

**KATHOLIEKE UNIVERSITEIT LEUVEN**

FACULTEIT SOCIALE WETENSCHAPPEN – FACULTEIT LETTEREN

OPLEIDING BEDRIJFSCOMMUNICATIE

**De liberale politicus: een man van  
zijn woord?**

Een keywordanalyse van het geschreven discours van  
Open Vld

Promotor: Prof. Dr. Ann Bertels

MASTERPROEF  
aangeboden tot het verkrijgen  
van de graad van  
Master in de  
bedrijfscommunicatie  
door  
**Laurens Bynens**

academiejaar 2011-2012

# Inhoudsopgave

1. Inleiding.....	2
2. Literatuurstudie.....	4
2.1. Gebruik van corpora: inleidende bedenkingen.....	4
2.2. Keywordanalyse.....	4
2.3. Collocaties, clusters en concordanties.....	5
2.4. Analyse van teksten: mogelijkheden en gevaren.....	7
2.5. Critical Discourse Analysis.....	8
3. Methodologie en hypothesen.....	10
3.1. Samenstelling van de corpora.....	10
3.2. Statistische (keyword)analyse met Wordsmith Tools.....	11
3.3. Hypothesen.....	12
4. Analyse van het Nationaal Nieuws.....	14
5. Analyse van de verkiezingsprogramma's.....	20
6. Analyse van het taalgebruik van vier Open Vld-senatoren.....	26
6.1. Senator Nele Lijnen.....	28
6.2. Senator Bart Tommelein.....	30
6.3. Senator Guido De Padt.....	31
6.4. Senator Alexander De Croo.....	32
6.5. Top 50 van de keywordslijst .....	33
6.6. Keywords uit de basisfilosofie.....	35
7. Conclusie.....	36
8. Dankwoord.....	39
9. Referenties.....	40
10. Biografie.....	43
11. Summary of the thesis.....	44

## 1. Inleiding

Het liberalisme behoort tot de belangrijkste politieke stromingen van de westerse wereld. De meeste Belgen zullen de liberale partij Open Vld wel kennen en hen ook kunnen identificeren als “de liberalen”, toch is het niet altijd even duidelijk wat een partij nu liberaal maakt. “*Waar staat Open Vld voor?*” is dan ook een goede vraag die de basisfilosofie<sup>1</sup> achter de Vlaamse Liberalen en Democraten inleidt op hun website. Vanuit onze intuïtie associëren we partijen die het liberalisme behartigen met woorden als *vrijheid*, *individualisme*, *ondernemen*, *verantwoordelijkheid*, enz. Ook in de beknopte basisfilosofie vinden we enkele van deze concepten terug. Een korte bloemlezing met enkele opvallende woorden in het vet:

*“Open Vld wil de burgers nog meer inspraak geven in onze democratie. Met de toevoeging “Open” voor de afkorting benadrukken de **Vlaamse liberalen** dat ze resoluut kiezen voor een **open** samenleving. [...] In een open samenleving hangt de toekomst af van wat elk **individu** –volgens zijn of haar mogelijkheden- ervan maakt. [...] Dit bewijst dat Open Vld’ers een groot vertrouwen hebben in mensen. Ze vinden dat elke mens zo vrij mogelijk moet zijn, bijvoorbeeld om zijn of haar eigen mening te vormen of om zelf te beslissen hoe hij of zij wil leven. Elke mens is ook **verantwoordelijk** voor zichzelf en medeverantwoordelijk voor zijn omgeving. [...] Iedereen heeft gelijke rechten en plichten. Het is de taak van de overheid om de vrijheid, de privacy, de veiligheid en het recht op **privé-eigendom** van haar burgers te beschermen. [...] Daarbij vindt Open Vld dat wie **initiatief** neemt zoveel mogelijk ruimte en zo weinig mogelijk tegenwerking verdient. Voor Open Vld vormt het vrij **ondernemerschap** de bouwsteen voor onze welvaart. De **overheid** legt slechts regels op als dat nodig is om de individuele **vrijheid** of het algemeen belang te waarborgen.”*

Onze intuïties worden bevestigd bij het lezen van de basisfilosofie van de partij. De Open Vld profileert zichzelf als een liberale partij met een gematigd Vlaams karakter die opkomt voor meer vrijheid, gelijkheid en individuele verantwoordelijkheid. Logischerwijze komen deze concepten dan ook naar voor in de teksten van de partij. Maar is dit wel zo? Schrijft de Open Vld en haar politici zo liberaal als ze denken? Dit vraagt om een grondige analyse. Geen daden maar woorden! Het liberale geschreven woord staat in deze meesterproef centraal. Deze is ontstaan omwille van twee redenen. Ten eerste heb ik persoonlijk een zeer grote interesse in de (Belgische) politiek. Daarnaast heb ik in het kader van mijn master Bedrijfscommunicatie een stage gedaan bij de Open Vld-fractie in de Belgische Senaat<sup>2</sup>. Hier zag ik hoe de teksten van de politici tot stand komen: soms worden ze geschreven door de politici zelf maar vaak schrijven hun medewerkers de teksten. Toch zijn deze voldoende liberaal getint, ook al zijn ze vaak van de hand van medewerkers die niet per se zelf liberaal denken (al is dit wel aannemelijk). Een zeer belangrijke opmerking hiermee verbonden is gemaakt door Stubbs (1996), die schrijft dat teksten altijd intertekstueel zijn en worden geschreven volgens routines en conventies (Fairclough 1992). Teksten worden gekenmerkt door een 'onbewuste achtergrond', een geheel van verwachtingen en conventies. Dit inzicht was essentieel bij het tot stand komen van deze meesterproef. Er bestaat, ook binnen de liberale ideologie, zoiets als een onbewuste achtergrond waartegen we nieuwe teksten schrijven. Intuïtieve bouwstenen worden gebruikt om een nieuwsartikel, opiniestuk, schriftelijke vragen aan ministers, enz. te schrijven, die ideologisch al dan niet een duidelijk standpunt innemen. Ook ikzelf moest tijdens mijn stage leren om een tekst voldoende 'liberaal' te maken. De keuze om een woord wel of niet te gebruiken, of de keuze voor een synoniem met een iets andere betekenis, kan de hele tekst beïnvloeden.

Door mijn studies kwam ik in contact met professor Ann Bertels die me wees op de mogelijkheden van corpuslinguïstiek (CL). Met computerprogramma's zoals *Wordsmith* is het mogelijk om statistische analyses uit te voeren op gigantische hoeveelheden tekst. Miljoenen woorden

<sup>1</sup> Op [www.openvld.be](http://www.openvld.be) vindt men onder de subcategorie *Partij* de knop *Filosofie*. Hier kan men de basisfilosofie van Open Vld lezen. De volledige link vindt u in de referentielijst.

<sup>2</sup> Ik besef dat sommigen misschien vrezen dat mijn betrokkenheid bij de partij mijn objectiviteit in het onderzoek zou kunnen beïnvloeden. Ik wil hier duidelijk stellen dat dit niet het geval is. Ik kom verder in de tekst nog terug op de objectiviteit van corpusonderzoek d.m.v. computerprogramma's.

kunnen worden verwerkt om onderliggende patronen te verduidelijken. Een van die mogelijkheden is het statistisch opsporen van keywords: woorden die, in vergelijking met een referentiecorpus, opvallend vaak voorkomen in een corpus van teksten. In deze meesterproef zal ik pogen te onderzoeken welke woorden de meest typische woorden of de keywords zijn van het geschreven discours van de Open Vld. Welke woorden gebruikt de partij zeer veel en weerspiegelen deze ook hun liberale karakter? Keywords zullen dus in deze meesterproef centraal staan. Ze zullen de leidraad zijn bij de analyses en het vertrekpunt om kritisch met de teksten om te gaan. De metafoor van Scott (2010) is hier zeker op zijn plaats: een *keyword* is als een *sleutel*, die de deur naar nieuwe inzichten opent. De computer ziet zelf geen patronen maar toont enkel de bestaande tekst op een nieuwe manier. Het is aan de mens om de patronen zelf op te sporen. Keywords zijn dus analytische hulpmiddelen waarmee we de tekst kunnen ontleden (Bondi 2010). Studies doorheen de 20ste eeuw probeerden vaste betekenissen te vinden voor woorden zoals *vrijheid*, *democratie* en *gelijkheid*, alsof er een vaste waarheid omtrent deze woorden bestaat. Politieke wetenschappers, zoals Chilton & Schäffner (1997) deze mensen spottend noemen, wilden komaf maken met linguïstische verwarring en ambiguïteit, beïnvloed door het logisch positivisme. De linguïsten die zich met deze materie bezig houden steunen op de methodes en theorieën van de linguïstiek en de corpuslinguïstiek, waardoor ze danig anders te werk gaan dan hun collega's van politieke wetenschappen.

Een bijzondere stroming tussen linguïstiek en sociale wetenschappen is Critical Discourse Analysis (CDA). Ik bespreek CDA hier kort maar kom er later nog op terug. CDA onderzoekt hoe in de politieke sfeer sociale (dynamische) betekenissen worden geproduceerd of veranderd en welke structuren, strategieën of andere processen hierbij een rol spelen (Van Dijk 2003). Fairclough en Wodak (1997) zien CDA niet als een koude (haast asociale) en objectieve sociale wetenschap (al vind ik het bijna negatieve gebruik van het woord 'objectief' in deze context wel enigszins gevaarlijk, hoewel de auteurs toch oproepen tot systematische en voorzichtige analyse). Het omgekeerde is volgens hen waar: CDA is geëngageerd en een vorm van interventie in de maatschappij. Het legt de (politieke) praktijken bloot die onzichtbaar zijn voor de meeste mensen. Dit terwijl politieke organisaties mijns inziens een grote verantwoordelijkheid hebben in onze democratie. Hun discours heeft te maken met en heeft een invloed op machtsrelaties, onder andere doordat ze bepaalde concepten of zaken als 'waarheid' trachten voor te stellen. Dit is al dan niet onbewust het doel van alle politieke partijen en dus geldt dit ook voor de Open Vld. Deze meesterproef wil niet het hele discours van de Open Vld ideologisch ontleden. Wel is het doel om aan de hand van een weldoordachte selectie van teksten een kritische keywordanalyse uit te voeren met de basisfilosofie als achtergrond.

## 2. Literatuurstudie

### 2.1. Gebruik van corpora: inleidende bedenkingen

Voor we beginnen met de analyse van de teksten van de Open Vld, wil ik eerst enkele theorieën in verband met tekstanalyse en corpuslinguïstiek bespreken om zo mijn eigen analyse te verantwoorden en theoretisch te kaderen. Er is doorheen de jaren heel wat literatuur geschreven over het gebruik van corpora en ideologische teksten. Op de volgende pagina's zal ik enkele van de belangrijkste auteurs en hun werken bespreken of toch tenminste aanhalen. Zo legt bijvoorbeeld Hunston (2002) op een heldere manier de doelen van corpora uit. Een corpus wordt volgens haar gedefinieerd op basis van zijn vorm en doel. Tegenwoordig ziet men een corpus steeds vaker als een collectie teksten die elektronisch worden opgeslagen. De inhoud wordt gepland (soms speelt toeval mee) en heeft een linguïstisch doel. Dit doel bepaalt de selectie van de teksten. Het gaat hierbij niet puur om het bewaren van teksten, zoals in een bibliotheek of archief. Het doel is om de data kwantitatief en kwalitatief te gebruiken. Waar corpora vroeger bijna exclusief gebruikt werden binnen het domein van het taalonderwijs, zijn de toepassingen tegenwoordig veel breder, bijvoorbeeld het analyseren van politieke teksten. Als een corpus een deel van de taal weergeeft, dan kunnen we met software deze taal onderzoeken op manieren die anders onmogelijk zijn.

Volgens Biber et al. (1998) heeft een corpusanalyse enkele essentiële eigenschappen: ten eerste is de analyse empirisch en worden bestaande patronen van feitelijk taalgebruik geanalyseerd. Ten tweede gebruiken we een grote verzameling van authentieke teksten (d.i. het corpus) als basis voor de analyse. Daarnaast maken we veelvuldig gebruik van computers, en hierbij passen we zowel automatische als interactieve technieken toe. Tot slot zijn de analysetechnieken zowel kwalitatief als kwantitatief. Deze combinatie is volgens mij een van de sterkste punten van corpuslinguïstiek, al staat dit standpunt hevig ter discussie in de literatuur. Deze problematiek wordt verder aangehaald in de literatuurstudie (zie hiervoor 2.4 en 2.5). Al deze punten zorgen er volgens Biber et al. (1998) voor dat er geen andere vorm van analyse zo betrouwbaar is. Ook Sinclair (1991) is optimistisch over de mogelijkheden van computertechnologie inzake tekstanalyse. Door analyse met de computer krijgen we zicht op zaken die vroeger onzichtbaar waren. Het is duidelijk geworden dat het niet in de natuur van een tekst zit om duidelijk afgelijnde regels te volgen die a priori bepaald zijn maar dat er ook veel flexibiliteit en innovatie is. De discipline wordt dan ook een stuk meer wetenschappelijk verantwoord wanneer we betekenis gaan bepalen aan de hand van formele, statistisch verantwoorde patronen in plaats van met onze intuïtie. Zo geeft bijvoorbeeld een woordenboek aan welke de verschillende betekenissen van een woord zijn, maar het is niet altijd even duidelijk hoe men het woord moet gebruiken om deze verschillende betekenissen weer te geven. Een thesaurus plaatst woorden met gelijkaardige betekenissen samen maar de classificatie is abstract en conceptueel, aldus Sinclair (1991). Het gebruik van een programma als *Wordsmith Tools* laat daarentegen toe om de concrete betekenis van keywords in hun specifieke context te bepalen.

### 2.2. Keywordanalyse

Paquot & Bestgen (2009) omschrijven keywords beknopt als “items met een ongewone frequentie in vergelijking met een referentiecorpus”. Scott (2001) merkt op dat een keyword dus een lexicaal item is dat mee duidelijk maakt waar een tekst ‘over gaat’. Dit laatste wordt ook wel *aboutness* genoemd (Scott 2000). Aboutness is echter subjectief: de lezer moet zelf beoordelen waar de teksten over gaan en of er een gemeenschappelijke aboutness is. Volgens Scott (2001) kunnen keywords ons dus helpen bij het bepalen van de aboutness. Als een woord relatief vaker voorkomt in een tekst of analysecorpus, dan is het in vergelijking met de ‘algemene taal’ of een referentiecorpus waarschijnlijk belangrijk binnen het analysecorpus. Dit inzicht staat centraal in deze meesterproef. Een andere definitie van keywords, meer van toepassing op sociale wetenschappen, wordt gegeven door Williams (1976). Hij ziet keywords als de typische woordenschat van een cultuur of maatschappij. Van politici wordt soms gezegd dat ze in een andere taal spreken. Dit komt volgens Williams omdat ze andere waarden hebben of omdat ze zaken anders beoordelen. Het gaat bij Williams niet over woorden uit een gespecialiseerd domein of over speciale woorden. Hij bekeek de afkomst van alledaagse

woorden (bv. *familie, ervaring, kunst, enz.*) en wat hiermee gebeurde doorheen de geschiedenis. Ook bij de Open Vld zijn de keywords woorden die deel uitmaken van onze dagelijkse woordenschat. Toch gebruikt de partij deze woorden vaak op diens eigen wijze, om er haar visie op onze samenleving mee op te bouwen.

Keywords hebben, zoals net aangehaald, een hoge frequentie in vergelijking met een referentiecensus. Sinclair wijst in zijn werken meermaals op het feit dat linguïsten het belang van woorden met een hoge frequentie onderschatten. Volgens hem zijn de woorden, enkel al door hun hoge frequentie, een dominante factor in de tekst, maar weinig woorden “hebben een betekenis onafhankelijk van de co-tekst” (Sinclair 2004). Een woord heeft verschillende betekenissen, maar de specifieke betekenis wordt bepaald door de context. We kunnen ons bij de analyse van keywords dus niet beperken tot één betekenis of tot een basisbetekenis. Bertels (2008) toonde immers ook aan dat woorden met een hoge keyness vaak allesbehalve monosemantisch zijn. Het doel van deze meesterproef is niet om bij elk keyword te bepalen wat nu de exacte betekenis er van is voor de liberalen (voor zover dit al mogelijk is). Wel willen we aantonen dat een keyword zijn conceptuele betekenis krijgt doordat het in contact staat met zijn talig en contextuele omgeving.

Stubbs (1996) relateert de opvatting van Sinclair: frequentie moet vooral correct geïnterpreteerd worden. Men veronderstelt dat patronen van herhaling en variatie de betekenis duidelijk maken, al is dit volgens Stubbs niet altijd correct. Het register (hieronder verstaat hij tekst type, stijl, genre enz.) wordt pas duidelijk als men kijkt naar de systematische variatie van de frequentie van linguïstische features over teksten heen. Een vergelijkende analyse van de verschillende teksten zorgt voor de juiste interpretatie (i.p.v. enkel oog te hebben voor absolute frequentie). Zo komen we te weten wat typisch en atypisch is en welke delen van de tekst significant (linguïstisch of ideologisch) zijn.

Hoey (2007) heeft het in deze context over 'lexical priming': telkens we een woord tegenkomen onthouden we volgens de auteur onbewust de context en co-context. Als we het woord zelf gaan gebruiken, zullen we de context zoals we die eerder zagen gaan nabootsen. Hiermee verbonden herkennen we en hergebruiken we volgende zaken: grammaticale patronen waarin het woord voorkomt, de grammaticale functies van het woord, betekenissen en semantische associaties ('semantische voorkeur' volgens Sinclair (2004)), pragmatische associaties ('semantische prosodie' volgens Sinclair), genres, stijlen en domeinen waartoe het behoort, tekstuele positie (begin of einde zin), enz. Deze 'priming' is een eigenschap van het individu, niet van het woord en gebeurt dus anders bij elke persoon. Betekenis ontstaat voor een individu wanneer hij/zij met het lexicaal item in contact komt. Daarom toont een corpus ook niet de directe ervaring van mensen met taal, maar het toont wel welke woorden op welke manier geprimed worden. Deze theorie is interessant voor deze paper, omdat uit de analyse moet blijken welke woorden op welke manieren collectief 'geprimed' worden door de mensen van de Open Vld. Een specifiek corpus toont volgens Hoey (2007) immers beter dan een algemeen corpus hoe bepaalde woorden geprimed worden binnen een specifiek genre of domein, in dit geval de politiek.

