

# Een soort (van) thesis

Een corpusgebaseerd onderzoek naar het gebruik van *soort* in binominale constructies

**Robbert De Troij**

Masterproef aangeboden binnen de opleiding  
Master in de Taalkunde

Promotor prof. dr. Freek Van de Velde

Academiejaar 2016–2017

92 888 tekens





Ik verklaar me akkoord met de code of conduct van de faculteit Letteren voor geloofwaardig auteurschap.



# Inhoudstafel

---

|                                                                             |            |
|-----------------------------------------------------------------------------|------------|
| <i>Inhoudstafel</i> .....                                                   | <i>i</i>   |
| <i>Lijst van tabellen</i> .....                                             | <i>iii</i> |
| <i>Lijst van figuren</i> .....                                              | <i>iv</i>  |
| <i>Voorwoord</i> .....                                                      | <i>v</i>   |
| <b>Inleiding</b> .....                                                      | <b>1</b>   |
| <b>1 Literatuurstudie</b> .....                                             | <b>3</b>   |
| 1.1 (Bi)nominale constituenten .....                                        | 3          |
| 1.1.1 Geen nabepalingen.....                                                | 3          |
| 1.1.2 Binominale NC's .....                                                 | 5          |
| 1.2 Eigenschappen van binominale NC's met <i>soort</i> .....                | 6          |
| 1.2.1 De <i>SOORT</i> -constructie.....                                     | 6          |
| 1.2.2 Vormelijke eigenschappen.....                                         | 7          |
| 1.2.3 Semantische eigenschappen.....                                        | 9          |
| 1.3 <i>Soort</i> in de formele literatuur .....                             | 10         |
| 1.3.1 Van der Lubbe (1968).....                                             | 10         |
| 1.3.2 De ANS (Haeseryn et al. 1997).....                                    | 11         |
| 1.3.3 Vos (1999) .....                                                      | 11         |
| 1.3.4 De SOD (Broekhuis & Den Dikken 2012) .....                            | 12         |
| 1.3.5 Argumentatie voor een geïntegreerd synchroon-diachroon perspectief .. | 13         |
| 1.4 Diachrone verklaringen .....                                            | 14         |
| 1.4.1 Het ontstaan van binominale NC's.....                                 | 14         |
| 1.4.2 Herstelstrategieën.....                                               | 16         |
| 1.4.2.1 Drie mogelijke oplossingen.....                                     | 16         |
| 1.4.2.2 Symptomen van heranalyse .....                                      | 16         |
| 1.4.3 Diachrone ontwikkelingen van <i>soort</i> .....                       | 19         |
| <b>2 Vraagstelling, hypothesen en methodologie</b> .....                    | <b>22</b>  |
| 2.1 Vraagstelling en hypothesen .....                                       | 22         |
| 2.2 Methodologie .....                                                      | 23         |
| 2.2.1 Corpora.....                                                          | 23         |
| 2.2.1.1 CGN .....                                                           | 23         |
| 2.2.1.2 Corpus Van Olmen .....                                              | 23         |
| 2.2.1.3 De Gids .....                                                       | 23         |
| 2.2.1.4 SoNaR.....                                                          | 24         |
| 2.2.1.5 Condiv en Moroccorp.....                                            | 24         |
| 2.2.2 Analyse.....                                                          | 25         |
| <b>3 Resultaten</b> .....                                                   | <b>26</b>  |
| 3.1 Hypothese I (toevoeging van een relator) .....                          | 26         |
| 3.1.1 Synchroon onderzoek.....                                              | 26         |

|          |                                               |           |
|----------|-----------------------------------------------|-----------|
| 3.1.2    | Diachroon onderzoek.....                      | 29        |
| 3.2      | Hypothese II (heranalyse) .....               | 33        |
| 3.2.1    | Verandering in congruentie.....               | 33        |
| 3.2.2    | Verlies van de determinator .....             | 36        |
| 3.2.3    | Uittreding uit de NC.....                     | 41        |
| <b>4</b> | <b>Bespreking</b> .....                       | <b>43</b> |
| 4.1      | Strategie 1: toevoeging van een relator ..... | 43        |
| 4.2      | Strategie 2: heranalyse .....                 | 44        |
| <b>5</b> | <b>Conclusie</b> .....                        | <b>47</b> |
|          | <i>English summary</i> .....                  | 49        |
|          | <i>Referenties</i> .....                      | 51        |

# Lijst van tabellen

---

|                                                                                                                                    |    |
|------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabel 1: Corpusaantallen.....                                                                                                      | 24 |
| Tabel 2: Output logistische regressie, met binominale constructies als referentieniveau<br>(0).....                                | 29 |
| Tabel 3: Proporties van binominale constructies en relatorconstructies per eeuw tussen<br>1500 en 2000 (corpus Van Olmen) .....    | 30 |
| Tabel 4: Proporties van binominale constructies en relatorconstructies per decennium<br>tussen 1837 en 1936 (corpus De Gids) ..... | 32 |
| Tabel 5: Output logistische regressie, met binominale constructies als referentieniveau<br>(0).....                                | 33 |
| Tabel 6: Output logistische regressie, met binominale constructies als referentieniveau<br>(0).....                                | 40 |

# Lijst van figuren

---

|                                                                                                                                                         |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figuur 1: Distributie van binominale constructies en relatorconstructies in Nederland en Vlaanderen (CGN).....                                          | 26 |
| Figuur 2: Distributie van binominale constructies (0) en relatorconstructies (1) volgens constructies zonder (0) en met (1) Franse N2.....              | 27 |
| Figuur 3: Distributie van binominale constructies en relatorconstructies volgens determinator (CGN) .....                                               | 28 |
| Figuur 4: Effectplots van de logistische regressie.....                                                                                                 | 29 |
| Figuur 5: Proporties van binominale constructies en relatorconstructies per eeuw tussen 1500 en 2000 (corpus Van Olmen) .....                           | 30 |
| Figuur 6: Proporties van binominale constructies en relatorconstructies per decennium tussen 1837 en 1936 (corpus De Gids) .....                        | 31 |
| Figuur 7: Logistische regressie met het jaartal als predictor voor het gebruik van een relator (corpus De Gids).....                                    | 33 |
| Figuur 8: Distributie van congruentie met N1 en N2 per register als symptoom van heranalyse (SoNaR).....                                                | 34 |
| Figuur 9: Distributie voor Nederland (SoNaR) .....                                                                                                      | 34 |
| Figuur 10: Distributie voor Vlaanderen (SoNaR) .....                                                                                                    | 35 |
| Figuur 11: Diachrone verandering in congruentie (Van Olmen) .....                                                                                       | 36 |
| Figuur 12: Distributie van constructies zonder (0) en met (1) determinator volgens register (Condiv) .....                                              | 37 |
| Figuur 13: Distributie voor Nederland (Condiv) .....                                                                                                    | 38 |
| Figuur 14: Distributie voor Vlaanderen (Condiv).....                                                                                                    | 38 |
| Figuur 15: Distributie van de aan- en afwezigheid van een determinator in Condiv en Moroccorp.....                                                      | 39 |
| Figuur 16: Aan- en afwezigheid determinator volgens aan- en afwezigheid relator in Moroccorp.....                                                       | 40 |
| Figuur 17: Interactie tussen corpus en aanwezigheid determinator bij ‘soort’ (Condiv en Moroccorp) als predictor voor het gebruik van een relator ..... | 41 |



# Voorwoord

---

Deze meesterproef vormt het sluitstuk van mijn driejarige bacheloropleiding in de Taal- en Letterkunde, gevolgd door een eenjarige master in de Taalkunde. Toen ik begin september dit engagement aanging met mijn promotor, prof. Freek Van de Velde, heb ik mij met zeer veel goesting op de materie gestort. Deze meesterproef vormt dan ook het resultaat van een maandenlang proces, waarbij verschillende mensen mij steeds gesteund en geholpen hebben waar nodig. Ik wens die mensen dan ook oprecht te bedanken.

In eerste instantie gaat mijn dank uit naar professor Freek Van de Velde, die met zijn colleges over de historische grammatica van het Nederlands mijn interesse in dit studiegebied geprikkeld heeft en mij vervolgens begeleid heeft bij het schrijven van zowel mijn bachelorproef als meesterproef. Daarnaast wil ik ook mijn familie bedanken, voor hun niet-aflatende steun, zeker wanneer het allemaal iets minder vlot ging dan verwacht, en mijn goede vrienden Maxime en Sam; de ene voor de vele sloten koffie die we samen binnen werkten tijdens broodnodige adempauzes, de andere voor de gewaardeerde hulp bij de Engelse samenvatting. Ten slotte ook jij bedankt, Amber, om me door dik en dun te steunen en mijn leven sinds dit jaar zoveel rijker te maken.

Robbert De Troij  
Leuven, juni 2017



# Inleiding

---

In zijn monografie over de nominale constituent argumenteert Van de Velde (2009) dat het Nederlands geen nabepalingen kent. Met andere woorden: nomina vermijden postmodificatie zonder ingeleid te worden door wat in de taaltypologie een *relator* genoemd wordt. Het Nederlands kent echter een handvol constructies die niet lijken te voldoen aan die voorwaarde, zoals binominale constituenten zonder voorzetsel, van het type *een paar boeken*. Toch lijken voor die constructies (diachroon) een aantal ‘herstelstrategieën’ werkbaar, waardoor in bepaalde gevallen synchrone variatie bestaat tussen constructies met en zonder een relator (cf. Joosten & Vermeire 2006). Desondanks is die problematiek in de beschrijvende literatuur nog onvoldoende uitgeklaard. Dat gegeven in combinatie met de specifieke, niet-kwantificerende betekenis van binominale constituenten met *soort* als eerste nomen (Schermer-Vermeer 2008) geeft interessant voer voor het onderzoek van deze meesterproef. Daarbij wordt gestreefd naar een – zoals voorgestaan door onder anderen Bybee (2010b: 105) – geïntegreerd synchroon-diachroon perspectief. Concreet wordt getracht een antwoord te formuleren op volgende onderzoeksvraag:

- i. **Op welke manier conformeren de binominale NC’s met *soort* als N1 zich aan de relatorverplichting, zoals beschreven in Van de Velde (2009)?**

Ter beantwoording van die vraag worden twee mogelijke herstelstrategieën grondig onder de loep genomen: (i) het (diachroon) toevoegen van een relator en (ii) heranalyse. De laatste strategie, heranalyse, is geoperationaliseerd door te kijken naar enkele symptomen die zich manifesteren in geattesteerd taalgebruik, met name verandering in congruentie met het werkwoord (bv. *dat soort mensen werkt* > *dat soort mensen werken*) en het achterwege blijven van de determinator ( $\emptyset_D$  *soort* N). Daarnaast zou (*een*) *soort* (*van*) ook als bijwoordelijke bepaling op discoursniveau kunnen fungeren (bv. *ik was soort van te laat*). De algemene hypothese luidt dat binominale constituenten met *soort* inderdaad beide herstelstrategieën gebruiken.

Voor het onderzoek is gebruikgemaakt van data uit verschillende historische en hedendaagse corpora van het Nederlands. Het Corpus Gesproken Nederlands (Van Eerten 2007) en de historische corpora van Daniël Van Olmen en De Gids leverden materiaal voor de eerste herstelstrategie; het SoNaR-corpus (Oostdijk et al. 2013), Condiv (Grondelaers et al. 2000) en Moroccorp (Ruetten & Van de Velde 2013) voor de tweede. De meeste data zijn kwantitatief geanalyseerd; het gebruik van (*een*) *soort* (*van*) als bijwoordelijke bepaling op zinsniveau is exploratief en kwalitatief benaderd.

De opbouw van deze meesterproef is als volgt. Hoofdstuk 1 vormt het theoretische kader waaraan het onderzoek is opgehangen. Het hoofdstuk bevat een introductie tot de

belangrijkste basisconcepten en -ideeën (1.1), een beschrijving van de constructionele eigenschappen van binominale constituenten met *soort* (1.2), gevolgd door een evaluatie van enkele belangrijke (synchroon-theoretische) naslagwerken (1.3), met ten slotte een diachroon perspectief (1.4). Vervolgens worden de vraagstelling, hypothesen en methodologie besproken in Hoofdstuk 2. Hoofdstuk 3 presenteert de resultaten, die besproken worden in Hoofdstuk 4. Ten slotte volgt een algemene conclusie in Hoofdstuk 5.

# 1 Literatuurstudie

---

## 1.1 (Bi)nominale constituenten

### 1.1.1 Geen nabepalingen

Tamelijk recent is in Van de Velde (2009: Hoofdstuk 3) uitvoerig de stelling verdedigd dat de Nederlandse nominale constituent (NC) geen nabepalingen op het kernnomen verdraagt. Die claim gaat echter lijnrecht in tegen wat traditionele referentiegrammatica's als de *Algemene Nederlandse Spraakkunst* (ANS; Haeseryn et al. 1997: 845vv.) aannemen – net als recenter ook de *Syntax of Dutch* (SOD; Broekhuis & Keizer 2012) –, met name dat nomina door zowel voorbepalingen als ook nabepalingen gemodificeerd kunnen worden. Van de Velde (2009) toont echter uitgebreid aan dat de NC doorheen zijn historische ontwikkeling systematisch en uitsluitend aan de linkerzijde van het kernnomen (N) is aangewassen, met de chronologische ontwikkeling van respectieve *slots* of 'kavels' voor attributieve bepalingen (A), determinatoren (D) en kopbepalingen (M). Van de ontwikkeling van parallelle kavels ter rechterzijde van N is daarentegen geen sprake.<sup>1</sup> (1) geeft het sjabloon van de NC aan het begin van de eenentwintigste eeuw, (2) geeft een concreet voorbeeld van een NC met M, D, A en N.

(1) [ <sub>NC</sub> M [ D [ A [ N ] ] ] ]

(2) *zelfs* (M) *de* (D) *rijke* (A) *burgers* (N) (Van de Velde 2009: 39)

Concreet komt de claim neer op het volgende principe: N en wat traditioneel 'nabepalingen' genoemd worden, onderhouden geen *dependentierelatie*, zoals de nominale kern en de 'voorbepalingen' dat wel doen.<sup>2</sup> Onder *dependentie* in de NC moet een hiërarchische relatie worden verstaan, die minstens vormelijk gemarkeerd is, bijvoorbeeld door middel van de woordvolgorde (Van de Velde 2009: 20).<sup>3</sup> De constituent is dan het collectief van woorden waartussen zulke dependentierelaties bestaan. *Constituentie* is met andere woorden het resultaat van *dependentie* (ibid.).

Verschillende argumenten worden aangewend om te verdedigen dat de NC geen nabepalingen verdraagt. Een eerste belangrijk argument is het feit dat nabepalingen veel gemakkelijker dan voorbepalingen gescheiden kunnen voorkomen van hun nominale

---

1 Het sjabloon van de NC is een *emergente* structuur (Hopper 1987), d.w.z. dat het op elk ogenblik slechts een momentopname is van een immer voortschrijdende taalverandering, waarbij de momentane structuur getekend en gevormd is door allerlei gebruiksevents uit het verleden (Van de Velde 2009: 155–156).

2 De term 'voorbepaling' wordt hier verder gebruikt om het contrast met 'nabepalingen' expliciet aan te geven.

3 Interpretatieve hiërarchie leidt daarentegen niet tot dependentie (Van de Velde 2009: 20).

kern; een verschijnsel dat ‘discontinuïteit’ heet.<sup>4</sup> Doordat de nabepaling en de nominale kern zo gemakkelijk uiteen geplaatst kunnen – en zoals in (3) soms *moeten* – worden, luidt de argumentatie dat nabepalingen syntactisch veel minder hecht met N verbonden zijn dan voorbepalingen.

- (3) *Daar wordt weer een nieuwe lading ingenomen, die op haar beurt dan weer aan boord van het zeeschip wordt gebracht.* (Haeseryn et al. 1997: 1397, geciteerd in Van de Velde 2009: 55)

Een volgend argument, dat tevens een centrale rol speelt in het onderzoek in deze meesterproef, is de zogenaamde *relatorverplichting*. Die houdt in dat nabepalingen niet onmiddellijk kunnen volgen op het nomen waarop ze een “interpretatieve betrekking” (Van de Velde 2009: 88) hebben, maar dat ze moeten worden ingeleid door een *relator*. Relatoren – een begrip uit de taaltypologie – zijn vrije of gebonden morfemen die een semantisch-syntactische relatie markeren tussen twee *relata*; in dit geval dus N en de nabepaling (Van Langendonck 2010: 411). Relatoren verbinden op die manier niet alleen beide relata, maar scheiden ze ook: door de linguïstische afstand die ze creëren tussen de relata, zorgen relatoren tegelijkertijd ook voor een conceptuele afstand (cf. Hai-man 1985). Concreet kunnen relatoren adposities, casusmarkeerders of markeerders van onderschikking zijn (Dik 1983: 273). Zo’n relator is in het Nederlands aanwezig in voorzetselconstituenten (4), beknopte bijzinnen (5) en relatiefzinnen (6) (Van de Velde 2009: 88).

- (4) *het huis op de hoek*  
 (5) *een huis om te kopen*  
 (6) *het huis dat u wilt kopen*

Het Nederlands herbergt daarnaast echter ook een resem nabepalingen die niet ingeleid lijken te worden door een relator. Dat is op het eerste gezicht natuurlijk problematisch in het licht van de zonet geschetste stelling, maar Van de Velde (2009) oppert voor die groep drie mogelijke ‘herstelstrategieën’, om alsnog aan de relatorverplichting te voldoen. De herstelstrategieën die genoemd worden, zijn:

- (i) toevoeging van een relator;  
 (ii) syntactische omkering;  
 (iii) heranalyse (*reanalysis*; zie Harris & Campbell 1995; Croft 2000; Hopper & Traugott 2003; Campbell 2013: 273–274).

---

4 Een ander argument ontleent Van de Velde (2009) aan de taaltypologie. Greenberg (1966) poneerde reeds het universele principe dat talen consequent zijn in het plaatsen van bepalingen bij een kern. Zo ontdekte hij bijvoorbeeld dat als een taal als dominante woordvolgorde VSO heeft, bepalingen naar alle waarschijnlijkheid zullen volgen op N.

Herstelstrategie (i) spreekt voor zich: waar de nabepaling niet wordt ingeleid door een relator, wordt een relator toegevoegd. In het geval van herstelstrategie (ii) kan van een woordgroep, als in (7), een samenstelling gemaakt worden, als in (8) (Van der Horst 2006: 348; cf. ook Joosten & Vermeire 2006).

(7) *Het **aantal deelnemers** is verdubbeld.*

(8) *Het **deelnemersaantal** is verdubbeld.*

Wanneer herstelstrategie (iii) wordt toegepast, wordt het oorspronkelijke hoofd van de woordgroep geheranalyseerd als een voorbepaling bij de oorspronkelijke bepaling, die op zijn beurt kern wordt. Op die laatste strategie wordt uitvoeriger ingegaan in §4.2.2.

