

Faculté de philosophie, arts et lettres

Het begin van de Nederlandse zin

Syntactische analyse van de relatie tussen
aanloop en eerste stuk in mededelende zinnen
op basis van een corpusonderzoek

Auteur : Arnaud Van der Meiren
Promotrice : Elisabeth Degand
Année académique 2018-2019
Master [120] en langues et lettres modernes, orientation germanique
Finalité didactique

Het begin van de Nederlandse zin

Syntactische analyse van de relatie tussen aanloop en eerste stuk
in mededelende zinnen op basis van een corpusonderzoek

Auteur : Arnaud Van der Meiren

Promotrice : Elisabeth Degand

Année académique 2018-2019

Master [120] en langues et lettres modernes, orientation germanique
Finalité didactique

In eerste instantie bedank ik mijn scriptie supervisor, mevr. Elisabeth Degand, voor haar verbeteringen, haar tips, haar feedback en haar aanmoedigen.

Ik ben ook mijn gezin, mijn vrienden en mijn vakgenoten erkentelijk voor hun waardevolle steun bijna drie jaar lang.

Inhoudsopgave

Inleiding	11
Lijst van afkortingen, conventies en symbolen.....	13
1. De Nederlandse zin: Eerste stuk en aanloop.....	15
1.1 Inleiding: de Nederlandse zin	15
1.2 Het eerste stuk.....	17
1.3 De aanloop	20
1.4 Syntactische distributie en informatiestructuur	25
1.5 State of the art: Linkerperiferie en eerste stuk van de Nederlandse zin.....	31
2. Methode	35
3. Nevenschikkende voegwoorden	47
4. Resultaten.....	51
4.1 NRC	51
4.2 CGN	55
4.3 Vergelijking	60
5. Interpretatie	65
5.1 Relevantie	65
5.2 Aanloop.....	66
5.3 Eerste zinsplaats.....	70
5.4 Discussie	73
Conclusie.....	75
Bibliografie	77
Appendix.....	79

Inleiding

Dagelijks maken mensen en dieren gebruik van taal om voor een gevaar te waarschuwen, iemand te bedanken, zijn zinnen en behoeften uit te drukken, de wereld rond zichzelf te leren kennen enz, d.i. om te communiceren. Talen kunnen uiterst divers zijn: sommige worden gesproken en/of geschreven, andere functioneren door middel van gebaren, sommige zijn SOV-talen (subject, object en verbum) zoals het Nederlands of het Japans, andere zijn VSO-talen zoals het Arabisch of het Iers, nog andere zijn SVO-talen zoals het Frans of het Mandarijn. Het is dus duidelijk dat elke taal zijn eigen systeem met zijn eigen regels bevat. De spreker moet dan keuzes maken naargelang de mogelijkheden binnen een zeker taalsysteem om een welbepaalde boodschap mee te delen. Bijvoorbeeld: als een Nederlandstalige spreker een vraag wil stellen, zal hij schriftelijk ofwel een vraagwoord gebruiken, ofwel met het werkwoord beginnen. Mondeling zal hij verder extra intonatie toevoegen.

Deze scriptie bestudeert de relatie tussen semantiek en syntaxis binnen het informatieoverdrachtproces. Daarvoor wordt de aandacht besteed aan de aanloop en aan het eerste stuk in mededelende zinnen. Die twee zijn preverbale elementen met eigen functies binnen de zin. Aan de hand van een corpusonderzoek wordt er een vergelijking gemaakt tussen geschreven en gesproken taal ten opzichte van de aan- of afwezigheid van de aanloop en het eerste stuk en de functies daarvan. Op basis van corporastatistieken worden dan de gegevens geïnterpreerd om de relatie tussen vorm en inhoud in de communicatie aan het licht te brengen.

De scriptie wordt verdeeld als volgt: het eerste hoofdstuk omvat de noodzakelijke theorie over de Nederlandse zin, de aanloop en het eerste stuk. Dit hoofdstuk bevat ook een sectie over het verband tussen syntactische distributie en informatiestructuur en laat enkele wetenschappelijke artikelen de revue passeren. Hoofdstuk 2 betreft de gebruikte methode voor deze scriptie. Dan worden nevenschikkende voegwoorden wat nader bekeken gezien het gebrek aan consensus wat hen betreft. Onderdeel 4 bevat de resultaten van het corpusonderzoek. Die worden dan in het volgende hoofdstuk geïnterpreteerd op basis van de vraag: welke impact hebben de aan- of afwezigheid van de aanloop en het eerste stuk en de elementen daarin op de betekenis van de zin? Ten slotte wordt het tijd voor een samenvatting en conclusie.

Lijst van afkortingen, conventies en symbolen

ANS	Algemene Nederlandse Spraakkunst
BC	bijwoordelijke constituent
bv.	bijvoorbeeld
c.	circa (ongeveer)
cf.	confer (zie)
CGN	Corpus Gesproken Nederlands
d.i.	dat is
DM	discourse marker
d.w.z.	dat wil zeggen
ENA	eigenaamagent
enz.	enzovoort
et al.	et alii (en anderen)
i.e.	id est (dat is)
NC	naamwoordelijke constituent
nl.	namelijk
NRC	Nieuwe Rotterdamse Courant
NSVW	nevenschikkend voegwoord
o.a.	onder andere
SOV	subject-object-verbum
SVO	subject-verbum-object
t.e.m.	tot en met
VB	voornaamwoordelijk bijwoord
VC	voorzetselconstituent
VR	verbale rest
VSO	verbum-subject-object
VTT	voltooid tegenwoordige tijd
VWBW	voegwoordelijk bijwoord
WC	werkwoordelijke constituent

1. De Nederlandse zin: Eerste stuk en aanloop

1.1 Inleiding: de Nederlandse zin

Deze scriptie heeft als doel de syntactische distributie binnen de gesproken mededelende Nederlandse zin te bestuderen, in het bijzonder de elementen die vóór de persoonsvorm – ook PV genoemd – voorkomen. Daarvoor is het eerst cruciaal om de Nederlandse zin van binnen en buiten te kennen.

Ten eerste, wat is een zin? Wat betekent dit concept? De betekenis van het woord ‘zin’ is immers complex te beschrijven, ook al heeft die bepaalde eigenschappen: een zin vormt een conceptuele en formele eenheid, die onafhankelijk uit één woord of een reeks woorden bestaat en die een onderwerp met ten minste één gezegde – dat informatie over het onderwerp geeft – verbindt. In geschreven taal kenmerkt de zin zich door een hoofdletter aan het begin en een punt, een vraagteken, een uitroepeten of gedachtenpuntjes aan het einde, terwijl de gesproken zin intonatiekenmerken bezit, bijvoorbeeld eindtonen of pauzes (Verbanck, 1985: 47).

In (1) is ‘Jan’ het onderwerp en ‘slaapt tot 12u op zaterdag’ het gezegde aangezien dit segment informatie levert over Jan en over wat hij doet.

(1a) Jan slaapt tot 12u op zaterdag.

De Nederlandse zin structureert zich rond twee verbale polen; de eerste pool of PV en de tweede pool of verbale rest (of VR). Vóór de eerste pool komt het eerste stuk voor, tussen de twee polen staat het middenstuk en na de tweede pool bevindt het derde stuk zich. De drie stukken moeten niet per se ingevuld worden. Schematisch kan de zin als volgt voorgesteld worden:

Eerste stuk – Eerste pool – Middenstuk – Tweede pool – Derde stuk

In enkelvoudige en nevenschiktelijke zinnen bestaat de eerste pool uit het vervoegde werkwoord en kan de verbale rest uit voltooid deelwoorden of infinitieven samengesteld worden. Soms kan het ook leeg blijven, zoals in zin (1). Dan is er sprake van verbale nul-rest.

(1b) Jan slaapt tot 12u op zaterdag \emptyset \emptyset

De verbale rest kan gemakkelijk bekleed en/of vastgesteld worden door de zin in de voltooid tegenwoordige tijd (VTT) te zetten:

(2) Jan heeft op zaterdag tot 12u geslapen \emptyset

In ondergeschikte zinnen wordt de eerste pool door een conjunctor vervuld, ie. hetzij een onderschikkend voegwoord, hetzij een betrekkelijk voornaamwoord. Dan kan de tweede pool alle werkwoordsvormen bevatten – vervoegd werkwoord, voltooide deelwoorden en infinitieven. Voorbeeld (3) kan in twee aparte zinnen verdeeld worden: de hoofdzin ‘ik weet’ en de bijzin ‘dat Jan op zaterdag tot 12u slaapt’. Laten we ons voor dit voorbeeld alleen op de bijzin concentreren. De eerste pool is hier het onderschikkend voegwoord ‘dat’ en de verbale rest bestaat uit het vervoegde werkwoord ‘slaapt’. In vergelijking met (1b) heeft het onderwerp in (3) zich opvallend van het eerste stuk naar het middenstuk verplaatst omdat het voorveld niet bestaat in de structuur van de bijzin.

(3) Ik weet dat Jan tot 12u op zaterdag slaapt.

Rond de twee verbale polen bestaan de eerste-, midden- en derde stukken uit zinsdelen, i.e. zelfstandig functionerende woorden of combinaties van woorden (soms zinnen) waarvan de grammaticale functie bepalend is voor de betekenis van de zin, bv. onderwerpen, voorwerpen, bepalingen enz. (Verbanck, 1985). Een cruciale regel in de Nederlandse syntaxis is dat **maximaal** één zinsdeel in het eerste stuk kan staan, dus vóór de persoonsvorm. Daarom hebben we het over de PV2-regel – omdat de persoonsvorm op de tweede zinsplaats moet staan.

(4) Op zaterdag slaapt Jan tot 12u.

(5) *Op zaterdag Jan slaapt tot 12u.

(1a) Jan slaapt tot 12u op zaterdag.

Terwijl (4) perfect grammaticaal is, bevat het eerste stuk in (5) twee verschillende zinsdelen, wat de zin ongrammaticaal maakt. ‘Op zaterdag’ is immers een prepositionele constituent die als bijwoordelijke bepaling van tijd fungeert en ‘Jan’ is hier een nominale constituent (NC) die als onderwerp dient. In dit geval bevindt de persoonsvorm zich op de derde zinsplaats en wordt de zin ongrammaticaal. Een andere mogelijkheid is de voorzetselconstituent (VC) in het middenstuk te zetten, zoals in (1a).

De enige twee uitzonderingen waar de persoonsvorm niet op de tweede zinsplaats lijkt te staan zijn imperatiefzinnen (6) en vraagzinnen zonder vraagwoord (7). In die gevallen blijft het eerste stuk leeg. Het lijkt erop dat de persoonsvorm zich naar de eerste zinsplaats beweegt, maar dan zou de eerste pool niet vervuld worden ondanks de aanwezigheid van een vervoegd werkwoord. Correcter is dus om aan te nemen dat het eerste stuk leeg blijft in deze gevallen.

(6) Ø. Ga maar zitten, alsjeblieft.

(7) Ø. Heeft Jan vandaag tot 12u geslapen?

Nu de algemene basisregels van de Nederlandse syntaxis doorgenomen zijn, is het tijd om nader te kijken naar wat zich vóór de eerste pool bevindt.

1.2 Het eerste stuk

Wat vóór de persoonsvorm staat in de Nederlandse zin wordt eerste stuk, eerste zinsplaats of voorveld genoemd. De eerste zinsplaats is uitvoerig door Van Belle et al. (2006) geanalyseerd. Zoals hierboven vermeld, bestaat het eerste stuk uit maximaal één zinsdeel en kan het eventueel leeg blijven. In (1b) en (4) uit de inleiding fungeerden de eerste stukken respectievelijk als onderwerp en bijwoordelijke bepaling van tijd maar in principe kan elk willekeurig zinsdeel vooraan staan, zij het op basis van zijn vorm of zijn syntactische functie, naargelang van de informatieve geleding van de zin. Uit Verscheldens (2015) redelijk beperkte contrastieve onderzoek naar het voorveld in het Nederlands en in het Spaans blijkt dat de eerste zinsplaats in c. 57% van de mededelende zinnen als onderwerp dient. Satellieten komen ongeveer om de vier mededelende zinnen in het eerste stuk voor (c. 24%), waaronder satellieten van tijd die in ongeveer de helft van de gevallen het voorveld bezetten (c. 12% van het gehele corpus). Er wordt toegevoegd dat lijdend voorwerpen in c. 13% van de zinnen in het voorveld

stonden. Een dergelijk percentage kan gemakkelijk verklaard worden aangezien Verschelden met een klein corpus krantenartikelen werkte en in c. 75% van de gevallen waar het direct object zich in het voorveld bevond, bestond dat uit een citaat van het type (8).

(8) “Ebola houdt zich niet aan grenzen”, beklemtoonde Chrystos [...]

Van al de elementen die in het eerste stuk kunnen voorkomen, zijn hier voorbeelden van de voornaamste (9 t.e.m. 12).

(9) Hem vraag ik het niet.

(10) Die film zal ik nooit vergeten.

(11) Aan haar zus heeft ze nooit gedacht.

(12) Woedend gaf ze hem een mep.

Hier oefenen de eerste zinsplaatsen de syntactische functies van meewerkend voorwerp (9), lijdend voorwerp (10), prepositioneel object (11) en predicatieve toevoeging (12)¹ uit. Gezien hun vooropplaatsing in de zin wordt het belang van de objecten benadrukt door een contrast te doen ontstaan. In gesproken taal wordt het contrast nog aangescherpt door het zinsaccent op het eerste stuk te laten vallen. Andere talen zoals het Frans (10b) en het Engels (10c) kunnen een beroep doen op gekloofde zinnen om het contrast duidelijker te maken. Deze procedés vallen onder het verschijnsel focalisatie (Van Belle, 2006: 14). Dergelijke zinnen houden in dat ook al zal ‘ik’ die film nooit vergeten, ‘ik’ is misschien andere films al vergeten.

(10a) Die film zal ik nooit vergeten. [Andere films ben ik echter al vergeten.]

(10b) C’est ce film que je ne vais pas oublier.

(10c) It is this movie that I will never forget.

Schriftelijk worden eerste stukken en persoonsvormen nooit door een komma gescheiden met uitzondering van bijzinnen, zoals in (13). Door een komma toe te voegen is de spreker verplicht om een intonatieve pauze tussen twee verbale groepen – de verbale rest van de bijzin en de persoonsvorm van de hoofdzin – in te lassen, wat het begrip van de uiting vergemakkelijkt.

¹ Voorbeelden (9) t.e.m. (12) worden uit Van Belle et al. (2013: 9) opgenomen.

(13) Als Jan te laat opstaat, zal hij weer zijn trein missen.

De voornaamste uitzondering waar meer dan één element op de eerste zinsplaats schijnt te staan, is als een voegwoordelijk bijwoord – vanaf nu VWBW – in de zogenaamde Wackernagelpositie komt te staan. In dit geval gaat het VWBW met een ander zinsdeel gepaard (14). De VWBWen die in deze gevallen net vóór de eerste pool kunnen staan, zijn degene die normaal in de aanloop staan, i.e. ‘echter’, ‘immers’, ‘integendeel’ en ‘trouwens’ (cf. 1.3. punt (d)). Toch bestaan er andere mogelijkheden met ‘daarentegen’, ‘dus’ en ‘toch’ bijvoorbeeld. Een dergelijk verschijnsel doet zich voor omdat de voegwoordelijke bijwoorden als het ware deel uitmaken van het voorafgaande zinsdeel.

(14) Wegens zijn ziekte is Frank thuisgebleven. Willem echter voelde zich beter en is naar school gegaan.

Onder de elementen die zich liever niet in de eerste zinsplaats bevinden, tellen Van Belle et al. (2006: 17-18) de volgende:

- Het niet-onderwerp onzijdige voornaamwoord ‘het’. In plaats daarvan wordt het aanwijzende voornaamwoord ‘dat’ gebruikt (15):

(15a) Apotheken verkopen het nu goedkoper dan vroeger.

(15b) *Het verkopen apotheken nu goedkoper dan vroeger.

(15c) Dat verkopen apotheken nu goedkoper dan vroeger.

- Reflexieve en reciproke voornaamwoorden. Wederkerige pronomina kunnen alleen in het voorveld staan als ze een contrast ten opzichte van een ander zinsdeel uitdrukken. Wat de wederkerende voornaamwoorden betreft, ze moeten gepaard gaan met het suffix –zelf, dat de pronomina versterkt (16):

(16a) Ze vervelen zich enorm voor het ogenblik.

(16b) *Zich vervelen ze enorm voor het ogenblik.

(16c) Zichzelf vervelen ze enorm voor het ogenblik.

- Onbepaalde nominale constituenten. Gezien hun informatieve waarde staan die elementen normaal gezien zo ver mogelijk in het middenstuk en nooit op de eerste zinsplaats. De structuur wordt dan vervangen door een er-zin (17):

(17a) *Een auto-ongeluk gebeurde vanmorgen in mijn straat.

(17b) Er gebeurde een auto-ongeluk vanmorgen in mijn straat.