### 2.3. Collocaties, clusters en concordanties

Keywords (en hun contextuele betekenis) zullen dus centraal staan in onze analyse. Toch moeten we volgens Firth (1957) ook naar de omgeving rond de keywords kijken. Al in 1935 wees hij linguïsten op het feit dat vorm en betekenis onafscheidelijk zijn. Tot dan wees de linguïstiek dit inzicht af, een van de redenen waarom men ook kritisch was ten aanzien van de corpuslinguïstiek. Dit was dan ook een grove fout, want vorm en betekenis zijn wel verbonden (cf. Firth 1949 & Mahlberg 2005), wat bewezen wordt door het analyseren van collocaties. We spreken van een collocatie als twee of meerdere woorden statistisch vaker voorkomen dan verwacht. Collocaties tonen aan welke associaties en connotaties een woord heeft, en dus ook welke intuïties of eigenschappen leven in een taal bij een bepaald woord (Stubbs 1996). Stubbs acht de connotatie van een woord immers even belangrijk (of nog belangrijker!) dan de referentiële betekenis van een woord: “Het is niet wat we zeggen, maar hoe we het zeggen”. Zo kunnen we contextueel en sociologisch naar teksten kijken aan de hand van lexicaal collocaties (Firth 1935). Uit de definitie van een collocatie volgt dat de context rond de kern

(het keyword) ook belangrijk is, of om het te zeggen met de beroemde woorden van Firth: “You shall know a word by the company it keeps” (Firth 1957). Betekenis wordt niet enkel gecreëerd door een woord op zichzelf maar wordt mee gevormd door zijn omgeving. Bertels (2008) deed een interessante ontdekking in haar onderzoek: polyseme substantieven kunnen semantisch sterk gedesambigueerd worden door de aanwezigheid van adjectieven, determinanten en werkwoorden. Hierbij speelt de context een belangrijke rol. Betekenis wordt dus bewust gecreëerd door woordkeuze maar ook onbewust door de keuze voor bepaalde collocaties. Scott & Tribble (2006) onderscheiden negen ‘scopes’ of spans, oftewel hoe breed de context rond een woord (hier de *node*) kan zijn. Enkele voorbeelden van scopes zijn een paar woorden links en rechts van de node, het niveau van de zin of van de tekst, en zelfs andere gerelateerde teksten, enz. Units van meerdere woorden kunnen interessant zijn als ze vaak voorkomen, want dan worden patronen zichtbaar hetzij op basis van pure frequentie, hetzij omdat beide concepten deel zijn van ons mentaal lexicon.

Verbonden met collocaties zullen we in deze meesterproef ook clusters onderzoeken. Clusters zijn groepen van woorden die in een welbepaalde volgorde zeer vaak samen voorkomen. De relatie tussen de woorden is vaster en sterker waardoor ze vaste verbindingen vormen. De sterkste soort cluster is het idioom maar andere clusters zijn ook mogelijk. Woorden trekken elkaar immers ook aan zonder dat het om een idioom gaat (Milizia 2010). In een corpus van de *Open Vld* lijkt het zeer aannemelijk dat de cluster *Alexander + Croo* vaak gaat voorkomen, omdat deze woorden in dit corpus een sterke relatie hebben.

Sinclair (1991) heeft het over semantische prosodie: prosodie die zich over meer dan een woord uitstrekt en dus op meerdere woorden betrekking heeft. Zo vond Stubbs (2009) dat het woord *oorzaak* in zijn Engels corpus vaak een negatieve semantische prosodie had en dat de collocaties vaak negatief zijn (ongeluk, schade, dood, enz.). In deze context ziet Stubbs (2007) semantische prosodie (wat hij dan weer ‘discourse prosody’ noemt) als de discursieve functie van een eenheid van woorden. De keuze voor een bepaalde woordvorm drukt ook de intentie en de houding van de spreker uit. Hij wijst ook op het belang van het gebruik van computertechnologie bij het zoeken naar collocaties: door bijvoorbeeld te kijken naar de absolute frequentie van elke collocatie, krijgen we volgens hem zicht op belangrijke culturele concepten, door Stubbs ook wel culturele keywords genoemd. Het gaat hierbij om vaste en half-vaste uitdrukkingen (collocaties, catch phrases, clichés en idiomen), die culturele informatie encoderen. Denk hierbij aan voorbeelden als *ten strijde trekken tegen*, *zich sterk maken*, *lik op stuk geven*, *open kaart spelen*, enz. In deze uitdrukkingen worden ervaringen over de cultuur gerepresenteerd en doorgegeven. Zo wordt ook ideologie in deze culturele patronen uitgedrukt. Moon (1994) beargumenteert hoe vaste uitdrukkingen een belangrijke functie hebben: ze geven socio-culturele schemata weer en sluiten aan bij gedeelde culturele waarden. Zulke uitspraken (waaronder bijvoorbeeld metaforen) sluiten nu eenmaal aan bij het gemeenschappelijk gedachtegoed. Hun vorm is stereotiep en hun inhoud is zo ‘vast’ dat mensen deze snel als waarheid aanzien. Stubbs (2007) wijst er wel op dat er vervolgens moet gekeken worden naar het voorkomen van zowel kern als collocaat apart: als beide zeer veel voorkomen, kan hun collocatie toeval zijn. Daarnaast kan een woord relatief weinig voorkomen maar tegelijk wel vaak bij een ander woord, waardoor het in deze context toch opvalt. Het is echter mogelijk dat patronen moeilijk opgespoord kunnen worden door inflectie van werkwoorden of door onderbreking van de collocatie.

Met computerprogramma's zoals *Wordsmith Tools* kunnen we collocaties nog verder onderzoeken door middel van een concordantielijst: willen we een keyword verder analyseren, dan geeft *Wordsmith* een lijst met alle voorkomens van dat woord in elke specifieke context. Naast deze concordantielijsten toont het ook de belangrijkste collocaten. Tevens laat het toe om de sterkte van de collocatie (dit is een combinatie van een *node* en een *collocate*) te berekenen en te sorteren per node. Hiermee krijgen we zicht op de manier waarop een woord gebruikt wordt in het corpus. Een concordantie brengt immers verschillende uitspraken samen van verschillende sprekers of schrijvers op verschillende tijdstippen. Het toont onderliggende, wederkerende patronen en “bevest” de taal zodat de analyse ervan mogelijk wordt (Stubbs 2009). Stubbs geeft toe dat hierbij enige subjectiviteit

aanwezig kan zijn, maar corpuslinguïstiek is voor hem toch een zeer betrouwbare vorm van onderzoek. De lengte van het citaat (dat het keyword of de node bevat) in een concordantie is meestal het universele KWIC formaat (keyword in context), waarbij het keyword in het midden staat en de lengte zo breed is als het blad papier. De volgorde van de citaten kan wisselen: men kan op tekstvolgorde of alfabetisch op het woord voor of na het keyword. Welke methode men kiest, benadrukt sommige feiten en verbergt andere (Sinclair 1991).

Een laatste theoreticus die we met betrekking tot collocaties aanhalen is Damon Mayaffre. Hij acht het gebruik van corpora essentieel voor de semantiek, want “woorden hebben geen betekenis en waarde behalve in de context” (Mayaffre 2005). Het corpus is de statistische norm om keywords en collocaties te vinden, “zonder dewelke elke telling zinloos is” (Mayaffre 2007). Een woord als *vrijheid* komt misschien niet veel voor in een corpus vol wetenschappelijke artikels over fysica maar in een corpus met liberaal getinte teksten verwachten we dit woord veel meer. Daarnaast is het corpus ook de linguïstische norm, aangezien er geen absolute taalnorm bestaat: “de norm *is* het corpus” (Mayaffre 2009). De linguïstische regelmatigheden en onregelmatigheden van een tekst zijn pas zinvol wanneer het corpus als referentie geldt. Het corpus vormt de 'gewone' context, terwijl collocaties, belangrijk voor de interpretatie van een woord, de 'minimale context' van een woord tonen (Mayaffre 2008). Collocaties laten volgens hem ook toe om belangrijke woorden op een toegankelijke en systematische wijze te onderzoeken.

#### 2.4. Analyse van teksten: mogelijkheden en gevaren

Het is fantastisch dat we met een programma zoals *Wordsmith Tools* semantische, taalkundige en zelfs ideologische patronen kunnen ontdekken in een corpus van duizenden woorden. We kunnen dan ook akkoord gaan met Mayaffre (2002), die vindt dat we zo op een geheel nieuwe manier met teksten omgaan: we kunnen het corpus globaal, statistisch en synthetisch bekijken maar evenzeer voor een lokale, particuliere lectuur kiezen omdat we met één muisklik de lokale context van het woord kunnen onderzoeken<sup>3</sup>. Deze systematische contextualisering (groot of klein) wordt mede mogelijk gemaakt door de principes van keywords, collocaties, concordanties, enz. Mayaffre (2002) stelt ook dat de computerlinguïstiek ons in staat stelt om intuïtieve, subjectieve of empathische interpretaties van het corpus te vermijden: de statistiek objectiveert in grote mate de tekst. Dit is misschien geen ondanbare gedachte, zeker omdat het in deze paper om ideologisch geladen tekstmateriaal gaat.

Mayaffre wijst ons echter ook op de gevaren van analyse d.m.v. computerprogramma's, omdat een dergelijke lezing van de teksten onze semantische interpretatie verandert. Mogelijk overschatten we het belang van bepaalde woorden of lezen we tekstuele intenties die de schrijver helemaal niet had. Dit terwijl we veel belangrijker zaken over het hoofd kunnen zien. “Hoge frequentie moet een gids zijn, niet een obsessie.” (Mautner 2009) Ook weinig voorkomende items kunnen interessant zijn en enkel naar frequentie kijken zegt niet alles. Het corpus beheerst dan de onderzoeker in plaats van omgekeerd. Soms vinden we zaken die we bewust wilden vinden, we forceren een “projectie van de hypothesen” op het corpus (Mayaffre 2002)<sup>4</sup>. De taal wordt bij analyses een zodanig evident gegeven, dat vaak elke vorm van voorzichtigheid inzake methodologie wordt opgegeven. Mautner (2009) merkt ook nog een ander mogelijk gevaar op: het corpus is vaak gigantisch groot en de contextuele informatie kan voor elk onderdeel niet verzekerd worden, zodat uitspraken in termen van het socioculturele mogelijk niet verantwoord zijn. Hiervoor moeten we ons hoeden. We zullen voor de belangrijkste keywords dus ook (beknopt) de contexten bekijken. Grote corpora bestaan volgens Mautner uit gedecontextualiseerde informatie. De specifieke context gaat verloren en dus ook een belangrijk deel van het effect van een tekst. Dit verklaart deels waarom CDA niet altijd even happig is

<sup>3</sup> Ik ga minstens akkoord met deze stelling maar deel niet zijn visie op de 'negatieve' invloeden van cultuur, maatschappij, onderzoeker, enz.

<sup>4</sup> Mautner (2009) deelt deze mening niet: door de combinatie van CDA en CL kan men een van de grootste kritieken tegen CDA overwinnen. Volgens critici zou CDA gebruik maken van een kleine selectie teksten om de vooropgestelde doelen en meningen van de onderzoeker te staven en diens vooroordelen over de ideologie of persoon die hij onderzoekt, te bevestigen. Grote referentiecorpora helpen dus tegen het gevaar van het over- of onderschatten van bepaalde sociaal relevante feiten.



om aan corpuslinguïstiek te doen (Mautner 2009) De auteur merkt in één adem ook wel op dat deze decontextualisering kan helpen bij het doen van algemene uitspraken, omdat je geen last hebt van alle 'contextuele bagage' bij het onderzoek: men analyseert gewoon de 'platte tekst'. De computer doet overigens natuurlijk niet alles: uiteindelijk moet de onderzoeker de data evalueren en interpreteren, maar de software kan wel helpen om nieuwe perspectieven te openen en betrouwbare methoden te hanteren.

## 2.5. Critical Discourse Analysis

Zoals eerder aangehaald, werken we in deze thesis met tekstmateriaal met een duidelijke ideologische waarde, daar het de woorden van een openlijk liberale partij zijn. Dit maakt dat we met de nodige voorzichtigheid te werk moeten gaan in onze analyses, om zo de liberale ideologie in de tekst<sup>5</sup> te vinden (in de mate dat ze aanwezig is). Ideologie is namelijk een manier om de werkelijkheid voor te stellen en te construeren. Het is een constructie van identiteit, verbonden met macht, zowel de identiteit van het individu als van de groep of gemeenschap. Volgens Kress (1985) wordt ideologie het sterkst gearticuleerd in taal, en dus is taalonderzoek de beste manier om een ideologie te onderzoeken. Ook Fowler (1991) deelt deze mening: taalanalyse legt het verborgen discours bloot dat we soms niet meer opmerken door de "habitualisatie van het discours". We begeven ons met deze studie dus ook op het terrein van de Critical Discourse Analysis (CDA), een multidisciplinaire studie van teksten in hun sociaal-maatschappelijke context. Corpuslinguïstiek (CL) en CDA kunnen perfect samenwerken om te achterhalen hoe taal 'werkt'. Het bewijs hiervan is dat Firth, een van de grondleggers van de CL, zijn onderzoek deed met een sociale oriëntatie. Hij plaatste betekenis in een sociologisch daglicht. Hij stuurde het onderzoek van keywords dan ook in die richting, omdat hij deze woorden als sociologisch belangrijk zag (Mautner 2009). Er is evenwel geen eenduidig theoretisch framework voor CDA (Van Dijk 2003). Afhankelijk van de wetenschapper is het onderzoek meer taalkundig<sup>6</sup> dan wel meer sociologisch<sup>7</sup>. CDA is een discipline die zich bezighoudt met onderzoek naar hoe onze maatschappij mee gevormd wordt door tekst in al zijn vormen. Men onderzoekt sociale structuren en machtsrelaties, hoe deze instituties taal vormen en zelf gevormd worden door taal<sup>8</sup> (Mahlberg 2005). CDA vraagt zich af hoe taalgebruik attitudes, meningen en zelfs ideologieën kan vormen, en hoe deze standpunten in het taalgebruik geëncodeerd zitten (Stubbs 1996). De inhoud van een corpus moet dus volgens Mahlberg gekozen worden, zonder te kijken naar de taal, maar door te kijken naar de communicatieve functie die de inhoud heeft in de gemeenschap waarin het corpus ontstaat. Als men de teksten kiest volgens duidelijke criteria, dan kan het corpus goede inzichten geven over de betekenis die het heeft in de maatschappij en duiden hoe de maatschappij zichzelf vormt. Mahlberg merkt echter ook op dat CDA en CL niet gemakkelijk te verzoenen zijn, omdat corpusanalyse het grote plaatje bekijkt en CDA eerder diepgaand naar verbindingen tussen teksten kijkt. Maar CL kan wel zeer nuttig zijn om teksten in verband met ideologieën te ontleden omdat ze subtiele betekenissen kan blootleggen. Dit lijkt mij, bijvoorbeeld aan de hand van analyses van keywords en concordanties, een realistische opgave.

---

<sup>5</sup> Binnen de CDA bestaat er onenigheid over het gebruik van de woorden 'tekst' en 'discours'. Kress (1985) wijst er op dat wetenschappers met een sociologisch-wetenschappelijke achtergrond vaak de term 'discours' verkiezen, terwijl linguïsten zoals Van Dijk voor 'tekst' kiezen (al wil ik hier niet mee zeggen dat Van Dijk de ogen sluit voor het sociologische, allerminst zelfs!). Kress (1985) tracht de twee termen duidelijk op te delen in twee aparte domeinen. Ik ga hier niet ingaan op deze discussie (en ik probeer ook allerminst om het onderscheid duidelijk te maken in dit werk) omdat ze ons te ver zou voeren maar maak de lezer enkel attent op het debat. De discussie over 'discours' wordt ook o.a. aangehaald door Guilhaumou (2004).

<sup>6</sup> Een goed voorbeeld is Orpin, die niet begrijpt waarom CDA zoveel bekritiseerd wordt omdat het niet taalkundig verantwoord zou zijn. Hoewel het een multidisciplinaire wetenschap is, "is CDA stevig in toegepaste linguïstiek gesitueerd" (Orpin 2005).

<sup>7</sup> Hierbij denken we aan de 8 principes van CDA van Fairclough & Wodak (1997): machtsrelaties zijn discursief, CDA behandelt sociale problemen, discours vormt maatschappij en cultuur, discours doet ideologisch werk, discours is historisch, discours is een vorm van sociale actie en de link tussen tekst en maatschappij wordt gemedieerd.

<sup>8</sup> Deze relatie wordt op een zeer heldere manier besproken door Phillips et al. (2004). Ook Searle (2005) beschrijft de rol van instituties in onze maatschappij.

Er zijn verschillende types CDA, elk met eigen theorieën en verschillende analyses die hiermee verbonden worden. De CDA die zich bezig houdt met gesproken tekst verschilt van analyses van krantenartikels. Een centraal concept is “macht”, wat Van Dijk (2003) verbindt met controle: (politieke) groepen hebben macht als ze de acties en gedachten van andere (politieke) groepen kunnen beheersen. Hiervoor is een machtsbasis nodig, waardoor men toegang heeft tot schaarse sociale hulpbronnen zoals kennis, geld, status, informatie, enz. Gedomineerde groepen (de macht over deze groepen varieert) kunnen deze macht weerstaan, bestrijden, accepteren of zelfs 'natuurlijk' vinden. Dit laatste lijkt mij het doel van elk tekstueel feit van politieke oorsprong.

Als we de verbindingen onderzoeken tussen teksten die een aspect van de maatschappij weerspiegelen, stelt zich volgens Mahlberg (2007) de vraag welke lexicale items belangrijk zijn en hoe men die kan opsporen. Ze vindt WordSmith Tools één van de belangrijkste tools, waarbij men door frequenties te vergelijken (met referentiecorpus) *aboutness* en stilistische features kan vinden. Scott & Tribble (2006) toonden zo bijvoorbeeld aan hoe artikels in de krant The Guardian niet de wereld an sich tonen maar wel de wereld zoals The Guardian die kiest te tonen (zo worden bijvoorbeeld enkel nieuwswaardige artikels gepubliceerd). Ze analyseerden ook welke persoonlijkheden belangrijk waren tussen 1996 en 2001 (Clinton, Blair, Bush, Gore, Pinochet, ...). Door deze analyse, met het gebruik van collocaties, kon men aantonen wie werd voorgesteld als de goede en slechte in de gebeurtenissen omtrent de aanslagen op 11 september 2001. Keywords kunnen gebruikt worden om te tonen hoe de wereld en de maatschappij worden voorgesteld. Sommige keywords zijn in een cultuur of maatschappij uitermate belangrijk, niet alleen voor dat deel van de maatschappij maar voor de hele samenleving (zoals eerder opgemerkt door Williams 1976).

We kunnen ons vinden in de stelling van Stubbs (1996) dat een vaste, stabiele betekenis niet bepaald kan worden alvorens men het woord in zijn gebruik bekijkt. De betekenis van een woord is flexibel en het gebruik kan veranderen als het in andere contexten voorkomt. Via analyse krijgen we ook zicht op de abstracte patronen die zich niet op het individuele niveau van de woorden bevinden, maar in de grammatica en in de combinatoriek. Via deze linguïstische middelen wordt immers, al dan niet onbewust, ook 'ideologie' overgebracht naar de ontvanger, aldus Stubbs. De analyse van de keywords en van hun grammaticale en lexicale patronen legt het ware discours bloot, en toont het verschil tussen verschillende ideologieën. Het hoeft niet zo te zijn dat dit altijd bewust gebeurt, maar men schrijft wel op één manier en niet op een andere. Men had, door anders te schrijven, de zaken ook anders kunnen voorstellen. Woordkeuze verandert hoe de wereld wordt voorgesteld en heeft een invloed op de relatie tussen de lezer, andere mensen en de maatschappij. Een woordkeuze gebeurt altijd in een institutionele context. Woordkeuzes hebben gevolgen: Danet (1980) onderzocht het taalgebruik van advocaten in een zeer beroemde rechtszaak waarin een dokter abortus had gepleegd. De arts werd veroordeeld voor *moord* op een *baby*, terwijl hij misschien was vrijgesproken als zijn advocaat er op had aangedrongen dat het ging om *abortus* op een *foetus*. Zo gaat het ook in de politiek met woordkeuzes: denk maar aan andere termen om bijvoorbeeld *illegalen*, *werklozen* en *bedrijfsleiders* te benoemen.

