

Woorden tellen

Linguïstische analyse visitatierapporten woningcorporaties Nederland 2015

Thomas Schillemans, Marija Aleksovska, Rixt Riemersma



Inleiding

Nederland kent ruim 350 woningcorporaties die gezamenlijk een kleine 2,5 miljoen 'woongelegenheden' verhuren (Ministerie BZK 2016). Deze corporaties worden sinds 2009 periodiek gevisiteerd binnen een stelsel dat wordt beheerd door de Stichting Visitatie Woningcorporaties Nederland (SVWN). De SVWN heeft daartoe een methodiek ontwikkeld en "ziet toe op de kwaliteit en de onafhankelijke uitvoering van visitaties" (www.visitaties.nl).

Binnen het stelsel waren in 2015 zes verschillende visitatiebureaus actief die gezamenlijk ruim 100 visitaties hebben uitgevoerd. De bureaus werken daarbij met de methodiek 5.0, waarin corporaties worden beoordeeld op vier perspectieven namelijk: 1) Presteren naar Opgaven en Ambities (PnO&A), 2) Presteren volgens Belanghebbenden (PvB), 3) Presteren naar Vermogen (PnV) en 4) Governance. Door de 'binding' aan de methodiek werken de verschillende bureaus binnen eenzelfde kader. Tegelijk zijn er ook wel verschillen; de bureaus hebben een andere achtergrond en een wat ander perspectief op het visiteren. Een rondje langs de websites van de verschillende bureaus is een rondgang langs geheel andersoortige bedrijfspresentaties, waar de bureaus zichzelf ook op geheel eigen wijze presenteren, bijvoorbeeld door te stellen dat "Wij, de mensen van Pentascope, ... de wereld mooier en leuker [willen] maken" of "Ecorys is een internationaal onderzoeks- en beleidsadviesbureau, ontstaan vanuit het Nederlands Economisch Instituut met als missie het economisch onderbouwen van strategische keuzes van publieke en private leiders."

Het feit dat er verschillen zijn is natuurlijk en logisch: wat zou het voor zin hebben om een visitatiestelsel te ontwikkelen waarin verschillende bureaus participeren als die vervolgens precies dezelfde rapporten schrijven? In dit rapport zoomen we nader in op de visitatierapporten over woningcorporaties en op verschillen binnen en tussen visiterende bureaus. We doen dit aan de hand van een linguïstische analyse van 103 visitatierapporten uit 2015. We maken daarbij gebruik van de geautomatiseerde tekstanalyse van LIWC (Linguistic Inquiry & Word Count). Dit programma is ontwikkeld voor de analyse van teksten van individuen maar kan ook worden gebruikt voor de analyse van rapporten. De software analyseert teksten op taalkundige kenmerken, bijvoorbeeld op de vraag hoe positief of negatief een rapport is verwoord of in welke mate er een meer of minder analytische redenering uit naar voren komt. Dit type kenmerken kan van belang zijn voor goede en effectieve visitatierapporten.

De kwaliteit van de tekst is één van de kwaliteitskenmerken van een visitatie, naast andere uiteraard. Visitaties hebben doorgaans tot doel om gevisiteerde organisaties te beoordelen maar vooral om hen op grond daarvan aan te sporen tot verbetering (Schillemans 2009). Politiek-bestuurlijke actoren weten vaak intuïtief dat woorden tellen; het is de toon die de muziek maakt. De gekozen bewoordingen en de aard van de analyse hebben invloed op de vraag hoe de boodschap overkomt en daarmee ook op de vraag wat de ontvanger van de boodschap ermee kan doen. Om die reden is er, zeker de laatste jaren, de nodige aandacht voor de taal van beleid en voor de framing van issues en door beleidsactoren (De Bruijn et al 2012). Tegelijk is deze analyse vaak nog sterk intuïtief en casus-specifiek.

Het aantrekkelijke van de geautomatiseerde tekstanalyse via LIWC is dat teksten kunnen worden geanalyseerd op een groot aantal taalkundig relevante kenmerken dat in de meest uiteenlopende contexten van belang kan zijn. Door tekstuele patronen in visitatierapporten (en andere beleidsteksten) bloot te leggen, door rapporten onderling met elkaar te vergelijken en ook te relateren aan andere kenmerken, kan de analyse van dit deel van publieke beoordelingsprocessen – de verwoording van de boodschap – worden verdiept en gepreciseerd. Daarmee ontstaat langzaam ook meer specifieke kennis waardoor we nog beter – en minder intuïtief – leren hoe beoordelingen steeds effectiever kunnen worden verwoord.

De analyse in dit rapport is een eerste proeve van deze benadering, waarbij we ons richten op alle visitatierapporten uit 2015. Dat is een beperkte basis voor de analyse; de resultaten zullen dus met enige terughoudendheid moeten worden gelezen. Deze analyse kost de nodige tijd vanwege het voorbereiden van de teksten op analyse. Ook is de omvang groot genoeg om statistisch sterk significante verbanden te laten zien; zeker op overkoepelend niveau.

De volgende vragen worden – steeds kort en in deze volgorde – beantwoord in deze rapportage:

1. **Kenmerken:** wat zijn de belangrijkste linguïstische kenmerken van de visitatierapporten 2015?
2. **Externe verschillen:** op welke linguïstische kenmerken verschillen de visitatierapporten wel – en niet – van andere beoordelingsrapporten?
3. **Interne verschillen:** op welke linguïstische kenmerken verschillen de recensies in de start van de visitatierapporten wel – en niet – van de hoofdtaksten van de rapporten?
4. **Onderlinge verschillen:** op welke linguïstische kenmerken verschillen de geanonimiseerde visitatiebureaus wel – en niet – onderling?
5. **Relatie tot tevredenheid:** welke linguïstische kenmerken zijn gerelateerd aan grotere tevredenheid van gevisiteerden met de visitaties?

De analyses in dit rapport zijn in een eerdere vorm gepresenteerd op de masterclasses voor visitatoren van SVWN in november 2016. Op grond van de besprekingen aldaar zijn enkele aanpassingen gedaan en zijn sommige zaken opnieuw geanalyseerd. De gegevens in deze rapportage komen dus niet helemaal overeen met de gegevens die tijdens de masterclasses zijn gepresenteerd.

Over LIWC

Het Linguistic Inquiry and Word Count (LIWC) programma is een geautomatiseerd kwantitatief tekstanalyseprogramma. Het programma scant ingevoerde teksten en geeft vervolgens aan welk percentage van het totaal aantal woorden in de tekst verwijst naar emoties -zoals angst, afschuw, amuseren- en cognitieve processen –zoals analyseren, bedenken en realiseren- en geeft specifieke informatie over de tekst, bijvoorbeeld het percentage voornaamwoorden of het percentage voorzetsels.