Er moet een onderscheid gemaakt worden tussen onbepaalde naamwoordelijke constituenten als in (17) aan de ene kant en generieke en categoriale naamwoordelijke constituenten aan de andere kant. Alle nominale groepen hebben een gelijke vorm, met name ze worden allemaal gebruikt met onbepaalde determinatoren, wat het gevaar op verwarring groter maakt. Functioneel werken ze wel verschillend. Onbepaalde NC's dienen om nieuwe informatie in het gesprek bekend te maken, zoals in (17b) bijvoorbeeld. Daarom zitten ze over het algemeen achteraan in het middenstuk. Categoriale en generieke naamwoordelijke groepen geven integendeel geen onbekende informatie aan. Ze verwijzen eerder naar concepten dan naar geïsoleerde onbepaalde eenheden. In (17) is 'een auto-ongeluk' een onbepaalde component aangezien het over één ongeluk gaat (datgene dat vanmorgen in mijn straat gebeurde). Over dat bepaalde verkeersongeval weet de geadresseerde niets. In (18) gaat het niet meer om één precies ongeluk. 'Een auto-ongeluk' verwijst naar alle auto-ongelukken en moet begrepen worden als welke auto-ongeluk dan ook. De naamwoordelijke constituent wordt in deze gevallen veralgemeend.

(18) Een auto-ongeluk veroorzaakt vertragingen van het verkeer.

Ook enkele typen concessieve en conditionele bijzinnen kunnen niet in het eerste stuk staan. Als ze voorop geplaatst worden, moeten ze vóór het eerste stuk staan, in de zogenaamde aanloop (cf. 1.3 punt (e)).

1.3 De aanloop

Andere elementen kunnen dus voor de eerste zinsplaats – in de aanloop of linkerperiferie (dubbele onderstreping) – en na het derde stuk – in de uitloop – voorkomen in de Nederlandse zin. Dergelijke items hebben een lossere band met de zin en worden dus niet

als zinsdelen beschouwd. Met andere woorden, de elementen die in de aanloop en in de uitloop staan vervullen geen grammaticale functie. Vanuit een syntactisch standpunt maken ze geen deel uit van de zin – ze behoren tot de periferie van de zin – en worden ze daarvan afgescheiden door middel van een pauze in gesproken taal en van een komma in geschreven taal). Het schema van de Nederlandse zin wordt nu:

Aanloop, - Eerste stuk – Eerste pool – Middenstuk – Tweede pool – Derde stuk, - Uitloop

(19a) Haar broer, die is bijzonder groot.

(19b) Hij is bijzonder groot, haar broer.

Aangezien deze scriptie gaat over elementen die vóór de persoonsvorm plaatsvinden, zal de uitloop buiten beschouwing gelaten worden. Welke rol of rollen speelt de aanloop en welke elementen kunnen daar optreden?

Volgens Vandeweghe (2013: 259-260) heeft de aanloop vier verschillende functies: een onderwerp hernemen of introduceren, een tekst structureren, bij de gesprekssituatie aanknopen en van expressiviteit dienen. Daarvoor zijn de volgende elementen in de linkerperiferie te vinden:

(a) Elementen met een anticiperende functie kunnen in de aanloop of in de uitloop staan. Deze constituenten lopen op een zinsdeel van de zin vooruit of achteruit. Ze worden gebruikt om bepaalde items vooraf te omschrijven en ze komen overeen met het onderwerp (19), het lijdend voorwerp (20), het meewerkend voorwerp (21), het voorzetselvoorwerp (22) en de bijwoordelijke bepaling (23). De geanticiperende elementen worden verplicht door verwijzwoorden in de concrete zin uitgedrukt.

(20) De brief, hij heeft die al gestuurd.

(20) Haar vrouw, die heeft Pieter bloemen gegeven.

(22) De wereldvrede, daarvan dromen we al lang.

(23) Vorig jaar, toen was hij bij ons gekomen.

(b) Tussenwerpsels (24-25) en vocatieven (26) bevinden zich meestal in de aanloop maar ze kunnen ook in de uitloop staan. Enkele interjecties fungeren ook als bijwoord en kunnen dus eveneens op de eerste zinsplaats of in het midden stuk voorkomen, bv. helaas, goddank, pats enz. De ANS (Haeseryn, 1997: 673) beschouwt ‘ja’ en ‘nee(n)’ als niet-noodzakelijk emotionele interjecties. Er wordt geen onderscheid gemaakt tussen ‘ja/nee’ als antwoord of reactie op een voorafgaande uiting (27a) en ‘ja/nee’ als teken dat iemand het gesprek volgt (27b). Beide typen komen in de aanloop voor, aldus de ANS.

(24a) Verdorie, wē zullen te laat aankomen!

(24b) We zullen te laat aankomen, verdorie!

(25a) Helaas, ik kan mijn sleutels niet meer terugvinden.

(25b) Helaas kan ik mijn sleutels niet meer terugvinden.

(25c) Ik kan helaas mijn sleutels niet meer terugvinden.

(25d) Ik kan mijn sleutels niet meer terugvinden, helaas.

(26a) Erik, ik heb je hulp nodig.

(26b) Ik heb je hulp nodig, Erik.

(27a) Heeft hij zijn huiswerk al gemaakt? Nee, dat heeft hij nog niet gemaakt.

(27b) Wat zal je volgend jaar doen? Ja, ik weet het nog niet.

(c) De constructies ‘wat X betreft’ en ‘wat X aangaat’ kunnen ofwel in de aanloop ofwel in het eerste stuk staan. Ze komen vaker op de eerste zinsplaats voor in de gevallen zonder verwijswaarden die naar de nominale constituent verwijzen (28a). Toch is het correct de twee constructies in de aanloop te bewaren (28b). Als het verwijswaard aanwezig is, worden ze in de aanloop geplaatst (28c).

(28a) Wat mij betreft mag hij zo meteen vertrekken.

(28b) Wat mij betreft, hij mag zo meteen vertrekken.

(28c) Wat mijn studie aangaat, die wordt binnenkort uitgegeven.

(d) Enkele voegwoordelijke bijwoorden bevinden zich in de aanloop, in het voorveld of in het middenstuk. De grammatica’s stemmen bijna totaal overeen met elkaar. Volgens Theissen (1984) kunnen ‘althans’, ‘daarenboven’, ‘desondanks’, ‘dus’, ‘evenwel’ en ‘nochtans’ zowel in de aanloop of op de eerste zinsplaats staan. Andere structuurwoorden staan bijna uitsluitend in de aanloop: ‘immers’, ‘integendeel’ en ‘trouwens’. Nog andere verschijnen

bijna exclusief in het eerste stuk: ‘bovendien’, ‘dus’, ‘intussen’, ‘niettemin’, ‘overigens’. Haast alle voegwoordelijke bijwoorden kunnen ook in het middenstuk geplaatst worden. In hun *Nederlandse grammatica voor Franstaligen* stellen Van Belle et al. (2006: 39) een tabel met 26 zinsverbindende bijwoorden en hun plaatsingsmogelijkheden voor. De voornaamste verschillen in vergelijking met de ANS betreffen ‘evenwel’, dat volgens Van Belle et al. niet in het voorveld kan staan, terwijl ‘overigens’ gemakkelijk in de aanloop kan, alsu Van Belle et al (2006 : 39).

Vandeweghe (2013: 271) zet ook dergelijke bijwoordelijke constituenten (vanaf nu BC) liever op de eerste zinsplaats of in het middenstuk dan in de aanloop. De linguïst zondert ‘echter’, ‘immers’, ‘integendeel’ en ‘trouwens’ af door hun onmogelijke plaatsing in het eerste stuk.

(e) Zes typen afhankelijke zinnen kunnen/moeten ook in de aanloop staan. Toegevende zinnen met ‘hoe/wat/wie ... ook’ (29) en (ook) ‘al’-bijzinnen (30) worden in de aanloop, in de uitloop en in het middenstuk geplaatst. Ze worden facultatief met het bijwoord ‘toch’ gebruikt. Toegevende zinnen zonder voegwoord komen uitsluitend in de aanloop voor. Ook die gaan niet verplicht gepaard met het verwijswoord ‘toch’ (31)². Veronderstellende (32) en voorwaardelijke zinnen zonder voegwoord (33) kunnen zich ook in de aanloop, in het middenstuk en in de uitloop bevinden en worden vaak door het verwijswoord ‘dan’ vergezeld. Het laatste type bijzinnen dat in de aanloop – moet – staan, zijn de zinnen die temporele contrasten laten ontstaan (34). Ze worden zonder verwijswoord gebruikt.

(29) Wie je ook bent, je mag hier niet roken.

(30) Ook al had ze niet gestudeerd, toch is ze voor haar test geslaagd.

(31) Was de reclame groot, het succes bleef maar klein.

(32) Mocht je toch nog willen komen, dan hoef je maar te bellen.

(33) Regent het morgen, dan gaan we niet wandelen.

(34) Waren er vroeger nog veel bossen, nu zijn het beduidend minder.

² Voorbeelden (31) en (34) worden uit de ANS (Haeseryn, 1997: 972, 930) opgenomen.

(f) Samenvattende uitdrukkingen zoals ‘kortom’, ‘in/met één woord’, ‘met andere woorden’, ‘nog eens’ en ‘om kort te gaan’ kunnen in de aanloop of in het middenstuk vallen (35). ‘Kort gezegd/samengevat’ is de enige uitzondering die ook op de eerste zinsplaats kan optreden. Dergelijke aanlopen worden niet door middel van een verwijswaord in de eigenlijke zin herhaald.

(35) Kortom, ze hebben bij de politie aangifte gedaan.

Er moet ook toegevoegd worden dat ‘in/met één woord’ en ‘met andere woorden’ ook in het eerste stuk kunnen staan. In die gevallen worden ze echter niet als samenvattende uitdrukkingen beschouwd. Ze verschuiven immers van een figuurlijke betekenis – kortom, om kort te gaan – naar een letterlijke betekenis (36-37).

(36a) In één woord, ik kan het niet zeggen = kortom, ik kan het niet zeggen.

(36b) In één woord kan ik het niet zeggen. [Ik heb meer woorden nodig.]

(37a) Met andere woorden, ik kan het niet zeggen = kortom, ik kan het niet zeggen.

(37b) Met andere woorden kan ik het niet zeggen. [Ik kan alleen maar die woorden gebruiken.]

Soms kunnen er verschillende elementen in de linkerperiferie voorkomen. Aansprekingen worden immers vaak vergezeld met groeten of ‘ja/nee’ gezegden (38-39). Interjuncties kunnen ook andere tussenwerpsels (40) of anticiperende elementen begeleiden (41).

(38) Dag Lise, je ziet er sprankelend uit vandaag.

(39) Ja dokter, ik zal stoppen met roken.

(40) O nee, ik heb opnieuw mijn portemonnee vergeten.

(41) Op die dag, hè, toen zijn we eindelijk vertrokken.

In tegenstelling tot voegwoordelijke bijwoorden maken nevenschikkende voegwoorden – vanaf nu ook NSVW genoemd – geen deel uit van de aanloop of van de zin, volgens de ANS (1997: 969). Terwijl de VWBW een grammaticale functie binnen de syntactische zin vervullen, oefenen de NSVW alleen maar een tussen-zinsfunctie uit en vallen ze buiten het zinschema. Ze hebben dus geen invloed op de woordvolgorde van de zin. Nevenschikkende voegwoorden

behoren ook niet tot de aanloop omdat ze normaal gezien geen schriftelijke of intonatieve onderbreking veroorzaken (Vandeweghe, 2013: 269). Toch voegt de linguïst toe dat NSWV zich wel in de aanloop kunnen bevinden als ze met een komma-intonatie gepaard gaan, in tegenstelling tot wat de ANS beweert.

Nevenschikkende voegwoorden worden gebruikt om twee of meer zinnen te verbinden en staan hierdoor ertussen (42). Er bestaan vier typen nevenschikkende voegwoorden: ze kunnen aaneenschakelend ('en', 'noch', 'alsmede', 'alsook'), tegenstellend ('maar', 'doch', 'of', 'ofwel', 'dan wel', 'dan'), redengevend ('want') of gevolgaanduidend ('dus') zijn (Haeseryn, 1997: 644). Ze worden in hoofdstuk 3 nader onder de loep genomen.

(42) Ik heb dat meisje al ontmoet maar deze kerel, die heb ik nooit gezien.

1.4 Syntactische distributie en informatiestructuur

Nu de elementen die in het eerste stuk en in de aanloop kunnen/moeten staan de revue zijn gepasseerd, is het tijd om te onderzoeken in welke situaties bepaalde elementen vóór de eerste pool staan en waarom ze daar in die gevallen plaatsvinden. Met andere woorden, dit hoofdstuk houdt zich bezig met de factoren die de woordvolgorde in het Nederlands regelen. Deze studie is immers niet alleen geïnteresseerd in wat in de aanloop en in het voorveld voorkomt, ze streeft eveneens ernaar het verband tussen syntaxis en semantiek vast te stellen en uit te leggen. Dit theoretische overzicht wordt ook op basis van naslagwerken gebaseerd voordat de literatuur getoetst zal worden door middel van een corpusonderzoek.

Het verschil in betekenis tussen (37a) en (37b) uit hoofdstuk 1.3. is bijzonder sprekend. Naargelang van haar plaatsing in de zin, ofwel in de aanloop, ofwel in het eerste stuk, kan de uitdrukking 'met andere woorden' figuurlijk of letterlijk begrepen worden. Als ze in de franje van de zin staat, komt 'met andere woorden' overeen met samenvattende uitdrukkingen zoals 'kortom' of 'in het kort'. Aan de andere kant betekent het woord voor woord 'door andere woorden te gebruiken' als het op de eerste zinsplaats staat.

(37a) Met andere woorden, ik kan het niet zeggen.

(37b) Met andere woorden kan ik het niet zeggen.

Een cruciaal gegeven dat nooit vergeten mag worden als het gaat om communicatie is het belang van de context. Zinnen moeten met betrekking tot hun omgeving geanalyseerd worden. De zender past immers zijn/haar uiting aan op basis van wat al gezegd is en van zijn/haar hypothese over wat de hoorder weet.

Schiffrin (1988) maakt een onderscheid tussen zogenoemde kennis en meta-kennis. Spreker/hoorder-kennis betreft wat een spreker weet en wat een hoorder weet; spreker/hoorder-meta-kennis betreft wat spreker en hoorder weten over hun eigen kennis en wat zij weten (of veronderstellen te weten) over de kennis die ze met elkaar delen. Omdat discourse betrekking heeft op het uitwisselen van informatie, zijn kennis en meta-kennis constant in beweging. Vandaar dat het nuttig is als sprekers de moeite nemen om bepaalde veranderingen in deze interactief ontwikkelende informatiestatus te signaleren. (Evers-Vermeul, 2010: 152)

In het Nederlands wordt het voornaamste principe dat het meedelen van informatie met de woordvolgorde verbindt het links-rechts-principe genoemd. Dit houdt in dat communicatief gezien minder belangrijke elementen vóór belangrijkere elementen staan. Met andere woorden, items die de informatieve waarde dragen hebben de neiging om zich achteraan in de zin te bevinden. In de gesproken taal worden ze bovendien benadrukt doordat deze elementen het zinsaccent krijgen (Haeseryn, 1997: 913).

(43) Gisteren heb ik **Jan** gezien.

Ten opzichte van het links-rechts-principe is ‘Jan’ in (43) het element met de meest communicatieve waarde, i.e. de focus. Het wordt immers beklemtoond (vet gemaakt) en het staat achteraan in de zin – vlak vóór de tweede pool. ‘Gisteren’ wordt hier topic genoemd, i.e. “datgene waarover gesproken wordt in de zin of datgene waarover de taaluiting in de eerste plaats gaat” (Coussé, 2008: 16). Hier kan er verondersteld worden dat het gesprek gaat over wat mensen gisteren hebben gedaan. Onbewust plaatst de spreker ‘gisteren’ aan het begin van zijn/haar zin aangezien dit het topic is – daarover gaat het – en ‘Jan’ aan het einde van de uiting. Wat belangrijk is met betrekking tot ‘gisteren’ is dat de spreker ‘Jan’ heeft gezien. Daarom is ‘Jan’ in (43) de focus.

Toch is het links-rechts-principe niet het enige principe dat de woordvolgorde van de Nederlandse zin beïnvloedt. Belangrijke constituenten kunnen immers ook vooraan in de zin geplaatst worden. Volgens Coussé (2008: 17) is het eerste stuk de zinsplaats bij uitsteking om een

zinsdeel voorop te plaatsen. Daardoor wordt het element op de eerste zinsplaats benadrukt en kan de hoorder er extra aandacht aan besteden (44)

- (44a) Ik heb het boek **AAN JAN** gegeven.
- (44b) **AAN JAN** heb ik het boek gegeven.
- (44c) **HET BOEK** heb ik aan Jan gegeven.
- (44d) **IK** heb het boek aan Jan gegeven.

De woordvolgorde in (44a) houdt zich aan de regel van het links-rechts-principe. ‘Ik’ is de topic en ‘aan Jan’ is de focus gezien het zinsaccent en de plaatsing aan het einde van de zin. In zinnen (44b-d) worden de gefocuste elementen prominenter door vooropplaatsing in het voorveld. Verder bewaren ze het zinsaccent. Elke zin kan met een bepaalde vraag overeenkomen, respectievelijk “aan wie heb je het boek gegeven?” (44b), “Wat heb je aan Jan gegeven?” (44c) en “Wie heeft het boek aan Jan gegeven?” (44d).