### 1.1.2 Binominale NC's

In het tweede volume over nominale groepen van hun recente, lijvige SOD wijden Hans Broekhuis en collega's een uitgebreid hoofdstuk aan de beschrijving van binominale NC's (zie Broekhuis & Den Dikken 2012: Hoofdstuk 4). Er zijn twee types binominale NC's. Enerzijds zijn er binominale NC's met een voorzetsel, als in (9). Die constructies vormen geen probleem, omdat het voorzetsel – in (9) *van* – dienstdoet als relator.

(9) *een schat **van** een kind*

Daarnaast bestaan er ook binominale NC's zonder voorzetsel. Die constructies zijn echter *wel* problematisch, omdat de bepaling de kern op het eerste gezicht 'relatorloos' lijkt te volgen, zoals in (10).

(10) *een paar [Ø] boeken* (Broekhuis & Den Dikken 2012: 574)

Relatorloze binominale NC's worden in de theoretische literatuur gedefinieerd als NC's die twee nomina in juxtapositie bevatten. Hoewel vanuit historisch perspectief N1 het syntactisch hoofd is en N2 de bepaling (Van de Velde 2009), is het tegenwoordig vaak onduidelijk welk van beide nomina als syntactisch hoofd van de constructie fungeert (Haeseryn et al. 1997: 875; Vos 1999; Broekhuis & Den Dikken 2012: 574vv.). Het eerste nomen van relatorloze binominale NC's (N1) drukt vaak een zekere hoeveelheid uit van de entiteit(en) waar het tweede nomen (N2) naar verwijst. De SOD noemt die constructies daarom *quantificational* of kwantificerend, en onderscheidt vijf subtypes op basis van de betekenis van N1: *quantifier nouns* (11), *measure nouns* (12), *container nouns* (13), *part nouns* (14) en *collective nouns* (15) (Broekhuis & Den Dikken 2012: 575).

- (11) *een hoop problemen*
- (12) *een kilo bonen*
- (13) *een doos pillen*
- (14) *een stuk cake*
- (15) *een groep studenten*

Ondanks de verschillen tussen die kwantificerende binominale NC's in (11)–(15), geldt voor alle woordgroepen wel dat N1 – c.q. *hoop, kilo, doos, stuk* en *groep* – globaal genomen telkens een (relatief) begrensde en tot op zekere hoogte conceptueel identificeerbare subset afbakt van entiteiten waar N2 naar verwijst – c.q. *problemen, bonen, pillen, cake* en *studenten* (cf. Vos 1999). Die kwantificerende betekenis is moeilijker terug te vinden in de voorbeelden in (16)–(18), waar een vorm van het substantief *soort* als N1 van de binominale NC optreedt. Dat geldt zeker wanneer N2 in het enkelvoud staat.

- (16) *die soort aap/apen*
- (17) *dat soort auto/auto's*
- (18) *een soort appel/appels*

Binominale NC's met *soort* als N1 wijken op verschillende vlakken af van andere binominale NC's als in (11)–(15) en kennen aldus heel eigen vormelijke en semantische eigenschappen. De relevantste eigenschappen worden kort besproken in Sectie 1.2.

## 1.2 Eigenschappen van binominale NC's met *soort*

### 1.2.1 De SOORT-constructie

Door hun vormelijke en semantische eenheden lenen binominale NC's met *soort* zich uitstekend tot een beschrijving volgens de principes van constructiegrammatica (Schermer-Vermeer 2008). Constructiegrammatica (CxG) behelst een veelkleurig palet aan benaderingen, maar alle vertrekken ze – zich afzettende tegen een strikte scheiding van lexicon en syntaxis (Hoffman & Trousdale 2013: 1) – van de *constructie* als fundamentele bouwsteen van linguïstische kennis. Constructies zijn aangeleerde, symbolische paren van vorm en functie/betekenis (Goldberg 1995, 2003, 2006).<sup>5</sup> Ze vertonen verschillende gradaties van schematiciteit (Hoffman & Trousdale 2013: 2), zitten ingebed in complexe taxonomische netwerken (Goldberg 1995; Verhagen 2005) en zijn onderhevig aan mechanismen als frequentie en analogie (Bybee 2010*b*).

---

5 Volgens Goldberg (2003: 219–220, 2006: 5) is elk talig patroon een constructie op twee voorwaarden. Ten eerste mag geen enkel aspect van de vorm of de functie/betekenis van dat patroon rechtstreeks afleidbaar zijn uit de samenstellende delen of andere, reeds bestaande constructies. Ten tweede moet – in het geval van 'voorspelbare' patronen – het patroon voldoende frequent voorkomen.



Vanuit een CxG-geïnspireerd kader beschrijft Schermer-Vermeer (2008) de vormelijke en semantische eigenschappen van binominale NC's met *soort*, die zij de *SOORT*-constructie doopt. In de volgende paragrafen worden achtereenvolgens de eigenschappen van de vormelijke (§1.2.2) en semantische pool (§1.2.3) van de constructie besproken.

### 1.2.2 Vormelijke eigenschappen

Het opvallendste vormelijke kenmerk van binominale NC's met *soort* is de mogelijkheid om telbare substantieven in het enkelvoud toe te laten in de N2-*slot* (Schermer-Vermeer 2008). Dat is niet mogelijk in kwantificerende binominale constructies: vergelijk het grammaticale voorbeeld in (19) met het ongrammaticale voorbeeld in (20).

- (19) *een soort probleem*  
 (20) \**een hoop probleem*

N2 mag echter in tegenstelling tot partitieve (23) en pseudopartitieve (24) NC's (cf. Pekelder 1993: 81) niet gedetermineerd zijn, zoals blijkt uit de ongrammaticaliteit van (21). N2 kan wel door een adjectief gemodificeerd worden (22) (Broekhuis & Den Dikken 2012: 621). In tegenstelling tot nomina in de N2-*slot* kunnen nomina in de N1-*slot* wel gedetermineerd zijn.

- (21) \**dat soort (van) de koekjes*  
 (22) *dat soort (van) lekkere koekjes*  
 (23) *een paar van de koekjes*  
 (24) *een paar van die lekkere koekjes*

Daarnaast kan de N1-*slot* slechts gerealiseerd worden met een lexicaal element uit een beperkte subset *slot-fillers*. Volgens het corpusonderzoek van Schermer-Vermeer (2008: 11) zijn die, in navolging van de prototypische leden *soort* en *type*, stuk voor stuk pas in de tweede helft van de vorige eeuw deel gaan uitmaken van de *SOORT*-constructie (zie ook Hoeksema 2011). De betreffende *slot-fillers* van de N1-*slot* zijn opgelijst in (25). *Soort* en *type* zijn als frequentste en meest prototypische nomina van de set vetjes weergegeven.

- (25) {**soort**, **type**, *genre*, *formaat*, *kaliber*, *kleur*, *kwaliteit*, *maat*, *model*, *merk*, *slag*}  
 (Schermer-Vermeer 2008: 5)<sup>6</sup>

---

6 Schermer-Vermeer (2008) vermeldt *slag* echter niet, wellicht omdat het (nog) geen enkelvoudige telbare N1 bij zich kan dragen. Van der Lubbe (1968: 137) daarentegen plaatst *slag* wel in hetzelfde rijtje als *soort* en *type*.

Aan een laatste vormelijk aspect is in de beschrijvende literatuur tot nog toe opvallend weinig aandacht besteed, met name de (schijnbaar) optionele insertie van een relator in de vorm van de prepositie *van* (cf. Joosten & Vermeire 2006; Van de Velde 2009). Vergelijk de parallele, geattesteerde voorbeelden van een constructie zonder relator in (26) met een constructie met relator in (27). Beide voorbeelden komen uit het Corpus Gesproken Nederlands (CGN; cf. Van Eerten 2007). De relatorloze constructie wordt verder ‘binominale constructie’ genoemd, de constructie met relator ‘relatorconstructie’.

- (26) *en uh best een drukke week gehad en uh ja een beetje bloedarmoede waarschijnlijk en uh **dat soort dingen** en dan te veel alcohol en dat was niet zo goed gevallen* (CGN, fn000629)
- (27) *maar van die vierenzeventig zijn er een stuk of vier niet hier uh ziek en schoolziek en **dat soort van dingen*** (CGN, fv4000605)

De mogelijke alternantie tussen (26) en (27) lijkt uniek te zijn voor het Nederlands. In andere verwante (Indo-Europese) talen is het gebruik van een relatorconstructie met soortnamen als N1 steeds verplicht; binominale constructies zijn ongrammaticaal, zoals blijkt uit (28)–(31).

- (28) Engels (Keizer 2007: 152)  
*a sort **of**/\*Ø kitchen*
- (29) Frans (Mihatsch 2016: 146)  
*une sorte **de**/\*Ø révisionnisme économique*
- (30) Spaans (Mihatsch 2007: 234)  
*una especie **de**/\*Ø camello*  
‘een soort (van) kameel’
- (31) Portugees (Mihatsch 2007: 233)  
*uma sorte **de**/\*Ø Deus*  
‘een soort (van) God’

Ondanks het feit dat de alternantie tussen relatorconstructies en binominale constructies met *soort* slechts sporadisch aan bod komt in de literatuur, komen minimale paren als in (26)–(27) wel veelvuldig voor in natuurlijk taalgebruik. Schermer-Vermeer (2008: 11) oppert dat relatorconstructies wellicht meer voorkomen in informeel en gesproken taalgebruik. Bovendien zouden ze volgens de SOD iets gewoonlijker zijn in collocaties met het onbepaald lidwoord [*een* N1] (Broekhuis & Den Dikken 2012: 635). Verder zou het kunnen dat relatorconstructies over het algemeen iets frequenter voor-

komen in zuidelijk Nederlands (cf. Schermer-Vermeer 2008: 12). Al die aannames berusten echter op introspectieve oordelen en enige vorm van verzadigende empirische evidentie ontbreekt. In §3.1.1 zullen die aannames empirisch getest worden aan de hand van synchroon corpusmateriaal uit het CGN.

### 1.2.3 Semantische eigenschappen

Generaliserend over verschillende theoretische invalshoeken, kunnen over het algemeen twee gebruiken van *soort* onderscheiden worden. Die twee gebruiken worden hier geïllustreerd aan de hand van twee geattesteerde voorbeelden, in (32)–(33).

- (32) *en daar draaiden ze dus uh ja ook af en toe rock-'n-rollplaten Elvis uh Jailhouse Rock weet je **dat soort nummers*** (CGN, fn000375)
- (33) *nee 't is inderdaad **een soort doe-het-zelfzaak** maar ook **een soort IKEA en een soort designwinkel*** (CGN, fn000282)

In de formele taalkunde wordt een onderscheid gemaakt tussen ‘lexicale’ en ‘functionele’ nomina (cf. Abney 1987). Lexicale nomina hebben een descriptieve inhoud, i.e. ze zijn in zekere mate referentieel. Van het gebruik van *soort* in (32) kan gezegd worden dat het lexicaal is. De binominale woordgroep verwijst in dat voorbeeld naar een specifiek subtype nummers dat zich op basis van bepaalde (niet-nader gedefinieerde) eigenschappen onderscheidt van een groter geheel van nummers. Die eigenschappen worden niet uitgedrukt door *soort* zelf, maar blijken uit de context, waardoor zowel de spreker als de toehoorder weet welk subtype muziek bedoeld wordt.<sup>7</sup>

In (33) daarentegen is er geen sprake van een referentieel gebruik van *soort*: de woordgroep refereert daar niet aan een concreet subtype doe-het-zelfzaken, IKEA’s of designwinkels, maar aan een winkel die in zekere zin op een doe-het-zelfzaak, IKEA en designwinkel lijkt. De taalgebruiker beschikt in zijn/haar idiolect niet over een lexeem of constructie waarmee hij/zij de bedoelde referent accuraat kan categoriseren. Hij/zij moet daarom een strategie aanwenden die zijn/haar taalgebruik ‘vager’ of ‘losser’ maakt (cf. Sperber & Wilson 1991). Concreet gebeurt dat door de bedoelde referent te categoriseren aan de hand van een andere referent die daar een bepaalde gelijkenisrelatie mee onderhoudt. Die andere referent wordt in de binominale constructie uitgedrukt in N2, maar ze wordt ‘afgezwakt’ met behulp van de *chunk* [*een soort (van)*] als een zogenaamde *hedge* (Lakoff 1973) in de N1-slot. Door zijn gebruik als *hedge* heeft *soort* een specifieke, subjectieve betekenis: het drukt de persoonlijke attitude (*stance*) van de spreker uit ten aanzien van de entiteit waarnaar N2 verwijst. Niet zelden is daarbij trouwens sprake van negatieve semantische prosodie, waarbij een impliciete minachting of afkeuring in het gebruik van de constructie ingebakken zit (Schermer-Vermeer 2008: 21). In voorbeeld

---

7 Min of meer synoniem met het lexicale *soort* zijn woorden als *zulke* en *dergelijke* (Haeseryn et al. 1997: 1148).

(33) heeft *soort* met andere woorden, in tegenstelling tot (32), geen descriptieve inhoud. In formele modellen van taalkunde spreekt men daarom van het ‘functionele’ gebruik van *soort* (Vos 1999: 55vv.).<sup>8</sup>

### 1.3 *Soort* in de formele literatuur

In deze sectie volgt een kort overzicht van wat een aantal traditionele (Van der Lubbe 1968; Haeseryn et al. 1997) en gespecialiseerde grammatica’s (Vos 1999), en daarnaast ook de recente SOD (Broekhuis & Den Dikken 2012), hebben geschreven over binominale NC’s met *soort* als N1. Het ‘afwijkende’ gedrag van de betreffende constructies ten opzichte van kwantificerende binominale NC’s (zie §1.1.2) is immers reeds langer bekend in de literatuur: verschillende grammatica’s hebben de constructies vermeld, maar vrijwel steeds lag de nadruk op een strikt synchrone beschrijving vanuit overwegend formeel georiënteerde invalshoeken.

#### 1.3.1 Van der Lubbe (1968)

Een vroeg voorbeeld van een beschrijving van *soort* in binominale NC’s komt uit Van der Lubbe (1968) structuralistische grammatica van het Nederlands. Hij categoriseert binominale NC’s met *soort* onder wat hij ‘subklasse *a*’ van de zogenaamde ‘primaire nabepalingen’ noemt. Die subklasse omvat alle nabepalingen waarbij het kernsubstantief een hoeveelheid aangeeft en de nabepaling “datgene [noemt] waarvan deze hoeveelheid is aangegeven” (Van der Lubbe 1968: 134). Subklasse *a* komt dus in grote lijnen overeen met de eerdergenoemde kwantificerende binominale NC’s van de SOD (zie §1.1.2). Van der Lubbe (1968) beschouwt *soort* en vergelijkbare substantieven als *genre* en *slag* als de syntactische kern van de NC. Voor hem fungeert N2 duidelijk als nabepaling. Daarnaast merkt hij in verband met de betekenis van de betreffende woordgroepen wel op dat *soort* zich niet helemaal als de andere primaire nabepalingen van subklasse *a* gedraagt, wanneer de nabepaling een N2 in het enkelvoud is:

Staat het bepalende substantief in het meervoud, dan duiden de genoemde woorden eveneens een zekere hoeveelheid aan; vandaar dat we deze combinaties nog tot deze subklasse rekenen; bij een enkelvoudige bepaling is het begrip ‘hoeveelheid’ heel wat moeilijker terug te vinden. (Van der Lubbe 1968: 137)

---

8 Volgens Schermer-Vermeer (2008: 20–21) is de NC in (33) geen exemplaar van de *SOORT*-constructie, maar van de verwante *een soort (van) X*-constructie. Die heeft, in tegenstelling tot de *SOORT*-constructie, maar een variabele *slot*: *X*. De *chunk* [*een soort (van)*] is niet variabel, wat impliceert dat de overige nomina in (25) nog niet kunnen voorkomen in de constructie. Ook in de *een soort (van) X*-constructie is het gebruik van een relator optioneel, maar zoals gezegd in §1.2.2 zou de variant met een relator iets frequenter kunnen zijn dan in de *SOORT*-constructie.

Desalniettemin rekent Van der Lubbe (1968) *soort* toch tot subklasse *a* en negeert hij verder de moeilijk kwantificeerbare betekenis bij enkelvoudige nabepalingen. Het bestaan van relatorconstructies blijft bij Van der Lubbe (1968) tevens geheel onbesproken.

### 1.3.2 De ANS (Haeseryn et al. 1997)

De *Algemene Nederlandse Spraakkunst* (ANS) wijdt geen aparte bespreking aan binominale NC's met *soort* en rekent ze dus impliciet tot de kwantificerende binominale NC's. In tegenstelling tot Van der Lubbe (1968) acht de ANS het onduidelijk welk substantief fungeert als syntactisch hoofd van de binominale NC, wanneer de woordgroepen onbepaald zijn (zie Haeseryn et al. 1997: 1147–1149). Op dat punt wijken binominale NC's met *soort* af van de overige kwantificerende binominale NC's:

Combinaties met *soort* onderscheiden zich van de voorgaande doordat hier zowel een enkelvoudige als een meervoudige persoonsvorm mogelijk is, ook al wordt *soort* door een bepaalde determinator [...] voorafgegaan. (Haeseryn et al. 1997: 1148)

De onduidelijkheid omtrent het syntactische hoofd van de woordgroep vertaalt zich volgens de ANS dus concreet in de tweevoudige manier waarop het onderwerp en de persoonsvorm kunnen congrueren wanneer *soort* in het enkelvoud staat en N2 in het meervoud, zie (34)–(35) (H = hoofd, B = bepaling; Haeseryn et al. 1997: 1148).

- (34)    [[*Dit soort*]<sub>H</sub> [*problemen*]<sub>B</sub>] **wordt** in taalrubrieken dikwijls behandeld.  
(35)    [[*Dit soort*]<sub>B</sub> [*problemen*]<sub>H</sub>] **worden** in taalrubrieken dikwijls behandeld.

De ANS verzuimt echter te vermelden wat de distributie van beide varianten in het Nederlands is. Op basis van corpusonderzoek zal ook daarop een antwoord worden gegeven in Sectie 4.2.

### 1.3.3 Vos (1999)

In haar formele grammatica over partitieve constructies deelt Vos (1999) *soort* en *type* in bij de *kind nouns* ('soortnamen'), onder de *Direct Partitive Constructions* (DPC's) – haar terminologie voor relatorloze binominale NC's. Soortnamen hebben geen kwantificerende betekenis, maar duiden op een kwaliteit van de referent (cf. ook Haeseryn et al. 1997). Net als Van der Lubbe (1968: 137) merkt Vos (1999: 67) op dat soortnamen de enige categorie van N1's zijn die een N2 in het enkelvoud toelaten. Vos (1999) maakt echter wel als eerste van de geraadpleegde grammatica's een onderscheid tussen verschillende gebruiken van *soort*, iets wat noch Van der Lubbe (1968), noch Haeseryn et al. (1997) doet. Zoals eerder aangehaald (§1.2.3), maakt zij een onderscheid tussen lexiciaal en functioneel gebruik van *soort*. Ondanks zeer uitgebreide beschrijvingen van de

vormelijke en semantische eigenschappen, ontbreekt ook bij Vos (1999) enige vermelding van de mogelijkheid een relator te gebruiken.