Er bestaat een lange academische traditie van systematische tekstanalyse. Het wordt bijvoorbeeld gebruikt in de psychologie, omdat de woorden die we gebruiken mogelijk iets verraden over onszelf. Dit zit al impliciet in de bekende uitdrukking *Freudiaanse verspreking*: de woorden uit de verspreking 'verraden' wat de spreker eigenlijk denkt, wil of voelt. Op dezelfde wijze kunnen teksten ons ook inzicht geven in de verlangens en behoeften van mensen (Parker 2014) of kan een systematische analyse van teksten helpen om leugenaars te ontmaskeren (Newman et al 2003). De toepassing van dit type analyses is zeer divers. De LIWC is oorspronkelijk ontwikkeld om expressie van emoties te analyseren, maar kan ook worden gebruikt om in algemene zin iets te kunnen zeggen over het woordgebruik in een tekst (Zijlstra et al., 2004:272). Een voorbeeld maakt de toepassing van de LIWC duidelijk. Cohn, Mehl en Pennebaker (2004) onderzochten discussiegroepen op het internet na de aanslagen van 11 september 2001 op het World Trade Center in New York. Met behulp van het LIWC-programma signaleerden ze een toename van negatieve emoties direct na de aanslagen. Daarnaast signaleerden ze meer sociale betrokkenheid en ze zagen dat er vaker werd gesproken van 'wij' in plaats van 'ik' dan in de weken voor de aanslagen. Na twee weken leek het woordgebruik weer genormaliseerd te zijn. Deze bevinding is interessant voor volgend trauma- en crisisonderzoek.

Ook in de internationale betrekkingen is er een bloeiende traditie van tekstanalyse. Het is daar relevant, omdat er over vreemde mogendheden – denk aan de Sovjet-Unie vroeger of momenteel IS – vaak heel weinig kennis beschikbaar is. Onduidelijk is dan wat die ander van plan is en hoe de ander gebeurtenissen interpreteert. De officiële boodschappen zijn vaak ook niet betrouwbaar omdat ze een propagandistisch doel dienen. Maar tekstanalytici kunnen op zoek naar de verborgen boodschappen in de teksten, naar de spreekwoordelijke freudiaanse versprekingen en impliciete taalkenmerken, die blootleggen wat de leiders van vreemde mogendheden echt denken. Een moderne vertaling hiervan is bijvoorbeeld het onderzoek van Van Esch (2014) en anderen naar de opvattingen en denkbeelden die politieke leiders laten doorschemeren in hun teksten over de eurocrisis. Ook is er onderzoek dat laat zien dat de taal van politieke leiders verandert voorafgaand aan de escalatie van politiek geweld, zowel nationaal als internationaal (Suedfeld & Tetlock, 2014: 598).

Geautomatiseerde tekstanalyse heeft, als alle dingen in het leven, voor- en nadelen. In het verleden werden teksten altijd manueel geanalyseerd door teams van onderzoekers die daarvoor eerst getraind moesten worden en ten dele dubbel werk moesten verrichten teneinde tot een betrouwbare codering van teksten te komen (Tetlock et al 2014). Tekstanalyse was mensenwerk en dat is tijdrovend en feilbaar. Precies op die punten biedt geautomatiseerde tekstanalyse uitkomst: het kost nog steeds tijd omdat teksten moeten worden klaargemaakt voor analyse maar uiteraard zeer veel minder tijd dan het handmatig doornemen van pagina na pagina. Ook is de betrouwbaarheid veel beter. Computers kunnen vastlopen, gehackt worden of kwijt raken maar ze maken in beginsel geen fouten. Het invoeren van dezelfde tekst leidt stevast tot dezelfde linguïstische analyse. Wanneer de tekstanalyse door mensen gebeurt zit er altijd enige ruis tussen twee codeurs. Geautomatiseerde tekstanalyse heeft dus als voordeel dat het veel betrouwbaarder en veel minder tijdrovend is.

Daar staat tegenover dat mensen teksten beter begrijpen dan computers en daarmee dus preciezer en subtieler kunnen analyseren dan computers. Mensen herkennen ironie en sarcasme, bijvoorbeeld, en snappen het verschil tussen “alle patiënten leven, het gaat geweldig” en “alle patiënten zijn overleden, het gaat geweldig”. Voor de computer is de betekenis van beide identieke bijzinnen gelijk, de menselijke beoordelaar herkent de enorme verschillen uiteraard. In meer wetenschappelijk zin: de externe validiteit van menselijke codeurs is beter, (zie ook Conway et al., 2014), zij zijn veel beter in staat om te begrijpen wat een tekst echt betekent.

Dit issue van externe validiteit betekent dat de resultaten van de analyse in dit rapport niet *at face value* moeten worden geïnterpreteerd en dat het belangrijk is na te gaan of uitspraken ook herkenbaar zijn voor betrokkenen. Tegelijk geldt daar ook weer bij dat de ernst van dit issue varieert per analytische categorie. Zo telt de software bijvoorbeeld het aantal woorden en daarvan mag gerust worden aangenomen dat het betrouwbaar gebeurt. Anderzijds analyseert het ook meer complexe constructies zoals het analytisch gehalte van redeneringen; daar is het risico groter.

We hebben voor de analyse in dit rapport gekozen voor het LIWC programma, omdat het in meerdere talen beschikbaar is en ook een Nederlandse vertaling heeft (Zijlstra et al., 2004). Bovendien analyseert het de teksten op een groot aantal kenmerken dat voor de kwaliteit van beoordelingsrapporten van belang kan zijn.

De analyse

Hoe werkt het LIWC programma? De gebruiker opent een tekst in het programma, het programma scant de tekst en vergelijkt ieder woord in de tekst met woorden in een speciaal woordenboek. De woorden in het woordenboek zijn ingedeeld in verschillende woordcategorieën. Vervolgens wordt het totaal aantal woorden berekend en welk percentage van dit totaal toebehoort aan de diverse categorieën. Het programma berekent dus het percentage werkwoorden, zelfstandig naamwoorden, etc., maar ook welk percentage woorden een positieve of een negatieve emotie heeft. Alle woorden in de tekst worden zo in een of meerdere categorieën geplaatst. Het woord ‘huiten’ wordt bijvoorbeeld geplaatst in de categorie ‘werkwoorden’, maar ook in de hoofdcategorie ‘emotionele processen’ en in een subcategorie daarvan namelijk ‘negatieve emoties’. Het Nederlandse LIWC woordenboek telt 6.568 woorden, verspreid over 66 woordcategorieën (Zijlstra et al., 2004:279).

Nadat een tekst is ingevoerd in het programma wordt deze automatisch geanalyseerd op alle beschikbare categorieën en daar komen getalsmatige scores uit die an sich weinig betekenis hebben. Ze verkrijgen betekenis door middel van vergelijking met andere, vergelijkbare teksten.

We hebben onze analyse beperkt tot een klein deel van de woordcategorieën die binnen het LIWC beschikbaar zijn, omdat ze lang niet allemaal relevant zijn voor de visitatierapporten. We presenteren ze in vier groepen: toonzetting, focus, cognitieve mechanismen en ‘samengestelde variabelen’.

Toonzetting

Het LIWC analyseert teksten op, wat het noemt, positieve en negatieve emoties en daarmee in essentie op de hoeveelheid positieve en negatieve kwalificaties in teksten.

- **Posemo** duidt op 'positieve emoties'. Een voorbeeld zou bijvoorbeeld zijn de zin 'Het is goed om te zien'. In de visitatierapporten wordt bijvoorbeeld gesproken over de 'dappere voorlopersrol' van een corporatie en een andere corporatie 'mag zich gelukkig prijzen'.
- **Negemo** duidt op negatieve emoties, bijvoorbeeld in de zin 'Het is jammer om te zien'. In deze categorie vinden we in de visitatierapporten bijvoorbeeld zinnen als 'dit wordt door partijen jammer gevonden', 'het lijkt alsof er een zekere angst aanwezig is', 'het is bijzonder jammer dat er weinig tot geen vooruitgang is geboekt' en 'het feit dat de corporatie op dit moment weinig investeringsruimte heeft wordt door een partij als zeer spijtig gezien'. Rapporten die hoog scoren op 'negatieve emoties', hebben dus een negatieve toon.