De vooropplaatsing is hier niet verplicht. Het zinsaccent is immers een voldoende factor om een zinsdeel te benadrukken en om er een focus van te maken. De combinatie van fonologische en syntactische factoren versterkt nochtans de nadruk op de focus, wat de boodschap duidelijker maakt.

Het is ook cruciaal een onderscheid te maken tussen, aan de ene kant, topic en bekende informatie en, aan de andere kant, focus en nieuwe informatie. Volgens Herring (1990: 164) bestaan er – ten minste – twee typen topics en twee typen focus³.

De zogenaamde *continuous topic* bewaart zijn status van de ene zin naar de volgende. Vaak wordt het door een voornaamwoord vervangen of is het zelfs gewoon afwezig. *Shifted topics* integendeel zijn onderbroken. Meestal zijn dit lexicale nominale constituenten maar ze kunnen ook ingeleid worden door uitdrukkingen zoals ‘qua X’, ‘wat X betreft’ of ‘X, aan de andere kant’ enz.

³ Deze communicatieve-functionele aanpak van grammatica is door de Praagse School ontwikkeld. De werken van Danes of Jakobson o.a. hebben een enorme invloed gehad op de taalkunde sinds de eerste helft van de 20^{ste} eeuw. Ook al wordt de Praagse School in deze scriptie om redenen van tijd en ruimte niet rechtstreeks vermeld, het is cruciaal hun doorslaggevende rol te erkennen.

Wat de focus aangaat onderscheidt Herring de *contrastive focus* van de *presentational focus* (1990 : 164). *Presentational focus* gaat om het introduceren van een referent binnen een meestal verhalende uiting, bijvoorbeeld “er was eens een machtige tovenaar.” *Contrastive focus* daarentegen benadrukt een specifiek zinsdeel in tegenstelling tot een vergelijkbaar item uit een vorige uiting. Een voorbeeld daarvan is (44b).

(44b) **AAN JAN** heb ik het boek gegeven. [Niet aan Iris.]

Herring leidt eruit af dat *continuous topics* wel met bekende informatie gepaard gaan, terwijl *shifted topics* nieuwe of tenminste opnieuw geactiveerde informatie bevatten. De informatie is ook nieuw bij *presentational focus* enerzijds en bekend (of tenminste gesuggereerd) bij *contrastive focus* anderzijds.

In hun werk *Syntax. Structure, meaning and function* onderscheiden Van Valin en LaPolla (1997) eveneens twee typen focus: de nauwe focus en de brede focus. In de nauwe focus constructie bestaat de focus uit één constituent, zoals in (45)⁴. Uit dit voorbeeld blijkt dat de nauwe focus met Herrings *contrastive focus* vergeleken kan worden: een specifiek zinsdeel – mijn wagen – wordt benadrukt in tegenstelling tot een vergelijkbaar item uit een vorige uiting – je fiets. De functie van de nauwe focus is een referent te bepalen, aldus Van Valin en LaPolla.

(45) Ik heb gehoord dat je fiets kapot was.

Mijn **WAGEN** is kapot.

Betekent dat dat de brede focus per se met de *presentational focus* zal overeenstemmen? Van Valin en LaPolla verdelen de brede focus weer in twee: de zin focus en de gezegde focus. De zin focus bevat de hele zin (46) – bijgevolg bestaat er geen topic – terwijl de gezegde focus (47) uit de hele zin behalve de topic bestaat en een commentaar betreffende deze topic uitdrukt. Net zo goed als de *presentational focus* wordt de zin focus gebruikt om een nieuwe referent binnen het gesprek in te voeren of om een gebeurtenis met een nieuwe referent bekend te maken (Van Valin & LaPolla, 1997: 206). De functie van de gezegde focus is nog anders. Het dient om commentaar te leveren op de topic of om een gebeurtenis te rapporteren.

⁴ Voorbeelden (45), (46) en (47) worden uit Van Valin en LaPolla (1997: 206-208) opgenomen en vertaald.

(46) Wat is er gebeurd?

MIJN WAGEN IS KAPOT.

(47) Wat is er met je wagen gebeurd?

Mijn wagen **IS KAPOT.**

Uit de werken van Herring en Van Valin & LaPolla blijkt dat bijna alle elementen uit de Nederlandse zin focus kunnen worden. Dit wordt door Ilse Magnus bevestigd (2008). In haar werk over de plaatsing van bijwoordelijke bepalingen in het Frans en het Nederlands beweert de linguïste dat al de elementen van de Nederlandse zin focus kunnen worden – zowel de zinsplaatsen die de focus kunnen bevatten als de elementen die zelf gefocust kunnen worden – met uitzondering van de aanloop en de uitloop. Zoals hierboven vermeld, gaat de pragmatische zinsopbouw gepaard met fonologische factoren, namelijk het zinsaccent. Dit geeft aan dat de spreker voor een zeker element heeft gekozen en dat dit element de kern van zijn boodschap vormt. Er moet ook toegevoegd worden dat een zin meerdere zinsaccenten kan bevatten. Magnus bewijst dat al de constituenten van de eigenlijke Nederlandse zin gefocust kunnen worden door te tonen dat de focus minder met syntaxis dan met zinsaccent heeft te maken. In de volgende zinnen blijft de woordvolgorde immers dezelfde, ook al schuift het accent op.

(48) Ik gaf Jan dat boek **GISTEREN.**

(49) Ik gaf Jan dat **BOEK** gisteren.

(50) Ik gaf Jan **DAT** boek gisteren.

(51) Ik gaf **JAN** dat boek gisteren.

(52) **IK** gaf Jan dat boek gisteren.

Verder stelt de auteur vast welke bijwoordelijke bepalingen in de aanloop kunnen staan, met name *situational* bijwoordelijke bepaling en *comments on the utterance act*. De eerste situeren de stand van zaken vanuit een temporeel of een plaatselijk standpunt of leggen een verband met een andere stand van zaken (oorzaak, doel, voorwaarde enz.). In dit geval behoren de bijwoordelijke bepalingen in de aanloop tot de elementen met een anticiperende functie die in de concrete zin door een verwijzwoord uitgedrukt worden. Van Kampen (2010: 192) legt uit dat dergelijke vooruitlopende constructies een topicverschuiving aanduiden. Uit zin (53) blijkt immers dat ‘vorig jaar’ in de voorafgaande zinnen geen topic was in tegenstelling tot in de voorbeeldzin. De *comments on the utterance act* worden gebruikt om de relaties tussen

gesprekspartners te verklaren, om zich te verontschuldigen voor het uiten van ideeën of gedachten en om zijn uiting te rechtvaardigen (54)⁵. In tegenstelling tot *situational* bijwoordelijke bepalingen stemmen de *comments on the utterance* act niet overeen met een precies type constituenten die in de aanloop kunnen staan.

(53) Vorig jaar, toen was hij bij ons gekomen.

(54) Eerlijk gezegd, Jan heeft die vraag niet beantwoord.

Volgens Kuiper et al. (2017) wordt de woordvolgorde echter niet alleen door semantische factoren beïnvloed. In hun artikel over de plaatsing van eigennaamagenten – of ENA’s – in sportverslagen beweren de auteurs dat commentatoren ook onder invloed van temporele druk hun zinsstructuur aanpasten. Ze baseren hun onderzoek op transcripties van live voetbal- en rensportcommentaar in verschillende talen – Engels en Perzisch voor voetbal en Engels en Japans wat de paardensport betreft. Daarnaast hebben de linguïsten ook data uit Engels onlinegeschrevenvoetbalcommentaar verzameld. De druk komt uit het feit dat eigennamen moeilijker terug te vinden zijn uit het geheugen dan soortnamen. In voetbal en in rensport moeten de verslaggevers immers vooral eigennamen meedelen (spelers en paarden) aangezien die meestal met nieuwe informatie overeenkomen. Verder verliezen de journalisten extra tijd om eerst de betrokken spelers of paarden te herkennen.

Uit hun onderzoek blijkt dat al de geanalyseerde talen de neiging hebben om eigennaamagenten naar de rechterkant van de linker periferie te verschuiven. Daardoor krijgen ze meer tijd om de agenten te bepalen en zich hun namen te herinneren. Commentatoren kunnen een beroep doen op diverse technieken om ENA’s uit te stellen, zoals passieve constructies (55), met een nevenschikkend voegwoord beginnen (56) of met een bijwoordelijke bepaling beginnen (57) (Kuiper et al., 2017: 11)

(55) Played off by Almeida.

Gespeeld door Almeida.

(56) And now Xavi.

En nu Xavi.

⁵ Voorbeeld (54) wordt uit Magnus (2008: 115) opgenomen.

(57) From the block Almeida unable to break.

Uit het blok Almeida niet in staat om door te breken.

Kuiper et al. leggen uit dat een dergelijke verschuiving van de ENA's niet verklaard kan worden vanuit een communicatief standpunt. Ze lichten bijvoorbeeld hun stelling toe aan de hand van de theorie volgens welke bekende informatie vóór nieuwe informatie komt. De resultaten van hun onderzoek tonen aan dat de eigennaamagenten – die als nieuwe informatie beschouwd worden – in 70% van de gevallen na een andere constituent voorkomen, in overeenstemming met de theorie. Hoe kunnen de andere 30% gevallen dan gerechtvaardigd worden? Dezelfde redenering geldt voor de nevenschikkende voegwoorden. Er bestaat immers geen relevant verschil tussen de gevallen waar ze aanwezig en afwezig zijn. Verstandig sluiten de auteurs niettemin af dat hun standpunt niet per se lijnrecht tegenover een communicatieve functionele opvatting van grammatica staat, in het bijzonder het verband tussen syntaxis en semantiek ten gunste van de communicatie.

1.5 State of the art: Linkerperiferie en eerste stuk van de Nederlandse zin

Ook al hebben veel linguïsten hun aandacht op de Nederlandse syntaxis gericht, het verband en de interacties tussen het eerste stuk en de linkerperiferie zijn zelden onderzocht. Jacqueline Evers-Vermeul roerde de aanloop en zijn invloed op de zinsbetekenis aan. Ze vergelijkt de twee volgende zinnen (2010: 172-173):

(58) De gemiddelden zijn verschillend. De t-toets toont immers een significant verschil aan.

(59) De gemiddelden zijn verschillend. Immers, de t-toets toont een significant verschil aan.

Terwijl immers [in 58] fungeert als accessibility marker en dus aangeeft dat informatie in de zin als toegankelijk voor de lezer wordt beschouwd, is immers hier [in 59] niet midden in de zin te vinden. De schrijver heeft ervoor gekozen om immers voorop te plaatsen, met als effect dat het argument dat volgt op immers benadrukt wordt. Deze nadruk zou ontbroken hebben bij plaatsing van immers midden in de zin. (Evers-Vermeul, 2010: 173)

De auteur legt uit dat naargelang waar het bijwoord ‘immers’ in de zin staat, zijn betekenis enigszins verandert. In (58) wijst ze aan dat ‘immers’ als een *accessibility marker* fungeert. Deze *accessibility marker* houdt in dat de schrijver/spreker verwacht dat de informatie al bekend is door de lezer/hoorder. Een mogelijke parafrase van (58) zou bijvoorbeeld zijn:

(60) De gemiddelden zijn verschillend, zoals afgeleid kan worden uit de t-toets.

Door het bijwoord in de aanloop te plaatsen, heeft de zender tot doel wat hij/ze net gezegd heeft te benadrukken en te toe te lichten. Toch houdt ‘immers’ zijn *accessibility marker* functie. Er kan dus afgeleid worden dat het plaatsen van ‘immers’ in de aanloop een extra functie geeft aan het bijwoord, namelijk de nadruk te leggen op het argument in het middenstuk.

Zoals voorgesteld in een functionele beschouwing van grammatica zijn semantiek, syntaxis en fonologie bijgevolg nauw verbonden. Een ander voorbeeld waarin vorm en betekenis samenhangen is de passief-constructie. Kuiper et al. (2017: 1) beweren dat:

passive constructions demote the agent argument and promote the theme argument because the speaker wishes to place the theme argument into focus or topic position (Payne, 1997, p. 204). Because the information structure of sentences often serves the purpose of providing given information earlier than new information (Halliday, 1985, p. 278), the topic of a proposition comes before a comment. Hence the functional possibilities afforded by the passive construction.

A functional theory of language can be formulated strongly as follows: all the options made available by the lexis, morphology and syntax of a language serve communicative ends, i.e. are semantically or pragmatically motivated

Met andere woorden, de agensconstructie (van het Latijnse werkwoord *agere*, doen) die vaak in actieve zinnen te vinden is, wordt in passieve zinnen niet meer gebruikt. In actieve zinnen voert het thema de handeling uit in plaats van die te ondergaan terwijl passieve zinnen niet op het agens – dat de handeling uitvoert – maar op de handeling zelf focussen. De agentia zijn immers frequent afwezig in passieve zinnen, bijvoorbeeld als die onbekend zijn, of als ze integendeel door de context expliciet zijn (61):

(61) De dief is vanmorgen aangehouden [door de politie].

Informatiestructuur, i.e. hoe informatie vormelijk verdeeld wordt in een zin, en woordvolgorde zijn in verscheidene publicaties aangepakt. Lambrecht (1996: 3) beweert dat:

Even though information structure is concerned with such psychological phenomena as the speaker's hypotheses about the hearer's mental states, such phenomena are relevant to the linguist only inasmuch as they are reflected in GRAMMATICAL STRUCTURE (morphosyntax, prosody). The importance of this caveat cannot be overemphasized. I take information structure to be a component of GRAMMAR, more specifically of SENTENCE GRAMMAR, i.e. I take it to be a determining factor in the formal structuring of sentences. Information structure is NOT concerned with psychological phenomena which do not have correlates in grammatical form.

Hiermee wordt bedoeld dat informatiestructuur niet alleen betrekking heeft tot het abstracte verband tussen de informatie die een spreker wil uiten en zijn inschatting van de wijze waarop de hoorder in staat is om wat net gezegd is tegenover een bepaalde context te verwerken. Dit verband moet ook een manifeste invloed uitoefenen op de woordvolgorde en de intonatiepatroon van de zin.

De nauwe band tussen syntaxis en betekenis zal de voornaamste benadering zijn waarvan het functioneren van de aanloop en het eerste stuk onderzocht zal worden. Deze scriptie zal de volgende vragen trachten te beantwoorden: worden de voorschriften betreffende de aanloop en de eerste zinsplaats in het Nederlands nageleefd? Wordt de PV2-regel toegepast? Hoe nauw zijn vorm en semantiek verbonden ten opzichte van de aanloop en het eerste stuk? Welk effect hebben hun aan- of afwezigheid op de betekenis van de zin?

2. Methode

Deze vormelijk-functionele analyse van het pré-werkwoordelijke deel van de mededelende Nederlandse zin baseert zich ten eerste op Nederlandse grammatica's, in het bijzonder de *Algemene Nederlandse Spraakkunst* (1997). Vandeweghes *Grammatica van de Nederlandse zin* (2013), Verbancks *Beknopte Nederlandse grammatica* (1985) en Van Belles *Het project Nederlandse grammatica voor Franstaligen* (2006) zijn ook geraadpleegd om mogelijke leemtes op te vullen en om de bronnen te vergelijken. Vervolgens werd er data met betrekking tot informatiestructuur en woordvolgorde verzameld uit wetenschappelijke artikelen en monografieën.

De concrete analyse van de aanloop en van het voorveld wordt door middel van corpusonderzoek uitgevoerd. Het eerste gebruikte corpus is het CGN (*Corpus Gesproken Nederlands*). Dit bevat c. 9.000.000 woorden en werd tussen 1998 en 2004 samengesteld. Het verzamelwerk bestaat uit opnames en transcripties van dagelijkse gesprekken (spontane dialogen, interviews, conferenties enzovoort). Het tweede corpus is samengesteld uit in het *NRC Handelsblad* uitgegeven nieuws artikelen. Het Handelsblad is een liberaal Nederlandse krant die uit de fusie tussen het *Algemeen Handelsblad* en de *Nieuwe Rotterdamse Courant* (NRC) geboren is. Aangezien deze scriptie de aanloop en het eerste stuk voornamelijk in mondeling taalgebruik onderzoekt, is het cruciaal om gesproken en geschreven materiaal met elkaar te vergelijken om de bevindingen in de juiste verhoudingen te zetten. In elk corpus werden er 200 relevante hits bewaard en geanalyseerd om de vergelijking te vereenvoudigen. De corpora zijn in Appendix A terug te vinden.

De corpora worden in verschillende stappen gebruikt: Eerst wordt de data verdeeld in zinnen, die individueel worden aangepakt. Voor elke zin worden de volgende vragen gesteld: Is de zin relevant met betrekking tot het onderzoek? (Is dit een zin? Is ze mededelend?) Indien ja, bestaat er een aanloop? Waaruit bestaat die? (Interjectie, tussenwerpsel, bijzin enz.?) Welke grammaticale functie oefent die uit? Bestaat er een eerste stuk? Waaruit bestaat dit? (Bijwoord, nominale constituent enz.?) Welke grammaticale functie bekleedt het?

Er zij op gewezen dat het corpusonderzoek met de hand gedaan wordt. Met andere woorden, er wordt geen beroep gedaan op software om de aanwezigheid en/of de syntactische functie van de aanloop en het eerste stuk te bepalen.

