

Faculteit der Letteren
Taalwetenschap

Studiejaar 2016-2017
11-07-2017

Discontinue Acl's

*Een corpusstudie naar de woordvolgorde in de
Latijnse Accusativus cum infinitivo*

Iris Faber

Inhoudsopgave

1. Inleiding	3
2. Accusativus cum Infinitivo (AcI)	5
2.1 AcI vs. A+I	5
2.2 Structurele analyses AcI	8
2.3 Conclusie	11
3. Woordvolgorde	13
3.1 Vrije woordvolgorde?	13
3.2 Discontinuïteit	15
3.3 Discontinue AcI's en A+I's	16
4. Methode	19
4.1 Materiaal	19
4.2 Onderzoeksopzet	21
4.3 Statistische analyse	24
5. Resultaten	26
5.1 Algemene analyse	26
5.2 Analyse AcI	26
5.3 Analyse A+I	30
6. Discussie	33
6.1 Conclusies en discussie AcI	33
6.2 Conclusies en discussie A+I	36
6.3 Algemene conclusies en discussie	37
7. Referenties	38

1. Inleiding

In het Latijn wordt gebruik gemaakt van zes naamvallen, waarvan de meest voorkomende de nominativus en de accusativus zijn. Deze worden meestal gebruikt om respectievelijk het subject en het object van een zin te markeren, zoals in (1).

- (1) *Luci-us* *vide-t* *Marc-um*
Lucius-NOM zien-3.SG Marcus-ACC
'Lucius ziet Marcus.'

De accusativus wordt echter niet alleen maar gebruikt voor objecten. In zinnen zoals (2) wordt de accusativus gebruikt om het subject van *venire* 'komen', *Marcus*, aan te duiden. Wanneer dit gebeurt hebben we te maken met complexe zinnen waar het accusatieve subject een onderdeel is van de ingebedde zin. Deze bijzin bestaat uit een nominaal deel in de accusativus en een werkwoord in de infinitieve vorm, en wordt daarom ook de *accusativus cum infinitivo* (AcI) genoemd.

- (2) *Luci-us* *dici-t* *Marc-um* *venire*
Lucius-NOM zeggen-3.SG Marcus-ACC komen.INF
'Lucius zegt dat Marcus komt.'

De AcI kan gevonden worden na werkwoorden die een waarneming (*verba videndi*), een gevoel of een mening (*verba sentiendi*), of een mededeling (*verba dicendi*) uitdrukken, en na *iubēre* 'bevelen', *vetare* 'verbieden' en *sinēre* 'toelaten'. De accusativus heeft de syntactische rol van subject bij het infinitieve werkwoord, dit gebeurt in 9% van de gevallen (Pinkster, 2015)¹.

In zin (2) is een AcI weergegeven waarbij het accusatieve naamwoord en het infinitieve werkwoord direct naast elkaar staan. De twee elementen kunnen echter ook uit elkaar komen te staan, wanneer er een element uit de matrixzin tussenin staat, zoals in zin (3).

- (3) [...], *homin-em quidem scire* *arbitr-or* *neminem*
 mens-ACC echter weten.INF denken-1.SG geen.enkele-ACC
'[...], ik denk dat geen enkel mens het echter weet.' (Cicero *Tusc. 1.99*, uit Spevak (2010))

¹ Het corpus dat gebruikt werd voor de analyse bestond uit: Cicero *Oratoria pro Milone* 1-15, *Epistulae ad Atticum* 1.1-5; Caesar *De bello Gallico* 1.1-10; Nepos *Atticus* 1-9; Plinius maior *Naturalis historia* 2.1-28; Plinius minor *De morte Pereg* deel 1, deel 2; Plautus *Captivi* 1-210; Vergilius *Aeneis* 1.1-1.183; Horatius *Carmina* 4.1-7; Ovidius *Ars amatoria* 1.1-198; Seneca *Phaedra* 1-235. Niet-uitgedrukte subjecten zijn meegeteld.

Het MATRIXPREDICAAT, *arbitror*, staat hier tussen de twee delen van de AcI – *hominem quidem scire* en *neminem* – in, waardoor de woordvolgorde in deze zin anders is dan in zin (2).

Er is nog niet veel onderzoek gedaan naar AcI's waarbij een element uit de hoofdzin tussen delen van de constructie in staat. Ik heb daarom een corpusstudie uitgevoerd om te achterhalen welke elementen hierop van invloed zijn. Voor het onderzoek is gebruik gemaakt van data uit de PROIEL treebank van Haug & Jøhndal (2008). Deze data komt uit werken van twee verschillende auteurs: *Epistulae ad atticum* van Cicero en *Commentarii de bello Gallico* van Caesar.

In hoofdstuk 2 zal ik verschillende structurele analyses van de AcI bespreken. Het is echter belangrijk om eerst te bespreken hoe de AcI zich gedraagt in verschillende contexten. Aan de hand van enkele voorbeelden zal ik in paragraaf 2.1 laten zien dat de term 'AcI' gebruikt wordt voor twee verschillende structuren. Daarnaast zal duidelijk worden waarom het belangrijk is om deze structuren van elkaar te onderscheiden. Vervolgens ga ik in hoofdstuk 3 in op de woordvolgorde in het Latijn, en de notie 'discontinuïteit', met name met betrekking op de AcI. In hoofdstuk 4 ga ik in op de methodiek van het uitgevoerde onderzoek en de opgestelde modellen, waarvan ik in hoofdstuk 5 de resultaten geef. Tot slot laat ik in hoofdstuk 6 zien wat deze resultaten betekenen voor het onderzoek, en voor de bestaande theorieën over de AcI.

2. Accusativus cum infinitivo

2.1 ACI vs. A+I

Bolkestein (1976a) beargumenteert dat de term ‘AcI’ eigenlijk meer de morfologische dan syntactische structuur van de bijzin benoemt. Er zijn namelijk twee verschillende syntactische structuren die beide uit een naamwoord in de accusativus en een infiniet werkwoord bestaan. Aan de hand van zin (1) en zin (2), wordt duidelijk in welk opzicht de structuren van elkaar verschillen, en welke gevolgen dat heeft voor het gedrag van de constructie in een zin.

Vooraf is het belangrijk om op te merken dat ik zowel de syntactische functie van woorden betrek in de bespreking, als de semantische functie. Om dit onderscheid duidelijk te maken volg ik het voorbeeld van Bolkestein, door de syntactische functies in hoofdletters te noteren (e.g. SUBJECT, OBJECT, COMPLEMENT, PREDICAAT) en de semantische functies telkens met een hoofdletter te beginnen (e.g. Agens, Patiens, Geadresseerde).

(1) *dic-unt me venire*
zeggen-3.PL 1.SG.ACC komen.INF
‘Ze zeggen dat ik kom.’

(2) *cog-unt me venire*
dwingen-3.PL 1.SG.ACC komen.INF
‘Ze dwingen mij te komen.’

In zin (1) zien we een ‘echte’ AcI; de accusativus en de infinitivus vormen hier één syntactisch geheel in de matrixzin. De structuur in zin (2) bestaat uit een los accusatief complement bij het MATRIXPREDICAAT *cogunt* en een aanvullingsinfinitivus, en zal ik daarom vanaf nu ook geen *accusativus cum infinitivo* meer noemen. In plaats daarvan volg ik het voorbeeld van eerdere auteurs (o.a. Pinkster, 1984; Schoof, 2004) en noem ik de tweede structuur een *accusativus + aanvullingsinfinitivus*, oftewel *A+I*.

Het is echter belangrijk om op te merken dat niet alle auteurs dit onderscheid maken. Zeker in oudere literatuur (vóór Bolkestein, 1976a) worden beide termen vaak aangeduid met de term ‘AcI’. In de volgende paragraaf zal dit zeker van belang zijn, omdat door onderzoekers soms verschijnselen worden aangehaald die te verklaren zijn door het onderscheid tussen de AcI en A+I.

Bolkestein vond meerdere argumenten voor het onderscheiden van zinnen (1) en (2). Veel van deze argumenten hebben te maken met de grammaticaliteit of welgevormdheid van de zinnen. Het probleem is echter: hoe bepaal je of een Latijnse zin wel of niet grammaticaal is? Er zijn geen moedertaalsprekers van het klassiek Latijn meer, dus voor het bepalen van

grammaticaliteit moeten we gebruik maken van andere materialen. Het ligt voor de hand om te proberen constructies te vinden in teksten, maar ook als je iets niet vind kun je er niet vanuit gaan dat het niet bestond. Bolkestein heeft daarom een thesaurus en lexica geraadpleegd en gebruik gemaakt van haar eigen intuïtie om te bepalen of een constructie ongrammaticaal is, of toevallig niet voorkomt in teksten. Er blijft echter ruimte voor fouten, dus onderstaande observaties worden niet als feiten gepresenteerd, maar als hypothesen.

Als eerste zien we een verschil in de structuren wanneer we het infinitieve werkwoord uit de zin verwijderen. De accusativus in een A+I kan zelfstandig voorkomen; de zin blijft dus grammaticaal wanneer het infinitieve werkwoord wordt verwijderd uit de zin. Dit is te zien in zin (4), waar *me* zonder na *cogere* zelfstandig voorkomt. Dit lijkt een 3-plaatsige structuur van (2) te suggereren: de accusativus is daar het OBJECT van *cogunt*, en vervult daarmee een onafhankelijke syntactische functie. Omdat het niet afhankelijk is van *venire* wordt de zin niet ongrammaticaal als *me* zelfstandig voorkomt. Bij een AcI leidt het verwijderen van het infinitieve werkwoord wel tot een ongrammaticale zin, zoals te zien is in zin (3). Wanneer we de accusativus zouden behandelen als het OBJECT van het MATRIXPREDICAAT, zou hetzelfde op moeten gaan als bij de A+I en zou de ongrammaticaliteit van (3) niet verklaard worden. Een 2-plaatsige analyse van zin (1) verklaart dit wel: als *me* wordt gezien als het SUBJECT van *venire*, en daar dus van afhankelijk van is, zou het verwijderen van *venire* leiden tot ongrammaticaliteit.

(3) **dic-unt me*
 zeggen-3.PL 1.SG.ACC
 *‘Ze zeggen dat ik.’

(4) *cog-unt me*
 dwingen-3.PL 1.SG.ACC
 ‘Ze dwingen mij.’

Bij de verba dicendi zien we een tweede verschil. Zowel bij de AcI's als A+I's zijn er werkwoorden die de productie van spraak aanduiden, zoals respectievelijk *dicere* en *admonere* in (5) en (6). Je zou daarom verwachten dat er een Geadresseerde toe kan worden gevoegd aan de zinnen (1) en (6). Bij de AcI is dit inderdaad het geval, maar bij de A+I niet, zoals te zien is in de zinnen (5) en (7). Omdat er maar één constituent in de zin kan zijn die de rol van Geadresseerde vervult, kunnen we concluderen dat zin (7) ongrammaticaal is, omdat in (6) al een Geadresseerde genoemd wordt: *me*. Het feit dat zin (5) wel grammaticaal is, laat zien dat *me* hier niet de rol van Geadresseerde vervult. De accusatieven in de AcI en de A+I hebben dus verschillende rollen in de zinsstructuur.

(5) *dic-unt ei me venire*
 zeggen-3.PL 3.SG.DAT 1.SG.ACC komen.INF
 ‘Ze zeggen aan hem dat ik kom.’