Focus

Een tweede categorie van analyse betreft de focus van het rapport. De focus kan retrospectief, contemporain en prospectief zijn, al naar gelang rapporten meer verwijzen naar het verleden, het heden of de toekomst.

- **Past** wil zeggen een focus op het verleden, bijvoorbeeld in een zin als "Het was een betrouwbare partner".
- **Present** wil zeggen een focus op het heden, bijvoorbeeld in een zin als "Het is een betrouwbare partner".
- **Future** wil zeggen een focus op de toekomst, bijvoorbeeld in een zin als "Het wordt een betrouwbare partner".

Cognitieve mechanismen

Onder cognitieve mechanismen valt een serie aan analysecategorieën die gaan over de wijze waarop in een tekst wordt geredeneerd. De tekstanalyse kan dan dienen als een hulpmiddel om inzicht te krijgen in de psychologische processen -emoties en gedachten- die achter woordkeuzes schuilgaan (Zijlstra et al., 2005: 56). Dit is lastiger te interpreteren maar tegelijk ook potentieel interessanter. We bespreken ze één voor één.

- **Cause** duidt op causale redeneringen in de trant van 'A komt door B'. Een voorbeeld uit één van de rapporten is: "het duurt lang voordat de onderhoudsdienst langskomt en dat komt doordat zij klachten niet mogen bundelen".
- **Insight** duidt op inzichtelijke redeneringen zoals "Het blijkt dat...". Een voorbeeld uit een visitatierapport is de volgende zin: "het blijkt dat er veel woningen gelegen zijn bij het woon-zorg complex".

- **Discrep** duidt op het benoemen van discrepanties zoals 'enerzijds is er waardering voor de eigen visie, anderzijds is een eerdere betrokkenheid van belanghebbenden gewenst'.
- **Tentat** gaat over tentatieve formuleringen zoals 'alhoewel wellicht niet geheel zuiver, heeft dit vermoedelijk ook het oordeel van veel belanghebbenden uiteindelijk medebepaald'.
- **Certain** gaat over stellige beweringen zoals 'de positie van de woningcorporatie is onmiskenbaar van belang in de regio'.
- **Cogmech** is vervolgens de *optelsom* van bovenstaande vijf categorieën.

Samengestelde variabelen

De samengestelde variabelen ten slotte zijn vier categorieën die, met uitzondering van de eerste, zijn samengesteld uit andere onderdelen. Ze hebben als doel om verder te reiken dan de directe taalanalyse en ook tot een inschatting te komen van de houding en de positie die de spreker of auteur van een tekst inneemt. Vier van deze categorieën lichten we eruit omdat ze relevant lijken voor visitaties.

- **Achieve** duidt op prestatiegerichtheid. Woorden als 'behalen', 'slagen', 'concurreren', of 'initiëren' duiden bijvoorbeeld op prestatiegerichtheid.
- **Analytic** is een samengestelde categorie die als doel heeft om de redeneerwijze van de auteur bloot te leggen. Een hoge mate van analytisch denken suggereert "formele, logische en hiërarchische denkpatronen". Een lage mate van analytisch denken suggereert meer vertellende, op persoonlijke ervaring gebaseerde denkpatronen. Het gaat dus om formeel-logisch versus persoonlijk-anekdotisch redeneren. In de analyse levert dit overigens regelmatig wat opvallende resultaten op, met name door de soms opvallende verschillen met de enigszins gerelateerde woordcategorie 'cognitieve mechanismen', die ons verrasten en niet altijd goed te interpreteren waren. Om die reden zal deze categorie in het vervolg vaak niet worden meegenomen in de rapportage.
- **Tone** brengt de positieve en negatieve emoties (zie 'posemo' en negemo') met elkaar in verband en weegt ze tegen elkaar af. Het is bijvoorbeeld voorstelbaar dat een visitatierapport bijvoorbeeld zowel opvallend veel positieve als opvallend veel negatieve uitspraken doet. In dat geval geeft deze categorie aan welke daarvan prevaleert.
- **Clout** ten slotte verwijst naar het zelfverzekerd innemen van een hogere sociale positie. Een zin als "Kom even hier", bijvoorbeeld, heeft dezelfde inhoud maar komt heel anders over – heeft veel meer 'clout' – dan de zin 'wil je alsjeblieft even hier komen?' Deze variabele is ingewikkeld maar ook van acuut belang voor visitatoren. Immers, zij spreken tegen de corporatie met als doel om te helpen in verantwoording en leren. Is het in dat geval beter om vanuit een meer gezagnemende positie – dus met meer clout – te spreken of komt het beter over om op meer sociaal gelijkwaardige basis te communiceren? Er zijn in theorie voor- en nadelen voor beide vormen van sociale positionering te bedenken voor beoordelaars als visitatoren.

In dit rapport worden 103 visitatierapporten uit 2015 geanalyseerd aan de hand van bovenstaande categorieën. Dat leidt tot een aantal heel duidelijke conclusies die, niettemin, omzichtig moeten worden behandeld, in verband met de genoemde intrinsieke beperkingen van de geautomatiseerde tekstanalyse. Zo is een beperking van het LIWC programma dat het geen inzicht geeft in de context waarin woorden worden gebruikt, hetgeen tot enige vertekening kan leiden in de resultaten (Zijlstra et al., 2004: 280). Een uitspraak als 'de woningcorporatie hoeft niet bang te zijn voor een slecht resultaat' zal door de lezer (terecht!) als positief (posemo) worden geïdentificeerd, maar wordt door LIWC begrepen als negatief (negemo) vanwege de woordkeuzes 'slecht' en 'bang'. Het blijft daarom belangrijk de resultaten van de tekstanalyse zorgvuldig te wegen en waar nodig aan te vullen met ander onderzoek en directe lezing van teksten.

De resultaten

In het resultaten-deel bespreken we de vijf deelvragen die voor deze rapportage zijn gesteld, t.w.:

1. **Kenmerken:** wat zijn de belangrijkste linguïstische kenmerken van de visitatierapporten 2015?
2. **Externe verschillen:** op welke linguïstische kenmerken verschillen de visitatierapporten wel – en niet – van andere beoordelingsrapporten?
3. **Interne verschillen:** op welke linguïstische kenmerken verschillen de recensies in de start van de visitatierapporten wel – en niet – van de hoofdteksten van de rapporten?
4. **Onderlinge verschillen:** op welke linguïstische kenmerken verschillen de geanonimiseerde visitatiebureaus wel – en niet – onderling?
5. **Relatie tot tevredenheid:** welke linguïstische kenmerken zijn gerelateerd aan grotere tevredenheid van gevisiteerden met de visitaties?