Zinnen die niet mededelend zijn, werden buiten beschouwing gelaten. Dit geldt ook voor vragende zinnen die dezelfde woordvolgorde als mededelende zinnen hebben, bv. “Jullie zijn naar Italië gegaan?”. Aan de hand van het vraagteken waren die zinnen makkelijk vast te stellen en uit te sluiten

Zinnen zonder persoonsvormen van het type “Aanval op de hoofdstad van Burundi” zijn ook weggelaten. Dergelijke zinnen waren te vinden in het *NRC*-corpus. Ze stemden immers met titels van krantenartikelen overeen, aangezien die soms geen werkwoord bevatten. Er werd beslist om zinnen zonder pv niet te behandelen omdat de twee polen van de zin de woordvolgorde in het Nederlands bepalen. Mag de verbale rest leeg blijven, er wordt in deze scriptie gemeend dat persoonsvormen verplicht waren wat mededelende zinnen betreft.

Om de eerste stukken en aanlopen te identificeren en te analyseren, werd er ook beslist de zinnen te segmenteren als volgt: Na leestekens – eindpunten, vraagtekens, uitroepetekens, gedachtepuntjes, dubbelpunten, kommapunten, komma’s – en na nevenschikkende voegwoorden die **nevengeschikte zinnen** verbinden. Dit betekent dat de gekoppelde items persoonsvormen moeten bezitten. De persoonsvorm van het tweede onderdeel kan ook elliptisch gebruikt worden (63). Anders kunnen de zinnen geen eerste stukken bevatten. Daarom worden nevenschikkingen zoals (64) vanzelfsprekend buiten beschouwing gelaten, in tegenstelling tot (63). Verder wordt er geen rekening gehouden met ondergeschikte zinnen. Hoewel afhankelijke zinnen wel een eerste pool inhouden, staan er geen eerste stukken vóór de voegwoorden die als eerste polen fungeren (65).

(62) Ik probeer haar op te bellen maar ze antwoordt niet.

(63) Ik hou van aardbeien maar zij [houdt] niet [daarvan/van aardbeien].

(64) Ik hou niet van kaas en \emptyset van vis.

(65) Ze wist niet \emptyset dat haar vader heeft moeten blijven zitten.

Bijvoorbeeld, een zin zoals “Wat in het Verdrag van Maastricht al werd voorzien, lijkt te worden bewaarheid: de monetaire grenzen mogen krap zijn getrokken, het laatste woord heeft de politiek en die neigt tot een zekere rekkelijkheid.” zal in de vier volgende segmenten worden verdeeld:

“Wat in het Verdrag van Maastricht al werd voorzien, lijkt te worden bewaarheid:”

“de monetaire grenzen mogen krap zijn getrokken,”

“het laatste woord heeft de politiek”

“en die neigt tot een zekere rekkelijkheid.”

Zoals eerder gezegd (hoofdstuk 1.3. (b)) beschouwt de ANS ‘ja’ en ‘nee’ als elementen die altijd in de aanloop staan, ongeacht hun functie – antwoord op een voorafgaande vraag, reactie op een voorafgaande zin en teken dat iemand een gesprek volgt. Na lang aarzelen werd er besloten om de theorie van de ANS in het corpusonderzoek te bewaren en toe te passen. Daardoor wordt de analyse vergemakkelijkt aangezien het complex kan zijn om de precieze functie van ‘ja/nee’ vast te stellen. Door al de voorkomens van ‘ja’ en ‘nee’ bij de aanloop te rekenen worden de risico’s op inconsistentie zoveel mogelijk beperkt.

Om de corpusonderzoeksmethode duidelijk te maken, worden de verschillende variabelen hierbeneden opgesomd. Alle worden van een uitleg en/of een voorbeeld voorzien. Meestal komen de voorbeelden uit de onderzochte corpora, maar ze kunnen ook uit naslagwerken komen of gewoon voor de gelegenheid verzonnen zijn. Al de mogelijkheden als reactie op de diverse variabelen zijn alfabetisch geordend. In sommige gevallen (bijwoordelijke bepalingen bv.) wordt de aard de constituent nog nader verduidelijkt. Een dergelijk onderscheid is immers relevant als de resultaten bekeken worden.

Variabele 1: Is de zin relevant voor het onderzoek? De twee enige mogelijkheden zijn **ja** of **nee**. Zoals eerder gezegd, moet de zin ten eerste mededelend zijn om als relevant beschouwd te worden. Ten tweede moeten ze ook een persoonsvorm bevatten (ook elliptisch). Ten laatste zijn de ondergeschikte segmenten niet relevant en werden dus buiten beschouwing gelaten.

Ja: “Liever was hij met het team van Kalev gekomen”. (NRC 134)

Nee: “Wanneer zijn ze getrouwd dan?” (CGN 33)

“[Dat water vindt hij te zwart.] En vooral te koud”. (NRC 154)

“Omdat ’t hartstikke spannend is en de dingen die nog gaan komen en zo”. (CGN 157)

Variabele 2: Wat is de vorm van het element in de aanloop? Op basis van de grammatica’s zijn er negen verschillende categorieën vastgesteld:

Adjectief:	“ <u>Sterker</u> , <u>de Duitse socialisten</u> <u>hebben</u> slecht een kans als [...]” (NRC 82)
BC:	“ <u>Trouwens</u> , <u>dat heb ik gehoord</u> ”. (CGN 78)
Bijzin:	“ <u>Was vorig jaar een tijd lang alle aandacht gericht op de top in Amsterdam</u> , [...] <u>nu gaat</u> het er om te <u>bepalen</u> welke lidstaten [...]”. (NRC 55)
NC:	“ <u>M’n ouders</u> , <u>die zijn</u> uh ...” (CGN 20)
Leeg:	“ <u>∅</u> <u>Je kan</u> dat eigenlijk zien als een vreemde valuta” (CGN 287)
Tussenwerpsel:	“ <u>Nou</u> , <u>daar weet</u> ik zelf niet veel meer van”. (CGN 62)
Vocatief:	“ <u>Erik</u> , <u>ik heb</u> je hulp nodig”.
Voorzetselconstituent:	“ <u>Met andere woorden</u> , <u>ik kan</u> het niet <u>zeggen</u> ”

Variabele 3: Wat is/zijn de syntactische functie(s) van het element in de aanloop? Hier wordt het neteliger aangezien het meer om subjectieve interpretatie gaat dan bij het vormelijke aspect. Bovendien kan éénzelfde aanloopelement twee of meer functies vervullen. Toch zijn er vier verschillende functies geselecteerd. Naast de klassieke en gemakkelijk te bepalen functie bijwoordelijke bepaling (van voorwaarde, van toegeving, van vergelijking enz.) bestaan er drie functies die wat abstracter zijn: elementen met een structurerende of sequentiële, een interactieve of interpersoonlijke en een retorische functie. Dergelijke functies worden voornamelijk doorkleine woorden zoals tussenwerpsels en bijwoorden vervuld.

In haar artikel *Towards an operational category of discourse markers: A definition and its model* (2017) onderscheidt Crible vier verschillende functies die discourse markers (DM’s) kunnen vervullen. Onder de term discourse marker neemt de auteur elementen zoals connectieven of *hedges* (*kind of, sort of, if you like ...*) maar wijst tussenwerpsels (behalve ‘yeah’ en ‘oh’) en responssignalen af. Toch zijn connectieven, interjecties en responssignalen

verwant aangezien ze allemaal deel uitmaken van de categorie ‘pragmatische markers’, aldus Crible (2017 : 108). Volgens haar functioneren DM’s en andere typen pragmatische markers op gelijke wijze. De bijzonderheid van DM’s is hun multifunctionaliteit. Die berust op drie principes: de discoursemarkerscategorie bevat elementen die veel verscheidene functies kunnen bekleden. Eénzelfde element kan diverse functies in diverse contexten vervullen. Eénzelfde element kan verschillende functies tegelijkertijd in dezelfde context vervullen. Dat betekent dat éénzelfde item in de aanloop gelijktijdig twee of meer functies kan uitoefenen.

Crible maakt een onderscheid tussen vier functionele domeinen voor discourse markers (2017: 107): ideationeel, retorisch, sequentieel en interactief. Het eerste domein betreft discourseverhoudingen tussen reële gebeurtenissen. Daartoe behoren de hierboven vermelde bijwoordelijke bepalingen met een makkelijk te bepalen functie: bepaling van toegeving, van voorwaarde, van tijd en van vergelijking bv. De auteur beschrijft het retorische domein als “discourse relations between epistemic and speech-act events, and metadiscursive functions”. Met andere woorden, ze verbinden de cognitieve wereld met de vorm van het gesprek. Een voorbeeld daarvan zijn conclusies. Verder kunnen ze ook een metadiscursieve functie uitoefenen, d.w.z. dat DM’s met een dergelijke functie iets over het gesprek zelf vertellen (bv. benadrukken of benaderingen). Sequentiële markers structureren gespreksitems vanuit een micro- en macrostandpunt. Ze kunnen een topicverschuiving of een *turn-taking* (gesprekken waar twee sprekers afwisselen en op mekaars uitingen reageren) aangeven. Het laatste domein is betrokken met het beheer van de interactie tussen de zender en de ontvanger van de boodschap. Deze categorie wordt ook interpersoneel genoemd.

Om wat concreter te zijn, elementen uit deze corpusanalyse die van het retorische domein deel uitmaakten, waren vaak tussenwerpsels die een aarzeling aangeven. Ze worden ook gebruikt om te tonen dat een bepaalde spreker het woord neemt of gaan een conclusie vooraf. Sequentiële of structurerende markers zijn vooral bijwoorden die clauses, zinnen en alinea’s met elkaar verbinden. Zulke elementen kunnen zich ook vaak op de eerste zinsplaats of in het middenstuk bevinden. Ze stemmen overeen met de in sectie 1.3. (c) vermelde items die in de aanloop kunnen staan. Er werd ook gekozen om de elementen met een anticiperende functie bij het sequentiële domein Onder te brengen. Het was immers noodzakelijk om zulke elementen met één van de vier categorieën te verbinden omdat anticiperende constituenten zich niet op hetzelfde grammaticale niveau bevinden als de andere domeinen. Verder worden ze gebruikt om op bepaalde elementen te focussen door hen een extrapositie te geven. Ze regelen

dus in zekere zin de verbanden tussen zinnen of tussen alinea's onderling. Ten laatste, interactieve pragmatische markers dienen te bewijzen dat het gesprek interactief blijft door een vraag te beantwoorden of door op een uiting te reageren. Hieronder staat een exhaustieve lijst van de diverse grammaticale functies die door aanloopelementen uitgeoefend kunnen worden:

- Bijwoordelijke bepaling: “Waren er vroeger nog veel bossen, nu zijn het er beduidend minder”. (temporele tegenstelling, ANS)
“Was vorig jaar een tijd lang alle aandacht gericht op de top in Amsterdam, waar het Verdrag van Maastricht moest worden aangepast, nu gaat het er om [...]” (toegeving, NRC 55)
“Zoals het ons opgedragen was, zo voerden we het werk ook uit”. (vergelijking, uit de ANS)
“Regent het morgen, dan gaan we niet wandelen”. (voorwaarde)
- Interactieve constituent: “[Grote ramp]. Nee ze waren hartstikke blij met mij”. (CGN 37)
- Retorische constituent: “Nou daar weet ik zelf niet veel meer van”. (CGN 62)
- Retorisch en interactief: “[Binnenkort krijgen we euro's volgend jaar.] Ja we komen uit de tijd van de guldens”. (CGN)
- Structurerende constituent: “Trouwens, dat heb ik gehoord”. (CGN 78)

Dit laatste voorbeeld toont aan dat het bijzonder complex is om de nauwkeurige functie(s) van pragmatische markers te bepalen. Vanuit het ene of andere standpunt kan hetzelfde segment van het interactieve of van het retorische domein deel uitmaken, bijvoorbeeld. Dit corpusonderzoek streeft ernaar zo objectief mogelijk te zijn maar taalwetenschap is geen exacte wetenschap, wat de kans geeft om verschillende interpretaties te bespreken.

Variabele 4: Wat is de vorm van het element in het eerste stuk? In sectie 1.2. is er vermeld dat items uit alle grammaticale functies op de eerste zinsplaats kunnen staan. Een rede daarvan is dat het voorveld een uitgelezen plek is om een zinsdeel te benadrukken en een focus van maken. Bovendien stellen Van Valin & LaPolla (1997 : 206) dat bijna alle elementen gefocust kunnen worden in de Nederlandse taal. Toch bestaan er beperkingen en kunnen bepaalde constituenten zich niet op de eerste zinsplaats bevinden, zoals reciproke en reflexieve (zonder het suffix -zelf) voornaamwoorden, het niet-onderwerp onzijdige voornaamwoord

“het”, onbepaalde constituenten en een aantal bijzinnen waaronder enkele gevolgzinnen, vergelijkende zinnen, toegevende zinnen en bijzinnen van voorwaarde of van veronderstelling. Hieronder een omvattende opsomming van de verschillende grammaticale categorieën die in het eerste stuk kunnen verkeren:

- Adjectief: “Opvallend was dat dit nauwelijks invloed heeft gehad op het beleid van [...]”. (NRC 75)
- BC: “Vroeger was ik echt uh had ik m’n haar al door de war heel erg”. (CGN 186)
- Bijzin: “Zoals de gewoonte is in de Unie, werd ook deze splijtzwam met een ondoorzichtig compromis afgedekt”. (NRC 74)
- Leeg: “[Daar weet ik zelf niet veel meer van] maar \emptyset vertelt m’n moeder heel vaak”. (CGN 63)
- NC: “De Israëlische club had een dag eerder Estland uitgeschakeld voor een toppositie”. (bepaald, NRC 158)
 “Turkije heeft aangeduid daarin niet plaats te zullen nemen”. (eigenaam, NRC 67).
 “Een appel moet je schillen voor je ’m opeet”. (generiek, ANS)
- Scheidbaar partikel: “Mee gaan dit jaar alle leerlingen uit de vierde en de vijfde klas”. (ANS)
- Telwoord: “Beiden menen dat deals volstrekt uit den boze zijn zolang er geen wettelijke regeling is”. (hoofdtelwoord, NRC 215)
 “De eersten krijgen een gouden medaille en 10.000 euro als beloning”. (rangtelwoord)
 “Tweemaal heb ik al gebeld maar ik sta nog in de deuropening te wachten”. (telbijwoord)
- Voornaamwoord: “Hij is ook verantwoordelijk voor het eerste team van Kalev en een jeugdselectie van de club”. (persoonlijk, NRC 130)
 “Want dat is niet meer zo’n hechte groep als die we vroeger ooit gehad hebben”. (aanwijzend, CGN 211)
 “Want iedereen is van tevoren goed gewaarschuwd”. (onbepaald)
 “Zichzelf vervelen ze enorm voor het ogenblik”. (wederkerend)

- “De mijne bevat minder suiker dan de meeste frisdranken van de distributiesector”. (bezittelijk)
- VB.: “En daarvan hoopt HCCR er voor het eind van dit jaar zo’n 2.000 te bewegen lid te worden”. (NRC 49)
- Voegwoordelijk constituent: “Als bondscoach ziet de oud-trainer van onder meer Donar [...] het doorvoeren van een verjonging binnen de selectie als zijn voornaamste taak”. (NRC 131)
- Voorzetselconstituent: “In mijn omgeving proefde ik daar weerstand tegen, vooral bij kleine ondernemers”. (NRC 29)
- WC: “Aan sport doen is belangrijk om gewicht te verliezen”.
- Zin in de directe rede: “Want in tweeduizend twee mag je met guldens en met euro betalen geloof ik”. (CGN 279)

Variabele 5: Wat is de syntactische functie van het element in de eerste zinsplaats? Naast het vormelijke vlak bestaat er in het Nederlands veel variatie met betrekking tot de functie van het item in het voorveld.