- (6) admon-ent me venire
aanmanen-3.PL 1.SG.ACC komen.INF
'Ze manen mij aan te komen.'
- (7) *admon-ent ei me venire
aanmanen-3.PL 3.SG.DAT 1.SG.ACC komen.INF
**'Ze manen aan hem mij aan te komen.'*

Vervolgens zien we een tweetal restricties op het gebruik van de infinitivus bij de A+I. De eerste restrictie betreft de tijd van de infinitivus: bij een A+I kan de infinitivus slechts in de tegenwoordige tijd staan; een infinitivus in de toekomstige of verleden tijd, zoals zin (9) is ongrammaticaal. Bij de AcI kan het PREDICAAT wel in een andere tijd gezet worden, zoals te zien is in zin (10). De tweede restrictie betreft de mogelijkheid tot passivisatie van de infinitivus: opnieuw is dit bij de AcI wel mogelijk (zoals in zin (10)), maar bij een A+I niet (zin (11)). Je kunt dit samen zelfs nog breder benoemen, zoals Pinkster (1984) doet: de infinitivus van een A+I moet beheersbaar zijn. Met 'beheersbaar' doelt hij erop dat de gebeurtenis controleerbaar (door het subject) moet zijn. Een passief werkwoord is niet controleerbaar, evenals een predicaat in de verleden of toekomstige tijd, dus vormen (9) en (11) een ongrammaticale zin. Ook zinnen als (12) zijn hierom ongrammaticaal: het PREDICAAT 'omkomen' is niet beheersbaar.

- (8) *dic-unt me ven-isse*
ZEGGEN-3.PL 1.SG.ACC KOMEN-INF.PERF
'Ze zeggen dat ik ben gekomen.'
- (9) **cog-unt me ven-isse*
dwingen-3.PL 1.SG-ACC komen-INF.PERF
**'Ze dwingen mij te zijn gekomen.'*
- (10) *dic-unt me mitt-i*
zeggen-3.PL 1.SG.ACC zenden-INF.PASS
'Ze zeggen dat ik gezonden word.'
- (11) **cog-unt me mitt-i*
dwingen-3.PL 1.SG-ACC zenden-INF.PASS
**'Ze dwingen mij te worden gezonden.'*
- (12) **cog-unt me perire*
dwingen-3.PL 1.SG-ACC omkomen.INF
**'Ze dwingen mij om om te komen.'*

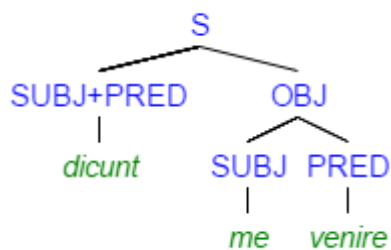
Er zijn niet alleen restricties te zien op de infinitivus, maar ook op de accusativus. Zoals in (14) te zien is kunnen niet zomaar alle accusativi het object worden van een infinitivus in een A+I.

De reden hiervoor heeft te maken met de ontdekking die we eerder de paragraaf hebben gedaan: de accusativus is de Geadresseerde. Als we dat feit erbij betrekken wordt het goed verklaard dat er restricties op de accusativus van toepassing zijn: het moet iets zijn dat de rol van Geadresseerde kan vervullen. Dit betekent dat de accusativus levend moet zijn (Pinkster, 1984), of specifieker: menselijk (Bolkestein, 1976a). Omdat de accusativus deze rol niet heeft bij de AcI, geldt deze restrictie daar ook niet. Een zin als (13) is daarom wél grammaticaal.

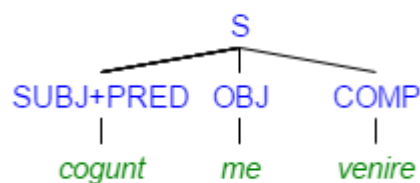
(13) *dic-unt curr-us venire*
 zeggen-3.PL wagen-ACC komen.INF
 ‘Ze zeggen dat de wagen komt.’

(14) **cog-unt curr-us venire*
 dwingen-3.PL wagen-ACC komen.INF
 *‘Ze dwingen de wagen te komen.’

De AcI en de A+I hebben dus duidelijk twee verschillende onderliggende structuren, waarop verschillende ‘regels’ van toepassing zijn. Veel auteurs concluderen in hun analyse dat voor het MATRIXPREDICAAT *dicere*, die een AcI oproept, er een 2-plaatsige structuur moet zijn (figuur 1) en een 3-plaatsige structuur voor het MATRIXPREDICAAT *cogere*, die een A+I bij zich heeft, (figuur 2). In de rest van deze scriptie wordt ook uitgegaan van deze analyses van de AcI en de A+I.



Figuur 1: Boomstructuur (3) ‘dicunt me venire’



Figuur 2: Boomstructuur (4) ‘cogunt me venire’

2.2 STRUCTURELE ANALYSES ACI

Veel traditionele grammatici analyseerden de AcI als de 2-plaatsige structuur in figuur 1. Ze presenteerden de hypothese van een diachrone ontwikkeling die de AcI als eindresultaat had. Zo gingen Kühner & Stegmann (1912), Leumann, Hofmann & Szantyr (1965) en Woodcock (1959) uit van het idee dat de AcI in eerste instantie geen syntactische eenheid was. Ze meenden dat er in het Latijn werkwoorden geweest zijn die zowel een losse accusativus als

complement kunnen nemen, als een accusativus én een infinitivus. Na verloop van tijd werd de accusativus steeds minder geassocieerd met het matrixwerkwoord, maar werd het meer gezien als het subject van de infinitivus. De accusativus en infinitivus zijn dus door de tijd heen tot een syntactische eenheid geworden. Toen dit proces eenmaal voltooid was kon de AcI ook als complement voorkomen bij werkwoorden geen losse NP's als object nemen, zoals (15).

- (15) *dic-unt me venire*
 ZEGGEN-3.PL 1.SG.ACC KOMEN.INF
 'Ze zeggen dat ik kom.'

Vaak zijn voor deze uitbreidingsanalyses door verschillende auteurs ongeveer dezelfde argumenten gegeven.

Om te beginnen zou een uitbreidingshypothese zowel de ongebruikelijke accusatieve naamval van het nomen als de infiniete vorm van het werkwoord verklaren. Als we aannemen dat de accusativus ergens in de geschiedenis als OBJECT gefunctioneerd heeft, zou de naamval ervan al een stuk logischer worden; de accusativus is per slot van rekening de gebruikelijke naamval voor OBJECTEN. Bovendien zijn complementen meestal infiniet (Bolkestein, 1979), dus ook de vorm van het werkwoord is met deze hypothese verklaard.

Nu werd de uitbreidingshypothese niet alleen gezien als handig, maar ook als waarschijnlijk: in het klassiek Latijn bestonden er namelijk nog steeds PREDICATEN die zowel een 2-plaatsige als 3-plaatsige structuur kunnen aannemen, stellen de auteurs. De 'verba sentiendi' kunnen namelijk voorkomen met zowel een los object als met een AcI, zoals te zien is in (16) en (17).

- (16) *fili-am audio*
 meisje-ACC horen-1.SG
 'Ik hoor het meisje.'
- (17) *fili-am cantare audio*
 meisje-ACC zingen.INF horen-1.SG
 'Ik hoor dat het meisje zingt.'

Tot slot werd naar passieve zinnen gekeken voor onderbouwing van de hypothese. In zin (18) is een voorbeeld gegeven van een persoonlijk passief. We zien dat de accusativus het onderwerp kan worden van een passief PREDICAAT. Het ligt daarom voor de hand om te denken dat deze ergens in de structuur een object is geweest. Bovendien kan de AcI in zijn geheel ook deelnemen aan passieve zinnen, zoals in (19). De AcI is hier in zijn geheel het subject van *dicitur* 'er wordt gezegd'. Dat spreekt duidelijk voor enige samenhang tussen de accusativus en de infinitivus.

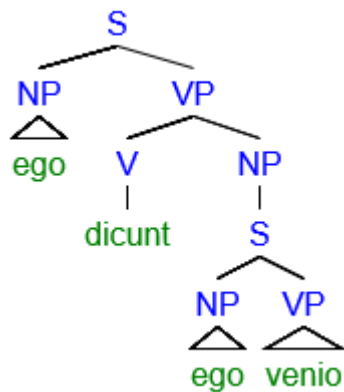
- (18) *vir-i ven-isse dic-untur*
 man-PL.NOM komen-INF.PERF zeggen-PASS.3.PL
 'De mannen worden gezegd te zijn gekomen.' / 'Men zegt dat de mannen zijn gekomen.'

- (19) vir-os ven-isse dic-itur
man-ACC.PL komen-INF.PERF zeggen-PASS.3.SG
'Er wordt gezegd dat de mannen zijn gekomen.' / 'Men zegt dat de mannen zijn gekomen.'

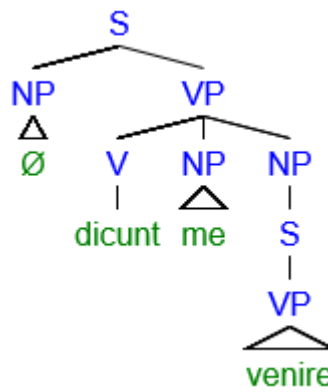
Bolkestein (1979) laat echter zien dat deze hypothese niet alles verklaart, en dat er bovendien voldoende bewijs is om deze te verwerpen. Zo hebben we in de vorige paragraaf al gezien dat er voldoende bewijs is om een structureel verschil tussen de AcI en de A+I aan te nemen. Aangezien de A+I als 3-plaatsig geanalyseerd moet worden, kan de AcI niet óók een 3-plaatsige structuur hebben en kan de accusativus niet het object van het hoofdwerkwoord zijn (geweest). De verba sentiendi zijn eveneens niet 3-plaatsig, ze zijn altijd 2-plaatsig, en kunnen zowel een nomen als een AcI of *ut*-zin als complement hebben. Daarnaast zijn er in 'gewone' zinnen bepaalde semantische restricties die door het MATRIXPREDICAAT hun complement worden opgelegd. Bij een AcI is de accusativus niet meer gevoelig voor deze restricties; het functioneert dus in ieder geval in de onderliggende structuur niet als een OBJECT van het PREDICAAT.

Daarnaast biedt de uitbreidingshypothese geen verklaring voor het verschil tussen de AcI en de A+I, of waarom een deel van de accusatieven en infinitieven één geheel is geworden, en een ander deel losse elementen in de matrixzin zijn gebleven.

Pepicello (1977) biedt daarom een andere verklaring van de AcI. Hij suggereert dat er raising plaatsvindt van de ingebedde zin naar de matrixzin. De accusativus begint dan als het onderwerp van de bijzin, waar het nominatieve naamval heeft zoals standaard is voor SUBJECTEN. Vervolgens verplaatst deze door raising omhoog in de boom naar de plaats voor het complement van het MATRIXPREDICAAT (zie de onderliggende structuur en oppervlaktestructuur in figuur 3 en 4).



Figuur 3: Onderliggende structuur



Figuur 4: Oppervlaktestructuur

Deze verklaring biedt een deel van dezelfde voordelen als de uitbreidingshypothese: doordat *me* het OBJECT is van *dicunt* in de oppervlaktestructuur wordt de accusatieve naamval verklaard. Bovendien is ook in andere talen gevonden dat wanneer het subject uit een bijzin verwijderd wordt, het PREDICAAT infiniet wordt. Daarnaast meent Pepicello dat de accusativus wel het object móet zijn, omdat het het SUBJECT kan worden van een passieve zin zoals (19).

Ook dit blijkt echter geen perfecte verklaring te zijn. Bolkestein stelt dat in meerdere gevallen er geen verklaring wordt geboden voor de accusatieve naamval. In de passieve zin (20) is *viros* in de onderliggende structuur het SUBJECT van *venisse*, en in de oppervlaktestructuur is het óók het SUBJECT, maar dan van *dicitur*. Waar de accusatieve naamval vandaag zou komen wordt met deze analyse niet verklaard. Hetzelfde geldt voor intransitieve werkwoorden, zoals *constat* ‘het staat vast’. Hier hoort namelijk ook geen OBJECT bij, dus kan de oppervlaktestructuur niet de naamval verklaren.