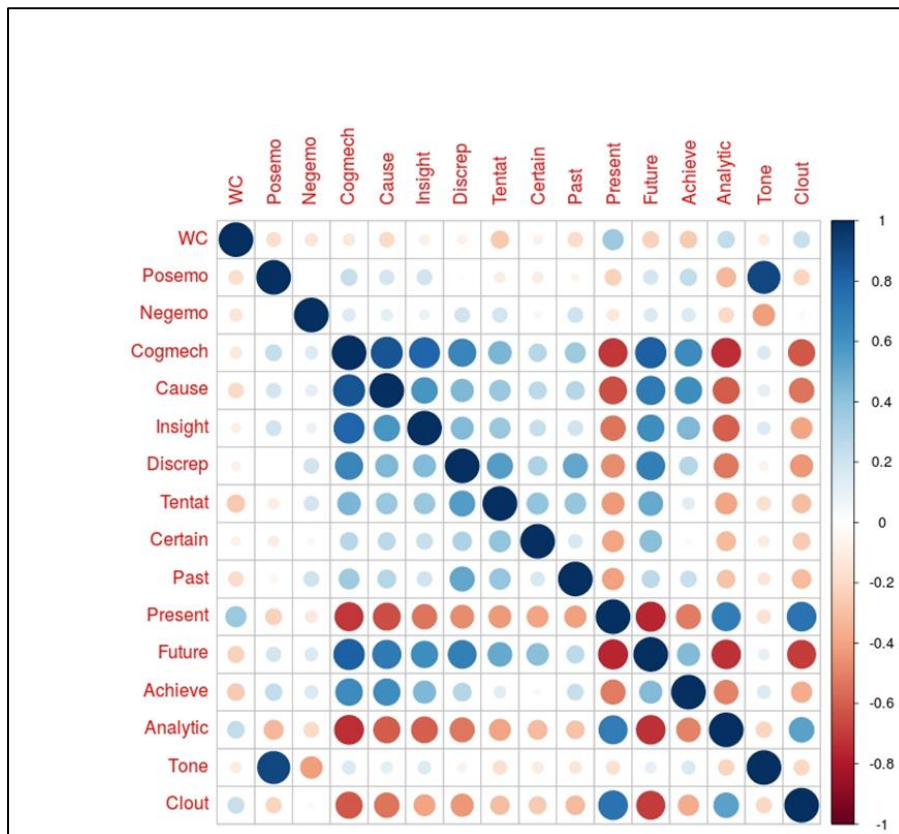
Kenmerken: gezaghebbende snapshots versus historische redeneringen

De analyse van de visitatierapporten levert een lange reeks getallen op die, zonder verdere referentie, weinig betekenis hebben. Wanneer we de scores vergelijken met de gemiddelde scores bij toepassing van LIWC in andere onderzoeken, zien we dat de waarden bij de visitatierapporten op bijna alle vlakken lager zijn dan bij de gemiddelden (zie voor de gemiddelden: <https://liwc.wpengine.com/compare-dictionaries/>). Dat zegt vervolgens niet zo veel, omdat niet duidelijk is waar die gemiddelden precies op zijn gebaseerd al is wel zeker dat het dan eerder over teksten van personen gaat dan over rapporten van groepen personen. Het enige punt waar de waarden een beetje hoger zijn is bij de tijdsfocus op heden, verleden én toekomst. Maar het is moeilijk beredeneren hoe dat zou kunnen en ook wat het zou kunnen betekenen dat de rapporten een nadrukkelijker tijdsbepaling kennen dan andere teksten die met LIWC zijn geanalyseerd.

Interessanter is het om te kijken binnen de rapporten. Zit er structuur en logica in? Hangen bepaalde kenmerken sterk samen?

Het antwoord op die vraag staat (verscholen) in figuur 1 en zal luiden dat er grosso modo twee typen rapporten lijken te bestaan.

Figuur 1 Correlaties woordcategorieën linguïstische analyse



Figuur 1 correleert de verschillende analysecategorieën met elkaar voor alle rapporten gezamenlijk. De blauwe cijfers duiden op positieve correlaties tussen zaken die dus sterk samenhangen. De rode getallen duiden op negatieve correlaties tussen zaken die dus niet vaak gezamenlijk voorkomen. Hoe hoger (en vetter gedrukt) het getal is, hoe sterker de twee zaken samenhangen.

Oppervlakkige beschouwing laat zien dat er heel duidelijke patronen in de teksten te herkennen zijn: er zijn een heleboel heel sterke correlaties. Een deel daarvan is echter enigszins artificieel. Op de diagonale as zijn bijvoorbeeld alle waarden '1', simpelweg omdat een factor dan met zichzelf wordt gecorreleerd. Ook zijn er een aantal logische correlaties. Om te beginnen tussen het overkoepelende begrip cognitieve mechanismen ('cogmech') en het rijtje 'cause' tot 'certain'. Dat is logisch, omdat 'cogmech' een samengestelde waarde is op basis van die andere. Dat kunnen we dus verder buiten beschouwing laten. Hetzelfde geldt, om dezelfde reden, voor de sterke samenhangen tussen 'tone', 'posemo' en 'negemo'.

Wanneer we de bevindingen op deze wijze versimpelen, blijven er een aantal opvallende patronen over.

Om te beginnen geldt voor de cognitieve mechanismen – dus de verzamelcategorie voor zaken als causaal denken of tentatief redeneren – dat die als geheel maar ook op onderdelen sterk positief samenhangen met een aantal andere factoren. Er is om te beginnen een heel sterke (0,82) correlatie met prospectief redeneren. In teksten in de toekomstige tijd wordt nadrukkelijk gebruik gemaakt van de verschillende cognitieve mechanismen.

Dat is op zich wellicht ook wel logisch, omdat voorstelbaar is dat het in toekomstgerichte redeneringen over mogelijke verbeteringen gaat. In dezelfde lijn lijkt ook begrijpelijk dat er ook een (iets minder) sterk positieve relatie is tussen alle vormen van cognitieve mechanismen en prestatie-focus. In tekstpassages over het al dan niet behalen van bepaalde doelen of prestaties wordt al snel gesproken in causale termen bijvoorbeeld, die het al dan niet behalen van die doelen verklaren.

Tegelijk zijn er rond de cognitieve mechanismen ook verschillende heel sterk negatieve correlaties. In de eerste plaats is er een sterk negatief verband met een contemporaine focus. In zinnen in de tegenwoordige tijd gebruiken de visitatiebureaus klaarblijkelijk meer beschrijvende dan cognitieve redeneringen. In de tweede plaats is er een sterk negatief verband tussen cognitieve mechanismen en 'clout'; het innemen van een hogere sociale positie. Dit suggereert dat in teksten waar een steviger sociale positie wordt ingenomen door visiterende bureaus, de redeneringen minder cognitief (dus minder causaal, etc.) zijn en vermoedelijk meer in een stellende vorm ("Het is verstandig"... "het is nodig...").

Een derde observatie is dat er een sterk negatief verband bestaat tussen een contemporaine en een toekomstgerichte focus (-0,75). Dat is op zich niet opvallend, daar de drie categorieën (heden, verleden, toekomst) elkaar op zinsniveau uitsluiten. Je kunt geen goede zin maken die halverwege van verleden naar toekomstige tijd verspringt (al gebeurt dat soms per ongeluk). Wat wel opvallend kan worden genoemd, is dat veruit het grootste verschil bestaat tussen heden en toekomst; en niet verleden. Sterker nog, een op het verleden gerichte focus hangt licht positief samen toekomstgerichte focus. En vervolgens hangen die twee groepen – dus contemporain versus historisch – weer op verschillende wijzen samen met cognitieve mechanismen enerzijds en 'clout' anderzijds.