- Belanghebbend voorwerp: “Voor de noodzaak van stabiliteit verderop in de straat bestaat inmiddels voldoende begrip”. (NRC 89)
- Bijwoordelijke bepaling: “Behalve dat ik me ziek voel, heb ik ook nog problemen met mijn burenen”. (beperking, ANS)
- “Voor zo’n snelle transactie was zijn mobiele telefoon altijd binnen handbereik”. (doel, NRC 105)
- “En meestal worden altijd alle verhalen aangehaald”. (frequentie, CGN 218)
- “Als bondscoach ziet de oud-trainer van onder meer Donar [...]” (gesteldheid, NRC 131)
- “Tot onze vreugde is juffrouw Betsy niet opnieuw tot gouvernante aangesteld”. (gevolg, ANS)
- “Tot tranen toe was de scheidende voorzitter bewogen”. (graad, ANS)
- “Langzamerhand is toch overeenstemming ontstaan dat de club zo groot mogelijk dient te worden”. (hoedanigheid, NRC 57)

	<p>“<u>Drie kilometers hebben</u> we gisteren in de Kampen gewandeld”. (maat)</p> <p>“<u>Aan de hand van statistisch materiaal werd</u> zijn hypothese bevestigd”. (middel, ANS)</p> <p>“<u>Volgens een legerwoordvoerder werden</u> tijdens de tegenaanval zo’n 100 rebellen gedood”. (modaliteit, NRC 11)</p> <p>“<u>Met m’n vader ging</u> ik naar uh zo’n amusementhal Flash”. (omstandigheid, CGN 122)</p> <p>“<u>Niet iedereen was</u> ingenomen met haar benoeming”. (ontkenning, ANS)</p> <p>“En voor m’n longontsteking <u>heb</u> ik ook in ’t ziekenhuis gelegen”. (oorzaak, CGN 49)</p> <p>“<u>Op een gedeeld eerste plaats staat</u> het concept van de uitbreiding van de Unie”. (plaats, NRC 62)</p> <p>“En op een gegeven moment uh <u>droogde</u> ik een beetje uit heb ben ik in ’t ziekenhuis opgenomen”. (tijd, CGN 60)</p> <p>“<u>Hoewel hij een contractverlenging niet geheel wil uitsluiten, kijk</u> Van Gent nu al uit naar eind april”. (toegeving, NRC 149)</p> <p>“<u>Zoals de gewoonte is in de Unie, werd</u> ok deze splijtzwam met een ondoorzichtig compromis afgedekt”. (vergelijking, NRC 74)</p> <p>“<u>Naargelang Frits ouder werd, werd</u> hij wijzer”. (verhouding, ANS)</p> <p>“Dus ‘tzelfde <u>als je nu in de winkel met een Duitse mark betaalt, krijgen</u> ze Nederlands geld voor terug.” (voorwaarde, CGN 289)</p>
Gezegde:	<p>“[Er is wel veel gepraat] maar <u>besloten is</u> er uiteindelijk niets”. (ANS)</p>
Handelend voorwerp:	<p>“<u>Door dat kind zijn</u> heel wat boeken verkocht”.</p>
Inherente constituent:	<p>“<u>In aanmerking komen</u> alle studenten die reeds kandidaat in de letteren en wijsbegeerte zijn”. (ANS)</p>
Lijdend voorwerp:	<p>“Of de ’t verleden <u>heb</u> je gewoon gehad”. (CGN 163)</p>
Medewerkend voorwerp:	<p>“<u>Voor alle kandidaten gezamenlijk komt</u> er een conferentie [...]” (NRC 65)</p>
Onderwerp:	<p>“<u>Van Gent heeft</u> een Estse vriendin bij wie hij een zoontje van twee jaar heeft”. (NRC 116)</p>

- Oorzakelijk voorwerp: “Dat Jan altijd maar dwars moet liggen ben ik het beu”. (ANS)
- Predicaat: “Opvallend was dat dit nauwelijks invloed heeft gehad op [...]”
(NRC 75)
- Structurerende constituent: “[guldens [...] die worden dan ingenomen] en dan krijg je d’r euro’s voor terug”. (CGN 286)
- Voorzetselvoorwerp: “Ervan afgezien dat hij daarmee een bekend Brits gezegde in de herinnering riep, lijkt de gedachte onlogisch en onpraktisch”.
(NRC 73)

Hieronder een voorbeeld om te tonen hoe het corpusonderzoek concreet gedaan wordt. Laten we de zin over het Verdrag van Maastricht opnieuw gebruiken: “Wat in het Verdrag van Maastricht al werd voorzien, lijkt te worden bewaarheid: de monetaire grenzen mogen krap zijn getrokken, het laatste woord heeft de politiek en die neigt tot een zekere rekkelijkheid.” In eerste instantie kan de zin in verschillende segmenten verdeeld worden, bijvoorbeeld als die nevenschikte clauses bevat. Hier krijgen we de vier volgende segmenten:

- “Wat in het Verdrag van Maastricht al werd voorzien, lijkt te worden bewaarheid:”
- “de monetaire grenzen mogen krap zijn getrokken,”
- “het laatste woord heeft de politiek”
- “en die neigt tot een zekere rekkelijkheid.”

Elk segment wordt apart behandeld. Er wordt geantwoord op de vijf vragen overeenstemmend met de vijf variabelen. Telkens worden de elementen die zich buiten de zin, in de aanloop, in het eerste stuk en in de rest van de zin bevinden, bepaald en in losse vakjes gezet. Verder krijgen de aanloop- en voorveldelementen een grammaticale categorie en een grammaticale functie:

Relevant	Vorm in de aanloop	Functie in de aanloop	Vorm in het eerste stuk	Functie in het eerste stuk	Buiten de zin	Aanloop	Eerste stuk	Rest van de zin
Ja	/	/	Bijzin	Onderwerp	/	/	Wat in het verdrag van Maastricht al werd	Lijkt te worden bewaarheid:
Ja	/	/	NC	Onderwerp	/	/	De monetaire grenzen	Mogen zijn getrokken,
Ja	/	/	NC	Lijgend voorwerp	/	/	Het laatste woord	Heeft de politiek
Ja	/	/	Voornaamw oord	Onderwerp	En	/	Die	Neigt tot een zekere reddelijkheid .

3. Nevenschikkende voegwoorden

In dit hoofdstuk worden nevenschikkende voegwoorden nader bekeken. Hun eigenschappen zijn immers moeilijk te bepalen en verschillen van één grammatica tot de andere. Eén punt waarover alle naslagwerken het eens zijn, is dat NSVW zeker niet op de eerste zinsplaats voorkomen. Volgens de ANS (Haeseryn, 1997 : 1074) oefenen NSVW geen grammaticale functie uit en kunnen bijgevolg niet als zinsdeel beschouwd worden. Ze gaan de concrete zin vooraf en hebben geen invloed op de woordvolgorde (66). Verbanck (1985) en Vandeweghe (2013: 268-269) voegen toe dat NSVW altijd een vaste positie tussen twee of meer gelijkaardige constituenten innemen om die te verbinden.

(66) Ik heb een sms'je gekregen en ik weet niet wat te antwoorden.

Zin (66) toont perfect aan dat nevenschikkende voegwoorden niet tot het voorveld behoren. Als dat zo was, zou het onderwerp 'ik' zich na de persoonsvorm in het middenstuk bevinden om zich aan de PV2-regel te houden (67)

(67) *Ik heb een sms'je gekregen en weet ik niet wat te beantwoorden.

Sommigen linguïsten beweren echter dat NSVW niet precies tussen twee constituenten staan en dat ze een nauwere band hebben met het tweede *conjunct*, i.e. de tweede constituent van de nevenschikking. Onder de argumenten die Ross (1967) oppert om zijn standpunt te motiveren, zijn hier de twee voornaamste:

(a) Als een nevenschikkinge zin in twee hoofdzinnen verdeeld wordt, gaat het NSVW altijd gepaard met het tweede constituent.

(68a) Ik heb een sms'je gekregen, en ik weet niet wat te beantwoorden.

(68b) Ik heb een sms'je gekregen. En ik weet niet wat te beantwoorden.

(68c) *Ik heb een sms'je gekregen en. Ik weet niet wat te beantwoorden.

(b) Evenzo is het nevenschikkend voegwoord op fonologisch vlak nauwer verbonden met de constituent waaraan het voorafgaat dan met het eerste lid van de nevenschikking. Dit wordt door Ross verklaard gezien de positie van intonatieve onderbrekingen, met name vóór de NSVW (1967 : 164-165). Daarom is (69a) de correcte wijze waarop de fonologie van (69) aangegeven moet worden.

(69) Peter, en Sandra, en Mathias hebben een sms'je gekregen.

(69a) ((Peter) (en Sandra) (en Mathias)) hebben een sms'je gekregen.

(69b) ((Peter) (en) (Sandra) (en) (Mathias)) hebben een sms'je gekregen.

(69c) ((Peter en) (Sandra en) (Mathias)) hebben een sms'je gekregen.

Uit deze twee voorbeelden blijkt immers dat NSVW wel innig verworven zijn met het tweede *conjunct*. Betekent dit echter dat de voegwoorden deel uitmaken van de zin? Aangezien ze in het eerste stuk uitgesloten zijn, zouden ze zich dan in de aanloop bevinden.

De ANS vermeldt vrij snel dat nevenschikkende voegwoorden niet tot de aanloop kunnen behoren maar geven er geen reden voor, anders dan dat ze tussen twee zinnen staan, wat gedeeltelijk door Ross weerlegd is. Volgens Vandeweghe kunnen ze ook niet in de aanloop staan vanuit een theoretisch standpunt (2013 : 269). Hij verklaart zijn mening door het feit dat er geen intonatieve pauzes bestaan tussen nevenschikkende voegwoorden en eerste stukken, terwijl aanlopen en voorvelden wel door een komma in geschreven taal en door een rust in gesproken taal gescheiden zijn. Er wordt niettemin toegevoegd dat de voegwoorden in de linkerperiferie kunnen voorkomen als de zender een komma-intonatie bij zijn uiting voegt. In (70a) zou dus *maar* in de aanloop staan, in tegenstelling tot (70b)

(70a) Maar, ik weet niet wat te doen.

(70b) Maar ik weet niet wat te doen.

Het wordt duidelijk dat de positie van de nevenschikkende voegwoorden nog voor problemen zorgt. Ze staan niet op de eerste zinsplaats. Ze lijken niet in de aanloop staan. Waar staan ze dan? Ze kunnen bijvoorbeeld met aansprekingen, met anticiperende elementen, met de 'wat X betreft' constructie of met enkele bijzinnen gekoppeld worden. In die gevallen wordt het moeilijk om de voegwoorden van de aanloop uit te sluiten (71) t.e.m. (74).

(71) [Elien, je mag hier komen zitten.] En Hans, je mag naast mij zitten.

(72) [Haar zusje is redelijk klein.] Maar haar broer, die is bijzonder groot.

(73) [Ik ben met het samenstellen van een corpus begonnen.] En wat mijn studie betreft, die wordt binnenkort uitgegeven.

(74) [Het aantal wolven is de laatste jaren flink afgenomen.] En waren er vroeger nog veel bossen, nu zijn het er beduidend minder.

In het corpusonderzoek zullen de nevenschikkende voegwoorden geen deel uitmaken van de aanloop. In de gevallen waar ze alleen vóór de eerste zinsplaats staan is het immers bijzonder complex om te bepalen of er een komma-intonatie bestaat in transcripties van gesproken taal. Natuurlijk past het om dezelfde methode te gebruiken voor de twee gehanteerde corpora, met name een corpus gesproken taal en een corpus geschreven taal. Bovendien maakt het geen verschil om de NSVW buiten beschouwing te laten als ze met andere constituenten gecombineerd worden aangezien er al een aanloop bestaat.

De Nederlandse grammatica's en de wetenschappelijke artikelen over coördinatie of over Nederlandse syntaxis geven vandaag de dag nog geen onmiskenbaar antwoord over de plaatsing van nevenschikkende voegwoorden. Deze vormen dus een kennisgebied dat naar nader onderzoek eist.

4. Resultaten

In dit hoofdstuk worden de analyseresultaten van het *Corpus Gesproken Nederlands* en het *NRC Handelsblad* bekend gemaakt. Hier gaat het om het pure, objectieve meedelen van gegevens. Deze zullen in het volgende hoofdstuk besproken en geïnterpreteerd worden. Om redenen van duidelijkheid worden de corpora apart aangepakt. Eerst beginnen we met het *NRC Handelsblad*. Daarna worden de gegevens van het *Corpus Gesproken Nederlands* weergegeven. Ten laatste worden de resultaten van beide corpora vergeleken voordat ze in hoofdstuk 5 bediscussieerd en verklaard worden.

4.1 NRC

Tabel 1. Pertinentie van de hits in het NRC

Relevant	Irrelevant	
200	16	216

Uit het geschreventaalcorpus werden er 216 hits geboekt en geanalyseerd. 200 van deze 216 uitingen werden als relevant beschouwd. De 16 irrelevante uitingen werden uitgesloten omdat ze geen persoonsvorm bevatten. Soms duiden ze de schrijver (75) van het artikel aan, soms ging het om vereenvoudigde titels (76), maar vaak hielden de hits gewoon geen PV2 in. Meestal geven deze uitingen extrainformatie betreffende voorafgaande zinnen (77).

(75) Door een onzer redacteurs. (NRC 23)

(76) Aanval op de hoofdstad van Burundi. (NRC 3)

(77) [Ik sloot deals.] Het liefst snelle deals waar veel geld mee te verdienen was. (NRC 104)

Slechts twee van de tweehonderd zinnen bevatten een ingevulde aanloop. Formeel bevat de linkerperiferie in het eerste geval een bijvoeglijk naamwoord terwijl de tweede aanloop een bijzin inhoudt. Hier kan al vastgesteld worden dat aanlopen in geschreven taal zelden ingevuld lijken te worden.

Tabel 2. Vorm van het element in de aanloop

Element in de aanloop	Hoeveelheid	Voorbeeld
Adjectief	1 (50%)	<u>Sterker, de Duitse socialisten hebben</u> slechts een kans [...] (NRC 82)
Bijzin	1 (50%)	<u>Was vorig jaar een tijd lang alle aandacht gericht op de top in Amsterdam, waar het Verdrag van Maastricht moest worden aangepast, nu gaat</u> het er om [...] (NRC 55)

Deze twee aanlopen hebben ook twee verschillende grammaticale functies. Aan de ene kant dient het bijvoeglijk naamwoord immers als structurerend connectief en aan de andere kant wordt de bijzin gebruikt als een bepaling van toegeving.

Tabel 3. Grammaticale functie van het element in de aanloop

Functie in de aanloop	Hoeveelheid	Voorbeeld
Bijwoordelijke bepaling	1 (50%)	<u>Was vorig jaar een tijd lang alle aandacht gericht op de top in Amsterdam, waar het Verdrag van Maastricht moest worden aangepast, nu gaat</u> het er om [...] (NRC 55)
Structurerend connectief	1 (50%)	<u>Sterker, de Duitse socialisten hebben</u> slechts een kans [...] (NRC 82)

Over eerste stukken in het *NRC*-corpus is er daarentegen veel meer te vertellen. Van de 200 relevante hits zijn er immers maar 8 die geen voorveld bevatten (4%). Lege eerste zinsplaatsen komen overeen met de zesde frequentste vorm uit het corpusonderzoek. De meest voorkomende grammaticale categorie in het eerste stuk is de naamwoordelijke constituent (of NC). Ze tellen voor de helft van de gevallen (50%) aangezien 100 op de 200 geanalyseerde eerste stukken nominale groepen bevatten. Dat komt bijna neer op het viervoud van de tweede frequentie; de voorzetselconstituenten met 27 hits (13,5%). Even frequent als de voorzetselconstituenten zijn de voornaamwoorden. Van de 27 gevallen zijn er 20 met persoonlijke voornaamwoorden (10%) en 7 met aanwijzende voornaamwoorden (3,5%). De

laatste constituent met meer dan 10% zijn de bijwoordelijke constituenten. Ook zij komen 20 keer voor in het geschreventaalcorpus. Buiten de bijzinnen die in 4,5% van de aanlopen aanwezig zijn, komen de resterende elementen niet meer dan in 1% van de gevallen voor. Tabel 4 vat deze resultaten samen en illustreert de verschillende gevallen.

Tabel 4. Vorm van het element in het eerste stuk

Vorm in het eerste stuk	Hoeveelheid	Voorbeeld
Naamwoordelijke constituent	100 (50%)	<u>De Israëlische club had</u> een dag eerder Estland uitgeschakeld voor een toppositie. (NRC 158)
Voorzetselconstituent	27 (13,5%)	<u>In mijn omgeving proefde</u> ik daar weerstand tegen, vooral bij kleine ondernemers. (NRC 29)
Voornaamwoord	20 (10%)	<u>Hij is</u> ook verantwoordelijk voor het eerste team van Kalev en een jeugdselectie van de club. (NRC 130)
	7 (3,5%)	[...] en <u>die neigt</u> tot een zekere rekkelijkheid (NRC 61)
Bijwoordelijke constituent	20 (10%)	<u>Dan kan</u> hij in ieder geval voor even zijn oude leven weer oppakken. (NRC 150)
Bijzin	9 (4,5%)	<u>Zoals de gewoonte is in de Unie, werd</u> ook deze splijtzwam met een ondoorzichtig compromis afgedekt. (NRC 74)
Leeg	8 (4%)	[Dan kan hij in ieder geval voor even zijn oude leven weer oppakken.] En <u>∅ gaan</u> varen. (NRC 151)
Voornaamwoordelijk bijwoord	2 (1%)	En <u>daarvan hoopt</u> HCCR er voor het eind van dit jaar zo'n 2.000 te bewegen lid te worden. (NRC 49)
Werkwoordelijke constituent	2 (1%)	<u>Ook spelevaren kan</u> een sleur worden (NRC 109)
Zin in de directe rede	2 (1%)	“Jammer”, <u>zegt</u> Van Gent. (NRC 136)
Adjectief	1 (0,5%)	<u>Opvallend was</u> dat dit nauwelijks invloed heeft gehad op het beleid van [...] (NRC 75)

Hoofdtelwoord	1 (0,5%)	<u>Beiden menen</u> dat deals volstrekt uit den boze zijn zolang er geen wettelijke regeling is. (NRC 215)
Voegwoordelijke constituent	1 (0,5%)	<u>Als bondscoach ziet</u> de oud-trainer van onder meer Donar [...] (NRC 131)

Qua grammaticale functie van de eerste zinsplaats-elementen in geschreven taal is de tendens ook duidelijk. 122 van de 200 eerste stukken dienen als onderwerp, wat voor 60% van de gevallen telt. De enige andere grammaticale functie die in meer dan 10% van de gevallen in het voorveld wordt vervuld is de bijwoordelijke bepaling. Deze functie houdt 50 hits in en zorgt voor 25% van het *NRC*. Binnen de bijwoordelijke bepalingen zet de tijdbepaling zich af tegen de andere bepalingen door zijn grotere frequentie. 26 op de 50 bepalingen geven immers een temporele verduidelijking aan (13% van het hele corpus), ver vóór plaatsbepalingen (4,5%), bepalingen van modaliteit (3%) en van toegeving (2%). Lege voorvelden staan hier voor de derde frequentste mogelijkheid (4%). Uit tabel 5 blijkt ook dat elementen uit de eerste zinsplaats in amper 3% van de hits een structurerende functie uitoefenen.

