking van structuren in literaire (niet-zakelijke) teksten op basis van de voorkomens van verschillende animacy combinaties in transitieve zinnen in de corpora volgens de lexicale analyse. Voorkomens zijn in procenten weergegeven (voor absolute aantallen zie tabel 3).

| | Literaire teksten | | | |
|------------|-------------------|------|-------|------|
| | M/NM | M/M | NM/NM | NM/M |
| Mulisch | 67 | 19,5 | 5,5 | 8 |
| Wipneus | 78,5 | 15,5 | 4,5 | 1,5 |
| Hiërarchie | M/NM | M/M | NM/NM | NM/M |

Samengevat blijkt dat met een indeling waarbij gekeken is naar de animacy van het lexicale item, er zowel een invloed is van doelgroep als genre. Gemarkeerde structuren komen het minst vaak voor in teksten voor kinderen, maar ook de literaire teksten voor volwassenen laten een sterke voorkeur zien voor ongemarkeerde zinnen. Bij de zakelijke en literaire teksten zien we een andere rangschikking van aantal voorkomens voor gemarkeerde structuren. Dit verschil komt vooral tot uiting in het hoge aantal zinnen met twee niet-menselijke argumenten in zakelijke teksten. Dit aantal is aanzienlijk lager in literaire teksten, waar zinnen met twee menselijke argumenten vaker voorkomen. Terwijl in zakelijke teksten juist die voor kinderen een patroon laten zien waarin markering van subject en object afhankelijk zijn, met een relatief hoog aantal voorkomens van zinnen met zowel een gemarkeerd subject als object, is het bij literaire teksten zo dat de teksten voor volwassenen een relatief hoog aantal zinnen laat zien waarin zowel subject als object gemarkeerd zijn wat betreft de animacy. In de discussie zal hier dieper op in worden gegaan.

4.3.2 Referentiële analyse

In figuur 3 en tabel VII is een overzicht gegeven van de resultaten van de referentiële analyse. Bij deze analyse werden nomina geclassificeerd op de animacy van de entiteit waarnaar ze verwijzen. De data laten zien dat alleen in zakelijke teksten de animacy van de lexicale items niet altijd overeenkomt met de animacy van de referent. Voor de volledigheid zijn zowel de voorkomens van zakelijke als van de literaire corpora weergegeven in tabel VII.

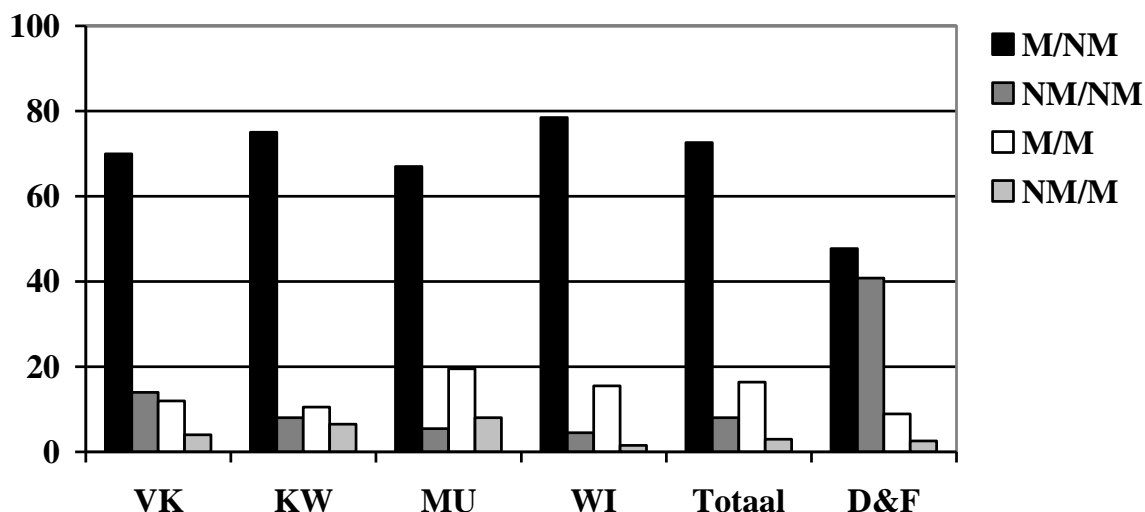


Figuur 3. Staafdiagram van voorkomens volgens de referentiële analyse van transitieve zinnen met de standaard vorm (M/NM) en afwijkende vormen samengenomen (M/M, NM/NM, NM/M) in de verschillende corpora