Dit alles suggereert, en we schrijven het heel voorzichtig, dat er twee typen rapporten lijken te zijn:

Gezaghebbende snapshots, met een sterke focus op het 'heden', relatief veel 'clout' en geringe nadruk op cognitieve mechanismen. Met snapshot wordt bedoeld dat de focus nadrukkelijker in het heden, in het hier en nu, ligt dan bij de andere rapporten.

Historische redeneringen, met een focus op verleden en toekomst, relatief weinig 'clout' en geringe nadruk op cognitieve mechanismen. Deze rapporten situeren zich nadrukkelijker in historisch perspectief.

Externe verschillen: visitaties als stevige, analyserende foto

Een tweede manier om betekenis te geven aan de resultaten van de tekstanalyse is door deze te vergelijken met de resultaten van andere beoordelingsinstanties. Hoe verhouden de op verantwoording en leren gerichte visitatierapporten zich tot rapportages van andere beoordelende instanties in de publieke sector?

Voor de vergelijking hebben we gebruik gemaakt van een *most different* vergelijking. We hebben de visitatierapporten vergeleken met performance-audits van een rekenkamer. Dat klinkt in eerste instantie misschien verwarrend maar is dat (hopelijk) in tweede instantie niet. Visitaties en audits zijn beiden onderdeel van stelsels van verantwoording en beoordeling in de publieke sector. Voor alle vormen van verantwoording en beoordeling geldt doorgaans dat zij zich tegelijkertijd richten op een controle-kant (voldoet de organisatie aan bepaalde maatstaven / eisen?) én op een verbeteringskant. Dat zie je duidelijk terug in de missies van betreffende organisaties. Zo formuleert de Algemene Rekenkamer als missie om op een viertal punten “te toetsen en te verbeteren” en heeft de Audit Dienst Rijk als missie dat het “Rijk is in audit en advies”. Ook in het visitatiestelsel voor woningcorporaties gaat het om “verantwoording afleggen en ... prestaties verbeteren.”

In de publieke sector zijn er beoordelaars die van huis uit meer aan de controlekant staan (auditors en rekenkamers) maar ook beoordelaars die meer aan de reflectie- en verbeteringskant zitten (zoals visitaties). Dat laat echter onverlet dat zij allen vormen van controle met verbetering combineren, allemaal willen bijdragen aan verbetering en de doorwerking van hun beoordelingen belangrijk vinden.

Tabel 1 laat zien hoe de gemiddelden voor de visitatierapporten zich verhouden tot de gemiddelden voor de performance audits van een rekenkamer.

Tabel 1 : Vergelijking visitatierapporten met Performance Audits

	Visitatie rapporten	Performance audits
Posemo	1.14	1.52
Negemo	0.34	0.42
Cogmech	3.21	4.51
Cause	0.65	1.13
Insight	1.25	1.61
Discrep	0.51	0.76
Tentat	0.64	0.88
Certain	0.42	0.56
Past	1.95	2.42
Present	8.80	3.30
Future	0.22	0.59
Achieve	0.59	0.85
(Analytical thinking)	27.13	22.41
Clout	2.61	0.14
Emotional tone	0.80	1.09
N	102	25

De analyse laat zien dat er heel duidelijke verschillen zijn wat betreft het “profiel” van de twee soorten rapporten. De blauwe kleur duidt op zaken waar de visitatierapporten ‘hoger’ scoren, de rode kleur juist weer op zaken waar de performance audits hogere waarden hebben.

De performance audits zijn.

- Sterker geschreven in termen van positieve emoties
- Maken veel nadrukkelijker gebruik van de verschillende cognitieve mechanismen
- Hebben een meer nadrukkelijke focus op prestaties
- Zijn nadrukkelijker geschreven in verleden tijd en, daarmee dus ook, nadrukkelijker retrospectief dan de visitaties, die kennelijk een meer contemporain, constaterende stijl hebben.

De visitatierapporten daarentegen zijn weer

- Sterker in analytische zin; scoren veel hoger op de waarde van analytisch denken
- Sterker ook in termen van “clout”, dat wil zeggen nadrukkelijker geschreven vanuit een stevige, sociaal hogere positie
- En vooral zijn ze zeer veel nadrukkelijker geschreven in tegenwoordige tijd.

Alles bijeen genomen lijkt het beeld te zijn dat de visitatierapporten, vergeleken met de performance audit rapporten, meer tijdsgebonden zijn, als stevige, analyserende foto's. Dit schrijven we voorzichtig, want de conclusie is gebaseerd op een beperkte hoeveelheid vergelijkingsmateriaal. Ook zou het een product van het relatief gereguleerde format van de visitatiemethodiek kunnen zijn. Desalniettemin is opvallend hoe groot de verschillen op sommige punten zijn.

Interne verschillen: meer uitgesproken, zekerder en meer toekomstgerichte recensies

De derde vraag voor de analyse is: hoe verhoudt de recensie in het visitatierapport zich linguïstisch tot het hoofdrapport zelf?

De recensie is een relatief nieuw onderdeel van de visitatiemethodiek. De recensie is bedoeld om “enerzijds kleur te geven aan de cijfermatige beoordelingen en anderzijds de boodschap van de commissie aan de corporatie helder en inspirerend te verwoorden” (SVWN 2014: 16). De recensie bevat de integrale beoordeling van de visitatiecommissie naast de meer op onderdelen weergegeven beoordeling in het rapport zelf.

De recensie lijkt op zijn minst enigszins op executive summaries bij andersoortige rapporten. De recensie is echt bedoeld als reflectie van de commissie op de corporatie. Het idee is dat de boodschap dezelfde moet zijn als die in het meer uitgebreide rapport maar wel op een andere manier wordt verwoord. Dit roept de vraag op of de tekst, los van de inhoud, ook linguïstisch gelijk- of juist andersoortig zou moeten zijn.

Tabel 2 laat zien hoe de recensies zich verhouden tot de hoofdteksten zelf.

Tabel 2: Vergelijking recensies met rapport zelf

	Recensies	Hoofdtteksten
Posemo	1.49	1.10
Negemo	0.44	0.32
Cogmech	3.71	3.17
Cause	0.55	0.65
Insight	1.67	1.21
Discrep	0.77	0.49
Tentat	0.70	0.64
Certain	0.63	0.40
Past	2.25	1.92
Present	9.46	8.73
Future	0.34	0.21
Achieve	0.60	0.59
Analytic	27.24	27.13
Clout	2.90	2.59
Tone	1.06	0.77
N	103	103

Tabel 2 is gemakkelijker te interpreteren dan figuur 1. Tabel 2 laat zien dat de recensies op bijna alle punten hogere waarden hebben dan de hoofdtteksten van de visitatierapporten.

Alle blauwgedrukte waarden zijn significant sterker voor de recensies dan voor de hoofdtteksten. Dat geeft aan dat de recensies op drie van de vier hoofdcategorieën nadrukkelijk hogere waarden hebben. Dat geldt in de eerste plaats voor de *toonsetting*. Recensies gebruiken zowel meer positieve als meer negatieve kwalificaties. Dit zou in de lijn van de bedoeling van dit onderdeel kunnen zijn; het is immers de bedoeling om een integraal en reflectief oordeel te vellen. Hetzelfde geldt voor de focus in de tijd die op alle drie de punten sterker is voor de recensies (heden, verleden én toekomst). Ten derde geldt het ook voor bijna alle cognitieve mechanismen. Er is slechts één punt waar de rapporten zelf wat sterker zijn en dat betreft causale redeneringen (als enige roodgedrukt).