Tabel 5. Grammaticale functie van het element in het eerste stuk

Functie in het eerste stuk	Hoeveelheid	Voorbeeld
Onderwerp	122 (61%)	<u>Van Gent heeft</u> een Estse vriendin bij wie hij een zontje van twee jaar heeft. (NRC 116)
Bijwoordelijke bepaling	26 (13%)	<u>Eind februari speelt</u> Estland tegen Nederland (tijd, NRC 140)
	9 (4,5%)	<u>Op een gedeeld eerste plaats staat</u> het concept van de uitbreiding van de Unie. (plaats, NRC 62)
	6 (3%)	<u>Volgens een legerwoordvoerder werden tijdens de tegenaanval zo'n 100 rebellen gedood.</u> (modaliteit, NRC 11)
	4 (2%)	<u>Hoewel hij een contractverlenging niet geheel wil uitsluiten, kijk</u> Van Gent nu al uit naar eind april. (toegeving, NRC 149)

	2 (1%)	<u>Langzamerhand is toch overeenstemming ontstaan dat de club zo groot mogelijk dient te worden.</u> (hoedanigheid, NRC 57)
	1 (0,5%)	<u>Voor zo'n snelle transactie was zijn mobiele telefoon altijd binnen handbereik.</u> (doel, NRC 105)
	1 (0,5%)	<u>Als bondscoach ziet de oud-trainer van onder meer Donar [...]</u> (gesteldheid, NRC 131)
	1 (0,5%)	<u>Zoals de gewoonte is in de Unie, werd ok deze splijtzwam met een ondoorzichtig compromis afgedekt.</u> (vergelijking, NRC 74)
Leeg	8 (4%)	[Dan kan hij in ieder geval voor even zijn oude leven weer oppakken.] En <u>∅ gaan varen.</u> (NRC 151)
Lijdend voorwerp	7 (3,5%)	<u>Dat water vindt hij te zwart.</u> (NRC 158)
Structurerend	6 (3%)	<u>Verder noemt hij de oprichting van HCCR een "stijlloze actie" [...]</u> (NRC 38)
Voorzetselvoorwerp	3 (1,5%)	<u>Daar heb ik plezier in.</u> (NRC 145)
Meewerkend voorwerp	1 (0,5%)	<u>Voor alle kandidaten gezamenlijk komt er een conferentie [...]</u> (NRC 65)
Predicaat	1 (0,5%)	<u>Opvallend was dat dit nauwelijks invloed heeft gehad op [...]</u> (NRC 75)
Retorisch	1 (0,5%)	<u>Eigenlijk is dat allemaal sleur.</u> (NRC 147)
Belanghebbend voorwerp	1 (0,5%)	<u>Voor de noodzaak van stabiliteit verderop in de straat bestaat inmiddels voldoende begrip.</u> (NRC 89)

4.2 CGN

In tegenstelling tot het *NRC* bevat het *Corpus Gesproken Nederlands* een groot aantal irrelevante clauses, 88 om precies te zijn. Deze zijn buiten beschouwing gelaten gezien hun interrogatieve waarde, soms zelfs zonder vraagteken (78) of de afwezigheid van een persoonsvorm, bijvoorbeeld in gevallen van antwoord of reactie op een voorafgaande zin. Tal van hits bestonden immers uit één interjectie of een combinatie daarvan (er zij op gewezen dat 'ggg' in het corpus betekent dat de spreker aan het lachen is).

(78) Nou wat moet 'k over mijn ouders vertellen. (CGN 19)

Tabel 6. Pertinentie van de hits in het CGN

Relevant	Irrelevant	
200	88	288

Net meer dan een achtste van de relevante hits hielden in het corpus gesproken taal een aanloop in, met name 26. Vanuit een vormelijk standpunt bestond de overgrote meerderheid van de linkerperiferieën uit tussenwerpsels (c. 76,92%). In de andere zeldzamere gevallen stond er een nominale constituent (c. 7,69%), een bijwoordelijke constituent (c. 7,69%) of een bijzin (c. 3,85%) in de aanloop. Het 26^{ste} en laatste geval bevatte twee aanloopelementen die elk hun eigen syntactische functie uitoefenen (c. 3,85%)

Tabel 7. Vorm van het element in de aanloop

Element in de aanloop	Hoeveelheid	Voorbeeld
Tussenwerpsel	20 (76,92%)	<u>Nou, daar weet</u> ik zelf niet veel meer van. (CGN 62)
Naamwoordelijke constituent	2 (7,69%)	<u>M'n ouders, die zijn</u> uh ... (CGN 20)
Bijwoordelijke constituent	2 (7,69%)	<u>Trouwens, dat heb ik gehoord.</u> (CGN 78)
Bijzin	1 (3,85%)	<u>als je wil omrekenen van een euro naar een gulden, dan moet</u> je dat delen door twee komma twee. (CGN 266)
Tussenwerpsel / naamwoordelijke constituent	1 (3,85%)	<u>Ja m'n vader die vaarde</u> op zee. (CGN 22)

Zoals eerder gezegd waren de verscheidene functies van de elementen in de aanloop complex te bepalen door hun subjectiviteit. Uit tabel 7 bleek dat de daar meest voorkomende constituenten in het corpus gesproken taal tussenwerpsels waren. Deze kunnen talrijke rollen vervullen, wat onze analyse bemoeilijkt. Daarin werden er 14 aanloopelementen met een retorische functie geboekt (c. 53,85%). Interactieve constituenten kwamen in haast een kwart van de gevallen voor (c. 23,08%) en structurerende aanlopen waren nog iets minder frequent (c. 15,38%). In één bepaald geval werd er gekozen om tegelijkertijd een interactieve en een

retorische rol toe te wijzen. De laatste hit betreft de eerder besproken dubbele aanloop. Aan de ene kant oefent het tussenwerpsel weer tegelijkertijd een interactieve en retorische functie uit en aan de andere kant werd de NC als structurerend beschouwd.

Tabel 8. Functie van het element in de aanloop

Element in de aanloop	Hoeveelheid	Voorbeeld
Retorisch	14 (53,85%)	<u>Nou daar weet</u> ik zelf niet veel meer van. (CGN 62)
Interactief	6 (23,08%)	[Grote ramp]. <u>Nee ze waren</u> hartstikke blij met mij. (CGN 37)
Structurerend	4 (15,38%)	<u>Trouwens, dat heb</u> ik gehoord. (CGN 78)
Retorisch + interactief	1 (3,85%)	[Binnenkort krijgen we euro's volgend jaar.] <u>Ja we komen</u> uit de tijd van de guldens. (CGN)
Retorisch + interactief / structurerend	1 (3,85%)	<u>Ja m'n vader die vaarde</u> op zee. (CGN 22)

Uit tabel 9 blijkt dat de frequentste grammaticale categorie die zich op de eerste zinsplaats in het *CGN* bevindt de voornaamwoorden zijn. In het corpusonderzoek duiken vooral persoonlijke (60) en aanwijzende (51) voornaamwoorden op zodat pronomina in meer dan de helft van de gevallen het voorveld bezetten (55,5%). In de tweede helft van de hits staan voornamelijk bijwoordelijke constituenten (24%) en naamwoordelijke constituenten (11,5%) in het eerste stuk. 2 van de 200 relevante clauses bevatten geen element op de eerste zinsplaats terwijl één eerste stuk constituenten uit twee grammaticale categorieën inhield, met name een naamwoordelijke constituent samen met een bijzin. Alle gevallen worden in tabel 9 weergegeven.

Tabel 9. Vorm van het element in het eerste stuk.

Element in het eerste stuk	Hoeveelheid	Voorbeeld
Voornaamwoord	60 (30%)	Want <u>dat is</u> niet meer zo'n hechte groep als die we vroeger ooit gehad hebben. (aanwijzend, CGN 211)
	51 (25,5%)	Ggg je <u>kan</u> 't nog makkelijker doen. (CGN 268)

Bijwoordelijke constituent	48 (24%)	<u>Vroeger was ik echt</u> uh had ik m'n haar al door de war heel erg. (CGN 186)
Naamwoordelijke constituent	23 (11,5%)	<u>Roy en ik denken</u> daar heel ander over. (CGN 145)
Voorzetselconstituent	7 (3,5%)	En <u>voor m'n longontsteking heb</u> ik ook in 't ziekenhuis gelegen (CGN 49)
Bijzin	5 (2,5%)	<u>Wat achter me ligt is</u> gebeurd en xxx. (CGN 156)
Zin in de directe rede	3 (1,5%)	<u>Want in tweeduizend twee mag je met guldens en met euro betalen geloof</u> ik. (CGN 279)
Leeg	2 (1%)	[Daar weet ik zelf niet veel meer van] maar \emptyset . <u>vertelt m'n moeder heel vaak.</u> (CGN 63)
Naamwoordelijke constituent / bijwoordelijke constituent	1 (0,5%)	En <u>twee jaar later toen kwam</u> ik (CGN 35)

De verdeling van de resultaten van de elementen in het voorveld vanuit een vormelijk en een functioneel standpunt in het corpus gesproken materiaal zijn redelijk gelijk. Ten eerste een constituent die in meer dan de helft van de gevallen voorkomt, hier onderwerpen (52,5%). Dan twee elementen die tamelijk frequent verschijnen, rond 25% en 10%. Deze twee grammaticale functies zijn de bijwoordelijke bepalingen (27,5%) en de lijdend voorwerpen (14%). Ten laatste enkele zinsdelen met zeldzame functies (2,5% of minder). Wat het eerste stuk met de NC en de BC betreft, dit moest maar één functie uitoefenen, aldus de grammatica's. Anders komen er twee zinsdelen voor de eerste pool, wat onmogelijk is. Uit het corpusonderzoek blijkt dat de NC en de BC samen de rol van bijwoordelijke bepaling van tijd spelen en dus de regel naleven.

Tabel 10. Functie van het element in het eerste stuk

Element in het eerste stuk	Hoeveelheid	Voorbeeld
Onderwerp	105 (52,5%)	Uhm <u>m'n moeder is</u> nou eenenvijftig geworden in december (CGN 6)

Bijwoordelijke bepaling	44 (22%)	En <u>op een gegeven moment</u> uh <u>droogde</u> ik een beetje uit heb ben ik in 't ziekenhuis opgenomen. (tijd, CGN 60)
	3 (1,5%)	<u>Daar had</u> je van die flipperkasten. (plaats, CGN 123)
	2 (1%)	En <u>voor m'n longontsteking</u> <u>heb</u> ik ook in 't ziekenhuis gelegen. (oorzaak, CGN 49)
	1 (0,5%)	En <u>meestal</u> <u>worden</u> altijd alle verhalen aangehaald. (frequentie, CGN 218)
	1 (0,5%)	[Je kan 't nog makkelijker doen. Delen door twee plus tien procent.] <u>Zo kan</u> 't ook. (hoedanigheid, CGN 270)
	1 (0,5%)	<u>Misschien</u> <u>zijn</u> we al op een uur. (modaliteit, CGN 223)
	1 (0,5%)	<u>Met m'n vader</u> <u>ging</u> ik naar uh zo'n amusementhal Flash. (omstandigheid, CGN 122)
	1 (0,5%)	En <u>toch</u> <u>wou</u> 'k altijd uh altijd graag spelen. (toegeving, CGN 54)
	1 (0,5%)	Dus 'tzelfde <u>als je nu in de winkel met een Duitse mark betaalt, krijgen</u> ze Nederlands geld voor terug. (voorwaarde, CGN 289)
Lijdend voorwerp	28 (14%)	Of <u>de 't verleden</u> <u>heb</u> je gewoon gehad (CGN 163).
Voorzetselvoorwerp	5 (2,5%)	Nee <u>daar wil</u> ik helemaal niet over praten. (CGN 201)
Structurerend	4 (2%)	[guldens die worden dan ingenomen] en <u>dan krijg</u> je d'r euro's voor terug. (CGN 286)
Leeg	2 (1%)	[Nou daar weet ik zelf niet veel meer van] maar <u>∅ vertelt</u> m'n moeder heel vaak (CGN 63)
Predicaatsnomen	1 (0,5%)	<u>Lestley</u> <u>heette</u> die. (CGN 81)

4.3 Vergelijking

Nu de resultaten van elk corpus apart behandeld zijn, is het tijd om die in perspectief te brengen door de gegevens van de schrijftaal en de spreektaal te vergelijken. Hierna zullen de divergentiepunten en de gelijkenissen besproken en geïnterpreteerd worden ten aanzien van de oorspronkelijke vragen rond het gebruik van de aanloop en het eerste stuk in het Nederlands.

Opvallend is ten eerste het verschil tussen het aantal irrelevante hits voor ieder corpus. Terwijl in de *Nieuwe Rotterdamse Courant* een totaal van 216 hits nodig waren om het doel van 200 pertinente zinsneden te verkrijgen, hebben we in het *Corpus gesproken Nederlands* alles samen 288 gevallen moeten analyseren teneinde de 200 werkbare hits te bereiken. Dat betekent dat de kans om onnuttig materiaal tegen te komen meer dan vijf keer groter is in het corpus gesproken taal. Dat betekent ook dat ongeveer één op de drie hits in dit corpus buiten beschouwing is gelaten.

Tabel 11. Relevantie per corpus

	NRC	CGN
Relevant	200	200
Irrelevant	16	88

Het is ook opmerkelijk vast te stellen dat de zaken voor elk corpus irrelevant zijn om verschillende redenen. Aan de ene kant bevat het *NRC* journalistieke titels zonder persoonsvorm om de boodschap te vereenvoudigen en duidelijk te maken of zijn er zinnen zonder eerste pool die extra informatie over een voorafgaande zin geven. Aan de andere kant zit het *CGN* vol met hits die uitsluitend uit interjecties bestaan et die bijgevolg geen concrete inhoud hebben.

Tabel 12. Vorm van het element in de aanloop per corpus

	NRC	CGN
Tussenwerpsel	/	20
NC	/	2
BC	/	2
Adjectief	1	/
Bijzin	1	1
Tussenwerpsel / NC	/	/
	2	26

Ten tweede, de twee corpora zijn opnieuw redelijk verschillend betreffende de aanloop, voornamelijk op het gebied van de hoeveelheid en dus van de frequentie daarvan. Terwijl het corpus geschreven taal amper 2 clauses op de 200 een aanloop bevat, omvat het tweede corpus 26 zinnen met een aanloop. Ook al moeten deze getallen gerelativeerd worden door de beperktheid van de corpora, ze zijn niettemin de moeite waard voor nadere analyse en bespreking.

Er kan vastgesteld worden dat één uit de twee aanloopelementen uit het *Handelsblad* terug te vinden is in het andere corpus, nl. een bijzin. Toch hebben ze niet dezelfde syntactische functie: de bijzin in geschreven taal dient als bijwoordelijke bepaling terwijl de bijzin in het *CGN* een anticiperende functie heeft. Deze wordt dus gebruikt om het gesprek te structureren. De andere grammaticale constituent die in het eerste corpus voorkomt is een bijvoeglijk naamwoord, onder de vorm van een vergelijkende trap. Deze dient twee zinnen te verbinden en structureert de tekst. Adjectivische aanlopen treden niet in het tweede corpus op. Zoals eerder vermeld bestaan de linkerperiferieën daar meestal uit tussenwerpsels (meer dan 75% van de gevallen).

Tabel 13. Functie van het element in de aanloop per corpus

	NRC	CGN
Retorisch	/	14
Interactief	/	6
Structurerend	1	4

Bijwoordelijke bepaling	1	/
Retorisch + interactief	/	1
Retorisch + interactief / structurerend	/	1
	2	26

Ten derde is de vergelijking tussen de twee corpora belangwekkender qua voorveld aangezien het *NRC* en het *CGN* respectievelijk 192 en 198 clauses met een eerste stuk bevatten. Uit het corpus geschreven taal blijkt dat precies de helft van de eerste stukken een naamwoordelijke constituent inhouden. Verschillend is dat in gesproken taal, waar amper iets meer dan 10% van de eerste zinsplaats elementen nominale groepen zijn. In dit corpus zijn de voornaamwoorden het meest voorkomend (55,5% bijna billijk verdeeld onder persoonlijke en aanwijzende voornaamwoorden). Ook al bezetten de pronomina het voorveld regelmatig in het *Handelsblad* (13,5%), deze zijn niettemin een stuk minder frequent in het eerste corpus. Het lijkt alsof de resultaten van de frequentie van NC's en voornaamwoorden in beide corpora omgewisseld werden. De andere syntactische constituenten die herhaaldelijk in het corpus schriftelijke taal terugkwamen zijn de voorzetselconstituenten (13,5%) en de bijwoordelijke constituenten (10%). Het verschil tussen de bevindingen van deze twee grammaticale categorieën in het *Corpus Gesproken Nederlands* is omgekeerd evenredig met de resultaten van het geschreven corpus. Terwijl het aantal voorzetselconstituenten op de eerste zinsplaats opvallend is afgenomen (3,5%), is de aanwezigheid van BC's vóór de eerste pool meer dan verdubbeld (24%). De resterende elementen treden te zeldzaam op om deze te vergelijken, ware het niet dat het *NRC* over het algemeen meer verscheiden is dan het *CGN*: Er komen 12 verschillende syntactische categorieën in het voorveld voor in het eerste corpus, tegen 7 in het tweede.

