

Het meten van tone of voice in Nederlandstalige geschreven tekst

Erasmus Universiteit Rotterdam
Erasmus School of Accounting & Assurance
Post-Master opleiding Internal Auditing & Advisory

Referaat ten behoeve van het slotexamen

Student: Haasnoot, J.P. (studentnummer 48345)

Begeleider: drs. M. (Maarten) Hoornweg RO

Deze versie van dit referaat is bestemd voor:
Examencommissie ESAA/IA&A
Maarten Hoornweg (begeleider ESAA/IA&A)
Michiel Boswinkel (begeleider DNB)

Abcoude, 29 april 2016

Het meten van tone of voice in Nederlandstalige geschreven tekst

Versiehistorie

Versie	Datum	Status
1.00	29-Apr-2016	Voor akkoord aan ESAA begeleider en doorgeleiding aan de examencommissie. Deze versie van 79 pagina's is t.o.v. de vorige versie 0.62: Ontdaan van commentaarballonnen en werkmarkeringen.
0.62	29-Apr-2016	Voor akkoord aan ESAA begeleider. Deze versie van 79 pagina's is t.o.v. de vorige versie 0.55: - aangepast nav het commentaar. De commentaarballonnen, en de wezenlijke teksttoevoegingen en tekstdoorhalingen zijn in deze versie nog blijven staan tbv het overzicht. De responses op de commentaarpunten zijn in een apart document toegelicht. - aangevuld met bijlage A en een voorwoord.
0.55	24-Apr-2016	Voor review aan ESAA-begeleider. Deze versie van 73 pagina's is t.o.v. de vorige versie 0.48: - aangevuld met een samenvatting - aangevuld met een beschrijving van het verloop (para 1.5) - aangevuld met een beschrijving van de evaluatie van het instrument o.b.v. verwachtingen uit de stakeholderanalyse (para 2.4 en 4.4) - aangevuld met H.6: beantwoording van de Odv'n en van de Cov - aangevuld met H.7: conclusies en aanbevelingen
0.48	18-Apr-2016	Voor review aan ESAA-begeleider. Deze versie van 63 pagina's is t.o.v. de vorige versie 0.38: - weer ontdaan van de alternatieve analyse. De bijbehorende validaties kostten te veel tijd; - aangevuld met hoofdstuk 5 (Evaluatie door experts) - aangevuld met paragraaf 4.3 (Duiding van meetresultaten) - aangevuld met toelichting op de beslistabel 1
0.38	11-Apr-2016	Voor review aan ESAA-begeleider. Deze versie van 66 pagina's is t.o.v. de vorige versie 0.23: - herzien v.w.b. de uitleg van validiteit (paragraaf 3.5); - gecompleteerd v.w.b. de resultaten uit de analyse o.b.v. het <i>four ears</i> -model (paragraaf 4.1); - aangevuld met (en integraal herzien voor) een alternatieve analyse o.b.v. schrijfrichtlijnen uit de praktijk - aangevuld met de eerste resultaten uit die alternatieve analyse (paragraaf 4.2.1) - aangevuld met bijlage H In de tekst zijn wezenlijke aanvullingen rechts in de kantlijn gemarkeerd.
0.23	06-Apr-2016	Voor review aan ESAA-begeleider. Deze versie van 56 pagina's is t.o.v. versie 0.17 uitgebreid met - conceptuele operationalisering in paragraaf 2.5 - invullingen voor resterende paragrafen H.3 - invulling paragraaf 4.1 - bijlagen F,G
0.17	29-Mrt-2016	Voor review aan ESAA-begeleider. Deze versie van 49 pagina's bevat een volledige beoogde inhoudsopgave en invullingen (gereed voor review) van de hoofdstukken 1 en 2, paragraaf 3.1, en bijlagen B, C, D, E.
0.16	27-Mrt-2016	Voor review aan meelezer. Deze versie van 47 pagina's bevat een volledige beoogde inhoudsopgave en invullingen (gereed voor review) van de hoofdstukken 1 en 2, paragraaf 3.1, en bijlagen B, C, D, E.

Voorwoord

Dit referaat is de weerslag van mijn onderzoek naar *tone of voice* in Nederlandstalige geschreven tekst. Voor zover de aanduiding van dat onderwerp nog geen ontzag inboezemt zal de toegepaste methode van *subjectivity analysis* dat toch doen: geautomatiseerd zoeken naar patronen die moeten duiden op iets subjectiefs in ongestructureerde dataverzamelingen. Dapperder dan zo'n zoektocht naar de Heilige Graal van de *Digital Humanities* is toch nauwelijks denkbaar.

Dapper of niet, mijn onderzoek, hoe spannend, leuk en boeiend ik het ook vond, is niet meer gebleken dan een houterige vingeroefening. Op een onderzoeksterrein dat voorlopig misschien wel meer pretenties heeft dan prestaties. Of, zoals Ewoud Sanders het verwoordde in zijn Woordhoek-column in het NRC Handelsblad van 31 maart 2016: "*De verwachtingen van automatische patroonherkenning zijn erg hooggespannen [...] maar mijn ervaring is dat handmatig slim en creatief zoeken in big data tot nu toe gemiddeld twee keer zoveel nuttige bronnen oplevert.*" Na afloop van het onderzoek deel ik die visie maar had ik mijzelf de lol van de poging niet willen ontzeggen.