Het overkoepelende beeld is dat de recensies bijna op alle punten meer uitgesproken zijn dan de onderliggende visitatierapporten. Dat zou wel in de lijn der verwachtingen liggen. Daarbinnen zijn er drie factoren waar de verschillen extra groot zijn. Dit suggereert dat recensies:

- Meer toekomstgericht zijn – dus zich meer richten op de vraag hoe de woningcorporaties in de toekomst (verder) kunnen verbeteren
- Stelliger zijn – dus met meer zekerheid zijn verwoord
- Meer discrepanties onderscheiden – dus vaker in termen van enerzijds, anderzijds en mitsen en maren spreken.

Onderlinge verschillen: interessante detailverschillen

De vierde vraag in de analyse betreft de vraag of de rapporten van de verschillende bureaus onderling verschillen.

Om die vraag te kunnen beantwoorden zijn de gemiddelde waarden voor de zes bureaus die in 2015 hebben gevisiteerd vergeleken met de gemiddelden steeds van de overige bureaus. We hebben de bureaus geanonimiseerd en in niet-alfabetische volgorde geplaatst en hernoemd tot B1 t/m B6.

Bij de analyse dient vooraf de kanttekening te worden gemaakt dat sommige bureaus in 2015 veel visitaties hebben uitgevoerd (Raeflex 38; Ecorys 35), andere een beperkt aantal (Procorp 11; Cognitum 10) en sommige zelfs een gering aantal (EY 6; Pentascope 2). De resultaten voor bureaus met meer visitatierapporten in 2015 zijn om die reden betekenisvoller dan die voor de bureaus met een gering aantal visitaties. Met name de resultaten voor Pentascope zijn kwetsbaar.

Tabel 3 hieronder laat weer de verschillen zien, waarbij opnieuw de roodgedrukte waarden duiden op significant lagere scores dan het gemiddelde en blauwgedrukte waarden op significant hogere scores.

Tabel 3: geanonimiseerde vergelijking visitatiebureaus

	B1	B2	B3	B4	B5	B6
Posemo	0.94	0.93	1.31	1.24	1.39	1.09
Negemo	0.35	0.30	0.33	0.37	0.25	0.36
Cogmech	3.18	3.30	3.05	3.48	3.80	3.32
Cause	0.64	0.71	0.60	0.65	0.84	0.62
Insight	1.18	1.20	1.24	1.41	1.42	1.33
Discrep	0.56	0.57	0.44	0.56	0.46	0.61
Tentat	0.64	0.80	0.61	0.78	0.36	0.70
Certain	0.38	0.55	0.41	0.50	0.31	0.35
Past	2.10	2.17	1.71	2.03	1.95	2.27
Present	9.10	8.36	8.61	8.75	8.76	9.80
Future	0.22	0.24	0.21	0.22	0.26	0.29
Achieve	0.61	0.53	0.58	0.51	0.90	0.48
Analytic	27.72	26.67	27.18	27.05	25.15	24.29
Clout	2.37	2.53	2.76	2.86	2.43	3.65
Tone	0.59	0.64	0.98	0.87	1.14	0.72

Tabel 3 hierboven laat alles bijeen zien dat ieder bureau een wat eigen taalkundig profiel heeft en, tegelijk, dat de onderlinge verschillen niet heel erg groot zijn. Dat lijkt in overeenstemming te zijn met het stelsel, waarin verschillende bureaus op basis van eigen accenten visiteren binnen een bestaand kader.

Als we op grond van deze bevindingen komen tot een voorzichtige typering van de bureaus ontstaan de volgende profielen.

B1 lijkt de meest analytische van de bureaus te zijn. Het kent relatief hoge waarden voor analytisch denken, voor de retrospectieve (en ook contemporaine) focus als ook voor het onderkennen van discrepanties. Parallel is het ook het minst emotionele bureau en drukt het vooral veel minder positieve kwalificaties uit dan in de andere visitatierapporten gebeurt.

B2 heeft – en dat is op geen enkele manier kwalificerend bedoeld – wellicht het meest gemiddelde profiel. De gemiddelde waarden voor de rapporten van dit bureau liggen vaak dicht bij de waarden van het gemiddelde überhaupt. Het valt wel op door een wat sterkere retrospectieve focus.

Ook de waarden voor **B3** liggen op veel punten dicht bij het gemiddelde. Daarbij is wel opvallend dat de redeneringen in de rapporten van dit bureau het minst gebruik maken van de cognitieve mechanismen en ook wat minder retrospectief zijn gesteld. Anderszins is dit het bureau dat het sterkst positieve kwalificaties verwoordt.

B4 en **B5** zouden de cognitieve bureaus kunnen worden genoemd; omdat zij beiden wat hogere waarden hebben voor cognitieve mechanismen. Bij die overeenkomst zijn er wel enkele duidelijke verschillen. Voor B4 geldt dat de zinsconstructies relatief vaak tentatief, zeker én inzichtelijk zijn. Dit betekent mogelijk een meer *constaterende* stijl van schrijven en redeneren, met zowel relatief veel tentatieve zinnen als relatief veel zekere zinnen. B4 lijkt daarmee steeds te willen verhelderen welke mate van zekerheid het heeft ten aanzien van de observaties. Voor B5 geldt in gelijke mate de focus op inzichtelijke redeneringen. Tegelijk valt dit bureau op door veel nadrukkelijker in termen van prestaties te schrijven en het sterkst causaal te redeneren. Het lijkt daarom een wat meer op doelbereiking gericht bureau te zijn (schrijven we heel voorzichtig).

B6 ten slotte valt op één punt op: het is het bureau dat het meest nadrukkelijk vanuit een wat hogere sociale positie ('clout') schrijft. Op de andere punten liggen de gemiddelde waarden overigens dicht bij de andere waarden.

Relatie tot tevredenheid: emotie en zekerheid

De vijfde en laatste deelvraag betreft de relatie tussen linguïstische kenmerken van de teksten en de tevredenheid met de evaluaties van gevisiteerde corporaties.

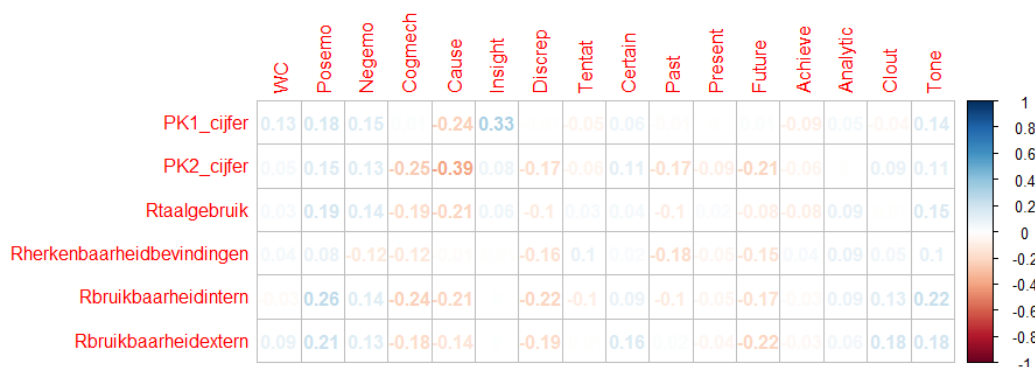
Na iedere visitatie ontvangen de gevisiteerde bureaus in het kader van de kwaliteitszorg een uitnodiging om de visitatie en het visitatieproces te beoordelen. In deze tevredenheidsmeting worden een aantal verschillende voor ons hier relevante vragen gesteld. Er zijn allereerst twee vragen die meer op de visitatie zelf betrekking hebben. In één vraag geven corporaties aan wat ze vinden van de geleverde prestatie door het bureau in relatie tot de prijs (PK1_cijfer in figuur 2). In de andere vraag geven corporaties aan hoe ze het resultaat van de visitatie als geheel beoordelen, ook in relatie tot de kosten (PK2_cijfer). Daarnaast zijn er een aantal vragen over het

visitatierapport zelf; namelijk over de externe bruikbaarheid, de herkenbaarheid van de bevindingen, de interne bruikbaarheid van het rapport als ook het taalgebruik.