### 3. Methodologie en hypothesen

#### 3.1. Samenstelling van de corpora

Een corpus samenstellen lijkt op het eerste zicht onproblematisch, maar in de literatuur stuiten we toch op een aantal opmerkingen. We zullen er hier nog enkele bespreken, voor zover dit probleem nog niet aan bod is gekomen in de vorige onderdelen. Door politiek discours te verbinden met termen als macht, conflict, controle, dominantie, enz. (zoals Van Dijk 2003), kan men volgens Wilson (2003) bijna elk discours politiek noemen, waardoor ook elke discoursanalyse politiek getint wordt. Daarom stelt hij voor om het onderwerp dermate af te bakenen: “formele/informele politieke contexten en politieke actoren; hierbij gaat het om politici, politieke instituten, overheden, politieke media en politieke ondersteuning die allen actief zijn in politieke omgevingen om politieke doelen te realiseren.” Hierdoor sluiten we volgens Wilson wel een groot deel van het politiek discours uit, namelijk het politiek discours van het dagelijkse leven. Dit is echter een offer dat we in het licht van deze verhandeling graag brengen. Mayaffre (2009) vindt dat een corpus homogeen moet zijn. De teksten moeten niet allemaal exact hetzelfde zijn maar ze moeten wel vergelijkbaar zijn. Wanneer men homogene corpora heeft samengesteld, kan men volgens hem beginnen met een contrastieve studie van de corpora.

Het referentiecorpus bestaat uit een groot aantal artikels van de website van De Standaard ([www.standaard.be](http://www.standaard.be)). De Standaard profileert zich als een kwaliteitskrant en de krant heeft ook dit imago. Inhoudelijk is er vooral aandacht voor politiek en socio-economisch nieuws. Sport en cultuur worden ook dagelijks bericht. In het verleden werd de krant duidelijk geïnspireerd door het Vlaamse en christen-democratische gedachtegoed. Tegenwoordig profileert de krant zich een stuk neutraler. Of zoals de redactie het zelf stelt: “De verdediging van de rechtmatige belangen van de Vlaamse bevolking en het respect voor de christelijke beginselen blijven - samen met de sociaal gecorrigeerde vrijemarkteconomie en de pluralistische democratie - dé basiswaarden voor de redactie van deze krant. Maar met evenveel nadruk stelt de redactie dat De Standaard ongebonden blijft tegenover elke politieke partij, elke ideologie en elke zuil.” (7 september 1999). Dit overwegende vormt de Standaard voor mij een goede bron voor een referentiecorpus. Zoals eerder aangehaald, vormt dit corpus de linguïstische en statistische maatstaf en standaard (what's in a name?) om het tekstmateriaal van de Open Vld mee te vergelijken. In totaal bestaat dit corpus uit 200.106 woorden. De artikels zijn verzameld in de periode oktober 2011 tot en met februari 2012. Diverse thema's komen aan bod in de artikels (politiek, economie, sport, cultuur, varia, enz.) en geen enkel thema overheerst het corpus. Zo belichaamt dit corpus in de vergelijking met de liberale teksten op een representatieve manier het concept 'standaardtaal' waardoor goede analyses mogelijk moeten zijn.

Het eerste corpus van de Open Vld bevat alle nieuwsartikels die in 2011 op de nationale website van de partij zijn verschenen. Dit 'nationaal nieuws' bevat berichten over acties en initiatieven van politici van de Open Vld, de visie van de partij op actuele gebeurtenissen en thema's, opiniestukken van politici, enz. Via dit kanaal kan de partij haar leden informeren en haar visie op de actualiteit duidelijk maken. Dit corpus bestaat uit 170.573 woorden. De inhoud zal, net als de andere corpora van de Open Vld, vanzelfsprekend sterker politiek getint zijn dan de artikels van De Standaard maar hierdoor is de keywordanalyse net mogelijk en levert ze interessante informatie op.

Het tweede corpus van de Open Vld bevat de verkiezingsprogramma's van de partij sinds 1994 (toen de partij nog VLD<sup>9</sup> heette) tot 2010. Het corpus bestaat uit 164.665 woorden. De programma's zijn geschreven voor de verkiezingen op federaal (Kamer en Senaat), Vlaams, Europees, gemeentelijk en provinciaal niveau. Deze teksten tonen volgens mij het sterkst de liberale ideologie, al zal onderzoek dit moeten uitwijzen. Dit corpus is zeer interessant voor een keywordanalyse, want met haar woorden moet de Open Vld trachten te kiezer te overtuigen.

Het derde en laatste corpus van de Open Vld bevat teksten van de websites van vier politici van Open Vld: Bart Tommelein, Nele Lijnen, Guido De Padt en Alexander De Croo. Allen zijn ze senator voor Open Vld. Alexander De Croo is daarnaast ook nog voorzitter van de partij. De keuze voor vier senatoren is

---

<sup>9</sup> De teksten van VLD en Open Vld zullen als gelijkwaardig worden behandeld. Partijen evolueren natuurlijk doorheen hun geschiedenis maar het mag duidelijk zijn dat de (Open) Vld bekend is gebleven als de liberale partij in Vlaanderen.

willekeurig, we hadden ook kunnen kiezen voor leden van de Kamer<sup>10</sup>. Door de beperkte ruimte in deze meesterproef houden we het enkel bij vier senatoren, zodat hun politiek profiel toch enigszins homogeen is. Elke senator heeft zijn of haar thema's waarrond zij in de dagelijkse politiek werken, maar op hun website kunnen ze ook opiniestukken schrijven over andere onderwerpen of gebeurtenissen. Het materiaal komt dus van hun persoonlijke website. Het corpus telt in totaal 105.103 woorden, dus elke senator heeft een aandeel van ongeveer 25.000 woorden. Na analyse moet duidelijk worden welke thema's belangrijk zijn bij deze senatoren.

### 3.2. Statistische (keyword)analyse met Wordsmith Tools

*Wordsmith Tools* laat ons toe om een corpus vanuit verschillende invalshoeken te analyseren. Ten eerste is er *Wordlist*, een programma dat op basis van een corpus een lijst samenstelt met de woorden en hun frequentie of aantal voorkomens. Men laadt één of meerdere tekstdocumenten in *Wordlist* en men krijgt een frequentielijst. Bovenaan deze lijst staat standaard het woord met de hoogste absolute frequentie. Men kan ook rangschikken op alfabetische volgorde. *Wordlist* is niet het meest belangrijke programma in deze meesterproef maar het is toch een voorbereidende stap bij de analyse van keywords. De keywords worden immers berekend op basis van frequentielijst, meer bepaald op basis van de vergelijking van de relatieve frequentie (absolute frequentie en corpusgrootte) in twee corpora. Bovendien zullen we bij een analyse niet enkel rekening houden met de keyness van een keyword maar ook met zijn absolute frequentie.

Ten tweede gebruiken we de toepassing *KeyWords*, om de keywords in het analysecorpus te vinden. Hiervoor wordt de frequentielijst van een analysecorpus vergeleken met de frequentielijst van een referentiecorpus. Men krijgt dan van het analysecorpus de lijst met keywords, die standaard op dalende keyness gerangschikt worden. Op die manier ziet men dadelijk welke de belangrijkste keywords zijn. De minimumfrequentie om opgenomen te worden in lijst van keywords is in onze analyses drie. We gebruiken log-likelihood als statistische maat voor de berekening van de keyness van de data en een p-waarde kleiner dan 0,000001 voor de betrouwbaarheid van de resultaten. Keywords vinden is niet de enige functie van deze toepassing. Met *plots* ziet men de verdeling van elk keyword doorheen het corpus. *Links* toont met welke andere keywords elk keyword verbonden is. Deze functie is zeer interessant omdat het ons de verbindingen en associaties bij elk keyword toont. In onze analyses is de minimumfrequentie voor een link één, de span van een link is één tot vijf woorden. Met *clusters* kunnen we vanzelfsprekend de clusters bekijken die voorkomen in het analysecorpus. De minimumfrequentie voor een cluster om opgenomen te worden in de lijst is in onze analyses twee. Bij de analyse van de keywords beperken we ons in de analyses tot de top 50. Een uitzondering wordt gemaakt voor woordfamilies: stel dat het woord *economie* in de top 50 staat en *economische* in de top 100, dan kan *economische* voor de volledigheid ook bekeken worden. Dit geldt ook voor lemma's: zowel *liberaal* als *liberale* worden onderzocht. De voorwaarde is dus wel dat minstens één woord uit de woordfamilie of uit het lemma in de top 50 staat.

Een belangrijke vraag is welke woorden nu kans maken om keywords te zijn. Deze keuze is essentieel omdat de keywords een centraal deel zullen zijn van de analyse. Wanneer de selectie correct gebeurt, kunnen we het corpus tot een beperkte lijst van keywords en onderliggende patronen reduceren om tot diepere inzichten te komen (Scott & Tribble, 2006). Paquot & Bestgen (2009) schrijven dat een woord minstens twee of drie keer voor moet komen in het corpus om relevant te zijn. Daarom hebben we de minimumfrequentie in *Wordsmith* ook ingesteld op 3. Volgens Halliday (1992) moeten enkel substantieven toegelaten worden in de selectie. Dit lijkt voor deze specifieke onderzoekssituatie niet gewenst omdat andere woorden ook de toon van het corpus kunnen aangeven en dus als signaal kunnen dienen, bijvoorbeeld werkwoorden als *klagen* of *steunen* (Scott, 2001). Er zal dus wel een filter ingebouwd moeten worden om het aantal keywords te beperken zodat de uiteindelijke set woorden ook een goede representatie van de tekst vormt (Berber Sardinha, 1999a). Daarom hebben we een *stop word lijst* gemaakt: een lijst van woorden die niet nuttig zijn bij de analyse en daarom ook niet opgenomen worden. De lijst is ontstaan op twee manieren. Ten eerste is er een woordlijst gemaakt op basis van de drie corpora samen. Van deze lijst zijn uit de top 60 van de frequentielijst die woorden genomen die niet nuttig leken voor de analyse; functiewoorden behoren immers tot de meest frequente woorden in een tekst of corpus. Ten tweede zijn er later woorden aan de lijst toegevoegd die hoog stonden in de lijst van keywords maar die niet nuttig waren. Men zou kunnen wijzen op

---

<sup>10</sup> We hebben gekozen voor deze vier senatoren omdat er op hun website genoeg tekstmateriaal te vinden was. Dit was niet het geval bij Martine Taelman en Rik Daems, de overige senatoren van Open Vld. Daarnaast heb ik mijn stage gedaan bij senator Nele Lijnen.

het gevaar van subjectiviteit, maar het gaat hierbij om woorden die geen gevaar opleveren voor onze analyses. De volledige lijst van die woorden is: *de, van, het, een, en, in, te, voor, op, dat, is, die, om, met, zijn, aan, ook, niet, worden, er, maar, door, als, of, dan, bij, dit, tot, deze, wordt, naar, hun, kan, over, ze, wil, nog, heeft, hebben, uit, geen, zich, al, zal, wat, zo*. Het resultaat was een nuttigere en interessantere lijst van keywords bij elk corpus.

De laatste toepassing van *Wordsmith Tools* die we zullen gebruiken, is *Concord*. Bij het gebruik van *KeyWords* kunnen we er voor kiezen om een specifiek keyword nog verder te onderzoeken met *Concord*. Dit programma geeft een lijst met elk voorkomen van het keyword in elke specifieke context. Het toont ook de belangrijkste collocaten met de frequentie en de positie van het collocaat ten opzichte van het keyword, dat op de centrale positie staat. De minimum frequentie voor collocaties is in onze analyses vijf. Met *Clusters* zagen we in *KeyWords* de belangrijkste clusters voor het hele corpus. In *Concord* bestaat deze functie ook maar hier zien we enkel de clusters voor het specifieke keyword. Deze functie is zeer handig om te onderzoeken welke semi-vaste uitdrukkingen er bestaan in het corpus met het keyword.

Zoals al bleek uit de literatuurstudie bestaat deze meesterproef echter niet alleen uit pure statistiek. De kritische, ideologische analyse (Critical Discourse Analysis) wordt verbonden met de statistiek. Al de bovenstaande functies van *Wordsmith Tools* zullen gebruikt worden om de analyses methodologisch en statistisch verantwoord uit te voeren. We zullen kijken welke woorden of concepten uit de basisfilosofie van de Open Vld terug te vinden zijn in de top 50 en welke collocaties, links, clusters, ... hiermee verbonden zijn. Aan het einde van elke analyse volgt dan een interpretatie van de resultaten en ook een opsomming van de keywords die niet te vinden zijn in de top 50. Zeer concreet onderzoeken we de volgende woorden (en de bijhorende woordfamilie of het bijhorende lemma) uit de basisfilosofie:

*individu*  
*initiatief*  
*liberaal*  
*open*  
*ondernemen*  
*overheid*  
*privaat*  
*verantwoordelijkheid*  
*Vlaanderen*  
*Vld*  
*vrijheid*  
*zelfstandig*

Deze selectie van deze 'belangrijke' woorden is op twee manieren gebeurd. Eerst hebben we enkele van de woorden geselecteerd die vetgedrukt zijn in de basisfilosofie op [www.openvld.be](http://www.openvld.be). Daarna hebben we nog enkele woorden toegevoegd die we intuïtief met de liberalen associëren. Het gaat hierbij om *individu, ondernemen, overheid, vrijheid* en *zelfstandig*. Al deze woorden komen wel voor in de basisfilosofie maar ze zijn niet vetgedrukt. Toch willen we ook deze onderzoeken. De analyse beperkt zich echter niet tot deze kleine maar essentiële groep woorden. We kunnen verwachten dat in elk corpus andere keywords onze aandacht zullen trekken. Gelet op het onderwerp van deze meesterproef kunnen we ons de vrijheid wel gunnen om ook andere keywords te analyseren. We zullen ook aan het einde van elke analyse kijken welke belangrijke woorden uit de basisfilosofie geen keyword zijn, want hun afwezigheid is minstens even interessant.

### 3.3. Hypothesen

Het is niet eenvoudig om zeer concrete hypothesen te formuleren bij dit soort ideologisch geladen tekstmateriaal. Toch durven we nu al met enige zekerheid te stellen dat er verschillen zullen zijn tussen de verschillende corpora. Daarom zal ook de interpretatie na de analyse deels vergelijkend van aard zijn. Bij deze vergelijking richten we ons op het liberale gedachtegoed maar we vertrekken vanuit de keywordlijsten. Meer specifiek zullen we de keywords toetsen aan de basisfilosofie van de Open Vld: hoeveel van de belangrijkste woorden uit de basisfilosofie komen terug in de keywordlijsten van de drie corpora? Hierover hebben we de volgende verwachtingen:

H1: Het corpus met de verkiezingsprogramma's is het sterkst 'liberaal' getint. We verwachten dan ook dat zowat alle belangrijke woorden uit de basisfilosofie voorkomen in de top 50 van deze keywordslijst: *individu, initiatief, liberaal, open, ondernemen, overheid, verantwoordelijkheid, Vlaams, Vld, vrijheid en zelfstandig*.

H2: Het corpus met de nieuwsberichten van [www.openvld.be](http://www.openvld.be) zal minder 'liberaal' getint zijn dan het corpus met de verkiezingsprogramma's. Er komen dus minder woorden uit de basisfilosofie voor in de top 50 van deze keywordslijst, aangezien nieuwsberichten over het algemeen minder wervend zijn geschreven.

H3: Het corpus met de nieuwsberichten van de vier senatoren is minder liberaal getint dan het corpus met de verkiezingsprogramma's. Er komen dus minder woorden uit de basisfilosofie voor in de top 50 van deze keywordslijst.

H4: Het corpus met de nieuwsberichten van [www.openvld.be](http://www.openvld.be) is ongeveer even liberaal getint als het corpus met de nieuwsberichten van de vier senatoren. Ongeveer evenveel woorden uit de basisfilosofie zijn keywords in beide corpora, aangezien ze documenten van hetzelfde genre bevatten.

Met 'liberaal getint' bedoelen we dus de aanwezigheid van de belangrijkste woorden uit de basisfilosofie als keywords in de top 50. Dit kunnen we duidelijk controleren door na te gaan in welke mate die keywords aanwezig zijn en op welke plaats ze staan in de ranking.

## 4. Analyse van het Nationaal Nieuws

Op figuur 1 hieronder ziet u de eerste 50 keywords, verkregen uit het tekstmateriaal van het nationaal partijnieuws op [www.openVld.be](http://www.openVld.be).

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	OPEN	673	0.39	124	0.06	504.09	0.00
2	VLD	522	0.30	93	0.05	398.79	0.00
3	TOMMELEIN	269	0.15	7		358.62	0.00
4	VLAAMSE	492	0.28	133	0.07	277.47	0.00
5	MINISTER	447	0.26	152	0.08	199.98	0.00
6	VLAAMS	366	0.21	105	0.05	195.37	0.00
7	HTTP	132	0.08	1		193.16	0.00
8	PARLEMENT	266	0.15	65	0.03	163.26	0.00
9	MOET	633	0.36	322	0.16	155.83	0.00
10	VOORSTEL	185	0.11	36	0.02	133.71	0.00
11	VLAANDEREN	249	0.14	76	0.04	125.04	0.00
12	GATZ	76	0.04	0		117.26	0.00
13	KEULEN	97	0.06	5		115.97	0.00
14	PARLEMENTSLID	117	0.07	12		115.58	0.00
15	REGERING	423	0.24	208	0.10	110.91	0.00
16	ROUVEROIJ	68	0.04	0		104.92	0.00
17	ONDERWIJS	121	0.07	18		101.89	0.00
18	COM	89	0.05	6		100.01	0.00
19	VIMEO	64	0.04	0		98.74	0.00
20	VANDAAG	243	0.14	91	0.04	96.67	0.00
21	WETSVOORSTEL	67	0.04	2		87.75	0.00
22	MOETEN	518	0.30	314	0.16	86.19	0.00
23	ALEXANDER	127	0.07	28	0.01	84.27	0.00
24	CROO	167	0.10	51	0.03	83.80	0.00
25	MARINO	60	0.03	1		83.61	0.00
26	VERHOFSTADT	90	0.05	11		82.98	0.00
27	DAEMS	59	0.03	1		82.10	0.00
28	OVERHEID	165	0.09	51	0.03	81.78	0.00
29	LAHAYE-BATTHEU	53	0.03	0		81.77	0.00
30	SAS	53	0.03	0		81.77	0.00
31	BLOCK	66	0.04	3		80.87	0.00
32	BORGHT	52	0.03	0		80.23	0.00
33	BELGIE	57	0.03	1		79.08	0.00
34	IRINA	51	0.03	0		78.68	0.00
35	BEHEERSOVEREENKOMST	51	0.03	0		78.68	0.00
36	KNOP	60	0.03	2		77.38	0.00
37	RUTTEN	63	0.04	3		76.52	0.00
38	VANHENGEL	49	0.03	0		75.60	0.00
39	MAGGIE	57	0.03	2		72.96	0.00
40	ONS	366	0.21	207	0.10	72.03	0.00
41	BART	153	0.09	50	0.02	71.49	0.00
42	FINANCIELE	46	0.03	0		70.97	0.00
43	TURTELBOOM	89	0.05	15		70.12	0.00
44	#	5061	2.91	4990	2.46	70.10	0.00
45	VERA	45	0.03	0		69.43	0.00
46	SOCIALE	168	0.10	62	0.03	68.07	0.00
47	VLD-KAMERLID	44	0.03	0		67.88	0.00
48	LIJNEN	63	0.04	5		67.68	0.00
49	MISBRUIK	82	0.05	13		66.81	0.00
50	COMMISSIE	169	0.10	64	0.03	66.25	0.00

Figuur 1

Er vallen een hoop interessante opmerkingen te maken over dit corpus, op basis van bovenstaande keywordslijst, maar we houden het bij de belangrijkste zaken. Met een frequentie van 673 en een keyness van 504,09 is *open* zeer duidelijk het belangrijkste keyword. Tweede is *Vld*, dat 522 keer voorkomt. Het zal niet verbazen dat de twee zeer veel voorkomen en bovendien meestal samen voorkomen. Toch komt *open* ook voor in combinatie met andere woorden (wat de grotere frequentie verklaart). Enkele voorbeelden: *de open brief*, *de open Monumentendag*, *de open*

*samenleving, open geest, open houden, open kaart spelen, open debat, open gesprek, open en neutraal internet, Open Vld-fractie, Open Vld-senatrice, Open Vld-afdelingen, Open Vld-senator, Open Vld-parlementsleden, Open Vld-voorstellen, enz.* Voor een liberale partij lijkt *open* een zeer belangrijk keyword, toch komt het niet zo veel voor buiten de verbinding met *Vld*. Merk op dat *samenleving* niet eens een keyword is voor dit corpus. Daarnaast wordt *open* aan ondermeer de volgende keywords gelinkt: *commissie, Bart, minister, Ceysens, moeten, sociale, besluit, ons, Croo, meerderheid, Vlaamse, we, Vld, regering, Vlaams, parlement, enz.*

In deze top 50 zitten ook een aantal eigennamen van politici. De belangrijkste politicus op basis van de keywordslijst is *Tommelein*. Senator Bart Tommelein heeft dus, in vergelijking met het referentiecorpus, een zeer belangrijk aandeel in dit corpus. Van alle politici heeft hij de hoogste absolute frequentie (269 voorkomens tegenover 167 voor Alexander en Herman De Croo samen) én de hoogste keyness. Bart Tommelein staat dus aanzienlijk hoger gerankt in de keywordslijst (plaats 3) dan Sven Gatz (12), Sas van Rouveroi (16), Alexander de Croo (24), Marino Keulen (25), Guy Verhofstadt (26), Rik Daems (27), Maggie De Block (31), Nele Lijnen (48), enz. Dit resultaat verbaast mij toch enigszins. Men zou verwachten dat voorzitter Alexander De Croo het 'belangrijkste' zou zijn. Natuurlijk komt De Croo als voorzitter meer voor in het referentiecorpus dan Tommelein (51 versus 7) en dit verklaart waarom zijn naam minder key is (plaats 24). Toch is ook de absolute frequentie van Alexander De Croo een flink stuk lager dan die van Bart Tommelein. Een mogelijke verklaring is een zeer praktische reden: uit ervaring weet ik dat de medewerkers van Tommelein zéér veel teksten over hem schrijven en publiceren.