Het MATRIXPREDICAAT hoeft dus niet een werkwoord te zijn dat een object bij zich draagt of naamval kan uitdelen. Sterker nog: het MATRIXPREDICAAT hoeft niet eens een werkwoord te zijn (Johndal, 2010). De AcI kan ook opgeroepen worden door een nomen, zoals *rumor* ‘het gerucht (dat ...)’. Ook hier bieden de hypothesen geen verklaring voor.

2.3 CONCLUSIE

Gezien de tegenwerpingen van Bolkestein is het erg onwaarschijnlijk dat één van bovenstaande analyses het probleem met de interne structuur van de AcI oplost. Wat echter in het kader van dit onderzoek meer van belang is, is de status van de AcI als geheel. Alle auteurs die ofwel de uitbreidingshypothese ofwel de raisinghypothese aanhingen, waren het erover eens dat de AcI geanalyseerd moet worden als één geheel. Ik maak dus geen definitieve keuze voor één van de analyses, maar zal bij het representeren van de interne structuur gebruik maken van de 2-plaatsige structuur in figuur (1). Dit heeft twee redenen. Ten eerste is naar mijn mening voldoende bewijs geleverd voor de hypothese dat de AcI één constituent in de matrixzin is.

Daarnaast zullen we zien dat de AcI op eenzelfde manier gerepresenteerd wordt in de gebruikte treebank, en om de overzichtelijkheid te bewaren volg ik hun voorbeeld.

De A+I is niet veel besproken in eerdere literatuur, maar kunnen we analyseren als controlestructuur, en dus syntactisch ook op die manier representeren. Ook hier volg ik het voorbeeld van de treebank en analyseer ik de A+I als een 3-plaatsige structuur.

Daarnaast is het erg belangrijk om het onderscheid tussen de AcI en de A+I goed in de gaten te houden. De twee structuren gedragen zich erg verschillend, en het ligt voor de hand om een verschil in gedrag te verwachten wat betreft de woordvolgordes. Om hier conclusies over te trekken moet echter eerst meer informatie over woordvolgorde in het Latijn gepresenteerd worden, wat ik zal doen in het volgende hoofdstuk.

3. Woordvolgorde

3.1 VRIJE WOORDVOLGORDE?

Het Latijn wordt vaak, in tegenstelling tot bijvoorbeeld het Nederlands en Engels, gezien als een taal met vrije woordvolgorde (Weil, 1844). Op school wordt verteld dat het niet uitmaakt waar een woord in een zin staat. Toch wordt ook gezegd dat een werkwoord vaak achterin de zin staat, dat attributieve adjectieven hun hoofd volgen, het onderwerp vaak vooraan in de zin staat en dat de woorden in een woordgroep zoveel mogelijk bij elkaar blijven. Er blijken dus toch meer regels te zijn dan we in eerste instantie dachten. Maar waar het Nederlandse 'Jan begroet Piet' betekent dat Jan het begroeten doet, kunnen Lucius en Marcus in de Latijnse vertaling van de zin, (1) en (2), zonder problemen op elkaars plek komen te staan, terwijl de zin dezelfde betekenis behoudt.

(4) *Luci-us Marc-um saluta-t*
Lucius-NOM Marcus-ACC begroeten-3.SG
'Lucius begroet Marcus.'

(5) *Marc-um Luci-us saluta-t*
Marcus-ACC Lucius-NOM begroeten-3.SG
'Lucius begroet Marcus.'

In het Latijn zijn er andere aanwijzingen die je vertellen welke rol een woord heeft in de zin. Zoals in het vorige hoofdstuk al besproken is, is de nominativus de standaardnaamval voor het SUBJECT van een zin, en de accusativus de standaardnaamval voor het OBJECT. Het maakt voor ons begrip van deze zin dus niet uit waar een woord geplaatst wordt, omdat we aan de morfologische informatie kunnen afleiden dat in zinnen (1) en (2) Lucius het SUBJECT is en Marcus het OBJECT. Waar in het Nederlands en Engels de volgorde van constituenten (syntactisch samenhangende woordgroep, zoals een naamwoord met en bijbehorende lidwoord en/of adjectieven) belangrijk is om de syntactische informatie te achterhalen die ze (meestal onzichtbaar) dragen, geven de naamvalsuitgangen die informatie in het Latijn. Devine en Stephens (2006) stellen daarom dat het Latijn niet zozeer een taal is met vrije woordvolgorde, maar dat de woordvolgorde vrij is om andere dingen uit te drukken dan syntactische functies.

Pinkster (1984) benoemt in zijn boek een aantal bevindingen over de Latijnse woordvolgorde, en doet zelfs een poging om deze te combineren in een 'standaardformule' (3) voor transitieve zinnen zoals zin (4). Hierin is te zien dat zinsverbindende constituenten een plaats vooraan in de zin innemen, satellieten (constituenten die niet bij het predicaat of één van de argumenten horen) een plaats rondom het tweede argument (het OBJECT) hebben en het finiete werkwoord aan het einde van de zin staat. P₁ is gereserveerd voor de constituent waar de zin over gaat

(Topic, zie 3.2) en wordt daarom vaak ingenomen door het eerste argument, als deze expliciet aanwezig is. Het eerste en het tweede argument duiden respectievelijk het subject en object van een zin aan. In zin (4) vinden we achtereenvolgend een connector (*nam*), subject (*quis*), object (*hoc*), satelliet (*non*) en finiet werkwoord (*intelligit*). Of het subject het Topic van de zin is of niet, en of deze op P₁ staat of op de plek van het eerste argument, laat ik open voor discussie. Wat echter meer van belang is, is dat de volgorde van (4) overeenkomt met formule (3). De woordvolgorde lijkt dus wél op een bepaalde manier samengevat te kunnen worden, en niet vrij te zijn.

(6) (connector) - P₁ - (connector) - {arg. 1} - (satelliet) - arg. 2 - (satelliet) - V_{finiet}

(7) *Nam quis hoc non intellig-it [...]*?

want wie.NOM dit.ACC niet begrijpen-3.SG

‘Want wie begrijpt dit niet [...]?’

Er zijn meer regels te vinden in de Latijnse woordvolgorde. Zo is het Latijn een taal met voornamelijk preposities (Spevak, 2010) en postnominale attributen (De Jong, 1983; Pinkster, 1984)), zoals te zien is in zin (5). Het adjectief *reliquus* wordt hier als nomen gebruikt (in het Nederlands vullen we ‘dingen’ aan, in het Latijn wordt het adjectief zelfstandig gebruikt), en wordt aan zijn rechterkant gemodificeerd door de bepalingen *rei publicae* ‘staat’ en *malis* ‘slecht’. De adpositie *de* ‘over’ staat links van de NP als geheel.

(8) *De reliqu-is re-i public-ae mal-is*
 over overig-ABL.PL zaak-GEN.SG openbaar-GEN.SG slecht-ABL.PL
lice-t-ne dicere?
 mogen-3.SG-VRAAG zeggen

‘Mag ik praten over de overige slechte dingen van de staat?’

(Cic. *Phil.* 1.14, uit Spevak (2010), p. 200)

Ondanks deze tendensen om constituenten in een vaste volgorde te laten voorkomen, zien we dat de regels verbroken kunnen worden. Er treden dan discontinuïteit op: constituenten die syntactisch bij elkaar horen staan in de zin niet bij elkaar, zoals de NP in zin (6). Het SUBJECT van de zin, *novi consules*, wordt hier onderbroken door de connector *deinde*. In de volgende paragraaf zal ik de notie discontinuïteit verder bespreken, en richten op de Acl.

(9) *Nov-i deine consul-es [...] ingressi hosti-um*
 nieuw-NOM.PL vervolgens consul-NOM.PL binnengaan vijand-GEN.PL
fin-es
 gebied-ACC.PL

‘Vervolgens gingen de nieuwe consuls het gebied van de vijanden binnen.’

(Liv. 8.17.1, uit Spevak (2010), p. 226)

3.2 DISCONTINUÏTEIT

Er kan bij de analyse van discontinue NP's een belangrijke onderscheiding gemaakt worden tussen twee soorten discontinuïteit (Snijders, 2012). De eerste soort discontinuïteit zijn NP's die 'verplicht' discontinu zijn: ze worden onderbroken door een element dat op zijn vaste plaats in de zin staat, zoals verbindingswoorden aan het begin van de zin (zie de formule in (3)). Het tweede soort discontinue NP's zijn de NP's die onderbroken worden door een 'vreemd' element, zoals voorbeelden (4), (5) en (6). Dit is het soort discontinuïteit waar meer speelt dan slechts het hanteren van de standaard woordvolgorde, en daarom is deze ook interessanter voor analyse. Deze bespreking van discontinuïteit richt zich daarom voornamelijk over het tweede soort discontinue zinnen.

Eerder onderzoek heeft uitgewezen dat in het Latijn voornamelijk pragmatische informatie (o.a. Pinkster, 1984; 2015; Spevak, 2010), en de informatiestructuur (Bolkestein, 2001; Snijders, 2012) van grote invloed zijn op de woordvolgorde in zinnen. Zo merkt Pinkster bijvoorbeeld op dat in hoofdzinnen het PREDICAAT vaak de zinsfinale positie heeft en de woordvolgorde SOV wordt aangehouden (zie (3) hierboven), en dat uitzonderingen hierop bijna altijd verklaard kunnen worden aan de hand van pragmatische informatie, zoals in (7).

- (10) *Apud Helvetios longe nobillissimus fuit et*
Bij+ACC Helvetiër-ACC.PL veruit belangrijkste-NOM zijn.PERF-3.SG en
ditissimus Orgetorix.
rijkste-NOM Orgetorix.NOM
'Bij de Helvetiërs was veruit de belangrijkste en rijkste man Orgetorix.'
(Caes. *Gal.* 1.2.1, uit Pinkster (1984), p. 225)

De zinsfinale positie wordt hier ingenomen door het SUBJECT van de zin, *Orgetorix*, waardoor deze extra nadruk krijgt. OBJECTEN en SATELLIETEN kunnen om dezelfde reden de zinsfinale positie innemen, zoals in (8).

- (11) *Pater nobis decessit a.d. VIII Kal. Dec.*
'Vader is overleden op 23 november.'
(Cic. *Att.* 1.6.2, uit Pinkster (1984), p. 226)
- (12) *Pater nobis a.d. VIII Kal. Dec. decessit.*
'Vader is op 23 november overleden.'

Met andere woorden: zin (8) en zin (9) geven antwoord op verschillende vragen: respectievelijk 'Wanneer is vader overleden?' en 'Wat gebeurde er met vader op 23 november?'. Zin (7) zou dan dus antwoord geven op de vraag 'Wie is bij de Helvetiërs veruit de belangrijkste en rijkste man?'. In al deze zinnen geeft het element op de zinsfinale positie nieuwe informatie, dit is het

element in de zin dat we Focus noemen. Elementen met Focus staan in het Latijn vaak achteraan in de zin. Spevak (2010) merkt ook op dat het element waar de zin over gaat – het Topic – vaak aan het begin van de zin staat. In zin (8) is *pater nobis* het element met Topic, en het staat inderdaad aan het begin van de zin. Nieuwe informatie komt in het Latijn dus meestal ná bekende informatie.

Andere factoren die van invloed kunnen zijn op de woordvolgorde zijn bijvoorbeeld het aantal constituenten (hebben we bijvoorbeeld te maken met een transitief of een intransitief werkwoord?) en de syntactische categorie van een constituent (een AcI gedraagt zich bijvoorbeeld anders dan een complementaire *ut*-zin). De invloed van de syntactische categorie van een constituent kan gezien worden bij auteurs zoals Caesar. Voor constituenten van verschillende categorieën gelden verschillende regels voor de plaatsing in een zin. Zo staat doelbepalingen in de vorm van een adpositie + complement meestal vóór het PREDICAAT, maar staan doelbepalingen in de vorm van een bijzin juist áchter het PREDICAAT, zoals in zin (10).