Tabel VII. Overzicht van voorkomens van transitieve zinnen volgens de referentiële analyse met de standaard vorm (M/NM) en de gemarkeerde, afwijkende vormen (NM/NM, M/M en NM/M) samengenomen in de verschillende corpora.

| | Referentiële analyse | | | |
|----------------|----------------------|--------------|------------|--------------|
| | Standaardmodel | | Afwijkend | |
| | Abs. | Rel. | Abs. | Rel. |
| Volkskrant | 140 | 70 | 60 | 30 |
| Kidsweek | 150 | 75 | 50 | 25 |
| Mulisch | 134 | 67 | 66 | 33 |
| Wipneus | 157 | 78,5 | 43 | 21,5 |
| Totaal | 581 | 72,63 | 219 | 27,38 |
| Dahl & Fraurud | -- | 47,7 | -- | 53,3 |

In tegenstelling tot de voorkomens volgens de lexicale analyse laat in de referentiële analyse ook *de Volkskrant* een duidelijk hoger aantal ongemarkeerde dan gemarkeerde structuren zien.



Figuur 4. Staafdiagram van voorkomens volgens de referentiële analyse van transitieve zinnen voor elk van de vier animacy combinaties (M/NM, M/M, NM/NM, NM/M) in de verschillende corpora

Tabel VIII. Overzicht van voorkomens van transitieve zinnen volgens de referentiële analyse voor elk van de vier animacy combinaties (M/NM, M/M, NM/NM, NM/M) in de verschillende corpora

| | Referentiële analyse | | | | | | | |
|------------|----------------------|-------|-------|------|------|-------|------|------|
| | M/NM | | NM/NM | | M/M | | NM/M | |
| | Abs. | Rel. | Abs. | Rel. | Abs. | Rel. | Abs. | Rel. |
| Volkskrant | 140 | 70 | 28 | 14 | 24 | 12 | 8 | 4 |
| Kidsweek | 150 | 75 | 16 | 8 | 32 | 10,5 | 2 | 6,5 |
| Mulisch | 134 | 67 | 11 | 5,5 | 39 | 19,5 | 16 | 8 |
| Wipneus | 157 | 78,5 | 9 | 4,5 | 31 | 15,5 | 3 | 1,5 |
| Totaal | 581 | 72,62 | 64 | 8 | 131 | 14,38 | 24 | 5 |
| D&F | -- | 47,7 | -- | 40,8 | -- | 8,9 | -- | 2,6 |

In figuur 4 en tabel VIII is te zien dat de classificatie op basis van de referent tot gevolg heeft dat veel (lexicaal gezien) niet-menselijke woorden toch in de categorie ‘menselijk’ komen. Hierdoor neemt het aantal voorkomens van NM/NM structuren in de twee zakelijke corpora af van 130 naar 44. Voor de NM/M structuren is er een afname van 20 naar 10 (vergelijk tabel VIII en tabel III). Het aantal standaardzinnen met een ongemarkeerd menselijke subject en niet-menselijk object is in deze referentiële analyse vrijwel gelijk aan dat van de literaire, niet-zakelijke corpora. Uit de multinominale logistische regressieanalyse blijken Genre noch Doelgroep van invloed te zijn ($R^2=0,052$).

Als gevolg van deze referentiële analyse laten zowel zakelijke als niet-zakelijke teksten een rangschikking zien zoals eerder gevonden is in de literaire, niet-zakelijke teksten, met een voorkeur voor de ongemarkeerde standaardvorm (M/NM) die het meeste voorkomt. Dan volgen structuren met twee menselijke argumenten (M/M), die op hun beurt weer vaker voorkomen dan zinnen met twee niet-menselijke argumenten (NM/NM). De laagste frequentie wordt gevonden voor de structuur met zowel een gemarkeerd subject als object (NM/M). Uitzondering hierop is, net als in de lexicale analyse, het literaire, niet-zakelijke corpus voor volwassenen (*Ontdekking van de hemel*) met een lager aantal zinnen met twee niet-menselijke argumenten (NM/NM).

In tegenstelling tot wat we zagen voor de lexicale analyse, is er bij een animacyclassificatie op basis van de referent geen aantoonbaar verschil tussen de verschillende corpora. In alle corpora komt de standaardvorm vaker voor dan de afwijkende vormen en komen structuren met twee menselijke argumenten vaker voor dan de andere afwijkende vormen.