#### 1.3.4 De SOD (Broekhuis & Den Dikken 2012)

De eerste grammatica die de alternantie tussen relatorconstructies en binominale constructies met *soort* vermeldt, is de SOD. Afgezien van grammaticaliteitsoordelen op basis van een eenvoudige zoekopdracht in *Google* wordt er verder echter niet erg veel aandacht aan besteed (Broekhuis & Den Dikken 2012: 635–636). Desondanks verschaft de SOD wel gedetailleerde informatie over de vormelijke en semantische eigenschappen van de woordgroepen met *soort*. Zo onderscheidt de SOD, naast de twee gebruiken van *soort* die in §1.2.3 zijn besproken, nog een derde constructie. In dat geval bakent *soort* een (biologische) soort (*species*) af, als in (36).

(36) *deze/die soort aap/apen* (Broekhuis & Den Dikken 2012: 631)

Op die manier zijn er volgens de SOD dus niet twee, maar drie verschillende constructies met *soort*: enerzijds zijn er twee ‘lexicale’ gebruiken van *soort* (37)–(38), anderzijds is er het ‘functionele’ gebruik van *soort* (39). De betekenis is telkens tussen aanhalingstekens weergegeven (Broekhuis & Den Dikken 2012: 631–632).

- |      |                               |                       |
|------|-------------------------------|-----------------------|
| (37) | <i>deze/die soort apen</i>    | ‘species’             |
| (38) | <i>dit/dat soort auto’s</i>   | ‘subtype’             |
| (39) | <i>een soort appel/appels</i> | ‘N2-achtige entiteit’ |

Met betrekking tot het syntactische hoofd van de binominale woordgroep is de SOD iets gedetailleerder dan de ANS. Volgens de SOD is *soort* in (37) steeds het hoofd van de NC en N2 steeds de (na)bepaling. Om die reden worden zinnen als (40), waar het werkwoord congrueert met N2, niet volledig aanvaardbaar gevonden (zie Broekhuis & Den Dikken 2012: 636–637).<sup>9</sup>

(40) *\*?Deze soort vogels zijn moeilijk te observeren.* (Broekhuis & Den Dikken 2012: 636)

---

9 Toch komen zulke gevallen meermaals voor in natuurlijk taalgebruik. Op negen observaties uit het SoNaR-corpus (Oostdijk et al. 2013) zijn er zes waarbij het werkwoord congrueert met N2. Het betreft allemaal observaties uit Vlaanderen. Meer nog, voorbeeld (i) is afkomstig uit een Vlaamse krant, dus er is in principe nog eens redactionele controle over gegaan. Door het zeer kleine aantal observaties kunnen hier natuurlijk geen harde conclusies aan gekoppeld worden.

(i) *Al **die** soort ongemakken **zullen** nu de kop opsteken.* (SoNaR)

In (38) kunnen zowel N1 als N2 syntactisch hoofd zijn (cf. Haeseryn et al. 1997: 1148). In (39) ten slotte is N2 steeds het hoofd en *soort* steeds voorbepaling.

### 1.3.5 Argumentatie voor een geïntegreerd synchroon-diachroon perspectief

“[S]ynchrony and diachrony have to be viewed as an integrated whole.” (Bybee 2010*b*: 105)

Hoewel de geraadpleegde grammatica’s elk op hun manier zeker en vast waardevolle inzichten verschaffen over de syntactische structuur en semantische eigenschappen van de constructies, geven ze geen van alle een afdoende verklaring voor het bestaan en de functie van de alternanties tussen binominale constructies (26) – hier herhaald in (41) – en relatorconstructies (27) – hier herhaald in (42) – in het hedendaags Nederlands.

- (41) *en uh best een drukke week gehad en uh ja een beetje bloedarmoede waarschijnlijk en uh **dat soort dingen** en dan te veel alcohol en dat was niet zo goed gevallen* (CGN, fn000629)
- (42) *maar van die vierenzeventig zijn er een stuk of vier niet hier uh ziek en schoolziek en **dat soort van dingen*** (CGN, fv4000605)

Door hun inbedding in voornamelijk formeel georiënteerde frameworks besteden de grammatica’s vooral aandacht aan de synchrone beschrijving van de *competence*, het taalsysteem.<sup>10</sup> Aspecten van variatie en verandering – c.q. informatie over de distributie en frequentie van de concurrerende varianten in (41)–(42) – blijven daarbij consequent buiten beschouwing of worden hoogstens terzijde als ‘ongrammaticaal’ of ‘gemarkeerd’ geboekstaafd.

Taalgebruiksgebaseerde (*usage-based*) theorieën van taalkunde daarentegen, die poneren dat onze kennis van taal voortdurend gekneed wordt door gebruik en ervaring, claimen dat een opwaardering van taalvariatie en een geïntegreerd synchroon-diachroon perspectief onontbeerlijk zijn voor een volledig begrip van grammaticale kennis (Kemmer & Barlow 2000; Bybee 2006, 2010*a*, 2010*b*; Bybee & Beckner 2015).<sup>11</sup> Eerder

---

10 In hun voorrede vermelden Broekhuis en collega’s dat – hoewel ze benadrukken zich niet expliciet te willen aansluiten bij deze of gene taalkundige theorie – hun beschrijvingen vooral gestoeld zijn op inzichten uit formele taalkunde: “[The *Syntax of Dutch*] gives a comprehensive overview of the relevant research on Dutch that not only presents the findings of earlier approaches to the language, but also includes the results of the *formal* linguistic research carried out over the last four or five decades that often cannot be found in the existing reference books.” (Broekhuis & Keizer 2012: ix; mijn cursivering). Zie ook Van de Velde (2014) voor een kritiek daarop.

11 Vergelijk ook Van de Velde (2017), die, gesteund door empirische evidentie, pleit voor ‘historisatie’ van de taalkunde, naar het voorbeeld van onder andere de biologie.

dan introspectieve oordelen moet daarom worden gebruikgemaakt van empirische taaldata en de objectieve toetsing van hypothesen als methodologische basis (Tummers et al. 2005: 235–236; zie ook Hilpert & Gries 2016 voor historische taalkunde).

Een voorbeeld van een argument dat pleit in het voordeel van een gecombineerd synchroon-diachroon perspectief, is de alternantie tussen de – in formele termen – ‘lexicale’ en ‘functionele’ gebruiken van *soort*, die aangehaald zijn in §1.2.3. Een uitsluitend synchroon en formeel perspectief slaagt er niet in te verklaren waarom *soort* zowel ‘lexicaal’ als ‘functioneel’ gebruikt kan worden. Bijkomende diachrone inzichten kunnen het plaatje zeker en vast verrijken. Een verklaring voor het naast elkaar bestaan van beide gebruiken van *soort* is te vinden in grammaticalisatietheorieën (onder anderen Hopper 1991; Hopper & Traugott 2003; Heine 2003), meer bepaald bij het concept van ‘gelaagdheid’ of *layering*. Zo is er sprake van synchrone gelaagdheid tussen de oudere bronconstructie (*source construction*) in (32) – hier herhaald in (43) – en de jongere doelconstructie (*target construction*) in (33) – hier herhaald in (44).

- (43) *en daar draaiden ze dus uh ja ook af en toe rock-’n-rollplaten Elvis uh Jailhouse Rock weet je **dat soort nummers*** (CGN, fn000375)
- (44) *nee ’t is inderdaad **een soort doe-het-zelfzaak** maar ook **een soort IKEA** en **een soort designwinkel*** (CGN, fn000282)

Die synchrone gelaagdheid is met andere woorden het gevolg van grammaticalisatie waarbij *soort* in bepaalde contexten meer en meer ‘verbleekt’ (*semantic bleaching*; Heine 2003: 579) is, waardoor het zijn lexicale betekenis verloren heeft. Het verbleekte *soort* is vervolgens als *hedge* gerekruteerd (cf. §1.2.3). Voor het Nederlands is dat nog niet goed in kaart gebracht, maar vergelijkbare evidentie uit andere talen is er wel; zie onder meer de studies van Denison (2002), Davidse et al. (2008) en Brems & Davidse (2010) voor het Engels.

## 1.4 Diachrone verklaringen

### 1.4.1 Het ontstaan van binominale NC’s

Als binominale NC’s indruisen tegen de relatorverplichting (zie Sectie 1.1), dringt de vanzelfsprekende vraag zich op waarom ze in het hedendaags Nederlands überhaupt (nog) bestaan. Om die vraag accuraat te kunnen beantwoorden, moet, zoals beargumentteerd in §1.3.5, de diachronie in rekening worden gebracht.

Gewoonlijk wordt aangenomen dat binominale NC’s ontstaan zijn uit de opeenvolging van een nomen en een oorspronkelijke (partitieve) genitief (zie bijvoorbeeld Stoett 1923: 102–103). De partitieve genitief maakte initieel geen deel uit van de NC: hij kon immers gemakkelijker gescheiden voorkomen van de constituent waarop hij betrekking heeft, als in (49). Dat kan als bewijsvoering worden voorgelegd voor het feit dat



partitieve genitieven gevoelig autonomer waren dan hun hedendaagse pendanten, c.q. N2's in binominale NC's (Van der Horst 2008: 575vv.; Van de Velde 2009: 104–105). De genitiefmarkeringen – in de voorbeelden gegeven in (45)–(49) (Stoett 1923: 102) telkens de -s – laten toe het interpretatieve verband tussen beide constituenten te expliciteren.

- (45) *een lepel honich-s*
- (46) *een dropel water-s*
- (47) *XV pont gout-s*
- (48) *een sorte goed-s*
- (49) *soe seget een yegelick **wat** hij **guets** gehoert hevet* (Van der Horst 2008: 575)

Dat geëxpliciteerde verband is echter verloren gegaan met het afslijten van casusmorphologie, ten gevolge van de deflexie in de Germaanse talen, die hevig gewoed heeft in het nominale domein (Van Loey 1970: 117–118; Van Bree 1987: 236–237; Van der Horst 2008: 143–144; Van de Velde 2009: 105). Zoals blijkt uit een studie van Brugse teksten door Weerman & De Wit (1999), moet de erosie van de genitieflexie zich min of meer voltrokken hebben tussen de dertiende en de vijftiende eeuw. Daarbij werd in veel gevallen de oorspronkelijke morfologische genitief vervangen door een perifrastische relatorconstructie, meestal met het voorzetsel *van*.<sup>12</sup> Hoewel over het algemeen de vervanging door een relatorconstructie het meest toegepaste mechanisme is, zijn daarin ook diachrone fluctuaties waarneembaar, die afhangen van de semantische en morfologische aard van de specifieke N1 (cf. Hoeksema 2014: 421–425).

Binominale NC's lijken daar – zoals vermeld in Sectie 1.1 – op het eerste gezicht een 'hinderlijke' uitzondering op te vormen, omdat de 'nabepaling' niet ingeleid wordt door een relator. Dat is vooral problematisch als duidelijk wordt dat relatorloze binominale constructies diachroon de overhand lijken te nemen over constructies met een relator, zoals Van de Velde (2009: 105) suggereert.<sup>13</sup>

---

12 Daarnaast zijn er nog wel gevallen in het Nederlands waarin het in onbruik geraakte morfologische materiaal van de genitief hergebruikt wordt in nieuwe contexten (Weerman & De Wit 1999; Scott 2011, 2014; Pijpops & Van de Velde 2014; Norde & Trousdale 2016).

13 Opvallend is ook de evolutie van constructies met hechte apposities, die in het Middelnederlands steeds ingeleid werden door een relator, zoals in (i). In het hedendaags Nederlands daarentegen lijkt zo'n constructie echter uitgesloten, zie (ii) (Weerman & De Wit 1999: 1163–1164).

- (i) *der stede **van** Brucghe*
- (ii) *de stad Ø/\***van** Brugge*

De relator is in feite overbodig geworden, omdat de eigennaam *Brugge* tegenwoordig fungeert als hoofd van de NC, zodat het patroon niet langer problematisch is in het licht van de relatorverplichting (zie Van de Velde 2009: 95–96 voor een uitgebreidere argumentatie daarvoor).

## 1.4.2 Herstelstrategieën

### 1.4.2.1 Drie mogelijke oplossingen

Zoals reeds vermeld in Sectie 1.1, zijn volgens Van de Velde (2009) verschillende strategieën werkzaam om een inbreuk op de relatorverplichting (diachroon) af te wenden. De mogelijke strategieën worden hier nog eens herhaald:

- (i) toevoeging van een relator;
- (ii) syntactische omkering;
- (iii) heranalyse.

Welke strategie toegepast wordt, blijkt bijvoorbeeld af te hangen van semantische factoren als referentialiteit (Joosten 2003; Joosten & Vermeire 2006). Daarnaast blijkt de gekozen strategie ook vanuit een ruimer perspectief taalafhankelijk. Als twee nauwe verwanten van het Nederlands, met name het Engels en het Duits, van naderbij bekeken worden, blijkt het Engels vooral te opteren voor (een combinatie van) strategieën (i) en (iii), terwijl het Duits meer gebruikmaakt van strategie (ii) (Van de Velde 2009: 103). Het Nederlands bevindt zich, zoals op verschillende gebieden van de morfosyntaxis (Van Haeringen 1956; Hüning et al. 2006), in een ‘tussenpositie’, doordat het zich van alle drie de strategieën kan bedienen.

In tegenstelling tot strategieën (i)–(ii) is strategie (iii) *an sich* echter het moeilijkst waar te nemen ‘aan de oppervlakte’. Daarom wordt in de volgende paragraaf (§1.4.2.2) dieper ingegaan op een aantal mogelijke ‘oppervlaktesymptomen’ van heranalyse, die interessant zijn in het licht van binominale NC’s met *soort* (cf. infra).

### 1.4.2.2 Symptomen van heranalyse<sup>14</sup>

Van de drie herstelstrategieën is heranalyse als zodanig dus het moeilijkst waarneembaar; uit een binominale NC ‘in isolatie’ is het immers zo goed als onmogelijk op te maken welk nomen als hoofd en welk als bepaling fungeert (cf. ook Secties 1.1 en 1.3). Als daarentegen de onmiddellijke context/co-tekst mee onder de loep wordt genomen, kunnen een aantal ‘symptomen’ van die heranalyse afgeleid worden. Concreet zijn die symptomen het gevolg van het feit dat N1 geleidelijk aan zijn nominale kenmerken verliest (zie Van de Velde 2009: 99).

**(a) Verandering in congruentie:** Een eerste symptoom, dat reeds in §1.3.2 en §1.3.4 is aangehaald, is het feit dat N2 – en niet langer N1 – congrueert met het werkwoord. Dat is exact wat aan de hand is met het collectivum *paar* (Van der Horst & Van der

---

<sup>14</sup> De symptomen die hier worden beschreven, zijn in feite het *gevolg* van heranalyse.

Horst 1999: 205). Wanneer het gebruikt wordt in de zin van ‘een (bij elkaar horend) tweetal’, is *paar* nog steeds het hoofd van de binominale woordgroep, als in (50).

(50) *Een paar goeie schoenen is duur.* (Van der Horst & Van der Horst 1999: 205)

Daarnaast is er echter het ‘nieuwere’ gebruik van *paar* in de zin van ‘een (klein) aantal’. In dat geval is N2 het hoofd en selecteert N2 dus het werkwoord, zoals in (51).

(51) *Een paar kinderen speelden nog op straat.* (Van der Horst & Van der Horst 1999: 205)

N1 wordt door heranalyse een (kwantificerende) voorbepaling bij N2. N2, de oorspronkelijke bepaling, wordt op zijn beurt als hoofd van de woordgroep geherinterpreteerd.

**(b) Verlies van de determinator:** Het verlies van nominale kenmerken van N1 gaat gepaard met zogenaamde *decategorialisatie* van N1. Paul Hopper (1991) omschrijft decategorialisatie als volgt:

Forms undergoing grammaticization tend to lose or neutralize the morphological markers and syntactic privileges characteristic of the full categories Noun and Verb, and to assume attributes characteristic of secondary categories such as Adjective, Participle, Preposition, etc. (Hopper 1991: 22)

Een rechtstreeks gevolg van decategorialisatie van N1 is het feit dat N1 diachroon steeds vaker zonder determinator kan verschijnen, of dat de determinator geselecteerd wordt door N2 (Van de Velde 2009: 99).<sup>15</sup> Een typisch voorbeeld van zo’n scenario is de ontwikkeling van de kwantificeerder *tal van*: waar de verbinding tot zelfs in de negentiende eeuw (53) nog gewoonlijk een lidwoord bij zich droeg, is dat in het hedendaags Nederlands helemaal uitgesloten (Van der Horst 2003). Vergelijk de voorbeelden in (52)–(54). Merk op dat in het Middelnederlandse voorbeeld (52) *tal* genitieflexie op zijn complement regeert.

(52) Middelnederlands (Stoett 1923: 102)  
*een tal groens heeringhs*

(53) Laatnieuwnederlands (Van der Horst 2003: 59)  
*een tal van vrouwen*

---

15 Vergelijk Mihatsch (2016: 137; mijn cursivering) over de Romaanse talen: “As the first noun loses its head status the determiner use gets fossilized and adjectival modification is then impossible. The second nominal constituent also gets reduced in the more or less grammaticalized constructions. In these cases the second nominal is a bare noun, which can be modified by adjectives, but cannot take any determiners, since the determiner introducing the binominal or the first nominal constituent then functions as a determiner of the second noun.”

- (54) Hedendaags Nederlands (*Het Nieuwsblad*, 2006; SoNaR)  
 [Ø t]al **van** vrouwen voelen zich ongemakkelijk als ze alleen rijden.

Opmerkelijk is het feit dat de complexe voorbepaling *tal van* het resultaat is van twee gecombineerde herstelstrategieën: enerzijds het toevoegen van een relator (*van*), anderzijds heranalyse met decategorialisatie en het verdwijnen van de determinator als gevolg. Let wel dat in dit geval de volledige collocatie *tal van* geherinterpreteerd is, waardoor de relator zelf in de syntactische structuur onlosmakelijk deel uitmaakt van de bepaling (Van de Velde 2009: 99).

Gelijkaardig is de ontwikkeling van woorden als *paar*, *heleboel* en *aantal* in binominale NC's (Van der Horst & Van der Horst 1999: 205–210; Van der Horst 2003: 58). In hun *Geschiedenis van het Nederlands in de twintigste eeuw* lijsten Van der Horst & Van der Horst (1999: 205–206) verschillende geattesteerde voorbeelden op waarin niet langer het onzijdige *paar*, maar N2 het lidwoord lijkt te selecteren. Omdat N2 in die constructies meervoudig moet zijn (cf. §1.2.2), is dat lidwoord *in casu* altijd *de*. Die ommekeer is al een tijdje aan de gang, zoals geïllustreerd wordt door (55), dat stamt uit 1934. In dat voorbeeld wordt het lidwoord *de* duidelijk niet door *paar*, maar door *rukwinden* geselecteerd.