Hoe, nu, verhouden de linguïstische kenmerken van de rapporten zich tot de tevredenheid van gevisiteerden als één van meer mogelijke kwaliteitsmaatstaven? Of, anders en wat provocatiever gesteld, hoe moet een rapport eruit zien waar de corporatie meer tevreden mee is?

Om die vraag te beantwoorden zijn de woordcategorieën van de verschillende rapporten gematcht met de tevredenheidsscores op zes dimensies. Deze analyse werd beperkt door het feit dat er helaas slechts voor 62 rapporten – dus net iets meer dan de helft – een tevredenheidsenquête beschikbaar was. Dit beperkt de resultaten en leidt er vooral toe dat de resultaten vaak niet heel sterk zijn. Er zijn wel significante verbanden, dat blijkt uit figuur 2, maar van een beperkte statistische sterkte.

Figuur 2: linguïstische kenmerken irt aspecten van tevredenheid



De kleurstelling van figuur 1 – vergelijk maar met figuur 2 – is flets en suggereert niet al te sterke verbanden. Er is een veel grotere sample nodig om sterkere verbanden te vinden.

Dat gezegd hebbende komen er wel enkele patronen naar voren.

Positieve evaluaties van visitaties, op alle zes de onderdelen, hangen samen met enkele kenmerken van de visitatierapporten.

- Positieve emoties in rapporten zijn verbonden met positieve oordelen in tevredenheidsenquêtes. De gemakkelijke verklaring daarvan zou kunnen zijn dat corporaties het leuker vinden om positief te worden beoordeeld en dat positieve emoties samenhangen met positieve scores.
- Echter, ook negatieve emoties zijn op de meeste punten significant gerelateerd aan meer positieve scores in de evaluaties. Dat zou moeilijker te rijmen zijn met de verklaring hierboven. Overkoepelend geldt überhaupt dat meer emotioneel geladen rapporten samenhangen met positiever

evaluatie resultaten. Dit suggereert eerder dat meer emotioneel, dus ook meer positief én negatief uitgesproken rapporten, iets beter worden gewaardeerd. Dat zou ook in lijn zijn met communicatiewetenschappelijke inzichten.

- Tot slot valt op dat er een positief verband is tussen de zekerheid waarin rapporten zijn gesteld en de tevredenheidsresultaten.

De positieve correlaties suggereren dus – en opnieuw, we zeggen dit alleen voorzichtig – dat meer emotionele en meer zekere rapporten samengaan met meer tevreden corporaties.

Aan de andere kant zijn er ook negatieve correlaties. De visitaties worden minder beoordeeld als de rapporten

- Overkoepelend veel cognitieve mechanismen (zoals causale redeneringen, tentatieve formuleringen, etc.) bevatten.
- Meer in het bijzonder geldt dit specifiek bij causale redeneringen – als onderdeel van de sub categorie cognitieve mechanismen – die op zich sterk negatief samenhangen met tevredenheid
- Toekomstgerichtheid. Rapporten die relatief sterk in toekomstige tijd zijn geschreven hangen samen met zwakkere scores in de tevredenheidsmeting.

Het is de vraag hoe deze negatieve verbanden kunnen worden geduid. Een mogelijke verklaring zou kunnen zijn dat er hier ook een samenhang is met de beoordeling. Het zou zo kunnen zijn dat corporaties die meer kritisch zijn beoordeeld ook negatiever zijn in hun antwoorden in de tevredenheidsenquête. En het is ook voorstelbaar dat meer kritische visitatierapporten meer causale redeneringen gebruiken (om te verklaren waarom dingen niet goed gaan) en tegelijk ook meer in toekomstige tijd zijn gesteld (om aan te geven dat en hoe die zaken kunnen worden verbeterd).

Tot slot is er nog één op zichzelf staand positief verband. Het is moeilijk te duiden maar wel relatief sterk; vandaar dat we het hier benoemen. Er is een duidelijk positief verband tussen de tevredenheid van corporaties met de geleverde prestaties van het bureau ook in relatie tot de prijs en het vóórkomen van inzichtelijke redeneringen.

Tot besluit: de bevindingen hier zijn uitdrukkelijk eerste resultaten. Er is een grotere sample nodig om tot meer zekere bevindingen te komen. Ook zijn er andere – mogelijk betere – kwaliteitsmaten voorstelbaar voor visitaties. Voor nu echter kan één beeld wellicht blijven hangen als werkhypothese die getoetst (en daarmee mogelijk bijgesteld of verworpen) kan worden: meer emotionele en meer zekere visitatierapporten lijken positiever te worden ontvangen.

Tot besluit

Deze rapportage is een eerste proeve van toepassing van de geautomatiseerde tekstanalyse via LIWC op de visitatierapporten van woningcorporaties. Het is een verkennend onderzoek geweest, waarin we hebben trachten te achterhalen of deze methodiek op zinvolle wijze kan worden toegepast op beoordelingsrapporten en vervolgens hebben verkend welke bevindingen dat oplevert.

Als test beschouwen we het zelf – maar de lezer mag daar uiteraard anders over denken – als geslaagd. De tekstanalyse werkt, laat stabiele patronen, verschillen en overeenkomsten zien, en deze kunnen ook worden gebruikt voor betekenisvolle gesprekken over visitaties met betrokkenen. Wat dat betreft smaakt de verkenning naar meer.

Inhoudelijk hebben we een aantal observaties kunnen doen. Dat doen we met veel terughoudendheid: het zijn vooralsnog constatering en eerdere hypothesen dan conclusies. Voorzien van deze mitsen en maren zijn onze belangrijkste observaties op de vijf gestelde vragen:

Kenmerken van rapporten. De analyse suggereert dat er twee profielen van rapporten zijn. Aan de ene kant zijn er rapporten die we voor nu *gezaghebbende snapshots* noemen. Het zijn rapporten die meer in tegenwoordige tijd zijn gesteld en stilliger en vanuit meer gezaghebbende positie ('clout') zijn geschreven. Aan de andere kant zijn er rapporten met meer *historische redeneringen*, die zowel meer in verleden als in toekomstige tijd spreken en waarin cognitieve mechanismen zoals causale redeneringen meer voorkomen.

Externe verschillen rapporten ten opzichte van de rapporten van een andere beoordelaar. Vergeleken met de performance audits van een rekenkamer valt op dat de visitatierapporten nadrukkelijker vanuit een gezaghebbende positie zijn geschreven, nadrukkelijker in tegenwoordige (constaterende) tijd staan en wat meer analytisch van aard zijn.