5. Discussie en conclusie

Animacy is een belangrijke bron van informatie voor taalbegrip. Met een vergelijkende corpusstudie waarin gebruik is gemaakt van zakelijke en literaire teksten, die bedoeld zijn voor kinderen of volwassenen, is gekeken naar de invloed van doelgroep en genre op het gebruik van animacy in transitieve zinnen. Zinnen met een subject dat hoger is in animacy dan het object, dus met een menselijk subject en een niet-menselijk object, hebben we daarbij beschouwd als de ongemarkeerde standaardvorm, in navolging van onder anderen Comrie (1989). Omdat het classificeren van lexicale items niet zo eenduidig is als wel wordt verondersteld hebben we twee verschillende analyses uitgevoerd, een lexicale analyse waarbij de animacy van het subject en object bepaald is aan de hand van het lexicale item, en een referentiële analyse waarbij gekeken is naar de animacy van de referent van dat lexicale item. In beide analyses werd gebruik gemaakt van een tweedeling in animacy, te weten menselijk en niet-menselijk. Dit om een vergelijking te kunnen maken met de bevindingen van een eerdere studie in het Zweeds, waarbij gebruik gemaakt is van een corpus met alleen zakelijke teksten (Dahl & Fraurud, 1996).

Verder werd verondersteld dat zinnen met twee menselijke argumenten vaker zouden voorkomen in literaire teksten dan in zakelijke teksten. Niet-zakelijke teksten gaan immers meestal over mensen. Daarnaast was de verwachting dat in alle corpora de frequentie van zinnen waarin beide argumenten gemarkeerd zijn wat betreft animacy (en het subject dus lager is in animacy dan het object) het laagst zou zijn. Wat betreft de invloed van doelgroep werd verwacht dat in teksten voor kinderen meer ongemarkeerde structuren zouden worden gebruikt dan in teksten voor volwassenen.

Alleen bij de lexicale analyse bleek er verschil te bestaan tussen de verschillende corpora in het gebruik van gemarkeerde en de verschillende ongemarkeerde animacycombinaties. Zowel doelgroep als genre bleek van invloed. Zoals verwacht, was het aantal ongemarkeerde structuren in beide corpora voor kinderen en in het literaire corpus voor volwassenen het hoogst. *De Volkskrant*, het zakelijke corpus voor volwassenen, liet bijna evenveel gemarkeerde als ongemarkeerde combinaties zien, een patroon dat ook gevonden werd door Dahl en Fraurud (1996) die eveneens een corpus hadden gebruikt met zakelijke teksten bedoeld voor volwassenen.

De lexicale analyse laat verschillende voorkeursrangschikkingen zien voor de twee genres. Dit verschil zit echter niet zozeer in de voorkomens van de standaardvorm, als wel in die van de afwijkende vormen. De voorkeursrangschikking van de verschillende structuren op basis van de voorkomens voor literaire teksten reflecteert de voorkeur voor ongemarkeerde subjecten, door een menselijk object te prefereren boven een niet-menselijk subject. Dit levert een hiërarchie op waarin zinnen met twee menselijke NP's minder gemarkeerd zijn dan zinnen met twee niet-menselijke NP's. Hiermee wordt bevestigd dat literaire teksten veelal over mensen gaan. In zakelijke teksten is de combinatie van een niet-menselijk subject met een niet-menselijk object veel populairder dan in literaire teksten en komen ze vaker voor dan zinnen met twee menselijke argumenten.

In tegenstelling tot de voorspellingen bleek in zakelijke teksten voor kinderen de combinatie met zowel een gemarkeerd subject als object even vaak voor te komen als zinnen met twee menselijke argumenten. Echter, in de referentiële analyse bleken de meeste van deze

structuren een subject te hebben dat weliswaar als lexicaal item niet-levend is, maar wel een mens of een groep mensen representeert. Dit bleek ook voor een groot aantal niet-menselijke subjecten en objecten in *de Volkskrant* te gelden. De gemarkeerdheids hiërarchie is voor zakelijke en literaire teksten uiteindelijk gelijk: er is een overduidelijke voorkeur voor de standaardvorm. Of teksten nu bedoeld zijn voor kinderen of voor volwassenen, en of ze nu zakelijk zijn of literair, de ongemarkeerde structuur met een levend subject en een niet-levend object komt het meeste voor.