Tabel 14. Vorm van het element in het eerste stuk per corpus

	NRC	CGN
Voornaamwoord	27	111
NC	100	23
BC	20	48
Voorzetselconstituent	27	7
Bijzin	9	5

Leeg	8	2
Zin in directe rede	2	3
Voornaamwoordelijk bijwoord	2	/
WC	2	/
Adjectief	1	/
Hoofdtelwoord	1	/
Voegwoordelijke constituent	1	/
NC / BC	/	1
	200	200

Ten laatste, de vergelijking tussen de resultaten van de functie van het aanloopelement in beide corpora toont meer gelijkenissen. In de twee oefenen de voorvelden elementen meestal de functie van onderwerp uit hoewel de frequentie van onderwerpen ietsje groter is in het *NRC* (61%) dan in het *CGN* (52,5%). Op de tweede plaats komen in beide corpora bijwoordelijke bepalingen voor. Hier zijn de frequenties bijzonder gelijk (respectievelijk 25% in het geschreven corpus en 27,5% in het andere). Verder zijn deze bepalingen in de twee datasets voornamelijk tijdsbepalingen, ook al is het verschil hier groter (13% van het hele corpus in het *Handelsblad* tegen 22% in het corpus gesproken taal). Dit betekent dus dat de kans groter is om andere categorieën bijwoordelijke bepalingen tegen te komen in het schriftelijke corpus, met name 12% van het hele corpus, dan in het *CGN*. In het eerste corpus komen onderwerpen en bijwoordelijke bepalingen zo veelvuldig voor dat de andere syntactische functies nauwelijks of niet opduiken. Het derde meest voorkomende geval in dit corpus zijn lege eerste stukken met 8 voorbeelden op 200 (4%). In het tweede corpus dienen de eerste stukken nog regelmatig als lijdend voorwerpen (14%). In dit geval behoren ze meestal tot de voornaamwoorden, namelijk 19 op de 28 casussen.

Tabel 15. Functie van het element in het eerste stuk per corpus

	NRC	CGN
Onderwerp	122	105
Bijwoordelijke bepaling	50	55
Lijdend voorwerp	7	28

Structurerend	6	4
Leeg	8	2
Voorzetselvoorwerp	3	5
Meewerkend voorwerp	1	/
Retorisch	1	/
Predikaat	1	/
Predikaatsnomen	/	1
Belanghebbend voorwerp	1	/
	200	200

Uit dit hoofdstuk blijkt dat de corpusanalyse afwijkingen en raakpunten tussen het *NRC* en het *CGN* heeft aangegeven. Na de resultaten van de vorm en de functie van het element in de aanloop en op de eerste zinsplaats voor elke corpus objectief te hebben gerapporteerd, wordt het nu tijd om deze resultaten te bespreken in het volgende hoofdstuk.

5. Interpretatie

Dit hoofdstuk beoogt het corpusonderzoek tegenover het theoretische kader te stellen. Essentieel is immers dat de resultaten uit hoofdstuk 4 in de geschikte context worden geplaatst. Door de aanloop en het eerste stuk aan de hand van een corpusonderzoek te bestuderen heeft deze scriptie voornamelijk tot doel meer inzicht te krijgen in de mechanismes die deze twee syntactische constituenten verbinden. Meer in het algemeen streeft deze studie naar een beter begrip van de verhouding tussen syntaxis en semantiek, d.w.z. tussen de informatieve waarde van een zin en zijn betekenis aan de ene kant, en de wijze waarop deze informatie vormelijk verdeeld wordt aan de andere kant. Er zij op gewezen dat de data sample die in deze analyse gebruikt werd beperkt is. Met andere woorden, dat betekent dat de interpretaties en de duidingen die in deze sectie geleverd zullen worden, gebaseerd zijn op een kleinschalig onderzoek. definitieve conclusies kunnen dus nog niet getrokken worden. Het doel van dit hoofdstuk is aan de hand van een bescheiden corpus tendensen te ontdekken en daaraan zin te geven voor zover mogelijk.

5.1 Relevantie

Wat de relevantie van de hits aangaat, er werd eerder vermeld dat het corpus geschreven taal veel minder onnuttig materiaal inhoudt dan het *Corpus Gesproken Nederlands*. Met onnuttig materiaal wordt er bijna uitsluitend ‘hits zonder persoonsvorm’ bedoeld. Zonder PV krijg je immers geen eerste stuk en geen aanloop. Toch moeten we voor ogen houden dat de persoonsvormen elliptisch kunnen worden gebruikt.

Gesproken taal – en informele conversatie in het bijzonder – kenmerkt zich door spontaneïteit. Dat betekent dat de sprekers gebruik maken van pauzes en aarzelingen om hetzij wat extra tijd te verkrijgen en de inhoud van hun discours op voorhand voor te bereiden of dit discours te rectificeren en aan te passen. In sectie 1.4. hebben we bijvoorbeeld gezien dat sportcommentatoren diverse technieken gebruiken om eigennaamagenten (ENA’s) verder in de zin te laten opduiken (Kuiper et al. 2017). Daardoor winnen ze meer tijd om de ENA’s te herkennen en de naam daarvan uit hun geheugen terug te krijgen.

In tegenstelling daartoe typeert geschreven taal zich voornamelijk door een langere denktijd voor het schrijven. Vooral in gevallen waar de schrijvers de kans hebben om aan de hand van een klad fouten te verbeteren, de structuur van de tekst te bedenken en te bewerken, kortom wijzigingen door te voeren. In krantenartikelen zoals die uit het *NRC Handelsblad* is dat vanzelfsprekend het geval. In dergelijke artikelen is het ook niet ongewoon om uittreksels uit gesproken interviews terug te vinden. Ook deze worden dan aangepast om het lezen te vergemakkelijken en de inhoud te benadrukken. Ondenkbaar zou zijn de geschreven transcriptie van een mondeling vraaggesprek met alle aarzelingen, interjecties en fouten bijvoorbeeld. Het verschil tussen de spontaneïteit van gesproken taal en de bedenktijd van schrijftaal kan dus gedeeltelijk het contrast tussen irrelevante hits van de twee corpora verklaren.

5.2 Aanloop

Laten we ons nu dieper in het corpusonderzoek zelf verdiepen. Alweer is het essentieel te verduidelijken dat dit corpusonderzoek op kleinschalig niveau gedaan werd, en dat de ideeën die hier voorgesteld worden, om bevestiging vragen, hoe doordacht ze ook mogen zijn. De in deze scriptie gebruikte methode bestaat erin de resultaten van het corpusonderzoek met de theoretische achtergrond te confronteren om over de noodzakelijke kennis te beschikken voor de interpretatie.

Eén van de belangrijkste verschillen tussen de twee corpora is het aantal aanlopen. Wat zijn de factoren en de omstandigheden die een dergelijk decalage kunnen verklaren? Ter herinnering: in het *NRC* zijn amper 2 ingevulde aanlopen opgedoken terwijl het *CGN* 26 aanlopen bevatte. Onder de elementen die in de aanloop kunnen staan hebben we elementen met een anticiperende functie, interjecties en vocatieven, enkele typen bijzinnen, enkele voegwoordelijke bijwoorden, samenvattende uitdrukkingen en de constructies ‘wat X betreft/aangaat’.

In zijn inleiding over de aanloop vermeldt de *ANS* (1997 : 971) dat de aanloopelementen die een zogenaamd ‘echt’ verwijswoord vereisen, “nagenoeg beperkt zijn tot gesproken taal”. De voorbeelden die de *ANS* gebruikt om de aanloop met een echt verwijswoord toe te lichten, bestaan uitsluitend uit aanlopen met een anticiperende functie. Verder voegt de *ANS* concreet

toe dat dit type aanloop het enige is dat altijd een verwijswoord vereist en dat het vooral in gesproken taal gebruikt wordt (1997 : 971). Dit wordt ook bevestigd door Van Belle et al. (2006 : 36). In zijn ideale NVT-grammatica voor Franstaligen specificeert Hiligsmann welke aanloopelementen met een verwijswoord gepaard gaan (2001 : 109). De auteur detailleert zelfs voor bepaalde gevallen of ze mondeling of geschreven gebruikt worden. Zo bevestigt Hiligsmann dat elementen in de linkerperiferie die op een constituent in de concrete zin vooruitlopen in gesproken taal verschijnen.

Het andere geval waar verwijswwoorden, volgens Hiligsmann (2001 : 109), herhaaldelijk optreden, zijn de voorwaardelijke en veronderstellende bijzinnen zonder voegwoord. Voor dit type aanloop bepaalt hij niet of ze in een gesproken of geschreven context voorkomen maar hij preciseert dat ze tot een formeel register behoren. Een dergelijk register houdt een aantal beperkingen in en wordt gebruikt in een specifieke context. Hieruit kan worden afgeleid dat zulke aanloopelementen niet in grote aantallen aan bod komen.

Naast voorwaardelijke en veronderstellende bijzinnen voegt Hiligsmann de toegevende bijzinnen zonder voegwoord en de bijzinnen die een (temporele) tegenstelling uitdrukken bij de formele aanloopelementen (2001 : 109). Er kan aangenomen worden dat dit type linkerperiferie ook weinig voorkomt gezien hun restrictieve kader. De resterende bijzinnen, met name de (*ook*) *al*-bijzinnen en de toegevende bijzinnen met “hoe/wat/wie ... ook”, behoren niet tot een typerend register volgens de auteur. De *ANS* vermeldt echter dat deze toegevende zinnen voornamelijk in geschreven taal optreden in de gevallen waar het verwijswoord (meestal *toch*) afwezig is (1997 : 972).

Over de gebruikscontext van de voegwoordelijke bijwoorden, de samenvattende uitdrukkingen en de ‘Wat X betreft/aangaat’ constructie zegt de *ANS* niets. Misschien komen ze evenveel in gesproken als in geschreven taal voor.

De *ANS* geeft ook, geen verklaring voor interjecties en vocatieven. Toch is het interessant deze nader te bekijken. In zijn artikel verzet Hiligsmann zich tegen de *ANS* en legt uit dat tussenwerpsels en aansprekingen in de aanloop grotendeels in mondeling taalgebruik voorkomen (2001 : 109). Als we denken aan het eerste corpus, dat uit krantenartikelen bestaat, kunnen we veronderstellen dat deze artikelen in theorie geen interjecties en/of aansprekingen bevatten. Vocatieven worden gebruikt om zich tot iemand of tot een groep geadresseerden te

wenden. Behalve de lezers hebben de journalisten niemand aan te spreken. In bepaalde vaste rubrieken of in kranten met duidelijke politieke kleuren kan dat gebeuren maar vrijwel altijd hebben de dagbladschrijvers als doel feiten zo objectief mogelijk te melden. In deze gevallen hoeven ze niet de lezers aan te spreken en te overtuigen. Wat de interjecties betreft, het lijkt logisch dat deze meer geschikt zijn voor het gesproken taalgebruik maar de grammatica's die geraadpleegd werden geven daarover geen informatie. Uit een artikel van Norrick (2007 : 165) blijkt niettemin dat deze tendens reëel is. De auteur legt immers uit dat “interjections are characteristic of spoken language (with its utterances or intonation units) rather than written language (with its proper sentences)”. Kortom, tussenwerpsels zijn meer typerend voor gesproken taal dan voor schrijftaal gezien hun nauwe verband met de taaluitingen en de intonatie.

Nu al de mogelijke aanloopelementen de revue zijn gepasseerd, beschikken we over materiaal dat partieel het verschil in het aantal aanlopen tussen de twee data sets kan pogen verhelderen. Er is immers vastgesteld dat de meerderheid van de elementen die in de linkerperiferie kunnen staan hoofdzakelijk tot gesproken taal behoren: interjecties, vocatieven, elementen met een anticiperende functie. In tegenstelling daartoe komen toegevende bijzinnen met “hoe/wat/wie ... ook” zonder verwijzwoord in de eigenlijke zin vooral in schrijftaal voor. Essentieel is nu te bekijken hoe al deze constituenten in beide corpora verdeeld worden en of deze resultaten de hierboven voorgestelde interpretatie bevestigen.

Ook al bevat het *NRC* amper twee ingevulde aanlopen op 200 hits, er zij op gewezen dat één van deze twee aanlopen als bijwoordelijke bepaling van toegeving dient. Deze bepaling drukt een temporele tegenstelling uit en zou dus in een formele context teruggevonden moeten worden. Het is dus niet verwonderlijk een dergelijk geval terug te vinden in een krantenartikel dat deel uitmaakt van een liberaal dagblad dat zich bijgevolg tot een elite richt.

Betreffende het tweede corpus, er werd aangetoond dat de meeste aanloopelementen bestonden uit tussenwerpsels, nl. c. 72% (cf. tabel 7). Dit is in overeenstemming met onze voorafgaande bespreking volgens welke interjecties vooral in gesproken taal gebruikt worden. Bovendien verklaart dit het groter aantal aanlopen in het *Corpus Gesproken Nederlands*. Aangezien linkerperiferieën theoretisch vooral in gesproken taal zouden voorkomen, verwachten we dus grotendeels elementen zoals vocatieven, interjecties of elementen met een

anticiperende functie. Hoewel er geen aansprekingen in het corpus bestaan, is de hoeveelheid interjecties verhoudingsgewijs aanmerkelijk ten aanzien van de overige constituenten.

Onder de vijf aanloopelementen uit het *CGN* die een structurende functie vervullen, gaan er vier vooraf aan een verwijswoord in de eigenlijke zin. Dat telt voor c. 15% van alle aanlopen in het verzamelwerk. De meerderheid van de aanlopen in dit corpus oefenen een retorische functie uit. Ter herinnering: deze worden gebruikt om een aarzeling aan te geven of om het woord te nemen. Alweer is het geen verrassing dergelijke functies in gesproken taal tegen te komen aangezien aarzelingen en pauzes in dialogen gangbaar zijn. In gesprekken nemen de sprekers het woord om de beurt. Ze moeten dus de aandacht van hun gesprekspartner trekken voordat ze met hun zin beginnen. Daarvoor kunnen ze gebruik maken van interjecties. In geschreven taal, en in de schrijvende pers in het bijzonder, is dat zelden het geval. Het heeft immers geen zin om bijvoorbeeld aarzelingen in de definitieve tekst te bewaren want het doel van journalisten is feiten kort en bondig te rapporteren. Dit geldt ook voor de aanloopelementen met een retorische functie. Deze dienen een vraag te beantwoorden of op een uiting of een situatie te reageren. Ze komen tevoorschijn in situaties waarin verscheidene gesprekspartners betrokken worden. Dat betekent dus dat de communicatie niet per se mondeling hoeft plaats te vinden. In een monoloog bijvoorbeeld is er – normaal gezien – geen interactie, ook al is hij gesproken. De onderzochte sample uit het *CGN* voldoet niettemin aan de criteria doordat het om een discussie tussen spreker A en spreker B gaat. Daarom duiken regelmatig sommige aanloopelementen met een retorische functie op. Men moet ook niet uit het oog verliezen dat wat de functies van de constituenten aangaat, in het bijzonder de functies in de aanloop, is de grens tussen de interactieve en de retorische functie niet altijd glashelder en waren de keuzes dus subjectief getint.

Deze analyse en vergelijking tussen aanloopelementen in beide corpora stemt nogmaals overeen met de bovenstaande interpretatie betreffende het verschil in het aantal ingevulde aanlopen tussen het *NRC* en het *CGN*. In het corpus gesproken taal, waar ongeveer één clause op acht een linkerperiferie bevat, worden deze meestal ingevuld door tussenwerpsels – elementen die gekoppeld zijn met spreektaal. Qua functie dienen deze aanloopelementen vooral als retorische (c. 53%) en interactieve constituenten (c. 23%). Zulke pragmatische functies komen grotendeels in situaties voor waar verscheidene gesprekspartners rechtstreeks betrokken zijn, namelijk eerder de gesprekken uit het tweede corpus dan de krantenartikelen uit het eerste. Al is dit eerste corpus meer dan beperkt qua aanlopen (amper 2 gevallen), we

kunnen toch vaststellen dat deze geen interactieve of retorische functie vervullen. Het eerste voorbeeld fungeert immers als een bepaling van toegeving – ze maken dus deel uit van het ideationele domein, aldus Crible (2017 : 107) – terwijl de tweede aanloop behoort tot het structurerende domein. Er kan verwacht worden dat dergelijke sequentiële constituenten vaker dan retorische of interpersonele items in krantenberichten gebruikt worden. Het is immers van groot belang de verslaggevers hun tekst structureren om verbanden te leggen tussen zinnen en alinea's onderling en zo een rode draad te bewaren.