- (55) **De eerste paar rukwinden** wierpen korte, forse golven op. (Van der Horst & Van der Horst 1999: 205)

Daarnaast zijn er ook al gevallen te vinden waarin groepen met *paar* (en analoog ook *heleboel* en *aantal*) reeds zonder determinator kunnen verschijnen, zelfs in geredigeerde registers – getuige voorbeeld (56) uit *NRC Handelsblad* van 1995.

- (56) Binnen [Ø] **paar jaar** ben je een hele rij insektensoorten kwijt. (Van der Horst & Van der Horst 1999: 206)

Ook de hogere telwoorden *honderd*, *duizend*, *miljoen* en *miljard*, die oorspronkelijk volwaardige substantieven waren – zie voorbeeld (57), *met* relator –, zijn geheranalyseerd als 'telwoordachtige' voorbepalingen bij N2 (Van de Velde & Weerman 2014; Van der Horst & Van de Velde 2016; Van de Velde 2017: 52–56).

- (57) *Partie ende nidicheit baet der hellen menich **millioen van** zielen* (Van der Horst & Van de Velde 2016: 412)

Dat de betreffende woorden geheranalyseerd zijn, mag blijken uit het feit dat het complement bij het telwoord steeds congrueert met het werkwoord (vergelijk symptoom a). Het omgekeerde scenario is ongrammaticaal (58).

- (58) *honderd vrouwen kwamen/\*kwam naar het feestje* (Van de Velde & Weerman 2014: 124)

Toch vertonen ze tegelijkertijd nog een zekere vormelijke affiniteit met substantieven, doordat ze, in tegenstelling tot de lagere telwoorden als *twee* etc., met het meervouds-suffix *-en* kunnen verschijnen, zie (59).

- (59) *Nederland rekt op de titel, de honderden supporters in de zaal, de miljoenen kijkers in hun huiskamer.* (Van de Velde & Weerman 2014: 124)

Evidentie voor het feit dat de oorspronkelijke substantieven gedecategorialiseerd zijn – c.q. dat ze dikwijls niet meer als substantieven aanvoeld worden –, blijkt uit de observatie dat ze in geschreven taalgebruik opvallend frequent gerealiseerd worden als *honderde*, *duizende* en *miljoene* (d.i. zonder eind-*n*). Dat verschijnsel wijst op heranalyse als attributieve bepaling bij N2 (Van de Velde & Weerman 2014: 125–127; Van der Horst & Van de Velde 2016: 415–416). Enkele voorbeelden van zulke geherinterpreteerde telwoorden zijn gegeven in (60)–(61).

- (60) *Ik wens je duizende jaren geluk.* (Van de Velde & Weerman 2014: 126)  
 (61) *Onderweg zagen we miljoene vogels!* (Van der Horst & Van de Velde 2016: 416)

### 1.4.3 Diachrone ontwikkelingen van *soort*

Het substantief *soort* is ontleend aan het Franse *sorte* ‘soort, klasse’, dat zelf teruggaat op het Latijnse *sors* (genitief *sort-is*) ‘lot, deel, aandeel, rang, manier’ (Philippa et al. 2009: 219; WNT, s.v. *soort*). Het *Woordenboek der Nederlandsche Taal* (WNT) en het *Etymologisch Woordenboek van het Nederlands* (EWN) vermelden dat *soort* oorspronkelijk zowel een kwaliteit aanduidde als een collectief van personen of zaken met die kwaliteit, zoals in (62)–(63) (zie ook Schermer-Vermeer 2008: 8). In (62) betekent *es comen van onser soorten* dus ongeveer ‘behoort tot onze kring’ (WNT, s.v. *soort*).

- (62) *Die bode die ic hebbe ghesent om de slotlen van der poorte die es comen van onser soorten.* (WNT, ca. 1340–1360)

De collectieve betekenis blijkt duidelijk uit (63), waar *sorten* naar een (gewapende) bende verwijst.

- (63) *Hi (Everaert Tserclaes) trac af des graven Lodewijcs baniere van der stat huus, met siere soorten.* (WNT, 1432)

Wanneer *soort* betrekking had op een andere constituent, stond die laatste in de genitief of werd die ingeleid door het voorzetsel *van* (cf. §1.4.1), zie (64)–(65).

- (64) *Dat somyge copen **ene geheele sorte gudes** al up, darvan dattet danne ander copplude wedder moten copen na erem willen.* (WNT, ca. 1430)
- (65) *Een Riem als een Spaey, d'welck wort gemaect van **een sekere soorte van hart hout*** (WNT, 1602)

Die observaties lijken te suggereren dat een eerste herstelstrategie voor binominale NC's met *soort* inderdaad het toevoegen van een relator is, wat Van de Velde (2009: 100) ook aanhaalt. Die hypothese wordt dan ook verder synchroon en diachroon uitgedragen in Hoofdstuk 3.

Vervolgens wordt nagegaan of en in welke mate de symptomen van heranalyse, die geschetst zijn in §1.4.2.2, overdraagbaar zijn op binominale NC's met *soort*.<sup>16</sup> Het eerste symptoom is (a) veranderende congruentie: niet langer N1 bepaalt het getal van het werkwoord, maar N2. De beschrijvende literatuur lijkt te suggereren dat die veranderende congruentie ook van toepassing is op binominale NC's met *soort* (zie §1.3.2 en §1.3.4). Een tweede mogelijk symptoom van heranalyse is een gevolg van de decategorialisatie van N1, namelijk (b) het verlies van de determinator. De vraag die daarbij gesteld wordt, is of *soort* in een binominale NC al voldoende als (identificerende) voorbepaling bij N2 wordt herkend om de (overbodig geworden) determinator achterwege te laten. Een derde mogelijk symptoom van heranalyse is tot dusver nog niet expliciet ter sprake gekomen, maar manifesteert zich wel al duidelijk in verwante talen als het Engels. Het patroon *een soort (van)* zou mogelijks ook buiten de NC kunnen treden (Van de Velde 2009: 100). Concreet komt het neer op het volgende: *een soort (van)* fungeert niet (langer) als voorbepaling bij N2, maar als bijwoordelijke bepaling op zinsniveau. Zo modificeert *een soort van* in (66) het werkwoord *werken*, waardoor de volle betekenis van dat werkwoord afgezwakt wordt (cf. §1.2.3).<sup>17</sup>

---

16 Een derde herstelstrategie – syntactische omkering (c.q. compositie) – kan ook van toepassing zijn op binominale constituenten met *soort*, waarbij (ii) concurreert met (i).

- (i) *een soort aap/apen*  
(ii) *een apensoort*

Die strategie is echter enkel acceptabel als *soort* in de 'biologische' zin van 'species' gebruikt is (Broekhuis & Den Dikken 2012). Niet-biologische lezingen van *soort*, als in *een soort systeem*, zijn immers niet syntactisch omkeerbaar: \**een systeemsoort* (cf. ook Joosten & Vermeire 2006). Die strategie wordt in deze meesterproef dan ook verder buiten beschouwing gelaten.

17 In zekere zin vergelijkbaar is het intensifiërende gebruik van *massa*('s) in substandaardvariëteiten van het Nederlands, voornamelijk in West-Vlaanderen en westelijk Oost-Vlaanderen (De Clerck & Coleman 2013), zie (i).

- (i) *Nog volk da **massa's** gaat buizen?* (De Clerck & Coleman 2013: 148)

- (66) *Morgenochtend ga zelfs ik (ambtenaar) **een soort van** werken!* (Van de Velde 2009: 100)

Het Engels lijkt daarin dus al (veel) verder te staan dan het Nederlands (zie onder anderen Kay 1984; Quirk et al. 1992: 246vv.; Aijmer 2002). Enkele voorbeelden van gegrammaticaliseerd gebruik als bijwoordelijke bepaling of *discourse particle* (Aijmer 2002) zijn (67)–(70). Een typisch gevolg van de grammaticalisatie van het patroon is ook het veelvuldig verschijnen van gecondenseerde vormen, als in (68) (cf. Lehmann 1985).

- (67) *He is **kind of** clever.* (Quirk et al. 1992: 246)  
(68) *John is **sorta** schizophrenic.* (Kay 1984: 161)  
(69) *As he was walking along, he **sort of** stopped and fell flat on his face.* (Quirk et al. 1992: 387)  
(70) *This is **sort of** between the two of us.* (Aijmer 2002: 185)

## 2 Vraagstelling, hypothesen en methodologie

---

### 2.1 Vraagstelling en hypothesen

In Hoofdstuk 1 is geprobeerd op een theoretische manier de onderzoeksvraag in te bedenken in reeds bestaande inzichten en een argumentatie op te bouwen die pleit voor een datagebaseerde aanpak (§1.3.5). De centrale onderzoeksvraag van deze meesterproef luidt:

- i. **Op welke manier conformeren de binominale NC's met *soort* als N1 zich aan de relatorverplichting, zoals beschreven in Van de Velde (2009)?**

Concreet wordt op een empirische, corpusgebaseerde manier onderzocht of binominale NC's met *soort* inderdaad gebruikmaken van de volgende herstelstrategieën (Van de Velde 2009):

- (i) toevoeging van een relator (HYPOTHESE I);
- (iii) heranalyse (HYPOTHESE II).

Vooraf §1.4.2.2 had als bedoeling Hypothese II operationaliseerbaar te maken voor corpusgebaseerd onderzoek. Daarbij is doelbewust gekozen om drie 'symptomen' van heranalyse onder de loep te nemen:

- 1. Verandering in congruentie, waarbij in meer en meer gevallen N2 het getal van de persoonsvorm bepaalt;
- 2. Verlies van de determinator door decategorialisatie van N1. Er wordt verwacht dat vooral in spontaan, niet-geredigeerd taalgebruik de determinator zou kunnen wegvallen – in contexten met redactionele controle (zoals krantenmateriaal) komen zulke 'fouten' wellicht veel minder/niet voor;
- 3. Uittreding uit de NC en optreden van de *hedge een soort (van)* als bijwoordelijke bepaling op zinsniveau, al dan niet ontleend aan het Engels.

In de volgende sectie (Sectie 2.2) wordt de methodologie toegelicht, aan de hand waarvan beide hypothesen onderzocht worden. In eerste instantie worden de geraadpleegde corpora kort geïntroduceerd (§2.2.1), waarna de werkwijze van de analyse wordt toegelicht (§1.1.1).



## 2.2 Methodologie

### 2.2.1 Corpora

De observaties voor de analyses werden gehaald uit verschillende corpora, zowel met historisch als hedendaags Nederlands materiaal. Voor de eerste hypothese werd gewerkt met materiaal uit enerzijds het synchrone Corpus Gesproken Nederlands (CGN; Van Eerten 2007), anderzijds werden ook observaties gehaald uit het diachrone corpus van Daniël Van Olmen en het corpus *De Gids*.

Voor de tweede hypothese werd vanuit methodologische overwegingen *voornamelijk* gewerkt met synchrone corpora: SoNaR (Oostdijk et al. 2013), Condiv (Gronde-laers et al. 2000) en Moroccorp (Ruetten & Van de Velde 2013). Opzoekingen in diachrone corpora leverden immers geen of te weinig treffers op om er degelijke statistische analyses op te kunnen uitvoeren (een terugkerend probleem in onderzoek naar taalverandering, cf. Tummers et al. 2005: 236). In de volgende paragrafen (§§2.2.1.1–2.2.1.5) wordt elk corpus kort geïntroduceerd.

#### 2.2.1.1 CGN

Het synchrone corpus dat materiaal leverde voor de eerste hypothese is het Corpus Gesproken Nederlands of CGN, dat ongeveer 9 miljoen woorden gesproken Nederlands uit zowel Nederland als Vlaanderen bevat (Van Eerten 2007). Het corpus is gelemmatiseerd en getagd voor *part-of-speech*informatie (POS) (Van Eynde 2004).

#### 2.2.1.2 Corpus Van Olmen

Het diachrone corpus van Daniël Van Olmen (ook vermeld in Van Olmen 2013: 267; Van Olmen & Van der Auwera 2014: 226) bevat een verzameling toneel- en fictieteksten, tot stand gekomen tussen 1500 en 2000.

#### 2.2.1.3 De Gids

Voor de eerste hypothese werd ook een corpus geraadpleegd dat honderd jaargangen van het algemeen cultureel en literair tijdschrift *De Gids* bevat, goed voor in totaal ongeveer 70 miljoen woorden (Van de Velde 2009: 5). De honderd jaargangen lopen van 1837 tot 1936.

#### 2.2.1.4 SoNaR

Het SoNaR-corpus bevat ongeveer 500 miljoen woorden geschreven hedendaags Nederlands uit zowel Nederland als Vlaanderen (Oostdijk et al. 2013). Het corpus is uitgebreid gestratificeerd voor register, met teksten uit kranten, boeken, discussielijsten, e-mails, sms'en, chatkanalen, tweets, etc. Daarnaast is het gelemmatiseerd en getagd voor POS-informatie.

#### 2.2.1.5 Condiv en Moroccorp

Het Condiv-corpus is net zoals het SoNaR-corpus zowel voor Nederland als Vlaanderen gestratificeerd voor register en bevat in totaal ongeveer 47 miljoen woorden – ongeveer 20 miljoen woorden uit Nederland en ongeveer 27 miljoen woorden uit Vlaanderen (Grondelaers et al. 2000). Het bevat tevens een diachrone component, maar daarvan is voor dit onderzoek geen gebruik gemaakt.

Het Nederlandse subcorpus met chatmateriaal (*internet relay chat* of IRC) kent daarenboven een etnische, Marokkaans gekleurde pendant in het recent gecompileerde Moroccorp (Ruetten & Van de Velde 2013). Op die manier kan het Marokkaanse Nederlands op een systematische manier vergeleken worden met het Nederlandse Nederlands, zoals dat gedaan is in onder anderen Van de Velde & Weerman (2014) en Pijpops & Van de Velde (2015).

Het totale aantal observaties dat uit elk corpus is geëxtraheerd, is weergegeven in Tabel 1.

*Tabel 1: Corpusaantallen*

| <b>Corpus</b> | <b>Aantal observaties</b> |
|---------------|---------------------------|
| CGN           | 2876                      |
| Van Olmen     | 683                       |
| De Gids       | 4599                      |
| SoNaR         | 1110                      |
| Condiv        | 3475                      |
| Moroccorp     | 1159                      |
| Totaal        | 13 902                    |

### 2.2.2 Analyse

De relevante data werden met behulp van reguliere expressies in AntConc (Anthony 2014) onttrokken aan de verschillende corpora; het SoNaR-corpus werd online geraadpleegd via <http://www.openonar.clarin.inl.nl/>. Valse treffers werden verwijderd en bijkomende variabelen werden toegevoegd in Excel. In het hoofdstuk met de resultaten wordt voor elke deelanalyse duidelijk toegelicht welk corpus gebruikt is en op welke manier de data geanalyseerd zijn.

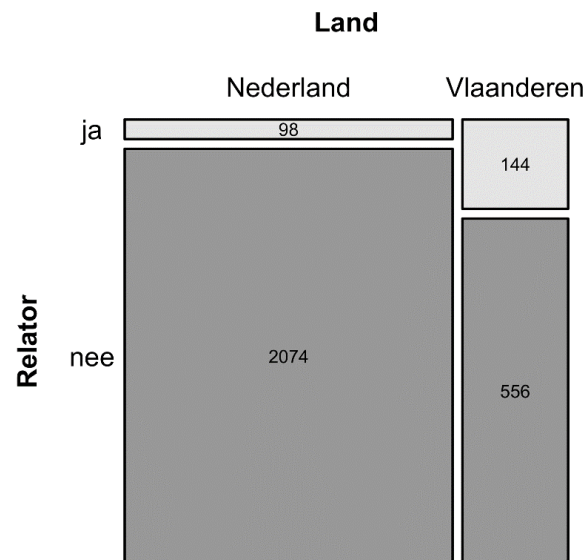
Voor de statistische analyses en bijbehorende visualisaties werd gebruikgemaakt van de *open-source* statistische software *R* (R Core Team 2016) en de pakketten *car* (Fox & Weisberg 2011), *dplyr* (Wickham & Francois 2016), *effects* (Fox 2003), *ggplot2* (Wickham 2009), *stringr* (Wickham 2016) en *vcd* (Meyer et al. 2006, 2016).

## 3 Resultaten

### 3.1 Hypothese I (toevoeging van een relator)

#### 3.1.1 Synchroon onderzoek

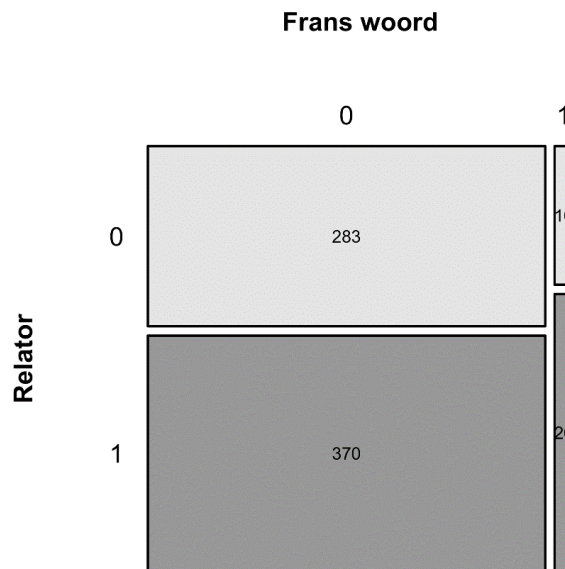
In eerste instantie werd een verkennende analyse uitgevoerd om na te gaan wat de contemporaine distributie van binominale constructies en relatorconstructies met *soort* is. Zoals aangetoond in Sectie 1.3 is daarover immers nog onvoldoende kennis voorhanden. De analyse werd uitgevoerd op 2872 observaties uit het CGN. Alle deelcorpora uit zowel Nederland als Vlaanderen werden mee opgenomen in de analyse. De distributie van beide varianten per land is gevisualiseerd met behulp van een mozaïekplot (Figuur 1). De oppervlakte van de ‘tegels’ van de mozaïekplot is een indicatie voor de hoeveelheid observaties in die tegel. In elke tegel is ook telkens het absolute aantal observaties gegeven. Zoals duidelijk valt af te leiden uit Figuur 1, zijn relatorconstructies over het algemeen (veel) minder frequent dan binominale constructies: de respectieve proporties zijn 8,43% en 91,57% (voor Nederland en Vlaanderen samen). De relatorconstructies komen over het algemeen dan weer wel significant meer voor in Vlaanderen dan in Nederland ( $p < 0,001$ ), de effectgrootte is matig (Cramér's  $V = 0,25$ ).



$$\chi^2(1) = 176.94; p = 0; \text{Cramér's } V = 0.25$$

Figuur 1: Distributie van binominale constructies en relatorconstructies in Nederland en Vlaanderen (CGN)

Een mogelijke hypothese voor het feit dat relatorconstructies meer voorkomen in Vlaanderen zou de invloed van het Frans kunnen zijn (Van de Velde, p.c.), waar relatoren verplicht zijn in binominale NC's met soortnamen (§1.2.2). Daarom werden de 683 observaties uit het corpus Van Olmen (om een eventueel tijdseffect niet uit te sluiten) door twee beoordelaars onafhankelijk van elkaar gecategoriseerd als Frans of niet-Frans, waarbij een hoge mate van betrouwbaarheid tussen de beoordelaars (*inter-rater reliability*) werd bekomen (Cohens  $\kappa = 0,89$ ;  $p < 0,001$ ).<sup>18</sup> Ook hier is de distributie van Franse of niet-Franse N2 per constructie gevisualiseerd met een mozaïekplot (Figuur 2). Er is echter geen statistisch significante associatie tussen beide factoren ( $p = 0,279$ ).