Onder de gemarkeerde structuren is er een voorkeur voor zinnen met een levend subject en komen zinnen waarin het object hoger is in animacy dan het subject het minst vaak voor. Het literaire corpus voor volwassenen vormt hier een uitzondering op. In *De ontdekking van de hemel* komen relatief veel zinnen voor met een gemarkeerd subject en object. Nadere inspectie heeft laten zien dat dit gaat om zinnen met een 'mentaal' werkwoord, zoals *deprimeren* of *bevallen*. Deze werkwoorden zijn uitzonderlijk omdat ze een menselijk (dus gemarkeerd) object vereisen (Lamers, 2001; Primus, 1999). Zoals eerder aangegeven, is bij de meeste transitieve werkwoorden een vereiste dat juist het subject levend is (Dahl and Fraurud, 1996, Hawkins, 1983). In teksten voor kinderen zullen deze mentale werkwoorden niet vaak gebruikt worden. En omdat het bij deze werkwoorden per definitie om het wel en wee van mensen gaat, is het eveneens niet te verwachten dat we ze vaak in zakelijke teksten zullen aantreffen.

Met deze vergelijkende corpusstudie is het universele patroon voor een voorkeur voor transitieve zinnen met een subject dat hoger is in animacy dan het object bevestigd. Ook bij afwijkende structuren blijkt de voorkeur voor een ongemarkeerd, menselijk subject sterk en komen zinnen met een levend object vaker voor dan zinnen met een niet-levend subject. Hoewel mensen blijkbaar graag over mensen praten, komt het niet veel voor dat ze het object prominenter maken dan het subject. Zinnen met twee niet-levende argumenten komen relatief ook vaak voor. Dit suggereert dat een zin met een niet-menselijk subject dat gelijk is in animacy met het niet-menselijke object minder gemarkeerd is dan een zin waarin zowel het subject als het object gemarkeerd zijn en het object dus hoger is in animacy dan het subject.

Dat er wel effecten gevonden zijn van genre en doelgroep in de lexicale analyse maar niet in de referentiële analyse, is niet zo verbazend. Literaire teksten bevatten vaak een heel andere onderwerpskeuze dan zakelijke teksten. Waar laatstgenoemde teksten praten over groepen mensen, generaliserend samengevat in abstracte begrippen, zitten literaire teksten veelal op het niveau van concrete personen. In dat verband is het dan ook niet verrassend dat de referentiële analyse een nivellerend effect heeft op de resultaten. De twee zakelijke teksten schuiven in de richting van de literaire. De vele afwijkende zinnen in *de Volkskrant* ondermijnen dus niet de conclusie dat mensen over mensen praten. Dit gebeurt ook in kranten, zij het op een indirectere manier. Dahl en Fraurud (1996) hebben alleen zakelijke teksten onderzocht en vonden een relatief hoog aantal gemarkeerde structuren. Juist zakelijke teksten laten echter bij de referentiële analyse een sterke verschuiving zien. Ons inziens is deze verschuiving ook te verwachten in de Zweedse data.

Met een hoog aantal ongemarkeerde structuren is het niet verbazend dat ook de invloed van doelgroep niet meer zichtbaar is bij de referentiële analyse. Waar in de lexicale analyse teksten geschreven voor kinderen een aanzienlijk groter aantal standaardzinnen lieten zien dan teksten voor volwassenen, is dat verschil verdwenen bij de toename van het aantal standaardzinnen in de referentiële analyse. Het is echter niet onwaarschijnlijk dat kinderen zakelijke teksten met meer zinnen die alleen indirect voldoen aan de ongemarkeerde standaardvorm moeilijker begrijpen dan zinnen waarin de animacy van het lexicale item ongemarkeerd is. Dit kan niet worden vastgesteld op basis van deze studie. Hiervoor moet een studie uitgevoerd worden die direct ingaat op gemak waarmee kinderen zinnen met verschillende animacycombinaties begrijpen.