- (13) *Sitti-us est [...] miss-us ut e-am provinci-am*
 Sittius-NOM zijn.3.SG gezonden-NOM om.te die-ACC provinvie-ACC
pertubare-t
 onrust.veroorzaken-3.SG
 ‘Sittius is gezonden om in die provincie onrust te veroorzaken.’
 (Cicero *Sul.* 56, uit Pinkster (1984), p. 211)

Ook de omvang en interne complexiteit van een constituent kunnen van invloed zijn op de woordvolgorde in een zin. Bij het laatstgenoemde is echter gevonden dat het bij de klassieke auteurs – zoals de auteurs die ik in dit onderzoek bestudeer – minder belangrijk is. Ze maken minder gebruik van extrapositie en laten lange constituenten graag staan waar ze thuishoren (Pinkster, 1984).

3.3 DISCONTINUE ACI'S EN A+I'S

‘Gewone’ finiete complementen hebben een CP met elementen posities aan de linkerkant in de boomstructuur (Jhndal, 2012), en het gebeurt zelden dat onderdelen van de finiete zin de matrixzin onderbreken. De AcI gedraagt zich echter niet zo als gewone finiete complementen. Jøhndal noemt het namelijk ‘typisch’ voor de AcI dat een deel van de AcI vóór het MATRIXPREDICAAT staat, en een deel erachter. Hij concludeert dat er weinig regels zijn betreffende de discontinuïteit, en dat meer scrambling zeker voor kan komen.

Bij zijn onderzoek naar backward control in het Grieks en Latijn vond Haug (2017) inderdaad dat tot 20% van de gevonden AcI's/A+I's in zijn dataset discontinu waren. Een voorbeeld van een discontinue A+I kan gezien worden in (10), waarbij deze onderbroken wordt door het

SUBJECT van de hoofdzin, *hostes* ‘vijanden’. Een A+I kan ook onderbroken worden door andere constituenten, zoals het PREDICAAT uit de matrixzin in (11). Het infinitieve werkwoord van de A+I bestaat hier uit twee delen: een werkwoordelijk deel *esse* ‘zijn’ dat functioneert als koppelwerkwoord en een naamwoordelijk deel *maximam laudem* ‘grootste verdienste’.

- (14) ... *ut iam ad vall-um castr-orum host-es accedere*
 ... dat nu naar wal-ACC legerkamp-GEN vijand-NOM.PL naderen-INF
aud-ere-nt
 durven-IMPF.CONJ-3.PL
 ‘... dat de vijanden nu naar de wal van het legerkamp durfden te komen’
 (Caesar *Gal.* 3.15.10, uit Haug (2017), p. 136)
- (15) *Publice maxim-am put-ant esse laud-em ...*
 gemeenschap grootste-ACC menen-3.PL zijn-INF verdienste-ACC
 ‘Als gemeenschap menen ze dat hun grootste verdienste is.’
 (Caes. *Gal.* 4.3.1, uit Haug (2017), p. 134)

Ook zijn er in eerdere onderzoeken voorbeelden gevonden van discontinue AcI’s (Spevak, 2010), zoals (12), waarin deze onderbroken wordt door het PREDICAAT uit de matrixzin, *puto* ‘ik meen’. Spevak vond dat bij de werkwoorden van denken bij meer dan één op de vier AcI’s het matrixwerkwoord de AcI doorbreekt (zie tabel 1). Als de werkwoorden van denken vergeleken worden met de werkwoorden van zeggen (tabel 2) is er een duidelijk verschil te zien in de frequenties van de verschillende volgorden: waar bij de werkwoorden van denken ongeveer de helft van de PREDICATEN achter de bijbehorende AcI’s staat, staat bij de werkwoorden van zeggen een gelijk percentage PREDICATEN vóór de AcI.

- (16) *De Buciliano, Sestium puto non moleste ferre.*
 over Bucilianus-ABL Sestius-ACC menen-1.SG niet vervelend verdragen
 ‘Over Bucilianus, ik meen dat Sestius zich niet aan hem ergert.’
 (Cic. *Att.* 15.17.2, uit Spevak (2010))

Tabel 1: volgorde PREDICAAT en AcI bij werkwoorden van denken (Spevak, 2010, p. 157)

Volgorde	Voorkomens	Percentage
AcI > PREDICAAT	73	53%
PREDICAAT > AcI	25	18%
PREDICAAT ingevoegd	39	29%
Totaal	137	100%

Tabel 2: volgorde PREDICAAT en AcI bij werkwoorden van zeggen (Spevak, 2010, p. 159)

Volgorde	Voorkomens	Percentage
AcI > PREDICAAT	29	28%
PREDICAAT > AcI	52	51%
PREDICAAT ingevoegd	22	21%
Totaal	103	100%

Spevak verklaart de verschillen tussen de twee soorten werkwoorden aan de hand van verschillen op het gebied van betekenis. Werkwoorden van zeggen dragen zelf vaak minder nieuwe informatie, en aangezien we al hebben gezien dat informatie in een zin vaak geordend wordt van bekend naar nieuw, is het logisch dat de AcI, die meestal Focus draagt, rechts van het PREDICAAT staat bij deze groep werkwoorden. Bij de werkwoorden van denken geeft het PREDICAAT echter de intentie van de auteur aan, en draagt daardoor vaker Focus. Dit zorgt ervoor dat het PREDICAAT vaker achter de AcI wordt geplaatst dan bij de andere categorie werkwoorden.

Deze bevindingen komen dus overeen met de verwachtingen van Jøhndal dat discontinuïteit regelmatig voorkomt bij de AcI. Hoewel dus door meerdere auteurs is aangetoond dat discontinuïteit inderdaad voorkomt, heeft geen van de auteurs een verklaring weten te vinden voor deze discontinuïteit. In een corpusstudie heb ik daarom geprobeerd te achterhalen welke factoren van invloed zijn op discontinuïteit van een AcI. Daarbij heb ik de volgende factoren meegenomen in een regressiemodel: de auteur, het soort MATRIXPREDICAAT, de beheersbaarheid van de infinitivus, de levendheid en lengte van de accusativus en of de constructie een AcI of A+I is. Daarnaast onderzoek ik exploratief andere omstandigheden rond de (discontinue) AcI's en A+I's, zoals de plaats van het MATRIXPREDICAAT, een eventuele restrictie op het aantal onderbrekende elementen, en door welke constituenten de constructies onderbroken kunnen worden.

4. Methode

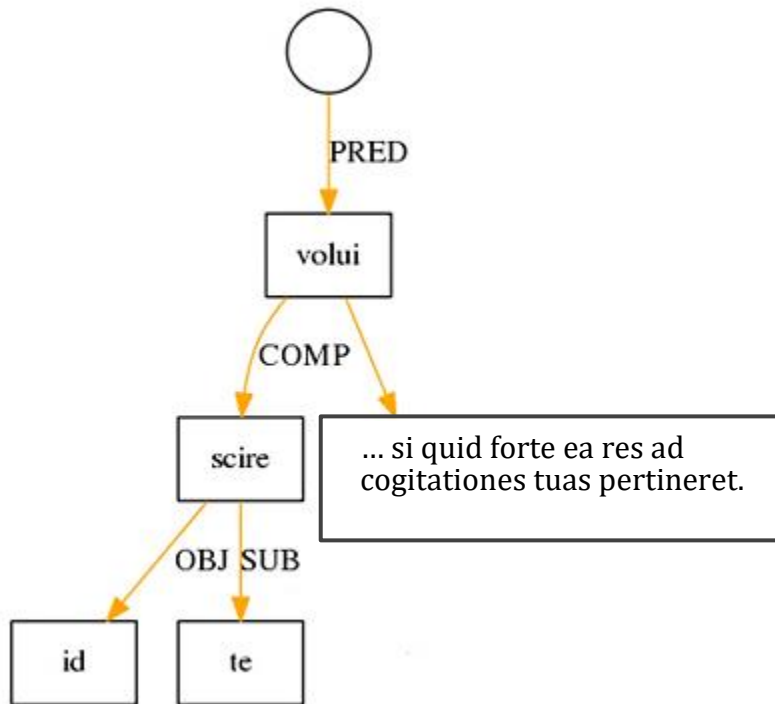
4.1 MATERIAAL

Ik heb voor dit onderzoek gebruik gemaakt van data in de Pragmatic Resources in Old Indo-European Languages (PROIEL) Treebank (Haug & Jøhndal, 2008). Deze treebank bevat teksten in oude Indo-Europese talen, waaronder het Latijn. Deze teksten zijn van morfologische en syntactische annotatie voorzien, en bij een deel van de teksten is ook de informatiestructuur geannoteerd. Voor de syntactische annotatie is gebruik gemaakt van een dependency-grammatica. In een dependency-grammatica wordt de syntactische structuur van een zin weergegeven als een set antisymmetrische relaties. Zoals Haug (2010) stelt: X is afhankelijk van Y als, wanneer Y uit de zin verwijderd wordt, X ook verwijderd moet worden. Bovendien geldt dat als X afhankelijk is van Y, Y niet afhankelijk kan zijn van X.

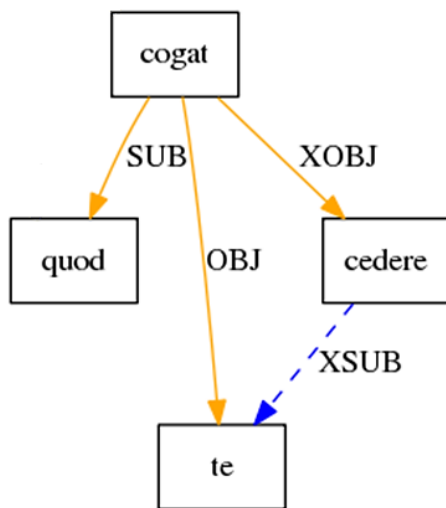
Eén van de dingen die in de PROIEL Treebank geannoteerd zijn is de relatie van een woord met zijn hoofd. Zo krijgt het matrixwerkwoord van een hoofdzin altijd de relatie PREDICATE (PRED in de boomstructuur), en subjecten en objecten worden respectievelijk aangeduid met SUBJECT (SUB) en OBJECT (OBJ). Eén van de voordelen van het gebruiken van deze treebank voor dit onderzoek, is dat de AcI en de A+I op verschillende manieren worden geannoteerd. De infinitivus van de AcI heeft de relatie COMPLEMENT (COMP) met zijn hoofd, en de accusativus heeft de relatie SUBJECT onder de infinitivus. In figuur 4.1 is de AcI uit zin (1) weergegeven in een boomstructuur. Bij de A+I zijn echter zowel de accusativus als de infinitivus afhankelijk van het matrixwerkwoord. De accusativus is hieraan verbonden via de relatie OBJECT, de infinitivus wordt geanalyseerd als een open complementspredicaat en krijgt de bijbehorende relatie XOBJ. Open predicaten zijn predicaten die geen eigen subject hebben, maar deze verkrijgen door coreferentie met een element in de matrixzin. Dat element kan zowel zichtbaar zijn, als geïmpliceerd door een werkwoord (zoals *ego* ‘ik’ geïmpliceerd is in *sum* ‘(ik) ben’). Een voorbeeld van de A+I-structuur is weergegeven in zin (2), met de bijbehorende boomstructuur in figuur 4.2.