Voor een liberale partij, die ideologische vrijheid boven alles plaatst, verbaast de zeer hoge frequentie (633 hits!) van het woord *moet*, dat op plaats 9 staat in de keywordslijst. *Moeten* staat op plaats 22, met een absolute frequentie van 518, en wordt hier ook besproken. Het zou kortzichtig zijn om aan te nemen dat het hierbij enkel om verplichtingen gaat. Laat ons daarom naar de verschillende gebruiken van het woord kijken. *Moet* en *moeten* komen voor in collocaties met de volgende woorden: *worden* ('*moet worden*' komt 30 keer voor, dit is zeer veel), *niet, nieuwe, meer, kunnen, mensen, Vlaanderen, Vlaamse, komen, VRT, maken, hervormingen, extra, betalen, nemen, Europese, gebeuren, kinderen, overheid, parlement, regering, krijgen, men, geven, ons, dringend, ...* We vinden bovendien de volgende links: *open, aantal, ons, VRT, minister, moet, overheid, Vld, besluit, Tommelein, middelen, Quickenborne Vlaamse, geven, regering, vandaag, commissie, meer, ...* Wordsmith geeft de volgende clusters: *Open Vld moet, moet er een, de overheid moet, de Vlaamse regering moet, moet dan ook, de Europese commissie moet, moet ervoor zorgen, ons land moet, ins ons land moeten, we moeten de.* Deze clusters bevatten toch veel verplichtingen. Enkel *moet ervoor zorgen* duidt een gevolg aan. Er zijn in de concordantielijst verschillende voorbeelden van *moet* als gevolg te vinden, toch overheerst de verplichting duidelijk. Een greep uit het aanbod:

*Dit moet u doen.*

*Ons land moet een duidelijk signaal geven.*

*Ons land moet opnieuw vooruit.*

*Het leger moet opleiden, niet opvoeden.*

*Wanneer gekozen wordt voor kernenergie, moet er een oplossing zijn ...*

*Men moet dringend weer aan tafel gaan zitten.*

*De overheid moet efficiënter optreden.*

*... parlement moet energiedebat naar zich toetrekken.*

*... VRT moet zichzelf reguleren.*

Op basis van deze voorbeelden willen we niet zeggen dat de Open Vld zich ontdoet van haar liberale (vrije) karakter, puur omdat *moet* en *moeten* zo sterk scoren als keywords. De Open Vld is net als alle andere politieke partijen in dit land gesitueerd in het dagelijkse leven en een actor in het politieke en het bestuurlijke landschap. Ook politici van de Open Vld moeten dus aandringen op actie van bepaalde instanties, bijvoorbeeld de regering, en ook zij moeten bepaalde eisen stellen ten aanzien van onze maatschappij. Toch is het verrassend dat exact deze verplichtende modale werkwoorden zo hoog scoren in de lijst van keywords.



Ook de V in Open Vld vinden we terug tussen de keywords en best hoog. *Vlaamse* vinden we op de vierde plaats, *Vlaams* staat op plaats 6 en op plaats 11 vinden we *Vlaanderen*. Zijn de liberalen in hun nieuwsberichten stiekem flaminganten? Wordsmith toont dat *Vlaams* gelinkt wordt aan volgende keywords: *voorstellen, minister-president, Alexander, Guy, ouders, stelt, Quickenborne, Gucht, immers, voorstel, Vlaanderen, private, werkstraffen, decreet, aanpak, Rouveroi, VRT, gemeenten, regering, financiële, Lijnen, zelfstandig, jongeren, moet, Somers, Lahaye-Battheu, Croo, Maggie, deze, parlamentslid, Turtelboom, inspectie, Gatz, Vld, resolutie, Tommelein, huisvesting, Borgh, Europees, Vlaamse, Daems, open, minister, Gysbrechts, Keulen, parlement, ...* Kijken we naar de concrete voorbeelden, dan kunnen we besluiten dat de partij heus niet sterk flamingant is: de meest gebruikte cluster is *het Vlaams Parlement* met 104 hits. Andere veel voorkomende clusters zijn *in het Vlaams* (68), *van het Vlaams* (27), *Vlaams minister van* (15), *zegt Vlaams parlamentslid* (15), enz. Het woord *Vlaams* komt dus vooral voor met woorden als *Parlement, minister, parlamentslid, gewest, volksvertegenwoordiger, ...* Over *Vlaanderen* valt niet zoveel te zeggen. Er zijn slechts twee clusters: *in heel Vlaanderen* en *er in Vlaanderen*. Ook bij de gelinkte keywords vinden we niets typerend: *onderwijs* en *meer*. Er zijn wel enkele interessante collocaties, zoals *Vlaanderen met mensen, Brussel, autonomie, actie, financiering, enz.*, maar deze collocaties komen slechts twee of drie keer voor en zijn dus weinig toonaangevend te noemen op een corpus van meer dan honderdduizend woorden. Het zou dus zeer kortzichtig zijn om te zeggen dat de partij zeer Vlaamsgezind is puur omdat de woorden *Vlaams(e)* en *Vlaanderen* keywords zijn. Er moet echter wel opgemerkt worden (en dit is best verrassend) dat de woorden *federaal* en *federale* geen keywords zijn (hun absolute frequenties zijn slechts 93 en 27). Dit is misschien omdat deze woorden niet altijd expliciet geschreven moeten worden wanneer men bijvoorbeeld over minister x of senator y schrijft.

We kunnen in de keywordslijst duidelijk nog een aantal thema's ontdekken die veel aan bod komen in de nieuwsberichten van de Open Vld. Zo komt het woord *regering* 423 keer voor en dit is een hoge score vergeleken met de 522 van *Vld*. *Regering* wordt gelinkt aan o.a. de volgende keywords: *Keulen, Marino, Annemie, Turtelboom, private, Guy, Gucht, Volcem, Alexander, voorstellen, we, openbare, voorstel, financiën, facturen, Quickenborne, minister, Vlaanderen, Zamouri, werkstraffen, Vld-senator, besluit, België, financiële, regering, Vlaams, Vanhengel, open, Vld, parlement, moet, deze, kunnen, huisvesting, Vlaamse*. De partij verbindt de regering dus met de financiën van ons land (*financiële* staat op plaats 46 van de keywords). Bekijken we de collocaties, dan scoort *Vlaamse regering* zeer hoog met 222 voorkomens. *Federale regering* komt slechts 27 maal voor. Verbonden met *regering* staat op plaats 32 *overheid*. Gelinkte keywords zijn *voorstellen, meer, werkende, Guy, werk, Willem-Frederik, wachtlijsten, regering, financiële, deze, besluit, werkstraffen, Zamouri, wetsvoorstel, Carina, huisvesting, moet en Vlaamse*. Net zoals bij *regering* is *Vlaamse overheid* (40) een frequentere collocatie dan *federale overheid* (16).

Nog net binnen de top 50 zien we het keyword *sociale*. De liberalen die het woord *sociale* in de mond nemen? Dit zijn de belangrijkste en meest opvallende collocaties: *zekerheid, huisvesting, partners, fraude, bescherming, zaken, commissie, Vlaamse, moeten, woning, maatregelen, broodnodig, betalen, leningen, ...* Er zijn geen noemenswaardige links. De clusters met *sociale* zijn de volgende:

*de sociale partners*  
*in de sociale*  
*van de sociale*  
*de sociale zekerheid*  
*de sociale huisvesting*  
*broodnodige en sociale*  
*dat de sociale*  
*een sociale woning*  
*en sociale maatregel*

Sociale zekerheid, sociale partners en sociale huisvesting zijn algemeen bekende uitdrukkingen in onze maatschappij en hebben niets te maken met socialisme an sich. *Sociale* is zeker geen zeer belangrijk keyword en ook deze beknopte analyse toont aan dat de nieuwsberichten niet bepaald een socialistische ondertoon hebben.

We hebben zeker en vast een aantal zeer interessante dingen kunnen vinden in de analyse van de nieuwsberichten van 2011. Door keywords als *Vlaamse*, *financiële*, *onderwijs*, enz. krijgen we zicht op de belangrijkste concepten en associaties die in deze teksten geconstrueerd worden. Toch kan men moeilijk de indruk krijgen dat dit nu een corpus van 'liberale' teksten is. Hoe goed toont de keywordanalyse aan dat het om de woorden van een liberale partij gaat in dit corpus? Niet echt duidelijk, zo blijkt. Het woord *liberale* staat pas op plaats 66 in de keywordslijst met een absolute frequentie van 87. *Ondernemers* staat op plaats 57 met een frequentie van 39, *ondernemen* is geen keyword met een absolute frequentie van 28. *Private* staat op 61 met 46 voorkomens. Pas op 160 vinden we *zelfstandig*. *Verantwoordelijkheid* is ook geen keyword en komt 46 keer voor. Dit geldt ook voor *initiatief*, dat 41 keer voorkomt. *Individu* (1 voorkomen!) en *individueel* (9 voorkomens) staan niet in de lijst. Een hoop 'liberale' concepten, die duidelijk vermeld worden in de beknopte basisfilosofie, vinden we dus niet meteen terug in de lijst van keywords. Hét grote woord van het liberale gedachtegoed, *vrijheid*, komt slechts 17 keer voor in het ganse corpus en is zelfs geen keyword. *Vrij* is er ook geen en komt 23 keer voor. Dit corpus toont ons niet bepaald een lijst van 'culturele keywords' (Stubbs 2007) die de cultuur van het liberalisme belichamen. Deze resultaten zijn toch enigszins verrassend maar misschien wijst dit er gewoon op dat de partij er voor kiest om zich niet al te sterk liberaal te profileren in haar nieuwsberichten. Het doel van dit soort berichten lijkt eerder praktisch informatief in plaats van sterk ideologisch. Centraal staat de informatie over de verschillende politici van de partij. Dit kunnen we, ter afsluiting van dit onderdeel, duidelijk zien wanneer we de clusters van het corpus bekijken. Omwille van praktische redenen heb ik een deel van de clusters weggelaten, die ons niets nuttig tonen. Het gaat hierbij om combinaties van het type *BE + HTTP* of *PORTRAIT + VIMEO*. Dit soort clusters toont wél dat de partij veel internetlinks en video's<sup>11</sup> op haar website plaatst (wat op zich wel een interessant feit is) maar het geeft ons geen ideologische informatie. Na het schrappen van dit soort clusters bekomen we dus volgende lijst (figuur 2):

---

<sup>11</sup> *Vimeo* is een online service die toelaat om video's online te plaatsen.

N	KW Cluster	Freq.
1	OPEN VLD	475
6	VLAAMSE REGERING	222
21	ALEXANDER [.] CROO	126
22	VLAAMS PARLEMENT	107
24	BART TOMMELEIN	101
41	VLAAMS PARLEMENTSLID	80
67	SEKSUEEL MISBRUIK	60
68	MARINO KEULEN	59
70	MAGGIE [.] BLOCK	55
71	SAS [.] ROUVEROIJ	53
72	DER BORGHT	52
73	ANNEMIE TURTELBOOM	51
74	EUROPEES PARLEMENT	50
75	IRINA [.] KNOP	49
76	SVEN GATZ	49
77	GUY VERHOFSTADT	48
78	VINCENT [.] QUICKENBORNE	48
79	WE MOETEN	48
80	MOETEN WE	45
81	VERA [.] DER	45
82	GWENDOLYN RUTTEN	44
83	VERA [..] BORGHT	44
84	NELE LIJNEN	42
85	BART SOMERS	40
86	VLAAMSE OVERHEID	40
87	MEER MENSEN	39
88	RIK DAEMS	39
89	DIRK [.] MECHELEN	38
91	CARINA [.] CAUTER	36
92	RAAD [.] STATE	36
93	VLAAMS VOLKSVERTEGENWOORDIGER	36
97	OPENBARE OMROEP	34
99	VLAAMS MINISTER	34
100	OPENVLD [] BE	33
159	VLD-VOORZITTER [.] CROO	32
160	VLD-VOORZITTER ALEXANDER	32
161	MINISTER [..] ZAKEN	30
162	OPEN VLD-KAMERLID	30
164	ZAKEN [.] TURTELBOOM	27
165	ZAKEN ANNEMIE	27
166	OPENVLD [] HTTP	25
167	OPEN VLD-VOORZITTER	24
168	REGERING [..] ZAKEN	24
170	GUY VANHENGEL	23
171	MEER [.] #	23
172	OPEN [...] CROO	23
173	WILLY [.] CLERCQ	23
174	OPEN [.] ALEXANDER	22
175	VOORSTEL [.] RESOLUTIE	22
178	MARLEEN VANDERPOORTEN	21
179	MINISTER [...] TURTELBOOM	21
180	MINISTER [...] ANNEMIE	21
181	SABIEN LAHAYE-BATTHEU	21
182	VLD VRAAGT	21
183	WILLEM-FREDERIK SCHILTZ	21

*Figuur 2*

De opmerking dat dit corpus vooral dient om de bevolking te informeren over het doen en laten van de Open Vld-politici, wordt hier nogmaals bevestigd. Verschillende clusters zijn immers simpelweg de namen van partijpolitici: Alexander De Croo, Bart Tommelein (wiens familienaam blijkbaar vaak alleen gebruikt wordt<sup>12</sup>), Marino Keulen, Maggie De Block, Sas Van Rouveroj, enz. Van alle clusters met meer dan 20 voorkomens (66 in totaal) bestaan er maar liefst 23 uit de naam en familienaam van een liberale politicus. Verder zien we in deze clusters dat *seksueel misbruik* een belangrijk onderwerp voor de partij is, maar ook de *openbare omroep*. Bij deze laatste vinden we slechts één nuttige collocatie: *Bart Tommelein*. Het is een gekend feit dat Tommelein zich veel bezighoudt met de openbare omroep en dit toont zich ook (weliswaar statistisch) in deze analyse. Voor het overige bestaat deze lijst voornamelijk uit een aantal functies binnen het politiek landschap, bijvoorbeeld *Vlaams minister*, *Vlaams volksvertegenwoordiger*, *Vld-voorzitter (De) Croo*, enz. Ook uit de clusters komen geen thema's naar voren (of het moest zo zijn dat de partij enkel berichtgeving doet over seksueel misbruik en de openbare omroep, al lijkt dit niet echt aannemelijk). Zoals eerder gezegd is dit corpus niet echt ideologisch geladen. Laat ons nu de verkiezingsprogramma's van de partij op de verschillende niveaus analyseren om te zien of deze 'liberaler' getint zijn.

---

<sup>12</sup> Men zou kunnen denken dat zijn naam vaak wordt voorafgegaan door zijn functie en dat de cluster *senator (Bart) Tommelein* wel veel voorkomt. Dit is ook niet het geval: *senator Tommelein* komt niet voor in dit corpus en *senator Bart Tommelein* maar 4 keer.

## 5. Analyse van de verkiezingsprogramma's

Figuur 3 toont de eerste 50 keywords uit het corpus bestaande uit verkiezingsprogramma's voor de verkiezingen op gemeentelijk, provinciaal, Vlaams, Federaal (Kamer en Senaat) en Europees niveau tussen 1994 en 2010. De partij heette tot 2007 *Vld*, dus dit corpus bevat ook teksten van vóór de naamsverandering. Een snelle blik op deze keywordslijst toont al dadelijk heel wat woorden die we met een liberale partij zouden associëren, iets wat bij het vorige corpus veel minder het geval was.