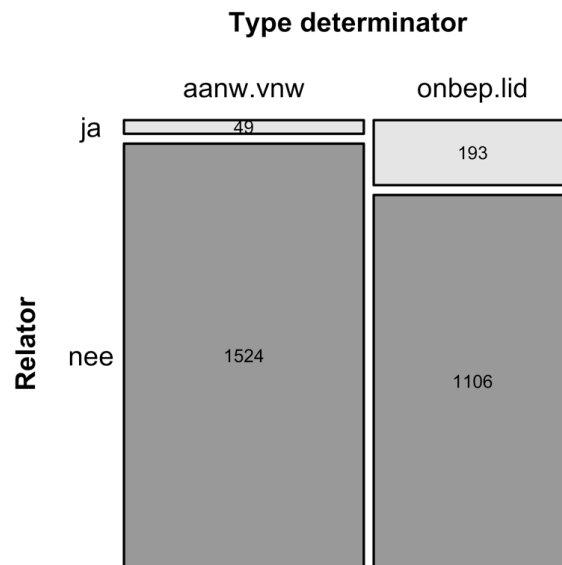


$$\chi^2(1) = 1.17; p = 0.279$$

Figuur 2: Distributie van binominale constructies (0) en relatorconstructies (1) volgens constructies zonder (0) en met (1) Franse N2

Daarnaast zou ook het type determinator een rol spelen: constructies met *een soort* ('onbep.lid') zouden iets frequenter met een relator gerealiseerd worden dan constructies met andere determinatoren – in dit onderzoek aanwijzende voornaamwoorden ('aanw.vnw') als *deze, dit, die, dat, zo'n*, enzovoort (Broekhuis & Den Dikken 2012: 635; zie §1.2.2). De 2872 observaties uit het CGN ( $n_{\text{aanw.vnw}} = 1573$ ,  $n_{\text{onbep.lid}} = 1299$ ; respectievelijk 54,77% en 45,23%) bevestigen die aanname. De associatie tussen het gebruik van een relator en het type determinator is statistisch significant ( $p < 0,001$ ) met een matige effectgrootte (Cramér's  $V = 0,21$ ). De gegevens zijn gevisualiseerd met een mozaïekplot in Figuur 3.

18 De ratings van de eerste beoordelaar waren geblindeerd voor de tweede.



$$\chi^2(1) = 127.14; p = 0; \text{Cramér's } V = 0.21$$

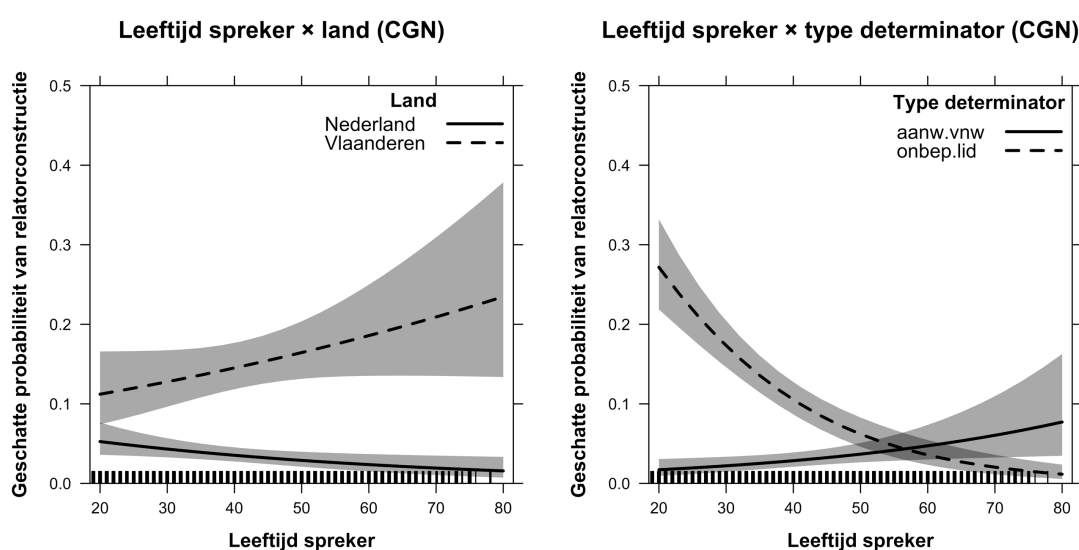
Figuur 3: Distributie van binominale constructies en relatorconstructies volgens determinator (CGN)

Een logistische-regressieanalyse (Gries 2009; Baayen 2011: 195–214; Spielman 2014) maakt duidelijk dat de leeftijd van de spreker interageert met zowel het land als met het type determinator.<sup>19</sup> In eerste instantie is het zo dat niet alleen de Vlamingen in het CGN meer relatorconstructies gebruiken dan de Nederlanders in het CGN, maar ook dat binnen de Vlamingen oudere Vlamingen meer relatorconstructies gebruiken dan jongere Vlamingen (zie het linkerpaneel van Figuur 4). Opvallend genoeg blijkt dat voor de Nederlanders niet te gelden. Voor jongere Nederlanders is de geschatte probabilliteit van het gebruik van een relatorconstructie iets hoger dan voor oudere Nederlanders. Verder suggereert het model dat de leeftijd van de spreker in interactie met het type determinator een effect heeft op het gebruik van een relator: jongere sprekers gebruiken significant meer relatorconstructies dan oudere sprekers wanneer *soort* een onbepaald lidwoord bij zich draagt (zie het rechterpaneel van Figuur 4). Voor aanwijzende voornaamwoorden daarentegen blijkt het gebruik van een relatorconstructie af te nemen bij jongere sprekers in vergelijking met oudere sprekers. De output van de logistische regressie is gegeven in Tabel 2. De interacties zijn gevisualiseerd in Figuur 4, telkens op een probabilliteitsschaal van 0 tot 0,5. De *model diagnostics* ( $C = 0,82$ ; Somers'  $D_{xy} = 0,64$ ) geven aan dat het model voorspellende kracht heeft (Baayen 2011: 204; Spielman 2014).

<sup>19</sup> Voor elke uiting werd de leeftijd van de spreker automatisch toegevoegd.

Tabel 2: Output logistische regressie, met binominale constructies als referentieniveau (0)

| Coëfficiënt                                                          | Logit | Betrouwbaarheidsintervallen |       | Significantie |
|----------------------------------------------------------------------|-------|-----------------------------|-------|---------------|
|                                                                      |       | 2,5%                        | 97,5% |               |
| intercept                                                            | -4,60 | -5,68                       | -3,61 | $p < 0,001$   |
| land = Vlaanderen                                                    | 0,11  | -0,76                       | 0,99  | $p = 0,803$   |
| type determinator =<br><i>onbep.lid</i>                              | 4,74  | 3,66                        | 5,89  | $p < 0,001$   |
| leeftijd spreker                                                     | 0,02  | -0,01                       | 0,04  | $p = 0,155$   |
| land = Vlaanderen $\times$<br>leeftijd spreker                       | 0,04  | 0,01                        | 0,06  | $p = 0,002$   |
| type determinator =<br><i>onbep.lid</i> $\times$ leeftijd<br>spreker | -0,08 | -0,10                       | -0,05 | $p < 0,001$   |



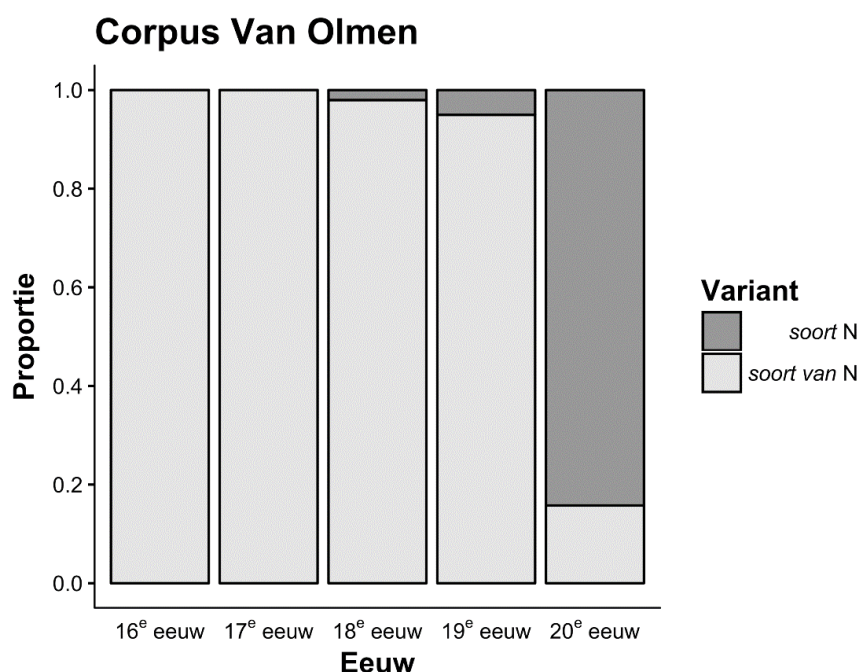
Figuur 4: Effectplots van de logistische regressie

### 3.1.2 Diachroon onderzoek

In een volgende fase van het onderzoek werd nagegaan of de over het algemeen lage proportie van relatorconstructies met *soort* een bepaalde fase in een lopende taalverandering weerspiegelt.<sup>20</sup> Daarvoor werd het corpus Van Olmen geraadpleegd, met als bedoeling te onderzoeken of een algemene, ruwe trend over de verschillende eeuwen heen te ontdekken valt. De relatieve proporties van binominale constructies (*soort* N; de donkergrijze balkjes) en relatorconstructies (*soort van* N; de lichtgrijze balkjes) per eeuw zijn gegeven in Figuur 5. De observaties situeren zich grofweg tussen 1500 en 2000. De absolute aantallen (n) en relatieve proporties (%) per eeuw zijn gegeven in Tabel 3.

<sup>20</sup> Cf. Van de Velde (2017: 47): “Many cases of synchronic variation are the result of competition between an innovative mutant encroaching on an obsolescent construction. In such cases, the synchronic skew in the proportion of one variant to the other is not arbitrary, but is a reflection of how far the change has progressed.”

De associatie tussen beide factoren is statistisch significant ( $p < 0,001$ ); er bestaat een duidelijke negatieve correlatie tussen het aandeel relatorconstructies en de tijd uitgedrukt in eeuwen (Kendalls  $\tau = -0,75$ ).



Figuur 5: Proporties van binominale constructies en relatorconstructies per eeuw tussen 1500 en 2000 (corpus Van Olmen)

Tabel 3: Proporties van binominale constructies en relatorconstructies per eeuw tussen 1500 en 2000 (corpus Van Olmen)

| Decen-<br>nium       | Variant            |         |                        |        | Totaal |
|----------------------|--------------------|---------|------------------------|--------|--------|
|                      | relatorconstructie |         | binominale constructie |        |        |
|                      | n                  | %       | n                      | %      |        |
| 16 <sup>e</sup> eeuw | 4                  | 100,00% | 0                      | 0,00%  | 4      |
| 17 <sup>e</sup> eeuw | 22                 | 100,00% | 0                      | 0,00%  | 22     |
| 18 <sup>e</sup> eeuw | 48                 | 97,96%  | 1                      | 2,04%  | 49     |
| 19 <sup>e</sup> eeuw | 264                | 94,96%  | 14                     | 5,04%  | 278    |
| 20 <sup>e</sup> eeuw | 52                 | 15,76%  | 278                    | 84,24% | 330    |
| Totaal               | 390                |         | 293                    |        | 683    |

Het oudste voorkomen van een woordgroep zonder relator in de dataset dateert uit de achttiende eeuw. Het betreft een vers uit *Het Surinaamsche leeven, tooneelschwyse verbeeld*, een toneelstuk van de achttiende-eeuwse auteur Don Experientia.<sup>21</sup> Het luidt als volgt (71):

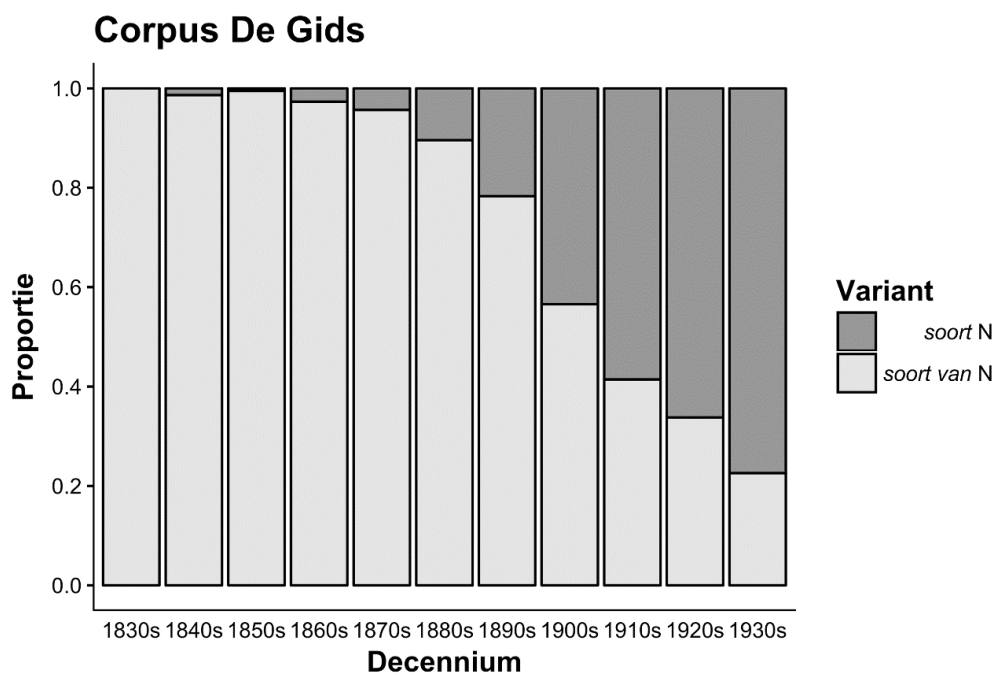
<sup>21</sup> De tekst van het stuk is voor de geïnteresseerde lezer integraal te raadplegen via [http://www.dbnl.org/tekst/expe001suri01\\_01/](http://www.dbnl.org/tekst/expe001suri01_01/).



(71) *Wat siet men menig kaalen Haan,  
 En **ander soort** Lantloopers;  
 Voor Adels en Barons hier staan,  
 Daar hun de honger dreef te gaan,  
 Tot hier, door Sielverkopers.* (Van Olmen, 18<sup>e</sup> eeuw)

Opvallend is de scherpe breuklijn tussen de negentiende en twintigste eeuw. Waar in de negentiende eeuw nog bijna 95% van de woordgroepen met *soort* constructies met een relator *van* waren, is dat in de twintigste eeuw nog slechts een goede 15% (Tabel 3). De gestage opkomst van relatorloze (binominale) groepen vanaf de achttiende eeuw lijkt in een stroomversnelling te zijn beland in de twintigste eeuw, met een drastische afname van relatorconstructies als gevolg.

Om die drastische afname van relatorconstructies duidelijker in kaart te brengen, werden de 4599 observaties uit het corpus De Gids geanalyseerd. De relatieve proporties van binominale constructies (*soort* N; de donkergrijze balkjes) en relatorconstructies (*soort van* N; de lichtgrijze balkjes) zijn gevisualiseerd in Figuur 6, de absolute (n) en relatieve (%) aantallen zijn gegeven in Tabel 4.



*Figuur 6: Proporties van binominale constructies en relatorconstructies per decennium tussen 1837 en 1936 (corpus De Gids)*

Tabel 4: Proporties van binominale constructies en relatorconstructies per decennium tussen 1837 en 1936 (corpus De Gids)

| Decen-<br>nium | Variant            |         |                        |        | Totaal |
|----------------|--------------------|---------|------------------------|--------|--------|
|                | relatorconstructie |         | binominale constructie |        |        |
|                | n                  | %       | n                      | %      |        |
| 1830s          | 47                 | 100,00% | 0                      | 0,00%  | 47     |
| 1840s          | 146                | 98,65%  | 2                      | 1,35%  | 148    |
| 1850s          | 190                | 99,48%  | 1                      | 0,52%  | 191    |
| 1860s          | 472                | 97,32%  | 13                     | 2,68%  | 485    |
| 1870s          | 595                | 95,66%  | 27                     | 4,34%  | 622    |
| 1880s          | 499                | 89,59%  | 58                     | 10,41% | 557    |
| 1890s          | 527                | 78,31%  | 146                    | 21,69% | 673    |
| 1900s          | 353                | 56,57%  | 271                    | 43,43% | 624    |
| 1910s          | 189                | 41,81%  | 263                    | 58,19% | 452    |
| 1920s          | 152                | 33,78%  | 298                    | 66,22% | 450    |
| 1930s          | 79                 | 22,57%  | 271                    | 77,43% | 350    |
| Totaal         | 3249               |         | 1350                   |        | 4599   |

De neerwaartse trend, die duidelijk doorschemert in Figuur 6, werd vervolgens gemodelleerd met een logistische regressie. Op die manier wordt (een deel van) de (logistische) S-curve volgens dewelke taalverandering typisch verloopt (cf. onder anderen Kroch 1989; Croft 2000: 183–190; Denison 2003; Blythe & Croft 2012), zichtbaar. Na de introductie van een nieuwe variant (*innovation* – het relatief ‘vlakke’ begin van de curve) volgt verspreiding (*propagation* – het steilere stuk van de curve) (Croft 2000: 185; Blythe & Croft 2012: 271). De logistische S-curve zal de *x*-as echter niet snijden, want taalverandering verloopt nooit ‘in een vacuüm’, maar tal van andere veranderingen en factoren spelen ook een rol, of zoals William Croft (2000) het verwoordt:

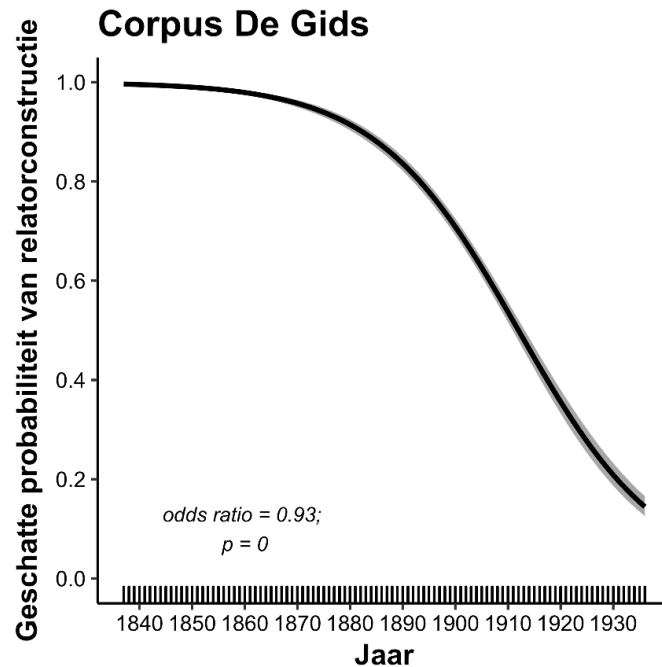
[T]he tail of the S-curve may be very long but never is achieved due to other interfering changes in the language. In other cases, the vestigial variants survive in surprising ways, as stereotypes of older or more traditional speakers, in remembered phrases, in passive community knowledge of the vestigial variant, and in the sporadic occurrence in one or two unusual speakers [...]. (Croft 2000: 185–186)

Zoals de synchrone analyses uit §3.1.1 bewijzen, is de relatorconstructie, ondanks de drastische afname aan het einde van de negentiende eeuw en de relatief lage proporties, inderdaad nog lang niet uitgestorven.