5.3 Eerste zinsplaats

Nu de aanlopen onder de loep zijn genomen, is het tijd om de voorvelden nader te bekijken. Uit tabel 14 bleek dat de eerste zinsplaatsen in beide corpora flink ongelijk zijn vanuit een vormelijk standpunt. In schrijftaal bestaan eerste stukken voor precies de helft uit naamwoordelijke constituenten. Voornaamwoorden komen regelmatig voor (13,5%) maar niettemin zijn deze aantallen ver van vergelijkbaar met die van de nominale groepen. Door in overvloed gebruik te maken van NC's op de eerste zinsplaats verwijzen schrijvers rechtstreeks naar hun topics. Semantisch worden deze onmiddellijk bekend gemaakt. Dit kan verwarring voorkomen en dus de overdracht van de boodschap vergemakkelijken. deze nominale groepen vooral uit eigennamen of definiete constituenten bestaan. Volgens de grammatica's staan onbepaalde NC's bijna nooit voorop in de zin om functionele redenen.

Opvallend is ook de frequentie waarmee voorzetselconstituenten zich in het eerste stuk bevinden in het *NRC*. Na de Nominale constituenten vormen de VC's de tweede frequentste syntactische categorie die vlak voor de eerste pool staat, op gelijke voet met de pronomina. Door die VC's nader te bekijken, kunnen we beseffen dat ze in het eerste corpus bijna uitsluitend dezelfde grammaticale functie hebben, namelijk bijwoordelijke bepaling (25 op 27). De *E-ANS* beschrijft deze als elementen die “een nadere omschrijving (dat wil zeggen een beperking of letterlijk: bepaling) bij het door het gezegde uitgedrukte [geven]” (Coppen, Haeseryn, de Vriend: 2002). Bepalingen houden zich dus bezig met het creëren van een context door extra informatie te geven. Ze stellen een kader vast waarin de hoofdhandeling plaatsvindt. In krantenverslagen is dit kader van uiterst belang want het biedt de mogelijkheid om de gebeurtenissen ruimtelijk en temporeel te plaatsen, de inzet te verklaren enz. Nu is de vraag: waarom staan deze bepalingen in het eerste stuk en niet in het middenstuk? In hoofdstuk 1.4.

is er vermeld dat de elementen die de focus krijgen in het Nederlands zo ver mogelijk achteraan zouden moeten staan. Door de bepalingen aan het begin van de zin te zetten wordt er vermeden met de handelingen uit het middenstuk samen te vallen. Dat kan voor een deel verhelderen waarom elementen met een bepalende functie regelmatig het voorveld bezetten in het corpus schrijftaal (25%).

Niet verrassend fungeren de meeste elementen op de eerste zinsplaats toch als onderwerp (61%). Dit was immers al het geval in Verscheldens (2015), waar c.57% van de eigenlijke zinnen met het onderwerp begonnen. De ANS (1997 : 937) legt uit dat ieder zinsdeel in theorie in het voorveld kan staan, naargelang de informatiestructuur van de zin. De constituenten die meestal vooraan in de zin staan zijn degene die al bekend zijn en die bepaald zijn (Van Belle et al. 2006 : 12). Krantenartikelen gaan vrijwel altijd gepaard met hoofdtitels, die aanwijzingen geven met betrekking tot de inhoud van de tekst, in het bijzonder het agens – degene die de handeling verricht. Ze maken dus de informatie uit het artikel op voorhand bepaald en/of bekend voor de lezers. In het artikel kan deze – nu – bekende informatie vóór de persoonsvorm opduiken. De grote aanwezigheid van onderwerpen op de eerste zinsplaats in het Nederlands is niettemin een al lang vastgesteld en verklaard verschijnsel.

Constituenten met een structurerende functie zijn misschien de syntactische functie die we in het eerste stuk meer verwacht hadden in het *NRC Handelsblad*. In sectie 5.2. werd de verdeling van de elementen in de aanloop besproken. Daarin hebben we voorgesteld dat sequentiële items gangbaar zouden zijn gezien hun vermogen de tekst te structureren. Terwijl informele gesprekken spontaan en ordeloos zijn, hebben journalisten een stuk meer tijd om hun artikel te plannen. Ze kunnen alle middelen aanwenden om hun tekst zo vlot mogelijk te maken. Eén van deze middelen is het gebruik van connectieven die verbanden leggen tussen de zinnen en de alinea's. Belangwekkend zouden de resultaten van een grootschalig corpusonderzoek zijn om na te gaan of structurerende elementen ook zo weinig in het voorveld voorkomen.

Vanuit een formeel standpunt waren de gegevens uit het *Corpus gesproken Nederlands* erg verschillend van het corpus geschreven taal. Enerzijds stonden er NC's in 50% van de gevallen, anderzijds komen voornaamwoorden het vaakst voor, te weten in 111 gevallen op 200 (60 persoonlijke voornaamwoorden en 51 aanwijzende voornaamwoorden). Dit stemt volledig overeen met de stelling van Van Belle et al. (2006 : 12), waarin de linguïsten de pronomina zetten onder de elementen die bij voorkeur op de eerste zinsplaats staan.

Gepronominaliseerde elementen zijn immers bekende informatie en kunnen dus natuurlijk aan het begin van de zin optreden. Waarom komen dan voornaamwoorden zo weinig voor in voorvelden in het eerste corpus? En waarom zijn NC's aanzienlijk minder frequent in het corpus spreektaal (11,5%)? Pronomina gebruiken maakt het mogelijk om hele naamwoordelijke constituenten niet te hoeven herhalen. Ze bevorderen zo een korte en snelle overdracht van het bericht in plaats van de zekerheid rechtstreeks begrepen te worden. In gesprekken wisselen de rollen van zender en ontvanger altijd en treden deze twee in interactie. De spreker bewaart de mogelijkheid om zijn voorafgaande uiting te verklaren of te corrigeren in geval van onbegrip. Verslaggevers uit de geschreven pers hebben deze gelegenheid niet en moeten zich dus vergewissen van de bevattelijkheid van hun stuk. Daarom maken ze misschien minder gebruik van voornaamwoorden ten gunste van de NC's die het voordeel hebben dat ze naar een concreet begrip verwijzen. Pronomina integendeel vervangen en verwijzen naar de nominale groepen in kwestie.

Net zoals in het *NRC* dienen de voorzetselconstituenten in het tweede corpus als bepalingen die de uiting contextualiseren. In het *CGN* zijn ze echter minder in aantal (3,5%). Ze worden op de tweede plaats vervangen door de Bijwoordelijke constituenten (BC), die in ietwat een kwart van de hits het eerste stuk bezetten. Ook zij vervullen voornamelijk de functie bijwoordelijke bepaling. Bij nadere beschouwing blijkt dat de BC's die in het voorveld staan het vaakst uit **één woord** bestaan. Dit fenomeen geldt voor beide corpora: 17 op 20 in de krantenberichten en al de 48 gevallen in de dialogen. In tegenstelling daartoe bevatten de VC's over het algemeen meer woorden: In gesproken taal bevatten de voorzetselconstituenten gemiddeld 3,29 woord. In geschreven taal stijgt dit gemiddelde tot 5,44 woord per constituent, met een piek van 18 woorden. Op basis van deze uiterst beperkt sample kunnen er toch twee verschijnsels vastgesteld worden. Ten eerste, voor dezelfde functie gebruikt spreektaal liever een grammaticale categorie die de hantering van minder woorden vereist. Ten tweede, voor dezelfde syntactische categorie en dezelfde functie, de eerste stukken uit het *CGN* lijken minder woorden te vereisen dan de krantenartikelen van het *NRC*. Dat lijkt het idee te bevestigen dat discussie zich meer op een bondige communicatie focust dankzij een reeks syntactische technieken.

Het laatste een tikkeltje betekenisvolle bespreking ten opzichte van de eerste zinsplaats in het corpusonderzoek betreft lijdend voorwerpen. Terwijl deze in het *NRC* nauwelijks aan bod komen (3,5%), komen ze wat regelmatiger voor in het *CGN* (14%). Door alweer de vorm

samen met de functie te examineren beseffen we dat 18 op 28 gesproken lijdend voorwerpen voorzetselconstituenten zijn. In tegenspraak daarmee bestaat uitsluitend één direct object uit een voorzetselconstituent in het corpus schrijftaal. Dat lijkt weer aan te sluiten bij de theorie van Van Belle et al. dat “persoonlijke en demonstratieve pronomina van nature geschikt [zijn] om als thema te fungeren.” (2006 : 12).

5.4 Discussie

Deze bespreking van de resultaten van het corpusonderzoek met betrekking tot het gebruik van de aanloop en het eerste stuk in gesproken en geschreven taal lijkt met de hoofdstellingen van de theoretische naslagwerken overeen te stemmen. De hoofdpunten van de interpretatie worden snel hieronder de revue gepasseerd: aanlopen lijken grotendeels in gesproken taal gebruikt te worden aangezien verscheidene vaak voorkomende constituenten die in de linkerperiferie kunnen staan karakteristiek tot spreektaal behoren (elementen met een anticiperende functie, tussenwerpsels, vocatieven). De frequentste functie van de aanloopelementen kan beïnvloed worden ten opzichte van het communicatiekanaal, d.i. geschreven of gesproken. In geschreven taal kan er verwacht worden dat de schrijvers meer gebruik maken van structurerende constituenten die samenhang binnen de communicatie brengen. In spreektaal vervullen de aanloopelementen daarentegen meestal interactieve en retorische functies, over het algemeen in situaties waar meerdere gesprekspartners betrokken zijn.

Wat de voorvelden betreft, er werd vastgesteld dat deze in meer dan de helft van de gevallen als onderwerpen dienden. Het eerste stuk is een plaats bij uitstek voor onderwerpen. Volgens het links-rechts-principe hebben de elementen die vooruit in de zin voorkomen over het algemeen minder communicatieve waarde dan de zinsdelen die achteruit staan. Onderwerpen geven aan wie de handeling verricht en zijn vaak al bekend. Bijgevolg zijn ze minder belangrijk vanuit een communicatief standpunt. Doorgaans is de handeling zelf die gefocust wordt. Deze bekende, herhaalde onderwerpen worden dan vaak door voornaamwoorden vervangen, aldus Herring (1990 : 164). Daarom vinden we enorm veel pronomina op de eerste zinsplaats, meer bepaald in spreektaal. Dit kan uitgelegd worden door een potentiële neiging om onze zinnen mondeling te verkorten en te vereenvoudigen. Het kan ook verklaren waarom zoveel voorveldelementen uit maar één woord bestonden in het CGN.

In – formele – geschreven taal zoals krantenartikelen is het vanzelfsprekend essentieel begrepen te worden. Daarom kunnen de schrijvers vaker nominale constituenten gebruiken om meer nuances en nauwkeurigheid te krijgen.

Conclusie

Kortom, uit dit kleinschalig corpusonderzoek betreffende de aan- en afwezigheid en de functie van de aanloop en de eerste zinsplaats in mededelende Nederlandse zinnen blijkt dat de praktijk in grote lijnen overeenkomt met de theorie. Verder werd er een poging gedaan om de processen die de basis vormen voor deze theorie onder de loep te nemen en op te vatten, met name de divergenties tussen schrijftaal en spreektaal. Er werden dan een reeks verschijnsels bepaald die potentieel een invloed hebben op het voorkomen en op het gebruik van aanloop- en voorveldelementen.

Nu daagt de volgende vraag op: Wat zijn de elementen die een concreet effect hebben op het preverbale segment van de Nederlandse zin? Wegens de beperktheid van de twee onderzochte corpora en het aantal variabelen zijn de invloedrijke fenomenen erg complex vast te stellen. Er werd bijvoorbeeld in de interpretatie vermeld dat de eerste stukken van het *NRC* redelijk veel voorzetselconstituenten bevatten. Daarna werd er vastgesteld dat vrijwel al deze VC's als bijwoordelijke bepalingen dienden. In het *CGN* integendeel komen weinig VC's op de eerste zinsplaats voor. Toch oefenen deze meestal ook de syntactische functie bijwoordelijke bepaling uit. In het corpus spreektaal bevinden zich daarentegen meer bijwoordelijke constituenten in de voorvelden. Net zoals de voorzetselconstituenten, spelen ook de BC's grotendeels de rol van bijwoordelijke bepalingen. Parallel werd er opgemerkt dat het aantal woorden bij de VC's en de BC's zeer constant bleef voor beide corpora. De meerderheid van de bijwoordelijke constituenten bestonden immers uit één woord (65 op 68 voor alle hits) terwijl voorzetselconstituenten gemiddeld 3,29 woorden in het *CGN* en 5,44 woorden in het *NRC* bevatten. Welke factor heeft dan de voornaamste invloed op de eerste stukken van beide datasets? Is het de grammaticale categorie, de grammaticale functie of het aantal woorden van het zinsdeel? Natuurlijk kunnen deze factoren elkaar beïnvloeden. In functie van de context kunnen ze zich misschien tegen elkaar verzetten of integendeel elkaar goed aanvullen.

Concluderend zou het bijzonder wetenswaardig zijn om dit onderzoek uit te breiden naar een groter aantal corpora en met extra variabelen rekening te houden (bv. meer hits per corpus, corpora van verschillende genres en stijlen, meer gesprekspartners, partners van verschillende leeftijden, uit diverse sociale achtergronden, enz.). Door meer gegevens vanuit verscheidene standpunten te analyseren krijgen we meer kansen om de middelen nader te

bepalen die concrete effecten hebben op de aanlopen, de eerste zinsplaatsen en hun respectieve functies. In verband daarmee zou een verder onderzoek naar nevenschikkende voegwoorden heel nuttig zijn, aangezien veel grammatica's en wetenschappelijke artikelen tegengestelde meningen hebben over NSVW, in het bijzonder over hun plaatsing binnen of buiten de zin. Misschien kan een dergelijk naslagwerk het verband tussen de aanloop en het eerste stuk gedeeltelijk verklaren.

Bibliografie

Coppen, P.A., Haeseryn, W., de Vriend, F. (2002). Bijwoordelijke bepalingen. Inleiding [Online]. De Elektronische ANS. Toegankelijk op : <<http://ans.ruhosting.nl/e-ans/20/10/01/body.html>>. [Geraadpleegd op 22 april 2019]

Crible, L. (2017). . Towards an operational category of discourse markers. *Pragmatic Markers, Discourse Markers and Modal Particles: New perspectives*, 186, 99.

Corpus Gesproken Nederlands. Toegankelijk op <http://lands.let.ru.nl/cgn/>

Corpus Nieuwe Rotterdamse Courant. Toegankelijk op <https://ivdnt.org/producten/vu-dnc/index.html>

Coussé, E. (2008). *Motivaties voor volgordevariatie. Een diachrone studie van werkwoordvolgorde in het Nederlands* (Doctoral dissertation, Ghent University).

Evers-Vermeul, J. (2010). Dus vooraan of in het midden?-Over vorm-functierelaties in het gebruik van connectieven. *Nederlandse taalkunde*, 15(2), 149-175.

Haeseryn, W. J. M., Romijn, K., Geerts, G., Rooij, J. D., & Van den Toorn, M. C. (1997). *Algemene Nederlandse Spraakkunst*.

Herring, S. C. (1990, August). Information structure as a consequence of word order type. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 16, No. 1, pp. 163-174).

Hiligsmann, P. (2001). De ideale NVT-grammatica voor Franstaligen, hoe zou die de aanloop moeten aanpakken?. *Contrastief onderzoek Nederlands-Frans*, 108, 97.

Kuiper, K., Bimesl, N., Kempen, G., & Ogino, M. (2017). Initial vs. non-initial placement of agent constructions in spoken clauses: A corpus-based study of language production under time pressure. *Language Sciences*, 64, 16-33.

Lambrecht, K. (1996). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents* (Vol. 71). Cambridge university press.

Magnus, I. (2008). The position of adverbials and the pragmatic organization of the sentence. *Languages and Cultures in Contrast and Comparison*, 175, 89.

Norrick, N. R. (2007). Discussion article: Pragmatic markers, interjections and discourse. *Catalan Journal of Linguistics*, 6, 159-168.

Ross, J. R. (1967). Constraints on variables in syntax.

Schiffrin, D. (1988). *Discourse markers* (No. 5). Cambridge University Press.

Theissen, S. (1984). De semantisch-syntactische valentie van enkele Connecterende Adverbiale Partikels (CAP's). *Studies over Nederlandse Partikels, Antwerp Papers in Linguistics*, (35), 90-101.

Van Belle, W., Lamiroy, B., Van Langendonck, W., Lahousse, K., Lauwers, P., Canegem-Ardijns, V., & Van Goethem, K. (2011). *Nederlandse grammatica voor Franstaligen*.

Vandeweghe, W. (2013). *Grammatica van de Nederlandse zin*. Maklu.

Van Kampen, J. (2010). Anaforische middelen voor topicverschuiving. *Nederlandse taalkunde*, 15(2), 189-209.

Van Valin, R. D., Robert, D., van Valin Jr, R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.

Verbanck, P. (1985). *Beknopte nederlandse grammatica*. Uitgeverij de Nederlandsche Boekhandel.

Appendix

De corpora zijn toegankelijk op de website:

https://drive.google.com/file/d/1LzJcLWZ8_RHM8hrdrSlo7VjzPBHSJ1Mf/view?usp=sharing

g

Ze werden ook per e-mail verstuurd en samen met deze scriptie op de website

<http://dial.uclouvain.be/memoire/ucl/> geüpload.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté de philosophie, arts et lettres

Place Blaise Pascal, 1 bte L3.03.11, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/fial