- (1) *Id te scire vol-u-i* (si quid forte ea res ad
het.ACC jij.ACC weten.INF willen-PERF-1.SG
cogitationes tuas pertineret.)
‘Ik wilde dat jij dit wist, (als die zaak toevallig betrekking had op jouw voornemen).’
(Cic. Att. 1.6.1)
- (2) (*Ego autem, cum sciam quam sis fortis, vehementius esse quiddam suspicor*)
quod te cogat cedere (et prope modum infringat.)
dat.NOM.N jij.ACC dwingen-3.PL opgeven.INF

‘(Ik vermoed echter, aangezien ik weet hoe dapper jij bent, dat er iets sterkers is) dat jou dwingt om op te geven (en bijna jouw grens breekt).’
 (Cic. Att. 1.19.9)



Figuur 4.1: voorbeeldannotatie AcI



Figuur 4.2: voorbeeldannotatie A+I

Uit het PROIEL-corpus zijn teksten van Caesars *Commentarii de bello Gallico* Cicero's *Epistulae ad Atticum* genomen en met behulp van een Perl-script zijn hier de AcI's en de A+I's uitgehaald². Bij de AcI's is dat gedaan door te zoeken naar infinitieve werkwoorden die de relatie COMP met hun hoofd hadden. De A+I's zijn gevonden door te zoeken naar infinitieve werkwoorden die de relatie XOBJ met hun hoofd hadden.

Tabel 1 laat zien dat de beide constructies redelijk gelijk verdeeld zijn over de auteurs, en er ongeveer evenveel AcI's als A+I's in de dataset zitten. Van de 1615 gevonden AcI's en A+I's, zijn willekeurig 400 zinnen genomen voor de analyse: 100 AcI's en 100 A+I's van beide auteurs.

Tabel 1: overzicht van de AcI's en A+I's in data

	AcI	A+I	Totaal
Caesar	325	285	610
Cicero	576	429	1005
Totaal	901	714	1615

4.2 ONDERZOEKSOPZET

De 400 overgebleven structuren zijn gedeeltelijk handmatig en gedeeltelijk semi-automatisch geannoteerd op onderstaande variabelen.

4.2.1 Discontinuïteit

Als afhankelijke variabele is geannoteerd of de AcI's en A+I's wel of niet discontinu zijn. Met discontinu doel ik op het voorkomen van elementen uit de matrixzin tussen de grenzen van de AcI en A+I. Die grenzen worden bij de AcI gemarkeerd door het infinitieve werkwoord in de boomstructuur. Alles wat bij de AcI hoort is daarvan afhankelijk, dus wanneer een element dat niet in de boomstructuur onder het infinitieve werkwoord staat wél tussen de delen van de AcI in de platte tekst staat, is deze discontinu. De grenzen van de A+I worden gemarkeerd door het accusatieve nomen in de boomstructuur, en alles wat daaronder staat, en het infinitieve werkwoord en alles daaronder.

² Caesar *Commentarii de bello Gallico* 1.1 t/m 1.54, 2.1 t/m 2.35, 3.1 t/m 3.27, 4.1 t/m 4.9 en 4.11 t/m 4.38, 5.1 t/m 5.7 en 5.49, 6.1 t/m 6.44, 7.3, 7.5 en 7.77 ; Cicero *Epistulae ad Atticum* boek 1, boek 2, boek 3, boek 4, boek 5.1 t/m 5.6 en 5.8 t/m 5.21, boek 6.1 en 6.9, boek 7.2

4.2.2 Auteur

Omdat Romeinse auteurs vaak een geheel eigen schrijfstijl, en dus ook woordvolgorde, hanteren is het belangrijk om in de analyse mee te nemen wie een tekst geschreven heeft. Zo schrijft Caesar in hoofdzinnen het PREDICAAT in 84% van de gevallen na het SUBJECT en het OBJECT, maar doet Cicero dat slechts in 50% van de gevallen (Pinkster, 1984). Hierbij is het belangrijk om op te merken dat er ook grote verschillen in woordvolgorde te vinden zijn binnen meerdere werken van één auteur. De toon van het werk – een ambtelijk verslag in opdracht van de Senaat bij Caesar, of een informele reeks brieven aan zijn beste vriend bij Cicero – zal hier ook een grote rol bij spelen. Ik zal daarom niet proberen de gevonden resultaten te generaliseren voor alle literatuur van Caesar en Cicero, of zelfs voor de gehele literatuur van de hele eerste eeuw voor Christus, maar stel alleen een theorie op over het gebruik van de AcI en de A+I in deze twee werken.

4.2.3 AcI of A+I

Zoals besproken in hoofdstuk 2 hebben de AcI en de A+I verschillende onderliggende structuren, waarbij de AcI een samenhangender geheel is dan de A+I (figuren 1 en 2). Omdat de delen van de A+I gemakkelijker te scheiden zijn ligt het voor de hand dat er meer discontinuïteit plaatsvindt bij de A+I dan bij de AcI.

4.2.4 Soort matrixpredicaat

Omdat er veel soorten PREDICATEN zijn die een AcI of A+I kunnen oproepen, maak ik hier in de analyse onderscheid tussen. De verschillende PREDICATEN die worden meegenomen in de analyse zijn: verba dicendi (werkwoorden die ‘zeggen’, ‘schrijven’ of ‘weten’ uitdrukken), verba sentiendi (werkwoorden die een waarneming uitdrukken), verba affectuum (werkwoorden die een mentaal proces aanduiden, zoals een gevoel, mening of behoefte), onpersoonlijke predicaten (zoals *constat* ‘het staat vast’), actiepredicaten (zoals ‘doen’) en predicaten die ‘kunnen’ uitdrukken.

4.2.5 Plaats matrixpredicaat

Omdat er veel verschillende meningen zijn in de literatuur over de plaats van het MATRIXPREDICAAT, en de standaardplaats hiervan ook bij de AcI en A+I nog niet vastgesteld is, neem ik deze variabele descriptief mee.

4.2.6 Verbaal/niet-verbaal matrixpredicaat

Zoals in hoofdstuk 2 al opgemerkt is, zijn niet alle MATRIXPREDICATEN verbaal, ze kunnen ook nominaal zijn, of een constructie van een adjectief met het koppelwerkwoord *esse* 'zijn'. Ik verwacht dat er te weinig data is in de verzamelde 400 zinnen om deze variabele mee te nemen in het model, maar zal de aantallen wel descriptief rapporteren.

4.2.7 Beheersbaarheid infinitivus

Omdat door Pinkster (1984) is opgemerkt dat een PREDICAAT altijd beheersbaar moet zijn in een A+I, kan deze variabele alleen voor de AcI-structuren worden meegenomen in de analyse. Ik heb de beheersbaarheid onderverdeeld in 2 variabelen: de tijd van de infinitivus en de vorm van de infinitivus (actief vs. passief). Deze variabelen zijn objectief te annoteren, wat de variabele 'beheersbaarheid' op zichzelf niet zou zijn, omdat ik geen moedertaalspreker van het Latijn heb en er dus geen intuïties over kan hebben. De tijd en vorm van een werkwoord geven daarom toch een deel van de beheersbaarheid aan, zonder dat er taalintuïties voor nodig zijn: zoals in hoofdstuk 2 besproken, zijn alle werkwoorden die niet in de tegenwoordige tijd staan niet beheersbaar, en alle werkwoorden die passief zijn, niet beheersbaar. Omdat ik hier de betekenis belangrijk vind, en niet de morfologische vorm, reken ik de deponentia (woorden met een actieve betekenis, maar een passieve vorm) tot de groep actieve werkwoorden. De variabele tijd heeft twee levels (tegenwoordige en niet tegenwoordige tijd), en de variabele vorm heeft twee levels (actief en passief), waardoor de overkoepelende variabele beheersbaarheid vier levels heeft: tegenwoordig actief, tegenwoordig passief, niet-tegenwoordig actief, niet-tegenwoordig passief.

4.2.8 Levendheid accusativus

Ook de levendheid van de accusativus kan slechts voor de AcI in de analyse worden meegenomen, omdat de accusativus in een A+I altijd levend is (Pinkster, 1984). Dit is daarom ook de reden dat er losse modellen worden opgesteld voor de AcI en de A+I, aangezien het toevoegen van een dergelijke variabelen voor beide constructies een vertekend beeld zou opleveren. Zoals genoemd in hoofdstuk 3 staan woorden met topic vaker vooraan in de zin, en omdat levende entiteiten eerder topic zullen zijn dan niet-levende entiteiten, zullen ze ook vaker vooraan staan dan niet-levende entiteiten. Hierdoor verplaatst de accusativus als het ware uit de AcI naar een plek meer vooraan in de hoofdzin en is de AcI discontinu. Levendheid kan in veel verschillende categorieën worden onderverdeeld, zo onderscheiden Zaenen en collega's (2004) er negen: HUMAN (één of meer mensen, inclusief niet-bestaande entiteiten die als mensachtigen worden gepresenteerd, zoals goden), ORG (een groep mensen wanneer ze zich als een groep presenteren), ANIMAL (niet-menselijke levende entiteiten, inclusief bacteriën en

virussen), PLACE (wanneer naar een plaats verwezen wordt als een plaats), TIME (uitdrukkingen die een tijdsperiode aanduiden), CONCRETE (concrete objecten/substanties, inclusief lichaamsdelen), NONCONC (niet-concrete dingen, zoals gebeurtenissen), MAC (intelligente machines, zoals computers en robots) en VEH (vervoersmiddelen). Ik acht de kans op grote verschillen tussen de categorieën in effectsterkte echter erg klein, bovendien blijven er, wanneer er veel wordt onderverdeeld, te weinig items per categorie over voor een betrouwbare analyse. Er worden daarom slechts twee levels gehanteerd: levend en niet-levend. Deze levels corresponderen respectievelijk met de categorieën HUMAN, ORG, ANIMAL en PLACE, TIME, CONCRETE, NONCONC, VEH van Zaenen en collega's.

4.2.9 Lengte van accusativus

De lengte van een constituent correleert positief met de hoeveelheid nieuwe informatie die deze brengt (Bard & Aylett, 1999; Bard et al., 2000). Nieuwe informatie is dus meestal langer dan bekende informatie. Aangezien pragmatische functies lastig te benoemen zijn – zeker in de brieven van Cicero, door de plotselinge overgangen naar nieuwe onderwerpen (Pinkster, 1984) – kan daarom de lengte van de accusativus worden gebruikt om de 'nieuwheid' ervan aan te duiden. Zoals in hoofdstuk 3 besproken is staat in het Latijn bekende informatie vaak vóór nieuwe informatie. De verwachting is daarom dat als de accusativus lang is, en dus nieuwe informatie brengt, er minder snel informatie uit de hoofdzin na de accusativus zal komen te staan en er minder discontinuïteit gevonden wordt. Daarnaast geldt volgens Spevak (2010) het principe van toenemende complexiteit: kortere constituenten komen, als het mogelijk is, vóór een langere constituent. Het kan dus beredeneerd worden dat wanneer de accusativus van de AcI of A+I erg lang is, er minder snel een element achter zal komen te staan. Hoewel Pinkster (1984) heeft gevonden, zoals in hoofdstuk 3 genoemd, dat dit principe vaak minder van invloed is bij de klassieke auteurs, kan het in samenwerking met de vorige aanname toch een rol spelen bij de hoeveelheid discontinuïteit. Bij een accusativus die bestaat uit een groot aantal woorden zal dus naar verwachting minder discontinuïteit optreden.

4.3 STATISTISCHE ANALYSE

Er is alleen data gebruikt waarbij zowel de accusativus als de infinitivus expliciet aanwezig zijn, zodat de grenzen van de structuren duidelijk zichtbaar blijven.

Omdat er een categorische uitkomstvariabele is, met meerdere predictorvariabelen, van verschillende niveaus, voer ik een logistische regressie uit. Van de predictorvariabelen die worden ingebracht zijn de meeste binair (zie paragraaf 4.2.2, 4.2.5 en 4.2.6). Eén variabele is

continu (zie 4.2.7), en één variabele is nominaal, met respectievelijk vier en vijf levels voor de AcI en de A+I. Hiervoor zijn dummyvariabelen ingebracht, waarbij de grootste groep MATRIXPREDICATEN als referentiecategorie is gebruikt. Aan de assumpties van lineariteit, onafhankelijke fouten, geen hinderlijke multicollineariteit en lineariteit van de logit is voldaan. Zoals gesuggereerd in Jaeger (2011) zijn voor elke variabele minstens 10 datapunten ingebracht.