De logistische regressiecurve is gevisualiseerd in Figuur 7, de output is gegeven in Tabel 5.

Tabel 5: Output logistische regressie, met binominale constructies als referentieniveau (0)

| Coëfficiënt | Odds ratios           | Betrouwbaarheidsintervallen |                       | Significantie |
|-------------|-----------------------|-----------------------------|-----------------------|---------------|
|             |                       | 2,5%                        | 97,5%                 |               |
| Intercept   | $1,65 \times 10^{61}$ | $3,34 \times 10^{57}$       | $1,06 \times 10^{65}$ | $p < 0,001$   |
| Jaar        | 0,929                 | 0,925                       | 0,933                 | $p < 0,001$   |



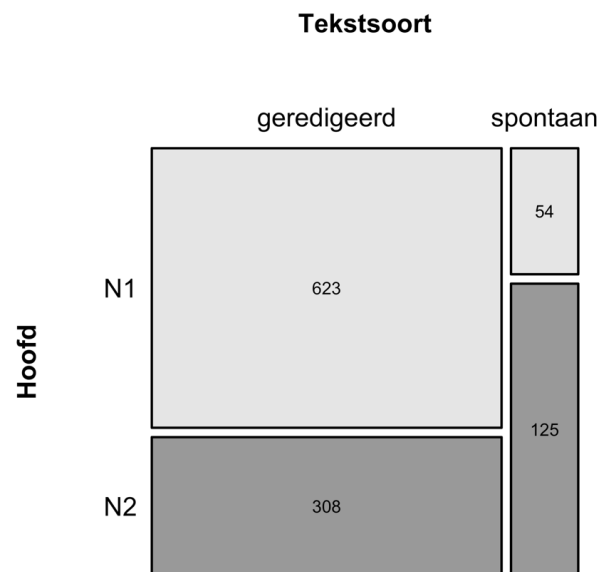
Figuur 7: Logistische regressie met het jaartal als predictor voor het gebruik van een relator (corpus De Gids)

## 3.2 Hypothese II (heranalyse)

### 3.2.1 Verandering in congruentie

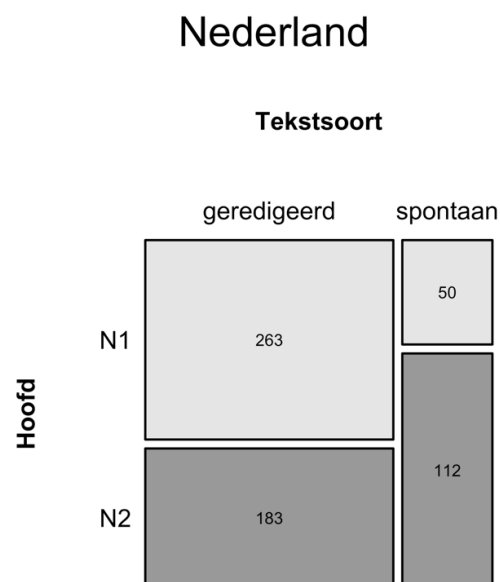
Het eerste symptoom van heranalyse, namelijk verandering in congruentie, werd getoetst aan de hand van de 1110 observaties uit het SoNaR-corpus (zie ook Voetnoot 8). Door de POS-tagging van het corpus kon immers gezocht worden op elk voorkomen van *soort* in een binominale NC die volgt op, of voorafgaat aan een vervoegd werkwoord. Over het algemeen zijn ‘niet-geheranalyseerde’ gevallen met N1 als hoofd het frequentst, met 680 gevallen (61,26%) tegenover 430 gevallen (38,74%) waarbij N2 met het werkwoord congrueert. De mozaïekplot in Figuur 8 toont de distributie van geheranalyseerde constructies (c.q. die gevallen waarin N2, en niet *soort*, congrueert met het werkwoord) en ‘niet-geheranalyseerde’ constructies, volgens register. De spontane registers bevatten materiaal uit chatgesprekken, blogs, sms’en en tweets. De geredigeerde registers bevatten krantenmateriaal. De associatie tussen constructievariant en register is statistisch significant ( $p < 0,001$ ), met een matige effectgrootte (Cramér’s  $V = 0,28$ ). Heranalyse onder de vorm van verandering in congruentie vindt dus (veel) meer plaats in spontaan

taalgebruik. Figuur 9 en Figuur 10 geven dezelfde distributie, voor respectievelijk Nederland ( $p < 0,001$ ; Cramér's  $V = 0,25$ ) en Vlaanderen ( $p < 0,001$ ; Cramér's  $V = 0,21$ ) apart.



$$\chi^2(1) = 85.22; p = 0; \text{Cramér's } V = 0.28$$

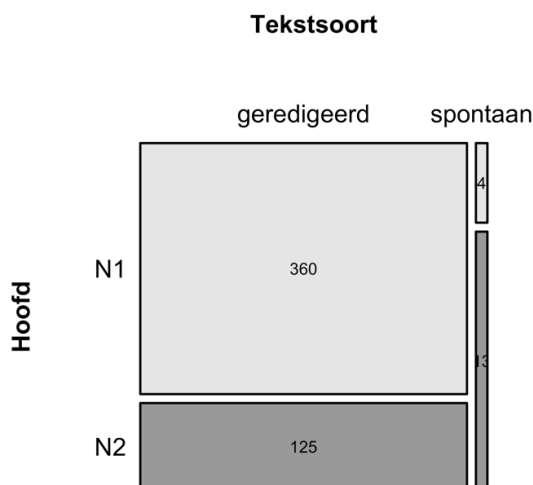
*Figuur 8: Distributie van congruentie met N1 en N2 per register als symptoom van heranalyse (SoNaR)*



$$\chi^2(1) = 37.58; p = 0; \text{Cramér's } V = 0.25$$

*Figuur 9: Distributie voor Nederland (SoNaR)*

## Vlaanderen



$$\chi^2(1) = 21.18; p = 0; \text{Cramér's } V = 0.21$$

Figuur 10: Distributie voor Vlaanderen (SoNaR)

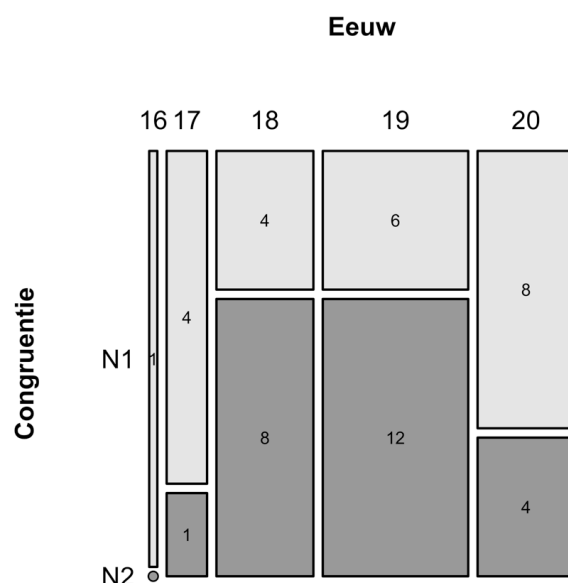
Hoewel in de overgrote meerderheid van binominale woordgroepen met *soort* in vroeger eeuwen het nomen *soort* zelf als hoofd diende, zoals in (72)–(74), zijn er in historische corpora sporadisch ook al voorbeelden te vinden waarin N2 congrueert met het werkwoord, zoals in (75)–(76).

- (72) *Maar, wat vraag ik na hun raauwe oordeelen, als **de wijzer soort van menschen**, met mijn gevoelens **overeenkomt**?* (Van Olmen, 17<sup>e</sup> eeuw)
- (73) *In goeden ernst, verstandige lezer, **die soort van ceremonyen mishaaft** my ten hooghsten* (Van Olmen, 17<sup>e</sup> eeuw)
- (74) *Dat het **een slegt soort van Menschen is**, welken niet waardig **is** te leeven, dat is zeker.* (Van Olmen, 18<sup>e</sup> eeuw)
- (75) *Soo ghy niet zijt van **dat soort van Gecken**, Die maer alleen na 't uyterlijk **sien** [...] (Van Olmen, 17<sup>e</sup> eeuw)*
- (76) *[...] dan kunnen zy wel zien, **wat soort van Winkelwaaren** 'er veil zyn in dit Ontleeders Pakhuis* (Van Olmen, 18<sup>e</sup> eeuw)

Figuur 11 toont een mozaïekplot met de distributie van congruentie met N1 (c.q. *soort* als hoofd) en N2 (de geheranalyseerde vormen) per eeuw in het corpus Van Olmen. In de tegels van de mozaïekplot zijn opnieuw telkens de absolute aantallen gegeven. Door het kleine aantal observaties is niet getest voor statistische significantie, maar de grafiek kan wel dienstdoen om de algemene trend in de data te visualiseren. Het aantal geheranalyseerde vormen neemt per eeuw duidelijk toe tot in de negentiende eeuw, maar in de twintigste eeuw zijn er plots opnieuw meer gevallen waarin N1 congrueert met het werkwoord. Die ommezwaai kan te wijten zijn aan het beperkte aantal observaties, en

een sluitende verklaring daarvoor kan hier om die reden niet gegeven worden. Een studie met meer observaties – ook van andere binominale NC's die de dubbele congruentiemogelijkheid kennen (cf. §1.4.2.2) – zou een duidelijk antwoord kunnen verschaffen.

## Corpus Van Olmen



Figuur 11: Diachrone verandering in congruentie (Van Olmen)

### 3.2.2 Verlies van de determinator

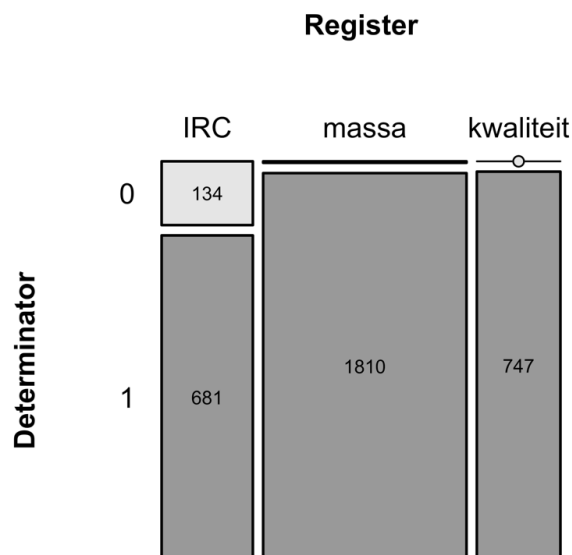
Voor het tweede symptoom van heranalyse, namelijk verlies van de determinator, werd in eerste instantie een analyse uitgevoerd op materiaal uit het Condiv-corpus geschreven Nederlands ( $n = 3475$ ), die vergelijkbaar is met de analyse op het materiaal uit het SoNaR-corpus.<sup>22</sup> Daarbij werd opnieuw de aan- of afwezigheid van een determinator afgezet tegen het register (kwaliteitskranten, massakranten en IRC), zoals gevisualiseerd in Figuur 12 (waarbij 1 staat voor aanwezigheid van een determinator en 0 voor een determinatorloze woordgroep). Ook hier is de associatie tussen beide factoren statistisch significant ( $p < 0,001$ ), met een behoorlijke effectgrootte (Cramér's  $V = 0,35$ ) en ook hier lijken spontane, niet-geredigeerde registers de voortrekkers in het proces van heranalyse. Figuur 13 en Figuur 14 geven dezelfde distributie, voor respectievelijk Nederland ( $p < 0,001$ ; Cramér's  $V = 0,33$ ) en Vlaanderen ( $p < 0,001$ ; Cramér's  $V = 0,35$ ) apart. De

22 Hier is bewust gekozen voor Condiv i.p.v. SoNaR om in een tweede stadium van het onderzoek expliciet te kunnen vergelijken met Moroccorp. Een *quick-and-dirty* zoekopdracht in SoNaR (collectie 'chats') leert dat determinatorloze vormen evengoed in SoNaR voorkomen. Een willekeurig voorbeeld daarvan is (i).

(i) *lol kwil gwoon ma zeggen er komt heel veel geld by kyken om **soort hotel** op te richte [...]* (SoNaR)

chatconversaties in (77)–(78) zijn twee voorbeelden van het gebruik van zulke determinatorloze woordgroepen met *soort*, respectievelijk uit een Nederlands en Vlaams chatkanaal.

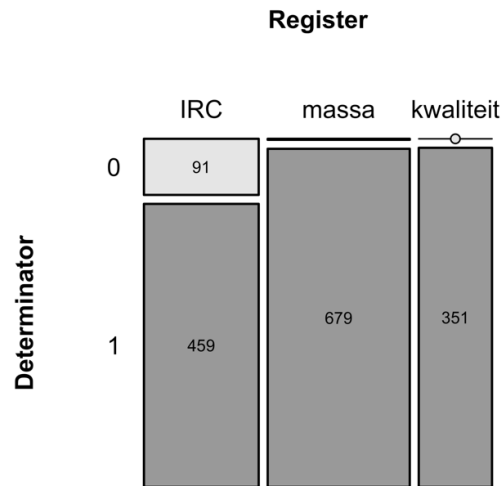
- (77) <DooGame> *ik heb es bij de multiplayer gekeken...*  
 <DooGame> *VETTE WAPENS!*  
 <BASE80> *ik heb ook al **soort minigun** gekregen*  
 (Condiv, IRC Nederland, #NL\_2)
- (78) <Boke> *slu grungie*  
 <grunge> *slu boke, dag,.....*  
 <poefelke> *t?*  
 <Sp|cer> *Boke :)*  
 <Boke> *spicer : is **soort groepen** [d]at ge op TW kunt aantreffen*  
 (Condiv, IRC Vlaanderen, #VLAAN\_1)



$$\chi^2(2) = 404.51; p = 0; \text{Cramér's } V = 0.35$$

*Figuur 12: Distributie van constructies zonder (0) en met (1) determinator volgens register (Condiv)*

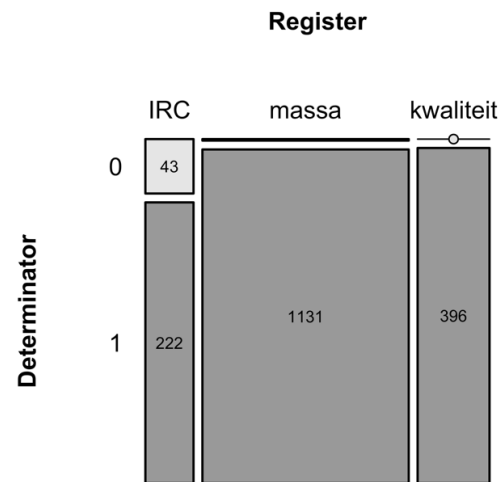
## Nederland



$$\chi^2(2) = 173.41; p = 0; \text{Cramér's } V = 0.33$$

*Figuur 13: Distributie voor Nederland (Condiv)*

## Vlaanderen



$$\chi^2(2) = 219.91; p = 0; \text{Cramér's } V = 0.35$$

*Figuur 14: Distributie voor Vlaanderen (Condiv)*

Vervolgens werden de IRC-data uit het Nederlandse deel van Condiv afgezet tegen data uit Moroccorp, een chatcorpus met Marokkaans-Nederlands (*'Moroccan-flavoured'*) materiaal (cf. Nortier & Dorleijn 2008; Grondelaers et al. 2015), zoals dat ook gedaan is in onder anderen Ruetten & Van de Velde (2013), Van de Velde & Weerman (2014) en Pijpops & Van de Velde (2015). Het zou immers zo kunnen zijn dat in etnisch Nederlands de determinator al frequenter wegvalt – i.e. dat de heranalyse van *soort* in etnisch



Nederlands al verder gevorderd is –, maar het kan evengoed een fenomeen zijn dat louter in ‘autochtoon’ Nederlands voorkomt. Daarvoor werden – na uitsluiting van valse treffers – 1023 observaties uit Morocccorp bekeken. Daarvan hadden 198 voorkomens geen determinator (19,35%), zie Figuur 15. Een chikwadraattest toont echter dat er geen significant verschil is tussen Nederlands IRC-materiaal uit Condiv en materiaal uit Morocccorp ( $p = 0,170$ ).

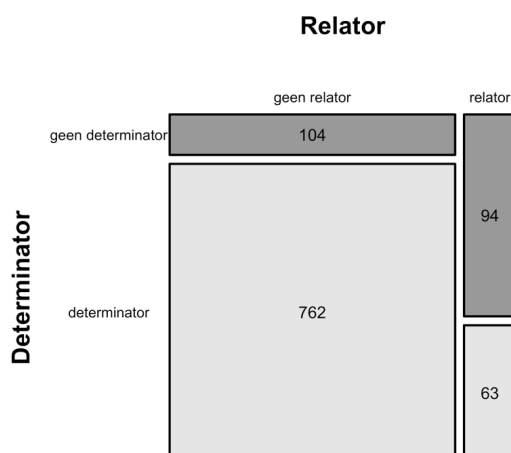
|              |                   | Corpus     |                |
|--------------|-------------------|------------|----------------|
|              |                   | Condiv-IRC | Morocccorp-IRC |
| Determinator | geen determinator | 91         | 198            |
|              | determinator      | 459        | 825            |

$$\chi^2(1) = 1.88; p = 0.17$$

*Figuur 15: Distributie van de aan- en afwezigheid van een determinator in Condiv en Morocccorp*

Bij afwezigheid van een determinator is er in Morocccorp wel een duidelijk verhoogd gebruik van relatorconstructies (59,87% van de gevallen) in vergelijking met constructies waarin wel een determinator aanwezig is (12,01% van de gevallen). De frequenties zijn gevisualiseerd in Figuur 16, de samenhang is statistisch significant ( $p < 0,001$ ), met een vrij hoge effectgrootte (Cramér's  $V = 0,44$ ).