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	VLD	1237	0.75	93	0.05	1426.06	0.00
2	MOET	1323	0.80	322	0.16	883.13	0.00
3	OPEN	851	0.52	124	0.06	772.09	0.00
4	MOETEN	891	0.54	314	0.16	422.12	0.00
5	OVERHEID	406	0.25	51	0.03	392.88	0.00
6	O	263	0.16	5		378.50	0.00
7	SAMENLEVING	225	0.14	12		280.52	0.00
8	BELEID	278	0.17	32	0.02	278.48	0.00
9	BURGER	190	0.12	5		264.46	0.00
10	SOCIALE	311	0.19	62	0.03	237.46	0.00
11	DAAROM	263	0.16	50	0.02	206.72	0.00
12	DIENT	168	0.10	10		204.57	0.00
13	ONDERWIJS	187	0.11	18		199.65	0.00
14	ONZE	411	0.25	142	0.07	198.69	0.00
15	BURGERS	197	0.12	30	0.01	174.64	0.00
16	KUNNEN	622	0.38	332	0.16	160.48	0.00
17	VLAANDEREN	266	0.16	76	0.04	155.08	0.00
18	UNIE	153	0.09	18		151.92	0.00
19	GEMEENTEN	133	0.08	13		141.25	0.00
20/LD-VERKIEZINGSPROGRAMMA		88	0.05	0		141.24	0.00
21	GEMEENTE	125	0.08	10		141.24	0.00
22	MENSEN	498	0.30	253	0.12	140.66	0.00
23	INZAKE	115	0.07	8		134.91	0.00
24	MIDDELEN	139	0.08	20		126.60	0.00
25	RESOLUTIE	113	0.07	12		116.59	0.00
26	VRIJHEID	103	0.06	9		113.38	0.00
27	BLADZIJDE	88	0.05	4		113.09	0.00
28	DIENEN	91	0.06	5		112.72	0.00
29	ARBEIDSMARKT	112	0.07	13		111.78	0.00
30	ADMINISTRATIEVE	93	0.06	6		111.13	0.00
31	WERKEN	240	0.15	87	0.04	109.85	0.00
32	BESTAANDE	105	0.06	11		108.86	0.00
33	PRIVATE	77	0.05	2		107.31	0.00
34	VIA	231	0.14	86	0.04	102.47	0.00
35	I	100	0.06	12		98.50	0.00
36	DIENSTEN	95	0.06	11		94.90	0.00
37	FISCALE	106	0.06	16		94.35	0.00
38	OCMW	85	0.05	8		91.40	0.00
39	KANSEN	99	0.06	14		90.88	0.00
40	KADER	99	0.06	14		90.88	0.00
41	SAMENWERKING	123	0.07	26	0.01	90.38	0.00
42	BEDRIJVEN	151	0.09	43	0.02	88.26	0.00
43	MEER	770	0.47	567	0.28	87.82	0.00
44	WELVAART	71	0.04	4		87.48	0.00
45	ARBEID	79	0.05	7		86.59	0.00
46	AANDACHT	103	0.06	18		84.95	0.00
47	ECONOMIE	133	0.08	35	0.02	83.15	0.00
48	VLAK	132	0.08	35	0.02	82.01	0.00
49	ONDERNEMINGEN	78	0.05	8		81.47	0.00
50	BELASTINGEN	121	0.07	29	0.01	81.39	0.00

*Figuur 3*

Het eerste woord is deze keer niet *open* maar wel *Vld*. De verklaring is logisch: dit corpus bestaat voor een deel uit verkiezingsprogramma's uit de periode dat de partij nog gewoon *Vld* heette. De keyness van *Vld* is dus de hoogste van alle woorden maar dit betekent natuurlijk niet dat het woord

ook de hoogste frequentie heeft van alle woorden (al is de frequentie van *Vld* wel hoger dan die van *open*). Ook al zit *open* in de naam van de partij sinds 2007, toch springt men niet bepaald creatief met het woord om. Er is slechts één cluster buiten *Open Vld* waarin *open* voorkomt, en dit is *open samenleving*. Een zeer belangrijke cluster weliswaar, maar men zou toch verwachten dat de partij het woord vaker en in meer contexten zou gebruiken.

Met een frequentie van 1323 en een zeer hoge keyness van 833,13 op plaats 2 in de keywordslijst valt het woord *moet* wel zeer erg op. Twee plaatsen lager staat *moeten*. Dezelfde vraag als bij het vorige corpus geldt ook hier: waarom komt het woord *moet* (of *moeten*) zo veel voor in teksten van een liberale partij, die in plaats van verplichtingen toch keuzes en vrijheid zou moeten verkondigen? Vanuit onze intuïtie associëren we dit woord niet direct met de liberale ideologie. Laat ons daarom deze keywords even van naderbij bekijken om te zien hoe ze gebruikt worden in de verkiezingsprogramma's. *Wordsmith* geeft een zeer lange lijst van keywords die gelinkt worden aan *moet* (dit heeft overigens de meeste links van alle keywords!) en *moeten*. Omwille van de enorme hoeveelheid links en omdat het duidelijker is, groeperen we de keywords<sup>13</sup> per thema. Deze klassering is subjectief, maar toont ons desalniettemin een aantal thema's en concepten die met elkaar verbonden zijn. Ten eerste zijn er een hoop keywords die we kunnen verbinden met de economie:

*ondernemen, private, jobs, werken, belastingen, fiscale, ondernemerschap, financieel, investeren, kmo's, zelfstandigen, werknemers, werklozen, verantwoordelijkheid, pensioen, ontwikkeling, vergrijzing, concurrentie, efficiënt, hernieuwbare, landbouw, zekerheid, betaalbaar, middelen, aanmoedigen, bijdragen, stimuleren, aanmoedigen, arbeid, arbeidsmarkt, verlagen, economie, ...* De partij toont zich in haar woorden ook als een liberale, progressieve partij: *vrijheid, vrije, innovatie, investeringen, ambitie, individuele, verhogen, aanmoedigen, aanpak, stimuleren, nieuwe, bevorderen, actief, moderne, kansen, ruimte, ondersteunen, verhogen, vorming, beter, ...*

Justitie, politie en veiligheid zijn ook vertegenwoordigd: *veiligheid, veiligheidsbeleid, politiediensten, criminaliteit, recht, justitie, misdrijven, overlast, ...*

Verder ontdekken we eerder 'sociale' woorden: *sociaal, ocmw, ontwikkelingssamenwerking, ouderen, bejaarden, senioren, gelijke, respect, duurzame, mensen, maatschappelijk, migranten, migratie, discriminatie samenleving, dienstverlening, welzijnshuis, zorgen, betaalbaar, maatschappelijk, ...*

Tussen de gelinkte keywords vinden we ook veel woorden in verband met gezinnen en kinderen: *kinderen, jongeren, scholen, leren, onderwijs, zorgen, opleiding, kinderopvang, ouders, leerkrachten, ontwikkeling, wonen, ...*

Tot slot zijn er ook nog een aantal keywords die we kunnen associëren met besturen en politiek: *democratisch, wettelijk, gemeentelijk, federaal, Vlaams, Vlaanderen, Europees, lokaal, gemeenschappen, gewesten, infrastructuur, leiden, legislatuur, bevoegdheden, vereenvoudigen, werking, behouden, land, autonomie, efficiënt, administratie, ambtenaren, wegwerken, overheid, samenleving, regelgeving, gemeente, burger, ...*

Vele van deze keywords zijn zeer essentiële en centrale begrippen in een verkiezingsprogramma. Ze zijn ook allemaal verbonden met *moet*, dat dus duidelijk een sleutelwoord is in de teksten. Zoals we al eerder opmerkten, is deze thematische opdeling (waarbij de categorieën vaak overlappen) puur subjectief en bevat ze dus geen sluitende bewijzen. Het is gewoon een oppervlakkige manier (die weliswaar gebaseerd is op statistisch materiaal) om de keywords rond de woorden *moet* en *moeten* te tonen en dus de concepten die ermee verbonden zijn, om zo een globaal beeld te geven. De hypothetische opmerking dat het vele voorkomen van *moet* en *moeten* afbraak doet aan het liberale karakter van de teksten, is hier niet terecht. In een verkiezingsprogramma moet een partij nu eenmaal wijzen op een aantal feiten die zij graag veranderd ziet en *moet* is dan een goed modaal werkwoord. Ook in de clusters zien we duidelijk dat de partij aandringt op actie of verandering en op wie er moet veranderen of wat er moet veranderd worden. Van hoge naar lage frequentie (met een minimum van vijf<sup>14</sup>) vinden we onder meer:

*de overheid moet  
voor open Vld moet*

<sup>13</sup> Bij gelijkaardige substantieven, adjectieven of werkwoorden (bijvoorbeeld *vereenvoudiging, vereenvoudigd* en *vereenvoudigen*) geef ik slechts één vorm, namelijk het lemma.

<sup>14</sup> De clusters met *moeten* hebben een frequentie lager dan drie, dus deze worden hier niet getoond.

*open Vld moet*  
*de Vlaamse regering moet*  
*in ons land moet*  
*het beleid moet*  
*de burger moet*  
*de Europese Unie moet*  
*ons land moet*  
*moet werk worden gemaakt*  
*werk moet lonen*  
*in 2020 moet*  
*het welzijnshuis moet*  
*sociale zekerheid moet*  
*Vlaamse regering moet*  
*de politie moet*  
*moet het parlement*

Op plaats 5 staat het keyword *overheid*. Welke keywords worden verbonden met de *overheid*? Het lijkt, op basis van deze lijst gelinkte keywords, dat de overheid een aantal vaste functies heeft. Ten eerste op economisch vlak: *ondernemingen, middelen, financieel, investeren, stimuleren, inzetten, samenwerking, werken, bedrijven, nodige, zekerheid, aanmoedigen, efficiënt, ...* Ten tweede vinden we 'sociale' woorden: *behoefte, mensen, zekerheid, kwaliteit, migranten, ouders, scholen, zorgen, onderwijs, ...* Ook het liberale en progressieve vinden we in de links: *vrijheid, burger, meer, samenleving, efficiënt, beter, nieuwe, open, actief, ...* Tot slot nog de 'politieke' keywords: *Vlaanderen, gemeenten, gewesten, werkbevoegdheden, federale, Vlaamse, ...* Toch is het aantal gelinkte keywords kleiner dan bij *moet* (78 versus 234) en wordt *moet* ook vier maal meer gebruikt in het corpus.

We hebben tot nu 'sociale' woorden als een categorie gebruikt om keywords onder te verdelen. Laten we nu eens kijken welke keywords gelinkt zijn met *sociale*, dat op plaats 10 staat in de keywordslijst: *vrijheid, veiligheid, zekerheid, mogelijk, belangrijk, betere, rol, maatschappelijk, tewerkstelling, kinderen, Europese, federale, ocmw, huisvesting, private, bijdragen, voordelen, noodzakelijk, voorzieningen, verlaging, culturele, beleid, financiering, recht, meer, rechten, dienen, noodzakelijk, ...* De belangrijkste clusters, in dalende volgorde van frequentie, zijn:

*De sociale zekerheid*  
*De sociale bijdragen*  
*De sociale voordelen*  
*De sociale partners*  
*De sociale economie*  
*De sociale fraude*  
*De sociale bescherming*  
*Sociale zekerheid moet*  
*De sociale woningen*  
*De sociale lasten*  
*De sociale voorzieningen*  
*De sociale woning*  
*De sociale rechten*  
*De sociale huisvesting*

Het mag duidelijk zijn dat ook in haar verkiezingsprogramma's de Open Vld geen 'socialistische' taal schrijft. Uitdrukkingen als *de sociale woning* en *de sociale fraude* zijn gewoon begrippen uit de politiek die alle partijen aangaan. Ook links als *tewerkstelling* en *huisvesting* tonen niet schokkends, enkel dat de partij zich ook bezighoudt met dagelijkse thema's van onze maatschappij.

De Open Vld profileert zich wél openlijk als de partij die de werkgever en werknemer wil ondersteunen. Het is daarom mogelijk interessant om het keywords *werken* (plaats 31) te onderzoeken om te ontdekken welke associaties dit heeft in de verkiezingsprogramma's. Uit de gelinkte keywords blijkt al dat *werken* een zeer belangrijke zaak is voor de partij. Het is iets wat de politiek moet stimuleren en ondersteunen: *aantrekkelijker, middelen, beleid, kansen, loopbaan, armoede, investeren, creëren, nodige, mogelijk, belastingen, alle, overheid, actief, meer, deeltijds, kwaliteit, ouders, beter, jobs, kinderen, leren*. Ook de clusters tonen enkele vertrouwde uitdrukkingen:

*weg te werken*  
*werken moet lonen*  
*samen te werken*  
*werken aan de*  
*wonen en werken*  
*uit te werken*  
*langer te werken*

Het initiatief voor nieuwe bedrijven moet voor de liberalen traditioneel vanuit het privé-initiatief komen. *Private* staat op plaats 33 in de keywordslijst. De links zijn: *maken, openbaar, meer, vervoer, bedrijven, zowel, rol, veiligheid, blijven, ondernemingen, moet, moeten, samenwerking, via, sociale, open, Vld, huurmarkt, sector*. De belangrijkste collocaten zijn: *huurmarkt, woningen, samenwerking, moet*. Dezelfde concepten vinden we ook bij de clusters: *de private sector, op de private, de private huurmarkt*. De burger moet duidelijk zelf ondernemen. *Ondernemingen* is een woord dat we terugvinden in de basisfilosofie van de Open Vld en ook in de lijst van keywords op plaats 49. Het aantal links is klein: *vlak, gemeente, overheid, maken, bestaande, tussen, private, moeten, meer, burgers, open, administratieve, kunnen, Vld, onze*. Bij de collocaties vinden we nog termen als *Vld, open, kunnen, administratie, burgers, meer, moeten*. Er is één cluster, namelijk *van onze ondernemingen*. Voor onze analyse is *ondernemingen* dus niet meteen het nuttigste keyword<sup>15</sup>, al contrasteert het ook niet met vorige bevindingen in dit corpus.

Hiermee verbonden loont het ook om te kijken naar de keywords *economie* en *economische*<sup>16</sup>. Dit laatste is geen zeer belangrijk keyword (plaats 69) maar het heeft wel veel gelinkte woorden, dus wordt het ook besproken in deze context. De belangrijkste links zijn: *ondernemerschap, ondernemen, investeren, actieve, versterken, mensen, samenwerking, belangrijk, innovatie, samenleving, reguliere, maken, lokale, ruimtelijke, welvaart, open, Vld, ordening, beter, nieuwe, moeten, meer, Vlaamse, sociale, onze, zorgen, beleid, steden, veranderingen, hogere, land, migratie, sociale, kunnen, culturele, duurzame, creëren, ontwikkeling, ...* Buiten *van de economie* en *de economie* zijn er geen clusters. Bij de collocaties vinden we ook niet veel nieuwe, interessante woorden buiten *sociale, Vlaamse, open, milieu, moeten, groene* en *welvaart*.

Ook het allerlaatste keyword van de keywordslijst heeft te maken met de economie. *Belastingen* staat nog net op de lijst (plaats 50) en in het algemeen vinden de liberalen dat deze omlaag moeten. Wordsmith toont volgende links: *mensen, verlaagd, welvaart, hoger, land, immers, drastisch, verlagen, moeten, alle, eigen, mogelijk, open, verlaging, leiden, financieel, verhogen, nieuwe, werken, lokale, moet, gemeentelijke, meer, kunnen, vereenvoudigen, Vld, arbeid*. Woorden als *verlaging* en *vereenvoudigen* bewijzen de algemene perceptie. De collocaties tonen hetzelfde patroon: *arbeid, verlagen, verschuiven, vereenvoudigen, wil, kunnen, verhogen, betalen, meer lokale, gemeente, moet, werken*. Ook de clusters bevestigen wat we al dachten:

*vereenvoudigen van belastingen*  
*verlagen verschuiven en*  
*de belastingen op*  
*verschuiven en vereenvoudigen van belastingen*  
*belastingen op arbeid*  
*arbeid naar belastingen*  
*verlaging van de belastingen*

<sup>15</sup> *Ondernemen* staat pas op plaats 232 in de keywordslijst en levert geen nuttige informatie op.

<sup>16</sup> Het keyword *bedrijven* wordt niet besproken omdat het geen interessante feiten oplevert.



*een verschuiving van de belastingen  
de gemeentelijke belastingen*

Op plaats 26 vinden we *vrijheid*, hét grote woord van alle liberalen. De gelinkte keywords liegen er niet om: *individuele, recht, overheid, veiligheid, sociale, kunnen, daartoe, iedereen, alle, Vld, open, mensen, flexibiliteit, burgers, moet, krijgen, moeten, meer*. Deze lijst kan men duidelijk liberaal interpreteren en dit geldt ook voor de collocaties: *meer, moeten, meningsuiting, persoonlijke, krijgen, mensen, leven, willen, flexibiliteit, open, burgers, Vld. Vrijheid van meningsuiting* is de enige zinnige cluster. Deze korte lijst van verbonden concepten toont ons desalniettemin dat de liberalen de vrijheid (van meningsuiting) van de burger willen vergroten.

Een liberale partij roept traditioneel op tot verantwoordelijkheid en ondernemen maar verbindt hier wel *kansen* mee. Mensen moeten de kans krijgen om zich te ontplooien om iets te maken van hun leven. *Kansen* staat ook in de top 50 van de keywordslijst (plaats 39) en vraagt dus ook kort onze aandacht. De links wijzen in de richting van het progressieve: *onderwijs, werken, creativiteit, maken, objectieve, samenleving, handicap, gelijkheid, werknemers, onze, burgers, maximale, diversiteit, arbeidsmarkt, moeten, iedereen, creëren, alle, open, Vld, krijgen, nieuwe, mensen, meer, gelijke, geven*. Een deel van deze woorden vinden we ook bij de collocaties. De clusters zijn *Open Vld wil kansen, kansen te geven* en *kansen geven om*. Dit gebruik van *kansen* (in de politieke context) is niet echt verrassend maar het bevestigt wel het karakter van de partij. *Kansen geven* is dus een cluster, *geven* (plaats 168) is echter ook een keyword. Ook hier zien we vertrouwde links: *economie, kansen, meer, mogelijk, bedrijven, eigen, gewesten, maken, bijkomende, voldoende, mogelijkheden, burgers, ons, werken, werknemers, Vlaanderen, ruimte, onze, nieuwe, alle, moeten, open, kunnen, Vld, moet, mensen*. De clusters zijn:

*kansen geven aan de  
kansen te geven  
kansen geven om  
kans geven op de  
het goede voorbeeld geven*

Tot slot willen we ook nog, zoals in het vorige hoofdstuk, nagaan wat de plaats is van de woorden *Vlaanderen* en *Vlaamse*. Dit corpus bestaat uit verkiezingsprogramma's voor de verschillende niveaus (waaronder het Vlaamse, maar ook het Europese, het federale en het provinciale niveau), en het is inderdaad zo dat *Vlaanderen* en *Vlaamse* niet in elk deel van het corpus even belangrijk zijn.<sup>17</sup> *Vlaanderen* staat met een 17de plaats zes plaatsen lager dan in het vorige corpus. *Vlaamse* is van de lijst verdwenen en vinden we pas terug op de 59ste plaats terwijl dit in het vorige corpus op de vierde plaats stond!<sup>18</sup> De belangrijkste gelinkte keywords bij *Vlaanderen* en *Vlaamse* (dit laatste heeft er meer) zijn: *ruimte, versterken, eigen, administratieve, provincies, werken, zorgen, ambtenaren, ontwikkelingssamenwerking, middelen, samenleving, steden, hoger, sector, beleid, onderwijs, scholen, inzetten, economie, bedrijven, overheid, sociale, mensen, landbouw, beter, bevoegdheid, essentieel, vervoer, maken, belangrijke, vlak, openbaar, mobiliteit, belangrijk, ons, manier, innovatie, samenwerking, legislatuur, nodige, binnen, resolutie, Europese, bevoegdheden, nieuwe, geven, land, verder, kunnen, open, moeten, meer, Vld, moet*. Er zijn geen interessante clusters voor *Vlaanderen*, bij *Vlaamse* vinden we het volgende:

*de Vlaamse overheid  
de Vlaamse regering  
de volgende Vlaamse  
Vlaamse regering moet  
de Vlaamse rand  
Vlaamse overheid moet*

<sup>17</sup> Bij het bekijken van de plots zien we dat *Vlaamse* en *Vlaanderen* veel meer voorkomen in één deel van het corpus.

<sup>18</sup> Zowel *Vlaanderen* en *Vlaamse* hebben een absolute frequentie van 266 (dit is minder dan in het vorige corpus) en staan op plaats 21 en 22 in de lijst van absolute frequenties in dit corpus. Toch is hun keyness beduidend lager.

*Vlaamse liberalen en  
de Vlaamse gemeenschap  
de Vlaamse liberalen*

Het mag duidelijk zijn dat dit corpus een stuk 'liberaler' getint is dan het vorige. Dit is ook niet verwonderlijk: de verkiezingsprogramma's zijn hét geschreven woord van de partij. Het zijn ook zeer belangrijke documenten want zonder stemmen verliest een partij haar macht. Zoals aangetoond, herkennen we de basisfilosofie duidelijker in de verkiezingsprogramma's dan in de nieuwsberichten. *Vrijheid* is een (vrij belangrijk) keyword, net als *samenleving*, *economie*, *ondernemingen*, *burger*, *werken*, enz. Toch moeten we ook hier vaststellen dat een paar andere woorden ontbreken in deze lijst. *Vrije* staat slechts op plaats 163 met een absolute frequentie van 67. *Verantwoordelijkheid* (absolute frequentie 70) vinden we pas op plaats 242 en wanneer we dit woord toch onderzoeken, levert het niets echt 'liberaals' op. De meeste constructies met dit woord gaan over de verantwoordelijkheid van regeringen en politici, slechts enkele hebben te maken met de verantwoordelijkheid van het individu. *Initiatief* is geen keyword en komt op meer dan 160.000 woorden slechts 20 keer voor. *Ondernemingen* is wel een keyword in de top 50, maar *ondernemen* staat pas op 232. *Ondernemerschap* vinden we op 67. Tot slot is *individu* geen keyword en dit is met maar 17 voorkomens ook logisch.