5. Resultaten

5.1 ALGEMENE ANALYSE

In tabel 5.1 zijn het aantal continue en discontinue AcI's en A+I's per auteur weergegeven. Van de onderzochte zinnen bleek bij 24,3% een element uit de hoofdzin voor te komen tussen de grenzen van de AcI of A+I.

Tabel 5.1: Aantal discontinue AcI's en A+I's per auteur

		Discontinu	Continu	Totaal
AcI	Caesar	9	91	100
	Cicero	34	66	100
	Samen	43	157	-
A+I	Caesar	13	87	100
	Cicero	41	59	100
	Samen	54	146	-
Totaal		97	103	400

Omdat aparte data-analyses uitgevoerd zijn voor AcI en de A+I bespreek ik deze ook los van elkaar. Hieronder volgt eerst een vergelijking tussen de AcI en A+I, waarna een statistische analyse en een descriptieve analyse van de AcI volgen in respectievelijk 5.2.1 en 5.2.2.

Vervolgens toon ik de resultaten van de A+I, wederom eerst die van het statistische model, in 5.3.1, en vervolgens de descriptieve variabelen in 5.3.2.

In tabel 5.2 zijn het aantal continue en discontinue per constructie weergegeven. Er is een Pearson's chi-square toets uitgevoerd. Er bleek geen significant effect te zijn tussen het soort constructie en of deze wel of niet discontinu is, $\chi^2(1) = 1,65$, $p = 0,199$.

Tabel 5.2: Aantal discontinue zinnen per constructie

	Discontinu	Continu
AcI	43 (21,5%)	157 (78,5%)
A+I	54 (27%)	146 (73%)

5.2.1 STATISTISCHE ANALYSE AcI

Er is een binaire logistische regressie uitgevoerd. Uit de analyses bleek dat discontinuïteit van de AcI voor 19,4% te verklaren is door de ingebrachte variabelen, $\chi^2(7) = 26,773$, $p < .001$. De

auteur bleek een significant hoofdeffect te hebben op of de AcI continu was of niet ($B(1) = 1.687, p < .001$). De $Exp(B)$ -waarde van 5.402 laat zien dat wanneer de auteur Cicero is, de kans op een discontinue AcI 5.4 keer groter was dan wanneer de auteur Caesar is. De beheersbaarheid van de infinitivus bleek geen significant effect te hebben op de discontinuïteit ($B(1) = 0.423, p = .265$), evenals de levendheid en de lengte van de accusativus (respectievelijk $B(1) = -.022, p = .959$ en $B(1) = .003, p = .975$). Ook bleek er geen significante invloed te zijn van het soort matrixpredicaat op de discontinuïteit. Zowel de verba dicendi als de verba sentiendi en de onpersoonlijke predicaten verschilden niet significant van de verba affectuum ($B(1) = -.119, p = .888$; $B(1) = -.459, p = .422$ en $B(1) = -1.975, p = .065$, respectievelijk).

Tabel 5.3: B-waarden en bijbehorende statistiek voor de predictorvariabelen

	B	SD	Wald	Sig.	Exp(B)	95% BI voor Exp(B)	
						Onder	Boven
Constant	-2.326	.633	13.518	< .001	.098		
Auteur	1.687	.427	15.606	< .001	5.402	2.339	12.474
Beheersbaarheid I	.432	.379	1.244	.265	1.526	.726	3.210
Levendheid A	-.022	.419	.003	.959	9.79	.430	2.226
Lengte A	.003	.112	.001	.975	1.004	.805	1.251
Verba dicendi	-.119	.433	.075	.784	.888	.380	2.075
Verba sentiendi	-.459	.571	.645	.422	.632	.206	1.936
Onpersoonlijke predicaten	-1.975	1.070	3.406	.065	.139	.017	1.130

5.2.2 DESCRIPTIEVE ANALYSE ACI

Van de 200 onderzochte AcI's bleken er 43 discontinue te zijn (zie tabel 5.1). In tabel 5.4 is weergegeven hoeveel onderbrekende elementen er in die gevallen zijn. De meeste gevallen van discontinuïteit bleken slechts uit één onderbrekend element te bestaan (86%), een paar AcI's hadden twee onderbrekende elementen (12%), en één AcI had zelfs drie onderbrekende elementen (2%). In Tabel 5.5 is uitgesplitst wat de functie van de onderbrekende elementen in de matrixzin was. Een ruime meerderheid van de discontinuïteit wordt veroorzaakt door een onderbreking van het MATRIXPREDICAAT (80%), de rest door SUBJECTEN (19%), ADJUNCTEN (14%) en VOEGWOORDEN (5%).

Tabel 5.4: Frequentie aantal onderbrekende elementen

Aantal onderbrekende elementen	Frequentie	Percentage	Percentage / discontinue zinnen
0	157	78,5	-

1	37	18,5	86,0
2	5	2,5	11,6
3	1	0,5	2,3

Tabel 5.5 Functie onderbrekende elementen

Functie	Frequentie	Percentage / aantal AcI's	Percentage / aantal discontinue AcI's
Predicaat	34	17,0	80,1
Subject	8	4,0	18,6
Adjunct	6	3,0	14,0
Voegwoord	2	1,0	4,7

In tabel 5.6 wordt de plek van het MATRIXPREDICAAT weergegeven. Bij de AcI's stond het PREDICAAT in het grootste deel van de gevallen achter de AcI (51%), in ongeveer een derde van de gevallen ervoor (30%) en een deel stond tussen delen van de AcI in (17%). In 4 gevallen ontbrak het MATRIXPREDICAAT in de zin. Dat was het geval omdat er een langer achtereenvolgend stuk was waarbij bijvoorbeeld eerder gezegde dingen werden herhaald in meerdere AcI's. Omdat het overbodig is om bij elke zin het PREDICAAT te herhalen, krijg je losstaande AcI's.

Tabel 5.6: Plaats matrixpredicaat

Plaats	Frequentie	Percentage
Voor AcI	60	30,0
In AcI	34	17,0
Na AcI	102	51,0
Ontbrekend (<i>ver voor AcI</i>)	4	2,0

Bij de AcI's is een deel van de PREDICATEN niet verbaal: in 7% van de gevallen bleek het predicaat nominaal te zijn, zoals te zien is in tabel 5.7.

Tabel 5.7: Aantal nominale en verbale matrixpredicaten

	Frequentie	Percentage
Nominaal	13	6,5
Verbaal	187	93,5

Onderstaande gegevens zijn opgenomen in het model van de statistische analyse. Bij de AcI waren er vier soorten matrixpredicaten te onderscheiden (tabel 5.8): de verba affectuum (die 44,5% van de data opmaakten), de verba dicendi (28,5%), de onpersoonlijke predicaten (21%) en de verba sentiendi (33%). De verba affectuum zijn duidelijk de grootste groep en is daarom

ook als referentiecategorie gebruikt bij het inbrengen van de dummyvariabelen in de regressieanalyse. Voor de vier losse AcI's die gevonden zijn, heb ik teruggekeken naar het MATRIXPREDICAAT dat de reeks AcI's introduceerde, en deze is geannoteerd bij de data. In alle vier de gevallen was dat een verbum dicendi. De actiepredicaten en werkwoorden van 'kunnen' kwamen bij de AcI niet voor.

Tabel 5.8: Frequentie soorten matrixpredicaten

Soort matrixpredicaat	Frequentie	percentage
Affectuum	89	44,5
Dicendi	57	28,5
Onpersoonlijk	21	10,5
Sentiendi	33	16,5

Zoals in tabel 5.9 te zien is, was de verdeling tussen beheersbare en niet-beheersbare infinitieven bijna gelijk verdeeld: 49,5% van de predicaten was beheersbaar, en 50,5 was niet-beheersbaar.

Tabel 5.9: Aantal beheersbare infinitieve predicaten

	Frequentie	Percentage
Beheersbaar	99	49,5
Niet-beheersbaar	101	50,5

Bij de accusativi van de AcI's is gekeken naar levendheid (tabel 5.10) en de lengte ervan (tabel 5.11). 34% van de accusativi bleek levend te zijn, tegenover 66% niet-levende accusativi. De lengte ervan verschilde van één tot twaalf woorden, met een gemiddelde van 1,80, waarbij een standaarddeviatie van 1,76 hoort.

Tabel 5.10: Levendheid accusativus

	Frequentie	Percentage
Levend	68	34,0
Niet-levend	132	66,0

Tabel 5.11: Lengte accusativus

Min.	Max.	Gemiddelde	SD
1	12	1,80	1.76

5.3.1 STATISTISCHE ANALYSE A+I

Er is een binaire logistische regressie uitgevoerd. Uit de analyses bleek dat discontinuïteit van de A+I voor 25,7% te verklaren is door de ingebrachte variabelen, $\chi^2(6) = 39.027$, $p < .001$. De auteur bleek een significant hoofdeffect te hebben op of de A+I continu was of niet ($B(1) = 1.418$, $p < .001$). De $Exp(B)$ -waarde van 4.127 laat zien dat wanneer de auteur Cicero is, de kans op een discontinue A+I 4.1 keer groter was dan wanneer de auteur Caesar is. De lengte van de accusativus bleek ook een significant effect te hebben op de discontinuïteit van de A+I ($B(1) = .216$, $p = .004$). De $Exp(B)$ -waarde van 1.298 laat zien dat wanneer de accusativus één woord langer wordt, de kans op discontinuïteit 1,3 keer groter wordt. Er bleek geen significante invloed te zijn van het soort matrixpredicaat op de discontinuïteit. Zowel de verba affectuum als de verba dicendi, de verba sentiendi en de actiepredicaten verschilden niet significant van de predicaten van 'kunnen' (respectievelijk $B(1) = -.856$, $p = .061$; $B(1) = -1.542$, $p = .071$; $B(1) = -2.250$, $p = .076$ en $B(1) = -.536$, $p = .255$).

Tabel 5.12: B-waarden en bijbehorende statistiek voor de predictorvariabelen

	B	SD	Wald	Sig.	Exp(B)	95% BI voor Exp(B)	
						Onder	Boven
Constant	-1.876	.444	17.858	< .001	.153		
Auteur	1.418	.401	12.513	< .001	4.127	1.882	9.051
Lengte A	.261	.090	8.336	.004	1.298	1.087	1.549
Verba affectuum	-.856	.457	3.503	.061	.425	.173	1.041
Verba dicendi	-1.542	.854	3.257	.071	.214	.040	1.142
Verba sentiendi	-2.250	1.267	3.154	.076	.105	.009	1.263
Actiepredicaten	-.536	.471	1.294	.255	.585	.232	1.473

5.3.2 DESCRIPTIEVE ANALYSE A+I

Van de 200 onderzochte A+I's bleken er 54 discontinue te zijn (zie tabel 5.1). In tabel 5.13 is weergegeven hoeveel onderbrekende elementen er in die gevallen zijn. De meeste gevallen van discontinuïteit bleken slechts uit één onderbrekend element te bestaan (63%), een aantal A+I's hadden twee onderbrekende elementen (22%), en acht A+I's hadden zelfs drie onderbrekende elementen (15%). In Tabel 5.14 is uitgesplitst wat voor functie die onderbrekende elementen hadden in de matrixzin. Een meerderheid van de discontinuïteit wordt veroorzaakt voor onderbrekend van het MATRIXPREDICAAT (61%), de rest door SUBJECTEN (35%), ADJUNCTEN (30%), NEGATIES (15%), INDIRECTE OBJECTEN (7%) en VOEGWOORDEN (4%).