## Moroccorp



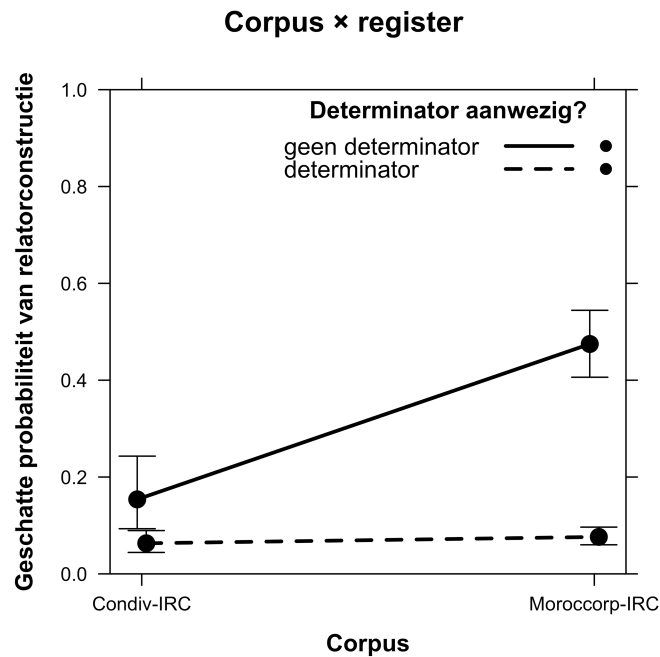
$$\chi^2(1) = 195.07; p = 0; \text{Cramér's } V = 0.44$$

Figuur 16: Aan- en afwezigheid determinator volgens aan- en afwezigheid relator in Moroccorp

Een logistische regressie ( $C = 0,720$ ; Somers'  $D_{xy} = 0,447$ ) toont dat de invloed van de aan- of afwezigheid van een determinator op het gebruik van een relatorconstructie interageert met het corpus, zie Tabel 6 en Figuur 17. Wanneer een determinator bij *soort* aanwezig is, is er zo goed als geen verschil, maar bij afwezigheid van een determinator is de geschatte probabiltiteit op een relatorconstructie significant groter in Moroccorp dan het Nederlandse IRC-gedeelte van Condiv.

Tabel 6: Output logistische regressie, met binominale constructies als referentieniveau (0)

| Coëfficiënt                                                | Logit | Betrouwbaarheidsintervallen |       | Significantie |
|------------------------------------------------------------|-------|-----------------------------|-------|---------------|
|                                                            |       | 2,5%                        | 97,5% |               |
| intercept                                                  | -1,70 | -2,32                       | -1,17 | $p < 0,001$   |
| corpus = Moroccorp-IRC                                     | 1,60  | 1,00                        | 2,27  | $p < 0,001$   |
| determinator aanwezig = ja                                 | -0,99 | -1,66                       | -0,28 | $p = 0,004$   |
| corpus = Moroccorp-IRC $\times$ determinator aanwezig = ja | -1,40 | -2,20                       | -0,63 | $p < 0,001$   |



Figuur 17: Interactie tussen corpus en aanwezigheid determinator bij 'soort' (Condiv en Moroccorp) als predictor voor het gebruik van een relator

### 3.2.3 Uittreding uit de NC

Een laatste symptoom van heranalyse is het gebruik van (*een*) *soort* (*van*) buiten de NC, als bijwoordelijke bepaling op zinsniveau (cf. §1.4.3). In het IRC-materiaal van Condiv zijn er, op enkele twijfelachtige gevallen na, geen voorbeelden te vinden van zulk gebruik van *soort*. De twijfelgevallen zijn observaties als (79), die als werkwoorden (*snowboarden*) maar ook als deverbatieve nominaliseringen (*het snowboarden*) kunnen worden gezien.

- (79) <edwin^>            *wilook nog es gaan wakeboarden.. lijk me ook wel es grappig*  
       <edwin^>            *soort **snowboarden** achter boot...*  
       <reempje>           *ow jaaah*  
       <reempje>           *heb ik eens op tv gezien*  
       (Condiv, IRC Nederland, #NL\_2)

In Moroccorp daarentegen komen gevallen met *soort* als bijwoordelijke bepaling op zinsniveau in vergelijking met Condiv behoorlijk veel voor: van de in totaal 1159 uitingen met *soort* uit Moroccorp zijn er 56 waarin *soort* niet in een binominale NC gebruikt is (4,83%). In (80)–(82) fungeert *soort van* telkens als bijwoordelijke bepaling bij een werkwoord, respectievelijk (*stoer*) *doen* in (80), (*aan het werk*) *zijn* in (81) en *leren* in (82), waarbij het de betekenis van die werkwoorden door zijn functie als *hedge* telkens afzwakt. De bijwoordelijke bepaling volgt op het werkwoord in (80) en (81); in

(82) gaat ze aan het werkwoord vooraf. In (83) daarentegen wordt een hele zin afgezwakt.

- (80) +\_MrTroublemaker *hahaha ja ze wou stoer doen **soort van** weet niet waarom*  
(Moroccorp)  
(81) +marcel *Amal24: ik ben aan hert werk, **soort van*** (Moroccorp)  
(82) chat-romeo *ben **soort van** aan het leren:p* (Moroccorp)  
(83) +SchattigMeisje *nabil\_1984: ja oke.. **soort van** kijk mij nouw normaal ben*  
*ik arrogant* (Moroccorp)

Daarnaast zijn er nog enkele gevallen waarbij *soort van* in zekere zin als een ‘kopbepaling’ fungeert (Van de Velde 2009), als in (84)–(86).<sup>23</sup> Het zijn echter slechts drie observaties die van de geraadpleegde corpora enkel in Moroccorp zijn teruggevonden. Twee ervan zijn dan ook nog eens van dezelfde chatter (Forlan), met name (85) en (86).

- (84) Mia\_West *als je **soort van** een goeie man bent* (Moroccorp)  
(85) Forlan *hij is **soort van** een slak* (Moroccorp)  
(86) Forlan *je bent **soort van** een man* (Moroccorp)

---

23 Mogelijks wordt *soort van* hier gerealiseerd in de kavel der kopbepalingen, omdat kopbepalingen typisch een (inter)subjectief karakter hebben (Van de Velde 2009: 293).

## 4 Bespreking

---

### 4.1 Strategie 1: toevoeging van een relator

In eerste instantie is een beeld geschetst van de contemporaine distributie van binominale constructies met *soort* zonder en met relator, als in respectievelijk (87) en (88).

- (87) *en met name **dat soort escapistische heimatliteratuur** bijvoorbeeld die deed het ontzettend goed* (CGN, fv400595)
- (88) *hij ontmoette toen andere academici in **een soort van centrum*** (CGN, fv800323)

Die informatie was tot op heden immers onderbelicht gebleven in de theoretische literatuur en grammatica's. Een analyse van gesproken Nederlands uit het CGN bevestigde de aanname dat varianten met een relator op dit moment sterk in de minderheid zijn, maar die **synchrone slice-of-time** zegt natuurlijk niets over de **diachrone** evolutie van de betreffende constructies. Verder komen de constructies met een relator over het algemeen iets meer voor in Vlaanderen dan in Nederland. Mogelijks zou de invloed van het Franse *une sorte de* N – met verplichte relator *de* – een rol kunnen spelen, maar de al dan niet Franse oorsprong van N2 bleek niet significant in samenhang met het type constructie.

Een logistische regressie suggereerde dat in interactie met het land ook de leeftijd van de spreker een rol speelt: oudere Vlamingen gebruiken meer relatorconstructies dan jongere Vlamingen, wat al een eerste indicatie is dat relatorconstructies – tegen de verwachtingen in – op hun retour zijn. De omgekeerde beweging zien we echter bij constructies met het onbepaald lidwoord *een*, als in (89).

- (89) *en dus ze komen terecht in **een soort van cirkelgang** of beter gezegd een spiraal-gang 't is **een soort van spiraal*** (CGN, fv400589)

De probabilistische kans op een relatorconstructie is dan veel hoger bij jongere sprekers dan bij oudere sprekers. Mogelijkerwijze is dat een gevolg van het feit dat het patroon *een soort van* door zijn frequent gebruik als *hedge* doorheen de tijd meer en meer als een eenheid gezien wordt en aldus steviger verankerd geraakt in het mentale lexicon.

In tweede instantie is nagegaan of binominale NC's met *soort* inderdaad een diachrone tendens vertonen om een relator toe te voegen (Van de Velde 2009). In dat geval zouden we een toename van relatorconstructies in de tijd moeten kunnen waarnemen. Historische data uit het corpus Van Olmen en het corpus De Gids moesten die verwachting staven. Analooq aan het *apparent time*-effect dat naar voren komt uit de CGN-data, vinden we dat relatorconstructies diachroon inderdaad het onderspit moeten delven.

Relatorloze constructies komen al sporadisch voor in de achttiende eeuw, en sinds ongeveer 1840 nemen ze in vrij snel tempo de bovenhand op relatorconstructies; aan het begin van de twintigste eeuw is proportioneel reeds meer dan de helft een constructie zonder relator. Bovendien verloopt die evolutie volgens een typische S-curve, met na de gestage opkomst een relatief snelle verspreiding vanaf ongeveer 1890.

Hoewel de strategie van het toevoegen van een relator diachroon op z'n retour is, kan worden gezegd dat het zeker geen 'uitgestorven' strategie is. De lage frequenties geven wel aan dat het niet de meest gebruikte herstelstrategie is.

## 4.2 Strategie 2: heranalyse

Zoals beargumenteerd in §1.4.2.2, is heranalyse niet altijd even gemakkelijk waar te nemen op het niveau van de individuele constituent. Daarom is gekeken naar drie symptomen in het ruimere discours die het gevolg zijn van heranalyse: (i) verandering in congruentie, (ii) verlies van de determinator en (iii) uittreding uit de NC.

Ten eerste is er verandering in congruentie met het werkwoord: niet langer *soort* congrueert met het werkwoord, maar N2, zie (90)–(91) – iets dat trouwens ook bij andere binominale NC's voorkomt (Van der Horst & Van der Horst 1999).

(90) *het soort grapjes is wel steeds hetzelfde* (SoNaR, spontane registers)

(91) *@Sophieflore bof jij even ! Dit soort mensen gaan dus ook naar de kamasutra-beurs* (SoNaR, spontane registers)

Geheranalyseerde vormen komen zowel in spontaan als geredigeerd taalgebruik voor, maar de analyse van observaties uit SoNaR toont aan dat spontane registers zoals chats, tweets en sms'en toch de voortrekkers zijn. Er werd daarbij enkel gekeken naar observaties waarbij enkelvoudig *soort* gevolgd wordt door een meervoudige N2. Het gebruik van een meervoudige persoonsvorm bij een binominale constructie met *soort* is duidelijk al stevig ingeburgerd en zelfs normatieve instanties zoals het *Genootschap Onze Taal* aanvaarden beide varianten in (90)–(91), waarbij ook verwezen wordt naar de ANS (zie §1.3.2). Meer nog, volgens het *Genootschap Onze Taal* is de geheranalyseerde vorm voor de meeste mensen zelfs de natuurlijkste:

'Dit soort situaties *zijn* te voorkomen' (met een meervoudige persoonsvorm) klinkt voor de meeste mensen het natuurlijkst. *Dit soort* heeft dan de betekenis 'zulke, dergelijke'. De redactie van ons tijdschrift *Onze Taal* kiest daarom doorgaans voor een meervoudige persoonsvorm in dit type zinnen. In *Onze*

*Taal* komen dus geregeld zinnen voor als ‘Dit soort woorden *zijn* vaak te horen’ en ‘Dat soort constructies *kunnen* lastig zijn.’<sup>24</sup>

Constructies waarbij N2 congrueert met het werkwoord komen al voor sinds de zeventiende eeuw. De data uit het corpus Van Olmen suggereren dat er een diachrone trend is waarbij congruentie met N2 frequenter wordt, maar door dataschaarste is niet getest voor significantie.

Een tweede symptoom van heranalyse dat onderzocht is, is het weglaten van de determinator, als in (92).

- (92) <RUSHy> *tom: je kan er toch ook [Ø]<sub>D</sub> andere soort stellingen ingooien als je zonodig stellingen wilt gooien* (Condiv, IRC Nederland, #CAFE\_2)

Het weglaten van de determinator in binominale constructies wijst op decategorialisatie (Hopper 1991) van N1 (c.q. *soort*), dat op die manier als een soort voorbepaling bij N2 wordt geheranalyseerd. Als N1 dan een voorbepaling wordt, wordt de ‘extra’ determinator voor N1 overbodig en kan hij worden weggelaten.

In zwaar geredigeerde registers zoals krantenmateriaal komen zulke vormen amper voor, maar in spontane registers, die dicht tegen gesproken taal aanleunen, zoals chatmateriaal, komen determinatorloze woordgroepen significant meer voor. De vraag of Marokkaans-Nederlands daarin al verder gevorderd is dan Nederlands-Nederlands werd echter niet positief beantwoord: er is geen statistisch significant verschil tussen beide corpora. Uit de analyse komt echter een andere opvallende observatie naar voren: in Moroccorp zijn er significant meer determinatorloze groepen die tegelijkertijd een relator hebben dan in Condiv, zoals in (93).

- (93) +Fes\_Rdam            *2 hrirra soms 3 ... msammen met Jam .. 5 loempiaas 3 pizzaas .. 2 bghiwaaat .. en dan Lavocaaa soort van milksake 2 glazen* (Moroccorp)

Nadere kwalitatieve inspectie van de observaties uit Moroccorp leerde dat het leeuwendeel van het grote aantal determinatorloze groepen met relator eigenlijk geen binominale woordgroepen zijn, maar bijwoordelijke bepalingen op zinsniveau. Qua semantiek bevinden ze zich eerder aan de functionele kant van het betekenispectrum; ze fungeren dus als een afzwakker of *hedge*. Zulke constructies komen echter niet voor in het zestien jaar oudere chatmateriaal van Condiv (noch in de overige geraadpleegde corpora), wat erop zou kunnen wijzen dat de constructie pas heel recent opduikt in het (Nederlands-)Nederlands, dat het een sjibbolet van Marokkaans-Nederlands is, of – en dat is de meest plausibele hypothese – dat het een calque uit het Engels is (cf. Aijmer

---

24 S.n. 2015. Dit soort situaties is / zijn te voorkomen. *Genootschap Onze Taal*. (<https://onzetaal.nl/taaladvies/dit-soort-situaties-is-zijn-te-voorkomen>) (Geraadpleegd op 24 mei 2017.)

2002). Hoe dan ook lijkt hun gebruik net als hun Engelse pendants beperkt tot informele registers (Aijmer 2002: 6).



## 5 Conclusie

---

In deze meesterproef is aan de hand van corpusgebaseerde analyses gepoogd een historisch gefundeerde verklaring te bieden voor de contemporaine distributiepatronen van binominale NC's met *soort* als N1. In de traditionele literatuur wordt N1 immers vaak als hoofd beschouwd en N2 als nabepaling of luidt het dat de status van hoofd niet eenduidig valt toe te kennen. Die eerste analyse is echter problematisch: Nederlandse NC's verdragen immers geen nabepalingen die onmiddellijk volgen op het nomen waarop ze betrekking hebben. Nabepalingen, zo argumenteert Van de Velde (2009), moeten immers bij voorkeur worden ingeleid door een relator – in het geval van binominale NC's met *soort* als N1 de prepositie *van*. Er lijken echter twee 'herstelstrategieën' werkbaar die ervoor zorgen dat de betreffende woordgroepen zich (diachroon) conformeren aan de relatorverplichting; enerzijds kan een relator worden toegevoegd, anderzijds kan de woordgroep zodanig worden geheranalyseerd zodat N1 wordt gezien als een voorbepaling bij N2.

Hoewel die herstelstrategieën wel hypothetisch gesuggereerd zijn in de literatuur (Van de Velde 2009; cf. ook Joosten & Vermeire 2006), zijn ze nooit empirisch getest. Daarom is vanuit een gecombineerd synchroon-diachroon perspectief geprobeerd beide herstelstrategieën in kaart te brengen. Verscheidene contemporaine en historische corpora van het Nederlands werden daarvoor geraadpleegd. De algemene onderzoeksvraag luidde:

**i. Op welke manier conformeren de binominale NC's met *soort* als N1 zich aan de relatorverplichting, zoals beschreven in Van de Velde (2009)?**

Om die onderzoeksvraag te beantwoorden, werd de eventuele toepassing van beide herstelstrategieën systematisch in kaart gebracht.

De eerste hypothese, namelijk dat relatorconstructies diachroon frequenter zouden worden, moest verrassend genoeg verworpen worden: het historische materiaal toonde immers dat relatorconstructies in de tweede helft van de negentiende eeuw proportioneel drastisch verminderen en dat binominale constructies zonder relator in snel tempo de bovenhand nemen. Die verandering verloopt tevens volgens een typische S-curve. Toch staan relatorconstructies, ondanks hun lage relatieve frequenties, op dit moment nog iets sterker in Vlaanderen dan in Nederland en in gevallen waarbij *soort* voorafgegaan wordt door de onbepaalde determinator *een*. De observatie dat relatorconstructies diachroon over het algemeen veel minder frequent worden, moet een aanwijzing zijn voor het feit dat relatoren niet (meer) vereist zijn. De logische bijkomende vraag

die dan gesteld moet worden, luidt of de andere herstelstrategie, met name heranalyse, ter compensatie productiever wordt.

Dat dat inderdaad het geval is, bewijzen drie ‘symptomen’ die het gevolg zijn van heranalyse en die reeds vaak de kop opsteken in voornamelijk informele registers. Ten eerste is het zo dat diachroon steeds vaker N2, het ‘nieuwe hoofd’, congrueert met het werkwoord, in plaats van N1. Hoewel de ‘niet-geheranalyseerde’ gevallen nog steeds in de meerderheid zijn, is er wel een significant verschil voor register: het leeuwendeel van de gevallen in spontane, informele registers zijn geheranalyseerde vormen. Toch vinden ook normatieve instanties als het *Genootschap Onze Taal* congruentie met N2 natuurlijker, wat erop wijst dat ze reeds sterk ingeburgerd zijn. Een verschijnsel dat nog niet algemeen gangbaar is, maar dat wel al frequent de kop opsteekt in informele registers, is het feit dat de determinator bij *soort* kan wegvallen. Omdat *soort* fungeert als een identificerende voorbepaling bij N2, is die determinator strikt genomen immers overbodig geworden. Een vergelijkende analyse met Marokkaans-Nederlands chatmateriaal uit Moroccorp gaf aan dat Nederlandse chatters niet significant verschillen van Marokkaanse chatters met betrekking tot het weglaten van de determinator. Verdere kwalitatieve analyse van de observaties uit Moroccorp legde wel een ander opvallend fenomeen bloot. Moroccorp-chatters gebruiken verschillende keren de chunk *(een) soort (van)* buiten de NC, op het niveau van het discours, als een discourspartikel met de functie van *hedge*. Dat fenomeen is niet teruggevonden in de andere geraadpleegde synchrone corpora, die alle meer dan tien jaar geleden samengesteld zijn. Mogelijks is het gebruik van *(een) soort (van)* dus een heel recent fenomeen en speelt analogie met het Engelse discourspartikel *sort of* een faciliterende rol.