Het lijkt ons duidelijk dat het classificeren van woorden op basis van hun lexicale animacy een vertekening geeft. Zelfs wanneer de referent niet gegeven is in de context, zal de lezer of luisteraar gebruik maken van zijn kennis van de wereld om de animacy van de referent af te leiden uit de context en zo tot een zinvolle interpretatie te komen. Dit is dan ook de reden dat zinnen als “het broodje ham wil afrekenen” en “Iran is geen kleuter, zegt Ahmadinejad tegen Europa” nauwelijks tot interpretatieproblemen leiden. Hoewel dit wellicht in mindere mate voor kinderen geldt, geeft de analyse waarbij classificatie plaatsvindt op basis van de referent, ons inziens beter de mentale representatie van de lezer en de schrijver weer dan wanneer er alleen gekeken wordt naar de animacy van het lexicale item zelf. Met deze animacyclassificatie wordt zowel in gemakkelijke als moeilijke en zowel in zakelijke als literaire teksten het universele patroon gevonden van een voorkeur voor zinnen waarin het subject hoger is dan het object in animacy.

Referenties

- Aissen, Judith. 2003. Differential Object Marking: Iconicity vs. Economy. . *Natural Language and Linguistic Theory* 21: 435-483.
- Asher, R.E., and Kumari, T.C. . 1997. *Malayalam*. London/New York: Routledge.
- Comrie, B. 1989. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- Dahl, Osten, and Fraurud, K. 1996. Animacy in Grammar and Discourse. In *Reference and Referent Accessibility*, eds. Thorstein Fretheim and Gundel Jeanette K., 47-64. Amsterdam: John Benjamins.
- Dahl, Osten. 2000. Egophoricity in Discourse and Syntax [1]. *Functions of Language* 7:37-77.
- de Hoop, H., and Lamers, M.J.A. 2006. Incremental distinguishability of subject and object. In *Case, Valency, and Transitivity*, eds. Leonid. Kulikov, Andrej. Malchukov and Peter. de Swart, 269-287. Amsterdam: John Benjamins.
- Dowty, David. 1991. Thematic Proto-Rules and Argument Selection [Sept]. *Language* 67:547-619.
- Hawkins, J.A. 1983. *Word Order Universals*. New York.
- Hawkins, Roger. 2004. The Contribution of the Theory of Universal Grammar to Our Understanding of the Acquisition of French as a Second Language [Nov]. *Journal of French Language Studies* 14:233-255.
- Hendriks, Petra, De Hoop, Helen, and Lamers, M.J.A. 2005. Asymmetries in language use reveal asymmetries in the grammar. Paper presented at *Proceedings of the fifteenth Amsterdam Colloquium*, Amsterdam.
- Hopper, P.J., and Thompson, S.A. 2002. Transitivity in Grammar and Discourse. In *Language*, ed. W. Bright, 251-299. Baltimore: Linguistic Society of America.
- Keenan, Edward. 1975. Variation in universal grammar. In *Analyzing variation in language* eds. Ralph Fasold and Roger Shuy, 136-148. Washington DC: Georgetown University Press.
- Lamers, M. 2005. Resolving subject-object ambiguities with and without case: evidence from ERPs. In *Case-marking: competition and variation*, eds. M. Amberber and H. de Hoop, 251-293: Elsevier.
- Lamers, M.J.A. 2001. Sentence Processing: using syntactic, semantic, and thematic information, Department of Linguistics, University of Groningen.
- Mulish, Harry. 1992. *Ontdekking van de hemel*. Amsterdam: Bezige Bij.
- Nunberg, Geoffrey. 1995. Transfers of meaning. *Journal of semantics* 12:107-132.

- Øvrelid, L. . 2004. Disambiguation of syntactic functions in Norwegian: modeling variation in word order interpretations conditioned by animacy and definiteness. . Paper presented at *Proceedings of the 20th Scandinavian Conference of Linguistics*, Helsinki.
- Primus, Beatrice. 1999. A Performance Theory of Order and Constituency [Aug]. *Linguistische Berichte* 179:368-379.
- Wijckmade, van B.J. 1970. *Wipneus en Pim helpen dokter Knippeling*. Maastricht: Schenk.
- Wijckmade, van B.J. 1977. *Wipneus, Pim en de vliegende stoel*. Maastricht: Schenk.
- Yamamoto, Mutsumi. 1999. *Animacy and Reference: A Cognitive Approach to Corpus Linguistics*.
- Zaenen, Annie, Carletta, Jean, Garretson, Gregory, Bresnan, Joan, Koontz-Garboden, Andrew, Nikitina, Tatiana, M., O'Connor, Catherine, and Wasow, Tom. 2004. Animacy encoding in English: why and how. Paper presented at *ACL Workshop on discourse annotation*, Barcelona.

Monique J.A. Lamers
Taalwetenschap
Radboud Universiteit Nijmegen
Postbus 9103
6500 HD Nijmegen