Tabel 5.13: Frequentie aantal onderbrekende elementen

Aantal onderbrekende elementen	Frequentie	Percentage	Percentage / discontinue A+I's
--------------------------------	------------	------------	-----------------------------------

0	146	73,0	-
1	34	17,0	63,0
2	12	6,0	22,2
3	8	4,0	14,8

Tabel 5.14: Functie onderbrekende elementen

Functie	Frequentie	Percentage / aantal discontinue AcI's	Percentage / aantal AcI's
Predicaat	33	61,1	16,5
Subject	19	35,2	9,5
Adjunct	16	29,6	8,0
Negatie	8	14,8	4,0
Indirect object	4	7,4	2,0
Voegwoord	2	4,0	1,0

In tabel 5.15 wordt de plek van het MATRIXPREDICAAT weergegeven. Bij de A+I's stond het PREDICAAT in het grootste deel van de gevallen achter de A+I (73%), in ongeveer een evenveel gevallen ervoor (13%) als tussen delen van de A+I in (15%). Bij de A+I zijn geen nominale PREDICATEN gevonden.

Tabel 5.15: Plaats matrixpredicaat

Plaats	Frequentie	Percentage
Voor A+I	26	13,0
In A+I	29	14,5
Na A+I	145	72,5

Onderstaande gegevens zijn opgenomen in het model van de statistische analyse. Bij de A+I waren er vijf soorten matrixpredicaten te onderscheiden (tabel 5.16): de actiepredicaten (die 28,5% van de data opmaakten), de verba affectuum (24%), de verba dicendi (8%), de werkwoorden van 'kunnen' (36,5%) en de verba sentiendi (3%). De werkwoorden van 'kunnen' zijn de grootste groep en zijn daarom ook als referentiecategorie gebruikt bij het inbrengen van de dummyvariabelen in de regressieanalyse. De onpersoonlijke predicaten komen bij de A+I niet voor.

Tabel 5.16: Frequentie soorten matrixpredicaten

Soort matrixpredicaat	Frequentie	percentage
Actie	57	28,5
Affectuum	48	24,0
Dicendi	16	8,0
Kunnen	73	36,5

Sentiendi	6	3,0
-----------	---	-----

De lengte van de accusativi van de A+I verschilde van één tot negentien woorden, met een gemiddelde van 1,96, waarbij een standaarddeviatie van 2,26 hoort.

Tabel 5.17: Lengte accusativus

Min.	Max.	Gemiddelde	SD
1	19	1,960	2,258

6. Discussie

Na een onderzoek te hebben uitgevoerd naar discontinuïteit bij AcI's en A+I's en waardoor deze wordt veroorzaakt zijn er verschillende dingen die opgemerkt kunnen worden. Aan de hand van de onderzochte data kan ik concluderen dat er geen significant verschil bestaat tussen de hoeveelheid discontinuïteit bij de AcI en A+I. Dit is zeker tegen de verwachting in. Door de verschillende structuren in zinnen met een AcI en een A+I (respectievelijk 2-plaatsig en 3-plaatsig) was het de verwachting dat bij de A+I meer discontinuïteit plaats zou vinden dan bij de AcI. Deze laatste is per slot van rekening één syntactisch geheel, terwijl de accusativus en het infinitieve werkwoord van de A+I losse onderdelen in de matrixzin zijn. Het is dan ook opvallend dat er geen significant verschil is gevonden. Na een analyse van de data denk ik dat dit gedeeltelijk te wijten kan zijn aan de samenstelling van de dataset. Hoewel de annotatiehandleiding duidelijk stelt dat wanneer iets geannoteerd kan worden met de relatie COMP (dit is de relatie die de AcI krijgt, evenals andere – finiete – zinscomplementen) dit altijd moet gebeuren, lijkt dit niet altijd te zijn gebeurd. Toen ik de data nader bekeek viel het me op dat een aantal van de constructies die als 'A+I' geannoteerd zijn, eigenlijk een AcI zijn. Er is dus niet door alle annotatoren consequent geannoteerd, wat ten koste gaat van de onderzochte data. Ik verwerp daarom nog niet mijn hypothese dat er een verschil is tussen de AcI en de A+I in hoeveelheid discontinuïteit, maar denk dat er opnieuw een onderzoek naar moet worden gedaan, waarbij elke constructie nogmaals gecontroleerd wordt. Deze data is niet betrouwbaar genoeg om op basis ervan een uitspraak te kunnen doen over dit onderwerp.

Waar wel met redelijke zekerheid iets over gezegd kan worden is de losse analyse van de AcI. Ook de analyse van de A+I bespreek ik, deze resultaten zijn alleen lastiger te interpreteren door de troebele data. In paragraaf 6.3 vergelijk ik de twee constructies en vat ik de belangrijkste bevindingen samen.

6.1 CONCLUSIES EN DISCUSSIE ACI

6.1.1 *Het model*

Uit de logistische regressie bleek dat de hoeveelheid discontinuïteit van de AcI significant beïnvloed wordt door de auteur van het werk, zoals was verwacht. De toets liet zien dat wanneer de auteur Cicero is, de kans op een discontinue AcI 5,4 keer groter is dan wanneer het een werk van Caesar is. Hiervoor zijn twee mogelijke verklaringen. Ofwel het verschil ligt aan de stijl van het werk – waar Caesar een erg formeel werk schrijft in opdracht van de Senaat, schrijft Cicero persoonlijke brieven naar zijn beste vriend en zwager –, ofwel het verschil ligt in de schrijfstijl van de auteur, onafhankelijk van het soort werk. Beide verklaringen zijn plausibel.

Cicero hanteert in zijn brieven een ander soort schrijfstijl dan in zijn overige werken (Albrecht, 2003). Er is bijvoorbeeld meer informeel taalgebruik, meer verkleinwoorden en interjecties. Ook gebruikt hij relatief gezien veel Griekse leenwoorden in zijn brieven. Dit is tegen zijn gebruikelijke schrijfwijze in, waar hij erg tegen het gebruik van leenwoorden is. In zijn brieven voelt hij echter minder de neiging om zijn eigen regels te hanteren; deze zijn per slot van rekening opgesteld voor formele proza, niet voor persoonlijke communicatie. Je kunt dus al aan andere voorbeelden zien dat Cicero een vrijere schrijfstijl aanhoudt in zijn brieven dan in zijn andere werken en minder geneigd is de regels te volgen. Daaruit volgt dat het erg goed mogelijk is dat het verschil in discontinuïteit verklaard kan worden door de toon van het werk. Om deze hypothese te toetsen zou er een nieuwe analyse uitgevoerd moeten worden, met meerdere werken van dezelfde auteur.

Het is echter ook een mogelijkheid dat het verschil echt bij de auteur ligt. Zo benoemt Albrecht onder andere dat Cicero erg bewuste keuzes maakte in zijn formuleringen. Hij bespreekt bijvoorbeeld zelfs een grammaticaal probleem met Atticus (*Ad Atticum* 7.3.10) en Pompeius vraagt hem om raad als een 'grammaticaal expert'. Maar hoewel hij dus erg puristisch kan zijn, schuwt hij aan de andere kant ook minder onlogische constructies dan Caesar (als het maar ten goede komt aan de levendigheid van de tekst en samenhang van het geheel). Over het algemeen schrijft Cicero meer op gevoel, dan star vast te houden aan regels zoals Caesar. Zo gebruikt Cicero meer anastrofes zoals *quo de* 'waarover' (in plaats van *de quo*, zoals je zou verwachten, aangezien het Latijn een taal is met linkshoofdige PP's) dan Caesar. Ervan uitgaande dat het een 'regel' is in de Latijnse grammatica om geen discontinuïteit te veroorzaken (de 'rule of domain integrity' van LIPOC), zou je een verschil in discontinuïteit dus ook kunnen verklaren door de verschillende schrijfstijlen van de auteurs. Zonder extra onderzoek kan niet vastgesteld worden of één van deze hypothesen, of beide hypothesen een verklaring bieden voor de gevonden discontinuïteit.

De overige variabelen die ingebracht zijn in het statistische model brachten daar tegen verwachting in geen significante bijdrage aan. De beheersbaarheid van de infinitivus is ingebracht door de tijd en vorm van het werkwoord samen te voegen. Het was van tevoren al duidelijk dat dit zeker geen vervanging was voor de beheersbaarheid, maar dat dit wel de beste methode was om deze objectief te meten. Het is echter een mogelijkheid dat de tijd en vorm van het werkwoord niet voldoende representatief zijn voor de beheersbaarheid, en daarom een vertekend beeld schetsen van de bijdrage ervan aan het model. Het zou ook kunnen dat de beheersbaarheid van het infinitieve werkwoord gewoon niet uitmaakt bij de woordvolgorde in zinnen met AcI's. Ik zie momenteel echter geen mogelijkheid om dit objectief te toetsen. Aangezien er geen moedertaalsprekers van het Latijn meer bestaan, is er geen goed beeld te krijgen van beheersbaarheid in deze taal, zonder deze te weergeven aan de hand van objectief te meten variabelen.

Zowel de levendigheid als de lengte van de accusativus leverden geen significante toevoegingen aan het model (met een p-waarde die zelfs erg dicht naar 1 naderde). Dit gegeven lijkt de

bevindingen van Pinkster (1984) te bevestigen: klassieke auteurs maken zich minder druk om de lengte van constituenten bij het schrijven van teksten. Hun keuzes voor de woordvolgorde in een zin hangt sterker samen met stilistische en pragmatische factoren.

Er zijn geen significanten verschillen gevonden tussen de hoeveelheid discontinuïteit bij de verschillende MATRIXPREDICATEN, maar met een p-waarde van net boven 0,05 (nl. 0,065) lijkt de dummyvariabele 'onpersoonlijke predicaten' me toch het bespreken waard. Hoewel je hier niet kan spreken van een significant effect, lijkt er toch een trend te zijn, namelijk dat de kans op discontinuïteit bij onpersoonlijke predicaten groter is ($Exp(B) = 0,139$) dan bij de andere predicaten. Het lijkt voor de hand te liggen dat áls er een verschil zou zijn, het dit verschil is, omdat dit soort predicaten het meest verschilt van de rest. Zowel de verba sentiendi als de verba dicendi en de verba affectuum zijn bijna allemaal 2-plaatsige predicaten, terwijl de onpersoonlijke predicaten 1-plaatsig zijn. De richting van de tendens is dan ook opvallend. De groep predicaten waarvan de meeste 1-plaatsig zijn zijn vaker discontinu dan de groep predicaten waarvan de meeste 2-plaatsig zijn: met minder vaste elementen in de zin, en dus minder elementen die de AcI zouden kunnen onderbreken, is de kans op discontinuïteit toch groter. Het lijkt ook in strijd met de observatie die Pinkster (1984) heeft gedaan. Hij vond namelijk dat de woordvolgorde in zinnen met 1- en 3-plaatsige predicaten op eenzelfde manier geanalyseerd kan worden als in zinnen met een 2-plaatsig predicat. Het lijkt me zeker de moeite waard om dit met een grotere dataset verder te onderzoeken.

6.1.2 Soort discontinuïteit en plaats MATRIXPREDICAAT

Uit de data-analyse bleek dat in de meeste gevallen van discontinuïteit (86%) er één onderbrekend element is, dat er vijf keer een discontinue AcI met twee onderbrekende elementen voorkomt (11,6%) en één keer met drie onderbrekende elementen (2,3%). Dit komt goed overeen met de data van Bolkestein (2001) over discontinue NP's; zij vond in 85% van de gevallen één onderbrekend element en in de overige 15% meestal twee onderbrekende elementen. Bovendien zag zij dat wanneer er twee onderbrekende elementen zijn, de elementen óf een PREDICAAT met een argument zijn, óf twee argumenten. Ook in deze data gebeurt dat op die manier: bij alle gevallen met meer dan één onderbrekend element is altijd één van de elementen het MATRIXPREDICAAT. De gelijkenis in gedrag tussen de AcI en de NP suggereert dat Johndal (2012) gelijk kan hebben in zijn analyse van de AcI als een NP-constituent.