Door de exploratieve aard van de analyse van *(een) soort (van)* als discourspartikel is grondiger onderzoek vereist. Meer corpusmateriaal zou een duidelijker beeld moeten geven van de distributie van het fenomeen – is het inderdaad zo dat het enkel voorkomt in etnisch Nederlands/Noord-Nederlands, of gebruiken Vlamingen het ook? Wat zijn de exacte discursieve eigenschappen van het discourspartikel? Is het een calque uit het Engels of heeft de Nederlandse variant ook eigen, idiomatische eigenschappen die het Engelse equivalent *sort of/kind of* niet heeft? Verder onderzoek zou een antwoord kunnen bieden op die vragen.

# English summary

---

This master's thesis aims to enrich the existing grammatical descriptions of binominals with *soort* 'sort/kind' as first noun, as in *een soort auto*, 'a kind [of] car', by providing corpus-based insights from a combined synchronic-diachronic perspective. Binominals are constituents containing two nouns, without it being clear which of them functions as the syntactic head. From a historical perspective, however, the first noun is clearly the head. In his monograph on the structure and history of the noun phrase (NP), Van de Velde (2009) demonstrates that Dutch NPs shun postmodification by a 'bare' constituent. He argues that a bound or free morpheme with a procedural function – called a relator – should always precede the postmodifying constituent to express the semantic coherence between both constituents. In the case of binominals with *soort*, this is the preposition *van* 'of', as in *een soort van auto*. Unfortunately, this variation is often largely ignored by existing reference grammars.

Analyses of synchronic material indicate that relator use in binominals with *soort* constitute a vast minority, with 'relatorless' binominals like *een soort* [Ø] *auto* being used in approximately 90% of the cases. It is found that factors that determine this synchronic relator use are the country, the determiner type, and the speaker's age (with several interactions between them). But this synchronic skew can, and in fact does often reflect a certain stage in an ongoing diachronic change. If this holds true for our case, we should expect a diachronic rise of relator use in binominals with *soort*, with constructions with *van* increasingly outnumbering constructions without *van*. However, a quantitative analysis of historical texts between 1500 and 2000 demonstrates that this is surprisingly not the case. In fact, the exact opposite appears to have happened, with relator use swiftly being reduced from the second half of the 19<sup>th</sup> century onwards.

This observation need not necessarily be problematic for Van de Velde's (2009) hypothesis, because he proposes a second 'repairing strategy' for binominals with *soort*. Speakers can reanalyse the binominal, so that the first noun loses its head status, thus becoming a kind of identifying premodifier of the second noun, which in turn becomes the syntactic head of the NP. Corpus analyses of three 'symptoms' of reanalysis indicate that this second repairing strategy is being used productively, especially in informal, 'spoken-like' registers such as chats. The three symptoms under scrutiny are the following:

- (i) the second noun as the new syntactic head of the NP triggers agreement in person and number on the verb, instead of the first noun;
- (ii) *soort* being reanalysed as premodifying N2 may cause the determiner to be dropped;
- (iii) the chunk (*een*) *soort* (*van*) is used outside the NP, as a hedging particle at the discourse level.

This last symptom only occurs regularly in a corpus compiled in 2013, and not in the other corpora, all of which were compiled more than ten years ago. This suggests it being only a very recent development, arguably facilitated by the English equivalent *sort/kind of*.

These insights into the construction's evolution and contemporary distribution add valuable information to the hitherto available knowledge about these particular constructions.

# Referenties

---

- Abney, Steven Paul. 1987. *The English noun phrase in its sentential aspect*. Cambridge, MA: The MIT Press. (Doctoraal proefschrift.)
- Aijmer, Karin. 2002. *English discourse particles: Evidence from a corpus*. Amsterdam: John Benjamins.
- Anthony, Laurence. 2014. AntConc (versie 3.4.3m) [computersoftware]. Tokio: Waseda University. (<http://www.laurenceanthony.net/>)
- Baayen, Rolf Harald. 2011. *Analyzing linguistic data: A practical introduction to statistics*. 5<sup>e</sup> dr. Cambridge: Cambridge University Press.
- Blythe, Richard A. & William Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88(2). 269–304.
- Brems, Lieselotte & Kristien Davidse. 2010. The grammaticalisation of nominal type noun constructions with *kind/sort of*: Chronology and paths of change. *English Studies* 91(2). 180–202.
- Broekhuis, Hans & Evelien Keizer. 2012. *Syntax of Dutch: Nouns and noun phrases*, vol. 1. Amsterdam: Amsterdam University Press.
- Broekhuis, Hans & Marcel den Dikken. 2012. *Syntax of Dutch: Nouns and noun phrases*, vol. 2. Amsterdam: Amsterdam University Press.
- Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4). 711–733.
- Bybee, Joan L. 2010a. Diachronic linguistics. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 945–987. 2<sup>e</sup> dr. Oxford: Oxford University Press.
- Bybee, Joan L. 2010b. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Campbell, Lyle. 2013. *Historical linguistics: An introduction*. 3<sup>e</sup> dr. Cambridge, MA: The MIT Press.
- Croft, William. 2000. *Explaining language change: An evolutionary approach*. Harlow: Longman.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Davidse, Kristin, Lieselotte Brems & Liesbeth De Smedt. 2008. Type noun uses in the English NP: A case of right to left layering. *International Journal of Corpus Linguistics* 13(2). 139–168.
- De Clerck, Bernard & Timothy Coleman. 2013. From noun to intensifier: *Massa* and *massa's* in Flemish varieties of Dutch. *Language Sciences* 36. 147–160.
- Denison, David. 2002. History of the *sort of* construction family. (Paper voorgesteld op de Second International Conference of Construction Grammar, Helsinki, 7 september 2002.)

- Denison, David. 2003. Log(ist)ic and simplistic S-curves. In Raymond Hickey (ed.), *Motives for language change*, 54–70. Cambridge: Cambridge University Press.
- Dik, Simon C. 1983. Two constraints on relators and what they can do for us. In Simon C. Dik (ed.), *Advances in Functional Grammar*, 267–298. Dordrecht: Foris.
- Fox, John & Sanford Weisberg. 2011. *An R companion to applied regression*. 2<sup>e</sup> dr. Thousand Oaks, CA: Sage.
- Fox, John. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8(15). 1–27.
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago/Londen: The University of Chicago Press.
- Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *TRENDS in Cognitive Sciences* 7(5). 219–224.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Greenberg, Joseph Harold. 1966. Some universals of language with special reference to the order of meaningful elements. In Joseph Harold Greenberg (ed.), *Universals of language*, 73–113. 2<sup>e</sup> dr. Cambridge, MA: The MIT Press.
- Gries, Stefan Th. 2009. *Statistics for linguistics with R: A practical introduction*. Berlin/New York: De Gruyter.
- Grondelaers, Stefan, Katrien Deygers, Hilde Van Aken, Vicky Van Den Heede & Dirk Speelman. 2000. Het CONDIV-corpus geschreven Nederlands. *Nederlandse Taalkunde* 4(5). 356–363.
- Grondelaers, Stefan, Paul van Gent & Roeland van Hout. 2015. Is Moroccan-flavoured Standard Dutch standard or not? On the use of perceptual criteria to determine the limits of standard languages. In Alexei Prikhodkine & Denis R. Preston (eds.), *Responses to language varieties*, 191–218. Amsterdam: John Benjamins.
- Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij & Maarten C. van den Toorn (red.). 1997. *Algemene Nederlandse Spraakkunst*. 2<sup>e</sup> dr. Groningen/Deurne: Martinus Nijhoff/Wolters Plantyn.
- Haiman, John. 1985. *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Heine, Bernd. 2003. Grammaticalization. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 575–601. Oxford: Blackwell.
- Hilpert, Martin & Stefan Th. Gries. 2016. Quantitative approaches to diachronic corpus linguistics. In Merja Kytö & Päivi Pahta (eds.), *The Cambridge Handbook of English historical linguistics*, 36–53. Cambridge: Cambridge University Press.
- Hoeksema, Jack. 2011. Het WNT: Een waarlijk nuttige tool? *Nederlandse Taalkunde* 16(2). 152–159.

- Hoeksema, Jack. 2014. De opkomst van *aan* als verbindend element in maatnomenconstructies. In Freek Van de Velde, Hans Smessaert, Freek Van Eynde & Sara Verbrugge (red.), *Patroon en argument: Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst*, 421–432. Leuven: Leuven University Press.
- Hoffman, Thomas & Graeme Trousdale. 2013. *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.
- Hopper, Paul & Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2<sup>e</sup> dr. Cambridge: Cambridge University Press.
- Hopper, Paul. 1987. Emergent grammar. *Proceedings of the thirteenth annual meeting of the Berkeley Linguistics Society*. 139–157.
- Hopper, Paul. 1991. On some principles of grammaticalization. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. 1: *Focus on theoretical and methodological issues*, 17–35. Amsterdam/Philadelphia: John Benjamins.
- Hüning, Matthias, Ulrike Vogl, Ton van der Wouden & Arie Verhagen. 2006. *Nederlands tussen Duits en Engels: Handelingen van de workshop op 30 september en 1 oktober aan de Freie Universität Berlin*. Leiden: SNL.
- Joosten, Frank. 2003. *Collectiva en aggregaatsnamen in het Nederlands: Begripsbepaling en typologie*. Leuven: KU Leuven. (Doctoraal proefschrift.)
- Joosten, Frank & Lea Vermeire. 2006. Collectiva en relationaliteit. *Nederlandse Taalkunde* 11. 23–58.
- Kay, Paul. 1984. The kind of/sort of construction. *Proceedings of the tenth annual meeting of the Berkeley Linguistics Society*. 157–171.
- Keizer, Evelien. 2007. *The English noun phrase: The nature of linguistic categorization*. Cambridge: Cambridge University Press.
- Kemmer, Suzanne & Michael Barlow. 2000. Introduction: A usage-based conception of language. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, vii–xxii. Stanford: CSLI.
- Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1. 199–244.
- Lakoff, George. 1973. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic* 2(4). 458–508.
- Lehmann, Christian. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20(3). 303–318.
- Meyer, David, Achim Zeileis & Kurt Hornik. 2006. The strucplot framework: Visualizing multi-way contingency tables with *vcd*. *Journal of Statistical Software* 17(3). 1–48.
- Meyer, David, Achim Zeileis & Kurt Hornik. 2016. *vcd: Visualizing categorical data. R package versie 1.4–3*.

- Mihatsch, Wiltrud. 2007. The construction of vagueness: ‘Sort of’ expressions in Romance languages. In Günter Radden, Klaus-Michael Köpcke, Thomas Berg & Peter Siemund (eds.), *Aspects of meaning construction*, 225–245. Amsterdam/Philadelphia: John Benjamins.
- Mihatsch, Wiltrud. 2016. Type-noun binominals in four Romance languages. *Language Sciences* 53. 136–159.
- Norde, Muriel & Graeme Trousdale. 2016. Exaptation from the perspective of construction morphology. In Muriel Norde & Freek Van de Velde (eds.), *Exaptation and language change: Current issues in linguistic theory*, 163–195. Amsterdam/Philadelphia: John Benjamins.
- Nortier, Jacomine & Margreet Dorleijn. 2008. A Moroccan accent in Dutch: A sociocultural style restricted to the Moroccan community? *International Journal of Bilingualism* 12(1/2). 125–142.
- Oostdijk, Nelleke, Martin Reynaert, Véronique Hoste & Ineke Schuurman. 2013. The construction of a 50-million-word reference corpus of contemporary written Dutch. In Peter Spyns & Jan Odiijk (eds.), *Essential speech and language technology for Dutch: Results by the STEVIN-programme*, 219–247. Berlijn: Springer.
- Pekelder, Jan. 1993. *Conventies en functies: Aspecten van binominale woordgroepen in het hedendaagse Nederlands*. Louvain-la-Neuve: Peeters.
- Philippa, Marlies, Frans Debrabandere, Arend Quak, Tanneke Schoonheim & Nicoline van der Sijs (red.). 2009. *Etymologisch woordenboek van het Nederlands: S–Z*. Amsterdam: Amsterdam: University Press.
- Pijpops, Dirk & Freek Van de Velde. 2014. A multivariate analysis of the partitive genitive in Dutch: Bringing quantitative data into a theoretical discussion. *Corpus Linguistics and Linguistic Theory* (ahead of print). 1–33.
- Pijpops, Dirk & Freek Van de Velde. 2015. Ethnolect speakers and Dutch partitive adjectival inflection: A corpus analysis. *Taal en Tongval* 67(2). 343–371.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1992. *A grammar of contemporary English*. 20<sup>e</sup> dr. Londen: Longman.
- R Core Team. 2016. *R: A language and environment for statistical computing*. Wenen: R Foundation for Statistical Computing. (<https://www.R-project.org/>)
- Ruette, Tom & Freek Van de Velde. 2013. Moroccorp: Tien miljoen woorden uit twee Marokkaans-Nederlandse chatkanalen. *Lexicos* 23. 456–475.
- Schermer-Vermeer, Ina C. 2008. De SOORT-constructie: Een nieuw patroon in het Nederlands. *Nederlandse Taalkunde* 13(1). 2–33.
- Scott, Alan K. 2011. The position of the genitive in present-day Dutch. *Word Structure* 4(1). 104–135.
- Scott, Alan K. 2014. The survival and use of case morphology in Modern Dutch. In Ronny Boogaart, Timothy Coleman & Gijsbert Rutten (eds.), *Extending the scope of construction grammar*, 107–137. Berlijn: De Gruyter Mouton.
- Speelman, Dirk. 2014. Logistic regression: A confirmatory technique for comparisons in corpus linguistics. In Dylan Glynn & Justyna A. Robinson (eds.), *Corpus methods*



- for semantics: *Quantitative studies in polysemy and synonymy*, 487–533. Amsterdam: John Benjamins.
- Sperber, Dan & Deirdre Wilson. 1991. Loose talk. In Steven Davis (ed.), *Pragmatics: A reader*, 540–549. Oxford: Oxford University Press.
- Stoett, Frederik A. 1923. *Middelnederlandsche spraakkunst: Syntaxis*. 's-Gravenhage: Martinus Nijhoff.
- Tummers, José, Kris Heylen & Dirk Geeraerts. 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1/2. 225–261.
- Van Bree, Cor. 1987. *Historische grammatica van het Nederlands*. Dordrecht: Foris.
- Van der Horst, Joop M & Freek Van de Velde. 2016. Miljoen. *Leuvense Bijdragen* 99/100. 410–424.
- Van der Horst, Joop M. & Kees van der Horst. 1999. *Geschiedenis van het Nederlands in de twintigste eeuw*. 's-Gravenhage: Sdu.
- Van der Horst, Joop M. 2003. Tal van informatie: De verandering van een telwoord. *Onze Taal* 72(2/3). 58–60.
- Van der Horst, Joop M. 2006. *Geldhoeveelheid of hoeveelheid geld?* Waarom de samenstelling soms beter is dan de groep. *Onze Taal* 75(12). 346–349.
- Van der Horst, Joop M. 2008. *Geschiedenis van de Nederlandse syntaxis*. Leuven: Leuven University Press.
- Van der Lubbe, Henricus F.A. 1968. *Woordvolgorde in het Nederlands: Een synchrone structurele beschouwing*. 3<sup>e</sup> dr. Assen: Van Gorcum.
- Van de Velde, Freek & Fred Weerman. 2014. The resilient nature of adjectival inflection in Dutch. In Petra Sleeman, Freek Van de Velde & Harry Perridon (eds.), *Adjectives in Germanic and Romance*, 113–145. Amsterdam: John Benjamins.
- Van de Velde, Freek. 2009. *De nominale constituent: Structuur en geschiedenis*. Leuven: Leuven University Press.
- Van de Velde, Freek. 2014. Nederlandse predeterminatoren als levend fossiel. *Nederlandse Taalkunde* 19(1). 87–103.
- Van de Velde, Freek. 2017. Understanding grammar at the community level requires a diachronic perspective: Evidence from four case studies. *Nederlandse Taalkunde* 22(1). 47–74.
- Van Eerten, Laura. 2007. Over het Corpus Gesproken Nederlands. *Nederlandse Taalkunde* 12(3). 194–215.
- Van Eynde, Frank. 2004. *Part-of-speech tagging en lemmatisering van het Corpus Gesproken Nederlands*. Leuven: Centrum voor Computerlinguïstiek, KU Leuven.
- Van Haeringen, Coenraad Bernardus. 1956. *Nederlands tussen Duits en Engels*. 's-Gravenhage: Servire.
- Van Langendonck, Willy. 2010. Iconicity. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 394–418. 2<sup>e</sup> dr. Oxford: Oxford University Press.

- Van Loey, Adolphe. 1970. *Schönfelds historische grammatica van het Nederlands: Klankleer, vormleer en woordvorming*. 8<sup>e</sup> dr. Zutphen: Thieme.
- Van Olmen, Daniël. 2013. The imperative of *say* as a pragmatic marker in English and Dutch. *Journal of Germanic Linguistics* 25(3). 247–287.
- Van Olmen, Daniël & Johan van der Auwera. 2014. “Over zo’n en zo meer”. In Freek Van de Velde, Hans Smessaert, Frank Van Eynde & Sara Verbrugge (red.), *Patroon en argument: Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst*, 215–228. Leuven: Leuven University Press.
- Verhagen, Arie. 2005. Constructiegrammatica en ‘usage based’ taalkunde. *Nederlandse Taalkunde* 10(3/4). 197–222.
- Vos, Hinderika M. 1999. *A grammar of partitive constructions*. Tilburg: Tilburg University. (Doctoraal proefschrift.)
- Wickham, Hadley & Romain Francois. 2016. *dplyr: A grammar of data manipulation*. *R package versie* 0.5.0.
- Wickham, Hadley. 2009. *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.
- Wickham, Hadley. 2016. *stringr: Simple, consistent wrappers for common string operations*. *R package versie* 1.1.0.
- Weerman, Fred & Petra de Wit. 1999. The decline of the genitive in Dutch. *Linguistics* 37(6). 1155–1192.
- WNT = onlineversie van De Vries, Matthias & Lammert te Winkel. 1882–1998. *Woordenboek der Nederlandsche Taal*. ’s-Gravenhage: Martinus Nijhoff. (<http://gtb.inl.nl/>)