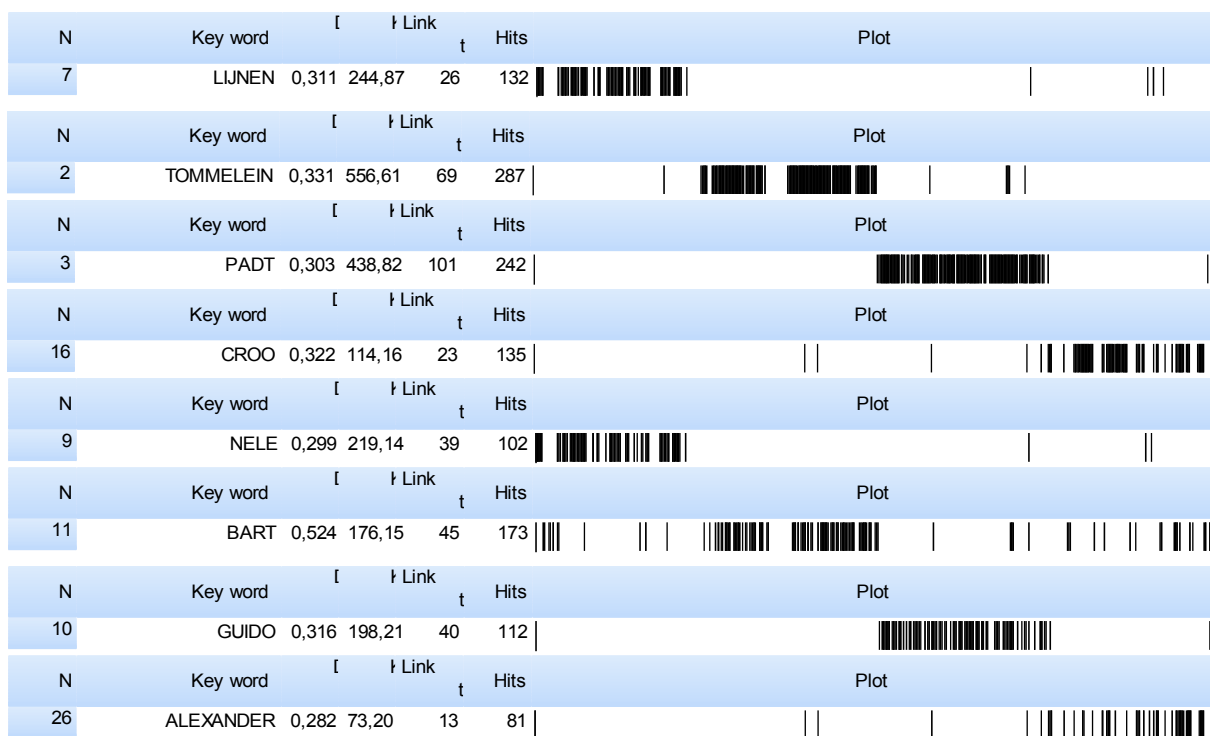
## 6. Analyse van het taalgebruik van vier Open Vld-senatoren

Onze laatste analyse behandelt het corpus met tekstmateriaal van vier Open Vld-senatoren: Nele Lijnen, Bart Tommelein, Guido De Padt en Alexander De Croo. Hieronder ziet u de keywords uit dit corpus (figuur 4). De methodologie voor dit corpus zal lichtjes anders zijn dan bij de vorige twee corpora. In deze analyse zullen de plots namelijk een belangrijke functie hebben omdat we op die manier kunnen nagaan welk keyword bij welke senator hoort. Het is immers zo dat elke senator wel handelt en schrijft vanuit de liberale ideologie (of dat vermoeden we toch) maar ieder heeft toch zijn of haar eigen thema's waarrond hij of zij actief is.

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	TOMMELEIN	287	0.27	7		556.61	0.00
2	PADT	242	0.23	11		438.82	0.00
3	OPEN	371	0.35	124	0.06	343.84	0.00
4	VLD	317	0.30	93	0.05	320.02	0.00
5	SENATOR	164	0.16	13		270.35	0.00
6	LIJNEN	132	0.13	5		244.87	0.00
7	HANDICAP	117	0.11	2		232.74	0.00
8	NELE	102	0.10	0		219.14	0.00
9	GUIDO	112	0.11	6		198.21	0.00
10	BART	173	0.16	50	0.02	176.15	0.00
11	VIVANT	80	0.08	0		171.86	0.00
12	WE	565	0.54	549	0.27	128.92	0.00
13	ONS	295	0.28	207	0.10	126.48	0.00
14	ONZE	228	0.22	142	0.07	115.82	0.00
15	CROO	135	0.13	51	0.03	114.16	0.00
16	SABAM	51	0.05	0		109.55	0.00
17	BASISINKOMEN	47	0.04	0		100.96	0.00
18	N-VA	46	0.04	0		98.81	0.00
19	WETSVOORSTEL	52	0.05	2		96.26	0.00
20	PERSBERICHTEN	43	0.04	0		92.37	0.00
21	OVERHEID	118	0.11	51	0.03	89.16	0.00
22	SENAAT	60	0.06	8		86.31	0.00
23	OOSTENDE	76	0.07	20		81.73	0.00
24	IMMERS	96	0.09	38	0.02	78.19	0.00
25	ALEXANDER	81	0.08	28	0.01	73.20	0.00
26	VRIJWILLIGERSWERK	34	0.03	0		73.03	0.00
27	ONGEVALLLEN	34	0.03	0		73.03	0.00
28	WERKEN	138	0.13	87	0.04	68.94	0.00
29	PARLEMENTAIRE	32	0.03	0		68.74	0.00
30	GESCHREVEN	50	0.05	8		67.55	0.00
31	KANSEN	59	0.06	14		67.08	0.00
32	WERKLOZEN	42	0.04	4		66.38	0.00
33	VRAAG	114	0.11	64	0.03	65.88	0.00
34	SOCIALE	112	0.11	62	0.03	65.79	0.00
35	UITKERING	43	0.04	5		64.47	0.00
36	MOET	317	0.30	322	0.16	64.42	0.00
37	WIJ	125	0.12	77	0.04	64.38	0.00
38	FRAUDE	53	0.05	11		64.31	0.00
39	KUNNEN	323	0.31	332	0.16	63.57	0.00
40	MINISTER	188	0.18	152	0.08	63.43	0.00
41	MOETEN	306	0.29	314	0.16	60.49	0.00
42	STAATSHERVORMING	52	0.05	12		59.96	0.00
43	KINDERBIJSLAG	31	0.03	1		58.52	0.00
44	BETREFT	62	0.06	21	0.01	56.84	0.00
45	MENSEN	257	0.24	253	0.12	56.64	0.00
46	PARKETTEN	26	0.02	0		55.85	0.00
47	SNELHEIDSOVERTREDINGEN	26	0.02	0		55.85	0.00
48	AANTAL	158	0.15	125	0.06	55.45	0.00
49	POLITIEK	85	0.08	43	0.02	55.13	0.00
50	REEDS	47	0.04	11		53.81	0.00

*Figuur 4*

Om aan te tonen dat het corpus wel degelijk uit vier duidelijk te onderscheiden delen bestaat (één deel per senator), bekijken we hieronder de plots met de namen van de senatoren (figuur 5). *Nele Lijnen* is de eerste senator in het corpus. Het tweede deel is van *Bart Tommelein*, gevolgd door *Guido De Padt*. Hier vierde deel van het corpus bestaat uit teksten van *Alexander De Croo*. Aan de hand van deze plots kunnen we dus duidelijk zien dat het corpus uit vier delen bestaat. We zullen daarna ook de rest van de top 50 van de keywordslijst van de Open Vld-senatoren analyseren. We bespreken eerste elke senator apart, inclusief zijn of haar thema's. Daarna onderzoeken we welke belangrijke woorden uit de basisfilosofie terug te vinden zijn in de top 50 van de keywordslijst van het senatorencorpus en welke woorden niet.



Figuur 5

Op de eerste<sup>19</sup> plaats in de keywordslijst staat het keyword *Tommelein*. Net als in het eerste corpus is Bart Tommelein dus een belangrijke eigennaam (de keyness is 556,61 en dit is een stuk hoger dan de 438,62 van *Padt*). Op de tweede plaats staat het keyword *Padt*, op plaats zeven staat *Lijnen*. Pas op de vijftiende plaats vinden we *Croo*. De voornamen van deze senatoren scoren ook hoog. Op plaats 8 staat *Nele* en op plaats 9 staat *Guido*. Bij deze senatoren zal hun voornaam dus vaak voorkomen samen met hun familienaam. Dit is minder het geval bij *Bart* (plaats 10) en *Alexander* (plaats 25). We kunnen dit ook statistisch bewijzen door naar de clusters *Voornaam + Familienaam* te kijken: *Bart Tommelein* komt 129 keer voor (= 44,9% van alle clusters met *Tommelein*), *Guido De Padt* 133 keer (= 54,9% van alle clusters met *Padt*), *Nele Lijnen* 100 keer (= 75,5% van alle clusters met *Lijnen*) en *Alexander De Croo* 73 keer (= 54% van alle clusters met *Croo*). Deze berekeningen tonen aan dat Nele Lijnen de combinatie *Voornaam + Familienaam* veel meer gebruikt dan Alexander De Croo, Guido De Padt of Bart Tommelein, hoewel Tommelein het belangrijkste keyword is in dit corpus. De namen van de mannelijke senatoren zullen dus meer gecombineerd worden met andere woorden (vermoedelijk hun functie binnen de partij). Dit wordt hieronder nog verder onderzocht per senator apart.

<sup>19</sup> De waarde van het cijfer naast een keyword in de lijst van plots is altijd één hoger dan de positie in de keywordslijst. Dit is omdat in de plotlijst *Overall* op de eerste plaats staat. Zo is bijvoorbeeld *Tommelein* het eerste keyword, in de lijst van plots staat het op de tweede plaats.

## 6.1. Senator Nele Lijnen

Bovenstaande plot (figuur 5) toont dat de teksten van Nele Lijnen eerst in het corpus staan. Er zijn zowaar geen gelinkte keywords voor *Lijnen*, enkel voor *Nele*: *wetsvoorstel*, *minister*, *Vld*. We vinden wel collocaten: *senator*, *open*, *Vld*, *Schiltz*, *vindt*, *Willem-Frederik*, *aldus*, *Vld-senator*, *zegt*, *wetsvoorstel*, *minister*, *haar*, *mensen*, *moet*, *nu*, *samen*, *daarom*, *we*, *meer*, *vijfpuntenplan*, *senaat*, *Hermans*, *handicap*, *personen*, *Vivant*, *sociale*, *Vautmans*, *stellen*, *pensioenopbouw*, ... Er zijn ook niet veel clusters:

*senator Nele Lijnen*  
*Nele Lijnen en*  
*Lijnen Open Vld*  
*Open Vld-senator Nele Lijnen*  
*Nele Lijnen aan*

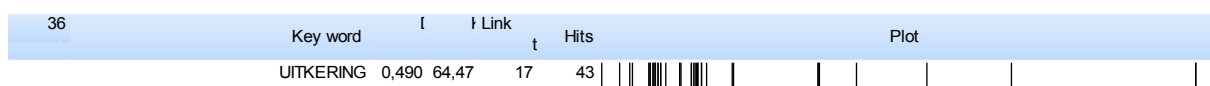
Deze beperkte set van resultaten toont ons wel twee 'sociale' thema's van de senator: Nele Lijnen houdt zich onder andere bezig met de gehandicapten en de pensioenen. Uit ervaring weet ik dat dit klopt, de senator is zeer veel bezig met deze materie. Dit deed en doet ze onder andere met Margriet Hermans en Willem-Frederik Schiltz.

*Sociale* is ook een keyword in dit corpus (plaats 34) en het hoort grotendeels bij Nele Lijnen, zoals te zien is in het plot (figuur 6). Er zijn twee links, namelijk *domiciliefraude* en *kunnen*. Uit ervaring weet ik dat domiciliefraude een thema is van Nele Lijnen. De belangrijkste collocaten zijn: *zekerheid*, *fraude*, *zaken*, *fiscale*, *werken*, *stelt*, *minister*, *domiciliefraude*, *media*, *partners*, *jobstudenten*, ... De enige cluster is *onze/de sociale zekerheid*.

Hier thematisch mee verbonden vinden op plaats 35 het keyword *uitkering*. Ook dit is duidelijk een thema van senator Lijnen (figuur 7). Er zijn geen nuttige links of clusters maar wel collocaten: *verliezen*, *vrijwillig*, *ontvangen*, *mensen*, *raken*, *sociale*, *bijverdienen*, *bovenop*, *moet*, ...



Figuur 6



Figuur 7

Al op plaats 7 in de keywordslijst vinden we het 'sociale' keyword *handicap*. De enige links zijn *kunnen* en *vrijwilligerswerk*. De collocaten zijn *mensen*, *personen*, *moeten*, *kunnen*, *jongeren*, *vrijwillig*, *aangeboren*, *werken*, *kansen*, *federale*, *jobstudenten*, *geven*, *doen*, *volwaardig*, ... De clusters tonen niets verrassend:

*met een handicap*  
*personen met een handicap*  
*voor personen met een handicap*  
*een handicap moeten*  
*een aangeboren handicap*  
*jobstudenten met een handicap*  
*jongeren met een handicap*

Ook het plot toont dat senator Lijnen het meeste over mensen met een handicap schrijft (figuur 8).



Figuur 8

Nele Lijnen was en is nog steeds lid van de organisatie Vivant. Vroeger vormde deze partij een

coalitie met de Open Vld, vandaag de dag is Vivant enkel nog een denktank. Toch staat het op plaats 10 in de keywordslijst. De enige links zijn *Open* en *Vld*. Collocaten zijn *Vld*, *open*, *basisinkomen*, *Roland*, *Duchâtelet*, *partij*, *meer*, *Nele*, *Lijnen*, *Somers*, *verkiezingen*, *beweging*, *balans*, *blijft*, *iedereen*, *aldus*, ... Roland Duchâtelet was ooit een belangrijk politicus en lid van Vivant. De partij kwam op voor een basisinkomen voor alle Belgen. Deze zaken vinden we rond het keyword *Vivant*. De enige cluster is *Vld en Vivant*. Uit het plot (figuur 9) blijkt ook dat Nele Lijnen verbonden is met Vivant. *Basisinkomen* is zelf ook een keyword (plaats 18) maar het heeft geen links. Er zijn niet zoveel collocaten (*iedereen*, *Vivant*, *kunnen* en *universeel*) en ook maar twee clusters: *basisinkomen voor iedereen* en *van een basisinkomen*. Het is wel duidelijk dat het basisinkomen een concept van Vivant is en dus ook van Nele Lijnen. Dit blijkt ook uit het plot (figuur 10).

N	Key word	f	Link	t	Hits	Plot
12	VIVANT	0,033	171,86	24	80	

Figuur 9

N	Key word	f	Link	t	Hits	Plot
18	BASISINKOMEN	0,254	100,96	9	47	

Figuur 10

Een mogelijk interessant woord is *werken* (plaats 28), een concept dat volgens het plot (figuur 11) voornamelijk bij Nele Lijnen hoort. De analyse valt echter tegen: er zijn geen nuttige links, collocaten of clusters. We vinden wel iets bij *vrijwilligerswerk* (plaats 26 - figuur 12). De links tonen opnieuw de 'sociale' kant van de senator (*uitkering*, *mensen*, *handicap*), net als de collocaten: *volwaardig*, *verliezen*, *kwaliteit*, *benutten*, *bijverdienen*, *wetsvoorstel*, ... Er zijn geen noemenswaardige clusters.

29	Key word	f	Link	t	Hits	Plot
...	WERKEN	0,690	68,94	33	138	

Figuur 11

28	Key word	f	Link	t	Hits	Plot
...	VRIJWILLIGERSWERK	0,416	73,03	17	34	

Figuur 12

## 6.2. Senator Bart Tommelein

De tweede politicus is volgens het plot Bart Tommelein. Zoals reeds gezegd is hij de senator met de hoogste keyness. Er zijn niet veel links verbonden met de keywords *Tommelein of Bart: betreft, immers, minister, open Vld, antwoord, Oostende*. De collocaties zijn de volgende: *open, Vld, senator, Vlaams, fractievoorzitter, minister, vindt, besluit, volksvertegenwoordiger, Oostende, tevreden, vraagt, Sabam, regering, liberaal, regionale, ik, Oostends, omroepen, fraude, parlement, hoopt, maatregelen, steunt, we, ...* De belangrijkste clusters zijn:

*Bart Tommelein*  
*Tommelein Open Vld*  
*Vlaams Volksvertegenwoordiger Tommelein*  
*Volksvertegenwoordiger Bart Tommelein*  
*Senator en fractievoorzitter Bart Tommelein*  
*aldus Tommelein*  
*senator Bart Tommelein*  
*Tommelein vindt het*  
*besluit Bart Tommelein*  
*vraag van Bart Tommelein*  
*Tommelein is tevreden dat*  
*fractievoorzitter en senator*  
*Tommelein wijst er op*  
*Bart Tommelein wil*

Deze clusters bevestigen de opmerking dat *Tommelein* veel voorkomt in combinatie met zijn verschillende functies. Daarnaast tonen ze ons ook waarover de senator schrijft in zijn berichten: Bart Tommelein geeft zijn mening (*aldus, besluit*), hij stelt vragen<sup>20</sup>, hij kan tevreden berichten over een nieuwsfeit, hij wijst een andere instantie of persoon op een feit en hij wil ook bepaalde zaken zien gebeuren.

Er zijn nog twee keywords die we duidelijk aan Bart Tommelein kunnen linken. Op plaats 16 staat de overheidsinstantie *Sabam*. Er zijn geen links, geen nuttige clusters en één nuttige collocaat, die we wel kunnen verbinden met het plot: *Tommelein*. Bart Tommelein heeft dus een aantal teksten geschreven in verband met Sabam (figuur 13). Op plaats 23 staat *Oostende*. De links zijn *Bart* en *open*, de collocaten zijn *Tommelein, stad* en *voorzitter* en de enige cluster is *Open Vld Oostende*. Ook hier toont het plot dat dit keyword toebehoort aan Bart Tommelein (figuur 14).

17	Key word	f	Link	t	Hits	Plot
...	SABAM	0,452	109,55	4	51	

Figuur 13

24	Key word	f	Link	t	Hits	Plot
...	OOSTENDE	0,336	81,73	17	76	

Figuur 14

<sup>20</sup> Senatoren stellen veel schriftelijke en mondelinge vragen aan de regering over het gevoerde beleid. Deze vragen zijn een zeer belangrijke activiteit binnen de senaat.