Ongeveer de helft (51%) van de MATRIXPREDICATEN staat na de AcI, wat lijkt te bevestigen dat PREDICATEN in het Latijn vaak achteraan in de zin staan. In 17% van de gevallen onderbreekt het PREDICAAT de AcI en in 30% gaat hij eraan vooraf.

6.2 CONCLUSIES EN DISCUSSIE A+I

6.2.1 *Het model*

De logistische regressie liet zien dat de hoeveelheid discontinuïteit ook bij de A+I significant beïnvloed wordt door de auteur van het werk: wanneer de auteur Cicero is, is de kans op een discontinue A+I 4,1 keer groter dan wanneer de auteur Caesar is. Aangezien zowel bij de AcI als bij de A+I zo'n groot effect is gevonden van de auteur, stel ik voor deze bevindingen op dezelfde manier te benaderen. Voor de analyse van deze variabele verwijs ik dus terug naar het begin van paragraaf 6.1.

Bij de A+I leverde de lengte van de accusativus wél een significante bijdrage aan de verklaring van de discontinuïteit: wanneer de accusativus één woord langer wordt, wordt de kans op discontinuïteit 1,3 keer groter. Dit is geheel tegen de verwachting in; waar een gebrek aan effect van de lengte van de accusativus bij de AcI nog overeenkomt met eerdere literatuur doet deze bevinding dat totaal niet. Het geheel wordt alleen maar opvallender wanneer wordt opgemerkt dat de langste gevonden A+I, die bestaat uit negentien woorden, niet alleen discontinu is, maar zelfs twee onderbrekende elementen heeft. Nu zou het kunnen zijn dat dit meetpunt de hele variabele een vertekend beeld geeft, door de uitzonderlijk lange accusativus, maar wanneer de analyse opnieuw uitgevoerd wordt zonder dit datapunt blijft de significantie behouden (zij het met een iets kleinere $Exp(B)$ -waarde van 1,2). Het significante effect is dus naar alle waarschijnlijkheid geen type-I-fout door uitschieters in de data en zal op een andere manier verklaard moeten worden. Bovendien moet er rekening gehouden worden met de opmerking aan het begin van dit hoofdstuk: de datapunten van de A+I worden verstoord door niet-consistente annotaties, waardoor tussen deze datapunten naar alle waarschijnlijkheid meerdere AcI's zitten. Nu hebben we al gezien dat de lengte van de accusativus voor de AcI geen significante variabele was, dus wanneer deze hier toch naar voren komt, ondanks de verstoorde data, is de kans aanwezig dat het effect groter is dan deze analyse laat zien. Om dit te achterhalen zou er echter ook zelf geannoteerd moeten worden, en ik zou ook aanraden een grotere dataset te nemen, omdat de spreiding in lengte vrij groot is.

Net als bij de AcI maken de verschillende MATRIXPREDICATEN geen significant verschil, maar is er bij een aantal (de verba affectuum, dicendi en sentiendi) wel een tendens te zien, met p-waarden van respectievelijk 0.061, 0.071 en 0.076.

6.2.2 *Soort discontinuïteit en plaats MATRIXPREDICAAT*

Bij de A+I bleek dat discontinue constructies meestal één onderbrekend element hebben (63,0%), maar dat ook twee (22,2%) en drie onderbrekende elementen (14,8%) voorkomen. Discontinue predicaten met meer dan één onderbrekend element komen dus vaker voor dan in de data van Bolkestein (2001). Er is bij de A+I ook veel variatie in de functie van het

onderbrekende element: zowel PREDICATEN en SUBJECTEN (samen 52 keer) als niet-argumenten (samen 30 keer) komen vaak voor. Ook dat komt niet overeen met de data van Bolkestein.

Bij de A+I staat het overgrote deel van de MATRIXPREDICATEN na de A+I (72,5%), en een aantal staan ervoor (13%) of onderbreken hem (14,5). Dit is ook in overeenkomst met de verwachting dat de meeste predicaten aan het eind van de zin staan.

6.3 ALGEMENE CONCLUSIES EN DISCUSSIE

Er valt een aantal dingen duidelijk op bij het vergelijken van de AcI en de A+I. Hoewel het verschil in hoeveelheid discontinuïteit niet significant is, is er wel een groot verschil op te merken in het soort discontinuïteit, en ook de plaats van het MATRIXPREDICAAT. Hoewel de PREDICATEN in beide constructies het vaakst daarachter staan, is dit een sterkere tendens bij de A+I dan bij de AcI. Een mogelijke verklaring daarvoor is dat de AcI-predicaten minder vaak een betekenisvolle bijdrage leveren aan de zin en daardoor meer naar voren staat (nieuwere informatie staat per slot van rekening na bekende informatie). Daarnaast valt op dat de functies van de onderbrekende elementen bij de A+I meer variëren dan bij de AcI, en er ook een voorkomen gevonden wordt van een meervoudig discontinue A+I waarbij geen van de onderbrekende elementen het PREDICAAT is. De A+I kan dus makkelijker gescheiden worden dan de AcI door meer soorten elementen. Dat komt overeen met de verwachting dat de accusativus en de infinitivus van de A+I makkelijker gescheiden kunnen worden, omdat ze niet één geheel vormen in de matrixzin. Deze hypothese blijft staan en zal in verder onderzoek uitgebreider getoetst moeten worden.

Vervolgonderzoek is dus zeker belangrijk. Verschillende hypothesen die nu niet definitief weerlegd of bewezen zijn kunnen met nieuwe data en nieuwe annotaties wél geanalyseerd worden. Voor deze analyses is echter in het PROIEL corpus helaas nog niet voldoende data beschikbaar, maar wanneer tijd aanwezig is om zelf te annoteren, zou meer toekomstig onderzoek zeker nuttig zijn voor onze kennis over discontinuïteit van AcI's.

7. Referenties

- Albrecht, M. Von (2003). *Cicero's style: A synopsis*. Leiden: Koninklijke Brill.
- Bard, E., & Aylett, M. (1999). The dissociation of deaccenting, Givenness, and syntactic role in spontaneous speech. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (Vol. 3, pp. 1753-6). San Francisco, CA: Congress organizers at the Linguistics Department, University of California, Berkeley.
- Bard, E.G., Anderson, A.H., Sotillo, C., Aylett, M., Doherty-Sneddon, G. & Newlands, A. 2000. Controlling the Intelligibility of Referring Expressions in Dialogue. *Journal of Memory and Language* 42, 1, 1-22.
- Bauer, B.L.M (1995). The emergence and development of SVO patterning in Latin and French: Diachronic and psycholinguistic perspectives. Oxford, New York: Oxford University Press.
- Bolkestein, A.M. (1976a). AcI- and ut-clauses with verba dicendi in Latin. *Glotta*, 54, 263-291.
- Bolkestein, A.M. (1976b). The relation between form and meaning of Latin subordinate clauses governed by verba dicendi. *Mnemosyne*, 39, 155-175, 268-300.
- Bolkestein, A.M. (1979). Subject-to-object raising in Latin? *Lingua*, 48, 15-34.
- Bolkestein, A.M. (2001). Possessors and experiencers in classical Latin. In: I. Baron, M. Herslund & F. Sorensen (eds.), *Dimensions of possession* (pp. 269-284). Amsterdam: John Benjamins Publishing Company.
- Devine, A.M. & Stephens, L.D. (2006). *Latin word order: structured meaning and information*. Oxford, New York: Oxford University Press.
- Haug, D. T. T. & Jøhndal, M. L. (2008). Creating a Parallel Treebank of the Old Indo-European Bible Translations. In: C. Sporleder and K. Ribarov (eds.), *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008)* (pp. 27-34).
- Haug, D.T.T. (2009). Computational and Linguistic Issues in Designing a Syntactically Annotated Parallel Corpus of Indo-European Languages. *Revue Traitement Automatique des Langues*, 50, 17-45.
- Haug, D.T.T. (2010). *PROIEL guidelines for annotation* [Manual]. Retrieved from <https://proiel.github.io/>
- Haug, D.T.T. (2017). Backward control in ancient Greek and Latin participial adjuncts. *Nat Lang Linguistic Theory*, 35, 99-159.
- Jaeger, T.F. (2011). Corpus-based research on language production: Information density and deducible subject relatives. In: E.M. Bender & J.E. Arnold (eds.), *Language from a cognitive*

perspective: Grammar, usage, processing (pp. 161-198). Stanford: Centre for the Study of Language and Information Publications.

- Jong, J. R. de (1983). Word order within Latin noun phrases. In: H. Pinkster (ed.), *Latin Linguistics and Linguistic Theory [Studies in Language Companion Series 12]* (pp. 131-144). Amsterdam: John Benjamins Publishing Company.
- Jøhndal, M.L. (2012). *Non-finiteness in Latin* (Doctoral dissertation, University of Cambridge, UK). Retrieved from <http://johndal.com/research/>
- Kühner, R. & Stegmann, C. (1912). *Ausführliche Grammatik der lateinischen Sprache, Satzlehre*, volume I & II. Hannover: Hahnsche Buchhandlung.
- Lakoff, R. (1968). *Abstract syntax and latin complementation*. Cambridge, Massachusetts: The M.I.T. Press.
- Leumann, J., Hofmann, B. & Szantyr, A. (1965). *Lateinischen Grammatik*, volume II: Lateinischen Syntax und Stilistik. München: Beck.
- Pepicello, W.J. (1977). Raising in Latin. *Lingua*, 42, 209-218.
- Pinkster, H. (1984). *Latijnse syntaxis en semantiek*. Amsterdam: B.R. Grüner Publishing Co.
- Pinkster, H. (1990). Evidence for SVO in Latin? In: Wright, R. (ed.), *Latin and the Romance Languages in the Early Middle Ages* (pp. 69-82). London: Routledge.
- Pinkster, H. (2015). *The Oxford Latin Syntax, Volume I: The Simple Clause*. Oxford: Oxford University Press.
- Schoof, S.C. (2004). *An HPSG account of nonfinite verbal complements in Latin* (Doctoral dissertation, Rijksuniversiteit Groningen, Netherlands). Retrieved from [http://www.rug.nl/research/portal/publications/an-hpsg-account-of-nonfinite-verbal-complements-in-latin\(1f0dab31-5032-4d5d-abcb-c584e489e5e7\).html](http://www.rug.nl/research/portal/publications/an-hpsg-account-of-nonfinite-verbal-complements-in-latin(1f0dab31-5032-4d5d-abcb-c584e489e5e7).html)
- Snijders, L. (2012). Issues concerning constraints on discontinuous NP's in Latin. In: M. Butt & T.H. King (eds.), *Proceedings of the LFG12 Conference*. Stanford: Centre for the Study of Language and Information Publications.
- Spevak, O. (2010). *Constituent order in classical Latin prose*. Amsterdam: John Benjamins Publishing Company.
- Weil, H. (1978). *The order of words in the ancient languages compared with that of the modern languages*. Revised edition of the translation by Charles Syper (1887). Amsterdam: John Benjamins Publishing Company.

Woodcock, E.C. (1959). *A new Latin syntax*. Cambridge: Harvard University Press.

Zaenen, A., Carletta, J., Garretson, G., Bresnan, J., Koontz-Garboden, A., Nikitina, T., O'Connor, M.C. & Wasow, T. (2004). Animacy encoding in English: why and how. In: *Proceedings of the Association for Computational Linguistics Workshop on Discourse Annotation* (pp. 118-125). Stroudsburg, Pennsylvania, USA: Association for Computational Linguistics.