Interne verschillen binnen de rapporten betreffen de verhouding tussen recensie en hoofdrapport. Analyse laat zien dat de recensies meer uitgesproken zijn op alle punten, hetgeen in overeenstemming is met de intenties. Meer in het bijzonder zijn ze zekerder/stilliger, geven ze vaker discrepanties aan en zijn ze ook toekomstgerichter.

Onderlinge verschillen betreft de verschillen tussen de bureaus. Analyse laat zien dat de bureaus op details interessant van elkaar verschillen. Bureaus lijken, binnen de kaders van de methodiek, eigen accenten te kunnen zetten.

Relatie tot tevredenheid is vooralsnog weliswaar te leggen, maar, gelet op de relatief kleinere sample voor dit deel van de analyse, wel met extra omzichtigheid. De belangrijkste werkhypothese die op grond van het materiaal valt te trekken is dat er een verband zou kunnen zijn tussen een meer emotionele toonzetting, in combinatie met meer zekere formuleringen, en tevredenheid bij gevisiteerde corporaties.

Voor de toekomst zou vervolgonderzoek de hier gevonden patronen verder kunnen verhelderen en verdiepen.

Allereerst zou het op het niveau van de rapporten zelf interessant zijn om de geautomatiseerde tekstanalyse te combineren met manuele tekstanalyse. Hoe komt het nu dat bepaalde rapporten, en bepaalde bureaus, op specifieke wijze uit de analyse komen? Op welke wijzen is dat te herkennen in de teksten zelf? Dit is een tijdrovende optie die wel nodig zou zijn om binnen de tekst echt te begrijpen hoe scores tot stand komen en ook inzicht kunnen geven in de vraag of en hoe rapporten anders geformuleerd zouden kunnen worden.

Vervolgens zou het meer op stelselniveau interessant kunnen zijn om de teksten te vergelijken in de loop der tijd. De visitatiemethodiek verandert van tijd tot tijd; levert dit ook op linguïstisch niveau andere patronen op?

Tot slot zou het, zeker in academische zin, interessant zijn om het laatste deel van de analyse, te weten de (cor)relatie van tekstkenmerken met tevredenheidsscores, fors uit te breiden. Om te beginnen zou het dan goed zijn om de analyse te doen op grond van een grotere matching tussen tevredenheidsenquêtes en visitatierapporten. De onderliggende vraag betreft de relatie tussen tekst en kwaliteit. Om dat verder te onderzoeken zou het goed zijn om de analyse van teksten uit te breiden naar andere beoordelingsrapporten in de publieke sector. Ook zijn andere, misschien ook betere, maatstaven voor 'kwaliteit' denkbaar dan tevredenheidsenquêtes. Te denken valt bijvoorbeeld aan categorisering van reacties op rapporten door gevisiteerde corporaties of de doorwerking van de rapporten in de tijd.

Referenties

- Bruijn, H. de, T. Schillemans en M. van der Steen. (2012). Framing framing. Betekenisgeving en besluitvorming over beleid. *Bestuurskunde*, 21(4):2-8.
- Cohn, M.A., Mehl, M.R., Pennebaker, J.W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science*, 15:687-693
- Conway, L. G., Conway, K. R., Gornick, L. J., & Houck, S. C. (2014). Automated integrative complexity. *Political Psychology*, 35(5):603-624.
- Esch, F. van (2014). *A matter of personality? Stability and change in EU leaders' beliefs during the Euro crisis*. In: D. Alexander & J. Lewis (eds.), *Making public policy decisions: Expertise, skills, and experience*, Londen: Routledge, pp. 53-76.
- Ministerie van BZK. (2016) *Cijfers over Wonen en Bouwen 2016*. Den Haag: BZK.
- Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and social psychology bulletin*, 29(5), 665-675.
- Parker, I. (2014). *Discourse Dynamics (Psychology Revivals): Critical Analysis for Social and Individual Psychology*. London: Routledge.
- Schillemans, T. (2009). *Visitaties als vorm van verantwoording*. In: Bovens, M. en T. Schillemans, 2009, *Handboek publieke verantwoording*. Den Haag: Lemma, p. 161 – 180.
- Suedfeld, P., & Tetlock, P. E. (2014). Integrative complexity at forty: Steps toward resolving the scoring dilemma. *Political Psychology*, 35(5):597-601.
- SVWN (2014), *Methodiek maatschappelijke visitatie woningcorporaties 5.0*. Zeist: SVWN.
- Tetlock, P. E., Metz, S. E., Scott, S. E., & Suedfeld, P. (2014). Integrative complexity coding raises integratively complex issues. *Political Psychology*, 35(5), 625-634.
- Zijlstra, H., Meerveld van, T., Middendorp van, H., Pennebaker, J.W., Geenen, R. (2004). De Nederlandse versie van de 'Linguistic Inquiry and Word Count' (LIWC). *Gedrag & Gezondheid*, 32(4):271-281.
- Zijlstra, H., Middendorp van, H., Meerveld van, T., Geenen, R. (2005). Validiteit van de Nederlandse versie van de Linguistic Inquiry and Word Count (LIWC). *Nederlands Tijdschrift voor de Psychologie*, 60:55-63.

Bijlage

Om de teksten geschikt te maken voor analyse zijn de volgende zaken uit de rapporten verwijderd:

Organisatie: Stichting Visitatie Woningcorporaties Nederland

- Voorkant
- Voorwoord visitatiecommissie (indien aanwezig)
- Inhoudsopgave
- Leeswijzer (indien aanwezig)
- Werkwijze (indien aanwezig)
- Bestuurlijke reactie □ reactie van RvC en directie corporatie op de visitatie, maakt daarmee dus geen deel uit van de visitatie zelf.
- Bronvermelding
- Bijlagen. De bijlagen bevatten:
 - Curricula Vitae van de visitatiecommissie
 - Position paper van de woningcorporatie in kwestie
 - Factsheet maatschappelijke prestaties
 - Checklist governancecode
 - Meetschaal
 - Lijst gerealiseerde prestaties, prestatietabellen of overzichten van prestaties (in samenwerking gemaakt met de woningcorporatie)
 - Lijst geïnterviewde personen (intern + extern)
 - Bronnenlijst
 - Onafhankelijkheidsverklaringen visitatiecommissieleden
 - Werkwijze*

*In sommige gevallen was de werkwijze ingebed in de introductie waardoor deze niet verwijderd kon worden.

Colofon

Dit onderzoek is uitgevoerd in overleg met de Stichting Visitatie Woningcorporaties Nederland.

Het onderzoek is onderdeel van het onderzoeksprogramma *Calibrating Public Accountability* dat dankzij een NWO-vidi beurs 2015 kan worden uitgevoerd. Het onderzoeksprogramma richt zich op de vraag hoe vormen van publieke verantwoording in de publieke sector kunnen worden verbeterd om publieke organisaties echt te helpen hun taken goed te kunnen uitvoeren.

Zie voor het project: <https://accountablegovernance.sites.uu.nl>

Thomas **Schillemans** is universitair hoofddocent bestuurskunde aan de Utrechtse school voor Bestuurs- en Organisatiewetenschap, Universiteit Utrecht.

Marija **Aleksovska** is assistent in opleiding aan de Utrechtse school voor Bestuurs- en Organisatiewetenschap, Universiteit Utrecht.

Rixt **Riemersma** is student assistent en student *research in public administration and organizational science* aan de Utrechtse school voor Bestuurs- en Organisatiewetenschap, Universiteit Utrecht.