### 6.3. Senator Guido De Padt

De derde politicus (tweede volgens hoogste keyness) is Guido De Padt. De gelinkte keywords bij *Padt* en *Guido* zijn: *kunnen, geschreven, Guido, wetsvoorstel, minister, vraag, open, persberichten*. De belangrijkste collocaten zijn: *senator, stelt, besluit, aldus, merkt, vraag, persbericht, kunnen, minister, pleit, cijfers, meer, partijgenoot, wetsvoorstel, efficiëntere, snelheidsovertredingen, probleem, maatregel, noodzakelijk, ...* We zien, net als bij Tommelein, een hoop woorden die verwijzen naar de taak van een senator om vragen of opmerkingen te maken bij het beleid van de regering. Dit zien we ook bij de clusters:

*Guido De Padt*  
*Senator Guido De Padt*  
*De Padt Guido*  
*Guido De Padt 0475 44 94 04<sup>21</sup>*  
*De Padt Open Vld*  
*stelt De Padt*  
*merkt De Padt*  
*aldus De Padt*  
*senator De Padt*  
*vraag van senator Guido De Padt*  
*Open Vld senator De Padt*  
*in persberichten geschreven*  
*De Padt wil*  
*stelt Guido De Padt*

Er zijn volgens het plot slechts twee keywords die we bij Guido De Padt kunnen plaatsen. Het gaat hierbij om *parketten* (figuur 15) en *snelheidsovertredingen* (figuur 16). Bij beide woorden is *politie* een gelinkt keyword en dit is ook het enige noemenswaardige wat vermeld kan worden bij deze keywords. Als stagiair in de senaat kwam ik te weten dat senator De Padt veel bezig is met thema's als de politie en snelheidsovertredingen. Ook volgt hij enkele juridische thema's op door zijn studies in de rechten.

47	Key word	t	Link	Hits	Plot	
	PARKETTEN	0,276	55,85	8	26	

Figuur 15

48	Key word	t	Link	Hits	Plot	
	SNELHEIDSOVERTREDINGEN	0,257	55,85	7	26	

Figuur 16

<sup>21</sup> De reden dat deze ietwat vreemde cluster in de lijst staat, is omdat deze senator zijn teksten altijd afsluit met zijn telefoonnummer.





### 6.5. Top 50 van de keywordslijst

We merkten al in de analyses van de namen van de senatoren dat zij bezig zijn met het schrijven van vragen, gericht aan de regering. Ook denken ze na (of leveren ze kritiek) over actuele thema's in ons land. Deze activiteiten zouden we dus ook terug moeten vinden wanneer we het keyword *senator* (plaats 5) onderzoeken in de top 50 van de keywordslijst van het senatorenkorps. De gelinkte keywords tonen alvast niet zoveel: *wetsvoorstel, parlementaire, open, Vld*. De collocaties bevatten volgende woorden: *open, Vld, Padt, Guido, Lijnen, Nele, vraag, Tommelein, Bart, fractievoorzitter, parlement, pleit, aldus, schriftelijk, liberale, daarom, samen, jaar, stelt, minister, volgens, verheugd, merkt, maken, Duchâtelet, jobstudenten, Alexander, vijfpuntenplan, tijd, leven, ...* De clusters zijn:

*Senator Guido De Padt*  
*Open Vld senator*  
*vraag van senator*  
*parlementaire vraag van*  
*senator en fractievoorzitter*  
*Open Vld-senator*  
*senator Bart Tommelein*  
*senator Nele Lijnen*  
*de liberale senator*  
*een schriftelijke vraag*  
*vraag van Open Vld*  
*fractievoorzitter en senator*

Woorden als *pleit, aldus, vraag, schriftelijk, merkt, stelt* en *volgens* zijn inderdaad verbonden met de activiteiten die daarnet aangehaald werden. Deze activiteiten zijn ook te vinden in de clusters. Uit de clusters wordt ook duidelijk dat Alexander De Croo niet veel schrijft over zijn job als senator, want de cluster *senator (Alexander) De Croo* komt niet voor in het corpus. Dit kunnen we ook bewijzen door te kijken naar het vierde deel van het plot bij *senator*: het woord komt nauwelijks voor in het deel van Alexander De Croo. Nele Lijnen en Guido De Padt, daarentegen, gebruiken het woord duidelijk veel meer (figuur 20).



Figuur 20

Het onderscheid tussen de voorzitter en de senatoren is ook merkbaar wanneer we kijken naar de verdeling van de keywords *we* (plaats 12) *ons* (plaats 13) en *onze* (plaats 14) (figuur 21). Vanuit onze intuïtie zouden we deze woorden vooral bij de voorzitter plaatsen, omdat hij als partijleider met deze woorden een saamenhorigheidsgevoel bij zijn leden kan creëren. Dit kunnen we ook onderzoeken door te kijken naar de plots van deze woorden. Zoals we zien gebruiken alle senatoren deze woorden op reguliere basis (en dit is ook maar logisch, want een politicus schrijft voor zijn achterban), maar vooral het vierde deel van het plot (dat van De Croo) is sterk zwart gekleurd. Ons vermoeden klopt dus: het is vooral de voorzitter die deze woorden veel gebruikt. In onderdeel 6.4. zagen we in de analyse van de teksten van Alexander De Croo verschillende keren het persoonlijke voornaamwoord *ik* voorkomen, onder andere bij de collocaten van *Alexander* en *Croo*. *Ik* is echter geen keyword. Het is verbazend dat we dit woord niet vaker zien opduiken, omdat politici in hun teksten ook zichzelf moeten 'verkopen'. Uit ervaring weet ik wel dat politici berichten vaak in de derde persoon schrijven, waardoor *ik* minder gebruikt wordt. *Ik* komt in het senatorenkorps 430 keer voor. Wanneer we het plot (figuur 22) bekijken, zien we dat vooral Bart Tommelein en Alexander De Croo dit persoonlijke voornaamwoord gebruiken.

N	Key word	f	Link	t	Hits	Plot
13						
14	WE	0,813	128,92	127	565	
15	ONS	0,745	126,48	56	295	
	ONZE	0,805	115,82	65	228	

Figuur 21

N	Plot
1	

Figuur 22

We hebben nu alle thema's van de vier senatoren uit de top 50 van de keywordslijst overlopen. In de keywordslijst komen echter ook nog woorden voor uit de basisfilosofie (voor zover die nog niet bij een senator besproken zijn). *Open* staat op plaats 3 en heeft met een absolute frequentie van 371 meer voorkomens dan *Vld* (vierde plaats – 317 voorkomens). *Open* dus nog anders gebruikt dan enkel in die cluster. Bekijken we de concordantielijnen, dan vinden we inderdaad een aantal andere gebruiken van het woord *open* (deze komen evenwel minder voor): *Open Vld-senator*, *Open Vld-fractieleider*, *Open Vld-voorzitter*, *open brief*, *open grenzen*, *open samenleving*, *open dialoog*, *open benadering*, *open en neutraal internet*. De gelinkte keywords bij *open* zijn: *moet*, *vraag*, *vivant*, *Oostende*, *Bart*, *wij*, *Tommelein*, *Vld*, *senator*, *Padt*. Ook bij de clusters zien we niets dat al het voorgaande tegenspreekt:

*Guido De Padt Open Vld*  
*van Open Vld*  
*Bart Tommelein Open Vld*  
*Open Vld wil*  
*Open Vld-senator*  
*Alexander De Croo Open Vld*  
*Open Vld Oostende*  
*Turtelboom Open Vld*  
*Nele Lijnen Open Vld*  
*vraag van Open Vld*  
*Open Vld is tevreden*

Het mag duidelijk zijn dat *open* toch voornamelijk bij Vld staat (de combinatie komt 288 keer voor), net zoals in de vorige corpora. Dit is duidelijk te zien op figuur 23, want hun verdeling is bijna identiek. De gelinkte keywords bij Vld zijn dan ook gelijkaardig: *we*, *Oostende*, *liberale*, *wij*, *Bart*, *Nele*, *open*, *Tommelein*, *Vivant*, *senator*. De collocaten en clusters zijn bijna dezelfde als bij *open*, omdat de twee zoveel samen voorkomen.

N	Key word	f	Link	t	Hits	Plot
4	OPEN	0,873	343,84	101	371	
N	Key word	f	Link	t	Hits	Plot
5	VLD	0,815	320,02	108	317	

Figuur 23

Een ander keyword dat ook in de vorige corpora voorkwam, is *overheid* (plaats 23). Het plot toont aan dat alle senatoren dit woord ongeveer evenveel gebruiken (figuur 24). Er zijn geen gelinkte keywords. De belangrijkste collocaten zijn: *federale*, *moet*, *Vlaamse*, *lokale*, *euro*, *aldus*, *ik*, *miljoen*, *handicap*, *werken*, ... Er zijn enkele clusters met *overheid*:

*de federale overheid*  
*van de overheid*  
*door de overheid*  
*de Vlaamse overheid*  
*de overheid moet*

22	Key word	f	f Link	t	Hits	Plot
	OVERHEID	0,850	89,16	23	118	

Figuur 24

De laatste keywords die we in deze meesterproef zullen behandelen zijn *moet* en *moeten*. Scoorden deze modale werkwoorden zeer hoog in de vorige corpora, hier staan ze slechts op plaatsen 36 en 41. De links zijn: *vandaag, maken, jobs, kunnen, jobstudenten, doen, mensen, we, minister, antwoord* en *geven*. Naast deze woorden vinden we bij de collocaties ook nog: *handicap, vrijwillig, blijven, minder, Vlaanderen, Padt, recht, kosten, eerst, sociale, betalen, jaar, ...* De clusters zijn: *met een handicap moeten, we moeten doen* en *die we moeten*. Ook in dit corpus lijken deze modale werkwoorden niet echt een verplichtend karakter te hebben.

### 6.6. Keywords uit de basisfilosofie

We hebben in deze analyse zeer veel keywords besproken. Toch valt het op hoe weinig keywords uit de basisfilosofie in deze keywordslijst zitten. We vonden wel *open, Vld* en *overheid*. Waar ergens op de lijst staan de andere keywords uit de basisfilosofie? Ten eerste ontbreken *individu* en *individueel*, die niet eens keywords zijn en een absolute frequentie hebben van 4 en 3. Initiatief is evenmin een keyword met een absolute frequentie van 28. *Liberale* staat wel in de lijst op 93 (absolute frequentie is 46) en *liberale* staat op 93 (absolute frequentie is 46). Met een absolute frequentie van 13 staat *ondernemers* op plaats 128. *Ondernemen* (absolute frequentie 18) en *onderneming* (absolute frequentie 2) staan niet in de keywordslijst. *Private* staat ook niet op de keywordslijst (absolute frequentie 13). *Verantwoordelijkheid* heeft een absolute frequentie van 29 en haalt de lijst ook niet, evenmin als *verantwoordelijk* (absolute frequentie 4). In de vorige corpora stonden *Vlaanderen* en *Vlaams* altijd hoog. Hier vinden we *Vlaanderen* pas op plaats 137 (89 absolute frequentie) en *Vlaams(e)* staat op plaats 143 (absolute frequentie 130). *Vrijheid* staat op plaats 151 (absolute frequentie 27), *vrij* is geen keyword (absolute frequentie 16). Tot slot vinden we *zelfstandige* op plaats 149 (absolute frequentie 17). We kunnen met betrekking tot dit corpus dus besluiten dat de keywordslijst vooral bestaat uit concepten die we kunnen linken aan de verschillende senatoren. Veel 'liberale' woorden uit de basisfilosofie ontbreken echter of scoren zeer laag op de keywordsranking.

## 7. Conclusie

Na een uitgebreide analyse van de drie corpora kunnen we deze meesterproef afsluiten met een algemene bespreking van de resultaten. We hebben de drie corpora elk afzonderlijk geanalyseerd met *Wordsmith Tools* maar het is ook nuttig om de corpora vergelijkend te aanschouwen. Hieronder ziet u figuur 25. We hebben de belangrijkste woorden uit de basisfilosofie (zoals beschreven in de methodologie) op de y-as geplaatst. Op de x-as staan de drie corpora. Corpus 1 bevatte het nieuws van [www.openvld.be](http://www.openvld.be), corpus 2 bestond uit verkiezingsprogramma's en corpus 3 bevatte teksten van vier senatoren. Per belangrijk woord uit de basisfilosofie overlopen we of het al dan niet een keyword is (welke positie het heeft in de keywordslijst) en ook wat de absolute frequentie van het woord is. Omdat we van bepaalde woorden meerdere vormen in acht willen nemen, staan tussen haakjes de resultaten voor die woordvormen. Bijvoorbeeld: *privaat* is geen keyword in corpus 2 maar *private* is dit wel op plaats 33. We zullen de resultaten uit deze tabel overlopen en daarna onderzoeken wat het effect hiervan is op onze hypothesen.

Keyword	Corpus 1 ( <a href="http://www.openvld.be">www.openvld.be</a> )		Corpus 2 (Verkiezingen)		Corpus 3 (Senatoren)	
	Keyword?	Abs. Freq.	Keyword?	Abs. Freq.	Keyword?	Abs. Freq.
individu	nee	1	nee	17	nee	4
individueel(e)	nee	9 (25)	nee (165)	8 (55)	nee	3 (12)
initiatief	nee	41	nee	20	nee	28
liberaal(e)	98 (66)	43 (87)	nee	13 (25)	107 (93)	24 (46)
open	1	673	3	851	3	371
ondernemen	nee	28	232	43	nee	18
ondernemers	57	39	202	23	128	13
ondernemingen	nee	24	49	78	nee	2
overheid	32	165	5	406	21	118
privaat(e)	nee	46 (4)	nee (33)	6 (77)	nee	1 (13)
verantwoordelijkheid	nee	46	242	70	nee	29
verantwoordelijk(e)	nee	46 (6)	nee	33 (12)	nee	4 (1)
Vlaanderen	12	249	17	266	137	89
Vlaams(e)	4 (6)	492 (366)	nee (59)	63 (266)	nee (143)	55 (130)
Vld	2	522	1	1237	4	317
vrijheid	nee	17	26	103	151	27
vrij(e)	nee	23 (9)	nee (163)	30 (67)	nee	16 (15)
zelfstandig(e)	169 (96)	19 (39)	249 (296)	20 (25)	nee (149)	2 (17)

Figuur 25

Vreemd genoeg is *individu* geen keyword in de geanalyseerde corpora. Ook de absolute frequenties liggen zeer laag. We zouden verwachten dat een liberale partij, die traditioneel de belangen van het individu (en niet zozeer van de groep) verdedigt, dit woord meer neerschrijft. Deze opmerking geldt ook voor *individueel(e)*. Een liberale partij zou het woord *initiatief* ook veel moeten gebruiken, daar zij op het politieke toneel voor de ondernemers spreekt. Toch is ook dit woord geen keyword in de geanalyseerde corpora. Hiermee verbonden staan *ondernemen*, *ondernemers* en *ondernemingen* in de tabel maar ook hier kunnen we hetzelfde concluderen: zijn de termen al keywords, dan zijn het niet bijster belangrijke. *Privaat(e)* scoort ook zeer slecht, want het is in geen enkel corpus een keyword. Het woord *liberaal(e)* is wel een keyword in twee van de drie corpora (vreemd genoeg niet in het corpus met de verkiezingsprogramma's), maar het staat nauwelijks in de top 100 en het is dus moeilijk een belangrijk keyword te noemen. Dit is *open* daarentegen wel, net als Vld. De logische reden is dat de partij haar eigen naam zeer veel gebruikt en dus komen de twee zeer veel (samen) voor. Een ander belangrijk keyword is *overheid*, dat vaak in clusters gebruikt wordt (*Vlaamse overheid*, *federale overheid*, ...). Het woord komt ook zoveel voor omdat de partij veel schrijft over wat de overheid moet doen voor de burgers. We zouden intuïtief denken dat de liberalen voor minder overheidsbemoediging en voor meer individuele verantwoordelijkheid pleiten. Toch is *overheid* een belangrijk keyword en

*verantwoordelijkheid* en *verantwoordelijk(e)* niet. Die twee laatste woorden zijn bijna nooit een keyword. Hiermee verwant staat ook *zelfstandig* best laag op de keywordslijsten. Wel vinden we het Vlaamse karakter terug in de keywordslijsten, want *Vlaanderen* en *Vlaams(e)* scoren vrij hoog met hoge absolute frequenties (vooral bij de teksten vanop [www.openvld.be](http://www.openvld.be)). Tot slot moeten we vaststellen dat ook dé liberale woorden bij uitstek, *vrijheid* en *vrij(e)* niet bepaald belangrijke keywords zijn. Enkel in het corpus met de verkiezingsprogramma's geraakt *vrijheid* in de top 50 en *vrije* is daar slechts een zwak keyword.

Bekijken we de tabel voor het eerste corpus, dan zien we dat slechts enkele van de termen in de top 50 van de keywordslijst staan. Het merendeel is zelfs geen keyword. Toch zou men verwachten dat de partij zich in haar nationale nieuwsberichten liberaal profileert. Deze teksten worden immers gebruikt om de leden te informeren over het doen, laten en denken van de partij. Onze analyse toonde aan dat dit corpus niet echt 'liberaal' getint is en vooral een praktisch en informerend karakter heeft. Zo zijn vele namen van Open Vld-politici belangrijke keywords. Deze opmerkingen gelden ook voor het derde corpus. Hier staan natuurlijk de vier senatoren centraal. De top 50 van het tweede corpus bevat meer woorden uit de basisfilosofie maar ook niet allemaal, zoals u hierboven kunt zien. We zullen nu kijken in welke mate onze hypothesen worden bevestigd door de corpusanalyses, want de resultaten van ons onderzoek zijn niet bepaald eenduidig.

H1: Het corpus met de verkiezingsprogramma's is het sterkst 'liberaal' getint. We verwachten dan ook dat zowat alle belangrijke woorden uit de basisfilosofie voorkomen in de top 50 van deze keywordslijst: *individu, initiatief, liberaal, open, ondernemen, overheid, verantwoordelijkheid, Vlaams, Vld, vrijheid en zelfstandig*.

Deze hypothese is deels correct. Het corpus met de verkiezingsprogramma's is wel degelijk het sterkst 'liberaal' getint. Dit bleek niet alleen uit onze uitgebreide analyse maar ook uit de bovenstaande vergelijkende tabel. Toch zijn zeker niet alle termen uit de basisfilosofie terug te vinden in de top 50 van de keywordslijst (bijvoorbeeld *liberaal*). Enkel *open, ondernemingen, overheid, private Vlaanderen, Vld en vrijheid* staan in de top 50. Toch vonden we ook bij andere keywords, zoals *belastingen, kansen, economisch* en *werken*, duidelijk associaties en verbindingen die 'liberaal' getint zijn.

H2: Het corpus met de nieuwsberichten van [www.openvld.be](http://www.openvld.be) zal minder 'liberaal' getint zijn dan het corpus met de verkiezingsprogramma's. Er komen dus minder woorden uit de basisfilosofie voor in de top 50 van deze keywordslijst, aangezien nieuwsberichten over het algemeen minder wervend zijn geschreven.

Deze hypothese klopt wel. De nieuwsberichten zijn eerder informatief dan ideologisch of wervend. Wanneer we kijken naar de bovenstaande tabel zien we dat het corpus met de verkiezingsprogramma's de meeste woorden uit de basisfilosofie in de top 50 van haar keywordslijst heeft staan. Dit laatste heeft er 7 en het corpus met nieuwsberichten van [www.openvld.be](http://www.openvld.be) maar 5. Ook bleek uit onze analyse dat het corpus met de verkiezingsprogramma's 'liberaler' is. De logische reden is dat de nieuwsberichten eerder informatief zijn en de verkiezingsprogramma's eerder wervend en ideologisch.

H3: Het corpus met de nieuwsberichten van de vier senatoren is minder liberaal getint dan het corpus met de verkiezingsprogramma's. Er komen dus minder woorden uit de basisfilosofie voor in de top 50 van deze keywordslijst.

Deze hypothese is ook correct en de opmerkingen bij de vorige hypothese zijn ook hier van toepassing. De enige drie 'liberale' keywords in de top 50 van het senatorencorpus zijn *open, Vld en overheid*. Het senatorencorpus is eerder informerend en bevat veel teksten die handelen over het denken en doen van de senatoren (daarom vinden we vaak woorden als *vindt, meent, aldus* en *wil*). De teksten uit het corpus met de verkiezingsprogramma's zijn eerder wervend en ideologisch.

H4: Het corpus met de nieuwsberichten van [www.openvld.be](http://www.openvld.be) is ongeveer even liberaal getint als het corpus met de nieuwsberichten van de vier senatoren. Ongeveer evenveel woorden uit de basisfilosofie zijn keywords in beide corpora, aangezien ze documenten van hetzelfde genre bevatten.

De laatste hypothese is niet helemaal correct. Er zijn in beide corpora wel 8 termen key maar in het

corpus met nieuwsberichten van [www.openvld.be](http://www.openvld.be) zijn er 5 keywords in de top 50. In het senatorencorpus zijn dit er maar 3. De teksten van beide corpora zijn wel eerder informerend en ze bevatten documenten van hetzelfde genre, zoals blijkt uit de analyse, maar in het eerste corpus zijn de woorden uit de basisfilosofie toch iets belangrijkere keywords.

Impliceren de resultaten van onze analyses dat de liberalen het liberalisme hebben opgegeven? Natuurlijk niet. Het gaat hier om een beperkte set geschreven teksten die we vergeleken hebben met krantenartikels uit een Vlaamse krant. De keywordslijsten die we zo verkregen hebben, tonen ons welke concepten het sterkst opvallen op een totaal van enkele honderdduizend woorden. Uit ons onderzoek bleek immers ook dat men bijvoorbeeld niet 'socialistisch' schreef. Zo vonden we bij eerder sociale keywords (zoals *handicap* of *uitkering*) geen 'socialistische' verbindingen. Het is logisch dat bijvoorbeeld een senator niet in elke zin woorden als *initiatief* of *ondernemen* gebruikt, laat staan dat hij of zij vaak een lofzang op de vrijheid gaat schrijven. De teksten die we gebruikt hebben zijn geen ideologische traktaten, maar teksten uit het dagelijkse, politieke leven. Hun doel is het informeren en overtuigen van leden en mogelijke kiezers. Het zou wel logisch zijn dat men deze keywords zou gebruiken om de teksten enigszins subliminaal een liberale ondertoon te geven. Onze analyses, die algemene patronen blootleggen, tonen jammer genoeg aan dat dit niet veel gebeurt.

Deze laatste opmerkingen maken wel duidelijk dat we de resultaten niet mogen negeren of onderschatten. De analyses bewijzen dat de Open Vld vele zeer belangrijke woorden uit de basisfilosofie niet vaak expliciet neerschrijft. We denken hierbij aan woorden als *vrijheid* of *verantwoordelijkheid*. Vooral de nieuwsberichten van [www.openvld.be](http://www.openvld.be) en van de senatoren bevatten zeer weinig keywords uit de basisfilosofie. Ik zou dan ook graag eindigen met een suggestie voor de partij, die op dit moment in een electorale crisis zit en ook ideologisch op zoek is naar een nieuwe richting. De Open Vld staat nog steeds bekend als de liberale partij van Vlaanderen maar ze zou zich ook in haar woorden nog sterker liberaal kunnen profileren. Dit zou men kunnen doen door de woorden uit de bovenstaande lijst explicieter te gebruiken. Politici staan misschien niet bekend als "mensen van hun woord", maar men kan tenminste beginnen met het gebruik van de juiste woorden.

## 8. Dankwoord

Deze meesterproef was er nooit geweest zonder de hulp van een aantal mensen. Ik wil dit kleine hoofdstukje dan ook gebruiken om hen van harte te bedanken. Eerst en vooral wil ik mijn promotor, Professor Ann Bertels, enorm bedanken. Ze heeft me niet alleen geholpen door haar uitgebreide vakkennis maar ook door er altijd voor mij te zijn als ik haar hulp nodig had. Stuurde ik een mailtje met een vraag, dan antwoordde ze vrijwel meteen. Stuurde ik een tekst ter correctie door, dan zat de verbetering nog diezelfde dag in mijn inbox. Ze gaf niet enkel constructieve kritiek maar ze vertelde me ook wanneer ze een stuk tekst goed geschreven vond. Dit gaf me moed om verder te doen. Ik besef maar al te goed hoeveel geluk ik heb gehad dat zij me heeft begeleid.

Ook wil ik mijn vriendin Mandy bedanken voor haar steun, haar liefde en haar advies. Meermaals wees ze me op het feit dat ik bepaalde constructies of woorden moest vermijden. Haar geduld was eindeloos wanneer ik nog maar eens begon te klagen over hoe zwaar de deadline wel niet was. Dankzij haar bemoedigende woorden heb ik die deadline toch bereikt.

Dank u Danny, Rutger en mama voor het nalezen van delen van mijn meesterproef. Jullie haalden er al die vervelende schrijffouten uit die ik zelf niet meer zag van vermoeidheid. Dank u mama en papa om me te steunen tijdens dit zware jaar. Dank u zus, Sara en Nele om me te helpen bij het verzamelen van al mijn tekstmateriaal, want zonder dat materiaal en zonder alle steun van jullie allemaal was deze meesterproef een grote flop geweest. Nu heb ik daarentegen een meesterproef waar ik trots op ben.



## 9. Referenties

- Berber Sardinha, A. (1999). Using KeyWords in text analysis: Practical aspects. *DIRECT Papers*, 42, 1-8.
- Biber, D. & Conrad, S. & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Danet, B. (1980). Baby or fetus? Language and the construction of reality in a manslaughter trial. *Semiotica*, 32 (3-4), 187-219.
- De Croo, A. (2011). *Nieuws*. [23.11.2011, alexanderdecroo.be: [www.alexanderdecroo.be/nieuws](http://www.alexanderdecroo.be/nieuws)]
- De Padt, G. (2011) *Pers*. [25.11.2011, guidodepadt.be: <http://www.guidodepadt.be/categorie/nl/pers/>]
- De Standaard. (2011). *De Standaard Online*. [3.11.2011, standaard.be: [www.standaard.be](http://www.standaard.be)]
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1): 61-74.
- Fairclough, N. (1992). *Discourse and Social Change*. Cambridge: Polity.
- Fairclough, N. & Wodak, R. (1997) Critical Discourse Analysis. In Van Dijk, T.A. (ed.) *Discourse as Social Interaction*. Londen: Sage Publications.
- Firth, J.R. (1935). The technique of semantics. *Transactions of the Philological Society*, 36-72.
- Firth, J.R. (1949). The semantics of linguistic science. *Lingua*, 1, 393-404.
- Firth, J.R. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in Linguistic Analysis*, Special Volume, Philological Society, 1-32.
- Firth, J.R. (1970). *Tongues of Men and Speech*. Londen: Oxford University Press.
- Fowler, R. (1991). *Language in the News: Discourse and Ideology in the Press*. Londen: Routledge.
- Guilhaumou, J. (2004). *Où va l'analyse de discours? Autour de la notion de formation discursive*. [12.2.2012, Revue-texto.net: [http://www.revue-texto.net/Inedits/Guilhaumou\\_AD.html](http://www.revue-texto.net/Inedits/Guilhaumou_AD.html)]
- Halliday, M.A.K. (1992). Some lexicogrammatical features of the Zero Population Growth text. In: Mann, W. C; Thompson, S.A. (eds.). *Discourse description - Diverse linguistic analyses of a fund-raising text*. Amsterdam, John Benjamins, 327-358.
- Hoey, M. (2007). Lexical Priming and literary creativity. In Hoey, M.; Mahlberg, M.; Stubbs, M.; Teubert, W. (eds.). (2007). *Text, Discourse and Corpora. Theory and Analysis*. Londen: Continuum.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Kress G. (1985). Ideological Structures in discourse. In Van Dijk, T.A. (ed.). *Handbook of Discourse Analysis. Volume 4. Discourse Analysis in Society*. Londen: Academic Press Inc.
- Lijnen, N. (2011). *SneL-nieuws*. [16.11.2011, nelelijnen.be: <http://nelelijnen.be/snelnieuws.cfm>]

Mahlberg, M. (2005). Lexical items in discourse: identifying local textual functions of sustainable development. In Hoey, M.; Mahlberg, M.; Stubbs, M.; Teubert, W. (eds.). (2007). *Text, Discourse and Corpora. Theory and Analysis*. Londen: Continuum.

Mayaffre, D. (2002). *L'Herméneutique numérique*. [10.04.2012, uottawa.ca: <http://www.uottawa.ca/academic/arts/astrolabe/articles/art0031.htm/Hermeneutique.htm>]

Mayaffre, D. (2005). Rôle et place des corpus en linguistique : réflexions introductives. *Texto!*, 10 (4).

Mayaffre, D. (2008). Quand "travail", "famille", "patrie" co-occurrent dans le discours de Nicolas Sarkozy. Etude de cas et réflexion théorique sur la co-occurrence. In Heiden, S. & Pincemin, B. (eds.), *Actes des 9es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2008)*. Lyon: Presses Universitaires de Lyon, 811-822.

Mayaffre, D. (2009). *L'analyse du discours assistée par ordinateur*. [ 24.02.2012, eprints.aidenligne-francais-universite.auf.org: [http://eprints.aidenligne-francais-universite.auf.org/19/1/pdf\\_Formation\\_Mayaffre\\_Alexandrie\\_dec09\\_.pdf](http://eprints.aidenligne-francais-universite.auf.org/19/1/pdf_Formation_Mayaffre_Alexandrie_dec09_.pdf)]

Mautner, G. (2009). Corpora and Critical Discourse Analysis. In Baker, P. (ed.) *Contemporary Corpus Linguistics*. Londen: Continuum.

Milizia, D. (2010). Keywords and phrases in political speeches. In Bondi, M. & Scott, M. (Eds.). *Keyness in Texts*. Amsterdam: John Benjamins Publishing Co, 127-146.

Moon, R. (1994). The analysis of fixed expressions in text. In Coulthard, M. (ed.). *Advances in written text analysis*. Londen: Routledge, 117-13

Open Vld (2011), *Filosofie*. [11.02.2012, Openvld.be: <http://www.openvld.be/Content.aspx?id=877773fb-e646-4742-b8cf-bd177254e356>]

Orpin, D. (2005). Corpus Linguistics and Critical Discourse Analysis. *International Journal of Corpus Linguistics*, 10 (1), 37-61.

Paquot, M. & Bestgen, Y. (2009). Distinctive words in academic writing: A comparison of three statistical tests for keyword extraction. In Jucker, A. & Schreier, D. & Hundt, M. (Eds.). *Corpora: Pragmatics and Discourse*. Amsterdam: Rodopi, 247-269.

Phillips, N. & Lawrence, T.B. & Hardy, C. (2004). Discourse and Institutions. *Academy of Management Review*, 29 (4), 635-652.

Scott, M. (2000). Focusing on the Text and Its Key Words. In Burnard, L. & McEnery, T. (eds.) *Rethinking language pedagogy from a corpus perspective*. Frankfurt: Peter Lang, 103-122.

Scott, M. (2001). Mapping the KeyWords to problem and solution. In Hoey, M. Scott, M. & Thompson, G. (eds.) *Patterns of Text: in Honour of Michael Hoey*. Amsterdam: Benjamins, 109-127.

Scott, M. (2010). Problems in investigating keyness, or clearing the undergrowth and marking out trails... . In Bondi, M. & Scott, M. (eds.). *Keyness in Texts*. Amsterdam: John Benjamins Publishing Co, 43-58.

Scott, M. & Tribble, C. (2006). *Textual Patterns. Key words and corpus analysis in language*

*education*. Studies in Corpus Linguistics 22. Amsterdam: Benjamins.

Searle J.R. (2005) What is an institution? *Journal of Institutional Economics*, 1 (1), 1–22.

Sinclair, J. McH. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J. (2004). *Trust the Text*. Londen: Routledge.

Stubbs, M. (1996). *Text and Corpus Analysis*. Oxford: Blackwell Publishers Ltd.

Stubbs, M. (2009). The search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics*, 30 (1): 115-137.

Tommelein, B. (2011). *Nieuws*. [18.11.2011, tommelein.be:  
<http://www.tommelein.be/2007/nl/default.asp>]

Vandermeersch, P. (1997). *AVV-VVK niet langer in kop De Standaard*. [16.02.2012, standaard.be:  
<http://www.standaard.be/Artikel/Detail.aspx?artikelId=dst9909070002>]

Van Dijk, T.A. (2003). Critical Discourse Analysis. In Schiffrin, D; Tannen, D.; Hamilton, H.E. (eds.): *The handbook of discourse analysis*, , Hoboken: Wiley-Blackwell, 352-371.

Williams, R. (1976). *Keywords*. Londen: Fontana Paperbacks.

Wilson, J. (2003). Political Discourse. In Schiffrin, D & Tannen, D. & Hamilton, H.E. (eds.): *The handbook of discourse analysis*. Hoboken: Wiley-Blackwell, 398-415.

## 10. Biografie

Laat ik beginnen bij het begin. Mijn naam is Laurens Bynens en ik ben geboren op 6 juli 1990.

Al enkele jaren heb ik een grote interesse voor vier zaken: muziek, geschiedenis, politiek en talen. Mijn passie voor muziek uit zich in mijn grote platencollectie. Ik heb er meer dan 400. Op het moment dat ik deze biografie schrijf, galmt een vinyl van Glenn Miller uit 1964 door mijn kot. In mijn boekenkast staan een hoop geschiedenisboeken en romans.

Na mijn humaniora (Latijn - Moderne Talen) moest ik kiezen tussen Geschiedenis en Taal- en Letterkunde. Ik verkoos mijn liefde voor de Duitse en Engelse taal, maar kon deze gelukkig combineren met enkele geschiedenisvakken. In mijn bachelorpaper, die voornamelijk over het fascisme tijdens het interbellum in Engeland ging, kon ik mijn liefde voor geschiedenis botvieren. Ook kwam de (fascistische) politiek aan bod in mijn bachelorpaper, maar mijn honger was niet gestild. Ik wilde zien hoe het er in het echt aan toeging in de politieke wereld.

Om die reden koos ik ervoor om tijdens mijn Master Bedrijfscommunicatie een stage te doen in de senaat. Hier werkte ik voor senator Nele Lijnen (Open Vld). Als communicatieverantwoordelijke was ik actief in de politiek en mijn kennis van talen kwam goed van pas. Ik was zeer gelukkig toen ik merkte dat het mogelijk was om ook mijn meesterproef een politiek tintje te geven. Met de hulp van professor Ann Bertels kon ik zowaar mijn stage, mijn liefde voor het geschreven woord én de politiek in een document samenvoegen. Het resultaat van deze samensmelting is deze meesterproef. Mijn stage en deze meesterproef waren beide onderdeel van een groter plan: een job krijgen in de politiek. Dit is gelukt: vanaf 1 juli begin ik als universitair medewerker van senator Nele Lijnen.

## 11. Summary of the thesis

For decades, the political party Open Vld has been known as the mainstream Liberal party of Flanders. Open Vld changed its name in 2007 (their name until then was just Vld) but the profile of the party has remained quite stable throughout its history. Like many other liberal parties in Europe, Open Vld wants to defend the rights and privileges of the entrepreneurs and those who work hard. They are in favour of less taxes in order to stimulate the economy. They want to give opportunities to the people but with more opportunities come more responsibilities. Like other liberal parties, they see themselves as the defenders of civil liberties. But since we live a democracy, Open Vld has to attract voters in order to gain power. One way to convince voters is by writing strong, liberal texts. The question at the beginning of my research was thus: how 'liberal' is the discourse of Open Vld? They may see themselves as the defenders of freedom but can we find this in their writings?

When we look at the concise philosophy of Open Vld ([www.openvld.be](http://www.openvld.be)), we notice that a number of words are made bold. These are words like *freedom*, *open*, *initiative*, *responsibility*, etc. These words have central function in this thesis, because we investigate how important they are in the liberal discourse. 'Important' is of course very relative: just because a politician uses a word a lot, does not mean that it is a word typical of a political ideology. Words like *the*, *or*, *if*, *and*, etc. are used very often yet they are not typical of any organization or person. This is why we needed a linguistic and statistical norm, a corpus of words that would function as a standard for our analysis. We chose to collect a number of articles from the Flemish newspaper *De Standaard*, which is known as a quality newspaper. *De Standaard* contains articles about politics, sports, culture, lifestyle, etc. We collected over 200.000 words from newspaper articles from different topics, in order to have a representation of the standard (what's in a name?) language. This corpus thus functions as our reference corpus. After this was done, we collected a number of liberal texts and divided them into three corpora.

The first corpus contains a number of articles published on the main website of the party ([www.openvld.be](http://www.openvld.be)). These texts are mainly informative since they inform the public about actions, thoughts and activities of Open Vld. The second corpus consists of electoral programmes written for the elections between 1994 and 2007. The third corpus is made up of texts written by four senators (Nele Lijnen, Bart Tommelein, Guido De Padt and Alexander De Croo) of Open Vld. Since we were investigating which words were typical in each corpus, we were performing what is called "keyword analyses". Keywords are words that are statistically important compared to a reference corpus. But, because we were also researching texts that are ideologically charged, we were also performing "critical discourse analyses". This analysis of ideological texts with the statistical methods of keyword analysis proved to be very effective, since we were able to expose general patterns which remain invisible if one reads the texts one by one. We decided that, in order to be called an important keyword, the word had to be in the top 50 of the keyword list of the corpus. Consequently, we researched which were the top 50 keywords of each corpus and which of them were 'liberal'. We also checked which 'liberal' words were not on the keyword list, because their absence is also very important. For our analyses we used a computer program called *Wordsmith Tools*.

The results of our research are far from one-sided. The liberal do not write as 'liberal' as one would assume. We found that in the first corpus (with texts from [www.openvld.be](http://www.openvld.be)), a lot of 'liberal' words are not keywords. Words like *individual*, *initiative*, *entrepreneurs*, *private*, *responsibility* and *freedom* were not keywords and those that were keywords were often not in the top 50. Only *Flanders*, *Flemish*, *government*, *open* and *Vld* were in the top 50, so the texts from this corpus could not be identified as very 'liberal'. Our explanation is that the texts in this corpus are more informative than ideological. Still, this result from our analysis was quite surprising. The second corpus (electoral programmes) came closer to our expectations: more 'liberal' words from the concise philosophy were found in the keywords list and more of them could also be found in the top 50 (*open*, *companies*, *government*, *Flanders*, *private*, *Vld* and *freedom*). The electoral programmes are very important because they are the texts that need to convince voters. They contain more 'liberal' keywords but still not all of them. The same remarks we made about the first corpus can be made for the third, which

contained texts written by 4 senators. A lot of 'liberal' keywords were absent and the content of the texts seemed to be rather informative instead of ideological. Each senator had his or her own topics. For example, when we looked at the keyword list we discovered that Nele Lijnen wrote a lot about handicapped people. We could clearly distinguish different topics and to which senator they belonged. This did, however, not mean that the corpus contained a lot of 'liberal' keywords. Only 3 of the selected keywords were in the top 50, namely *open*, *Vld* and *government*. The fact that *open* and *Vld* are ranked so high in each corpus is because the party uses its own name very often. Our hypotheses, which assumed that a lot of 'liberal' keywords would be present in the keyword lists, were thus disproved. This does not mean that our results are useless: they have shown us that Open Vld could write more 'liberal' through the use of certain keywords.