In de voorbereiding en uitvoering van het onderzoek heb ik van velen hulp gehad. Om te beginnen ben ik dank verschuldigd aan collega's bij De Nederlandsche Bank, die hebben geholpen bij het samenstellen van lexicons of bij het selecteren van toezichtbrieven of bij het meedenken over het schrijven van dit referaat zonder de vertrouwelijkheid van toezichtinformatie te schenden. Verder wil ik in het bijzonder Michiel Boswinkel, Jolanda Breedveld, Maarten Hoornweg, Miranda Snel, en Tigran Spaan bedanken voor hun procesmatige en inhoudelijke medewerking tijdens de afgelopen maanden. En, *last but not least*, degenen uit mijn privé-omgeving de me de afgelopen jaren met raad, daad en veel geduld hebben gesteund.

Samenvatting

In dit onderzoek is een meetinstrument ontwikkeld om *tone of voice* te meten in Nederlandse teksten. Dat instrument is toegepast op toezichtbrieven van DNB.

Het meetinstrument kan worden beschouwd als een samenstel van bouwblokken, elk met een specifieke, voor dit onderzoek passende, invulling. Elk bouwblok kan echter ook op andere manieren worden ingevuld, al naar gelang een specifieke toepassing dat verlangt.

Voor het samenstel van invullingen voor het onderzoek bij DNB geldt het volgende:

- het geheel werkt, dat wil zeggen er is een werkende aaneenschakeling mogelijk van inhoudelijk consistente en technisch op elkaar aansluitende componenten, vanaf conceptuele operationalisering tot aan met meetresultaten onderbouwde uitspraken over verandering van *tone of voice*;
- de evaluatie aan de hand van referentie-uitkomsten is niet goed uitgevoerd kunnen worden omdat de voor dit onderzoek beschikbare referentie-uitkomsten slechts beperkt vergelijkbaar zijn met de in dit onderzoek gevonden meetresultaten. Dit euvel lijkt in een andere toepassing in beginsel oplosbaar te zijn;
- de validiteit van de resultaten van de doorzoekingen is vastgesteld op het laagste niveau van afzonderlijke treffers;
- de controles op de validiteit zijn cruciaal maar, vanwege het handmatige en dus arbeidsintensieve karakter ervan, ook een wezenlijke tekortkoming van dit instrument als het over praktische toepasbaarheid gaat. Een efficiënt alternatief voor die controles dient zich niet direct aan en vergt nader onderzoek;
- er zijn in dit onderzoek enkele andere verbeterpunten geïdentificeerd die in een andere toepassing in beginsel oplosbaar lijken te zijn.

Inhoudsopgave

Voorwoord.....	3
Samenvatting	4
1	Inleiding 8
1.1	Setting the stage 8
1.2	Aanleidingen 9
1.3	Probleemstelling..... 10
1.3.1	Doelstelling 10
1.3.2	Vraagstelling 10
1.4	Onderzoeksopzet 11
1.5	Verloop van het onderzoek en afwijkingen van het onderzoeksontwerp 12
2	Conceptuele operationalisering van tone of voice..... 13
2.1	Voorlopige definitie van het begrip tone of voice..... 13
2.1.1	Internetartikelen, academische publicaties en bedrijfsspecifieke style guides.... 14
2.1.2	Kritische beschouwing en een voorlopige definitie van tone of voice..... 16
2.2	Analytische concepten uit de linguïstiek..... 17
2.2.1	Linguïstiek, corpustaalkunde, computationele taalkunde, subjectivity analysis . 17
2.2.2	Private state als basis voor een analytisch kader voor tone of voice..... 18
2.2.3	Subjectiviteit uit de private state en het niet-feitelijke van tone of voice 18
2.3	Theoretische modellen voor communicatie..... 19
2.3.1	Het conduit (of: transitional) model..... 20
2.3.2	Brown en Levinson: politeness theory en face 21
2.3.3	Schulz von Thun: four ears-model..... 22
2.3.4	Het relationele aspect uit het four ears-model als kader voor tone of voice..... 23
2.4	Synthese en een conceptuele operationalisering van tone of voice..... 24

3	Onderzoeksmethode en technische uitvoering	28
3.1	Nogmaals: subjectivity analysis	29
3.1.1	<i>Hoofdstappen van subjectivity analysis</i>	29
3.1.2	<i>Lexicon: bestand van signaalwoorden</i>	29
3.2	Corpora van toezichtbrieven prepareren	32
3.3	Lexicons van signaalwoorden samenstellen.....	33
3.4	Analyses uitvoeren in WordSmith	34
3.5	Validiteit toetsen en zo nodig trefferfrequenties corrigeren.....	36
3.5.1	<i>Signaalwoordvaliditeit en treffervaliditeit</i>	36
3.5.2	<i>Twee kwalitatieve validiteitscriteria</i>	37
3.5.3	<i>Concordantieanalyse om de validiteit te beoordelen</i>	38
3.5.4	<i>Meerdere iteraties van doorzoeken, valideren en corrigeren</i>	40
3.6	Interpretatie van de meetresultaten.....	40
4	Meetresultaten	42
4.1	Eerste iteratie met initiële lexicons en validiteitbeoordeling.....	42
4.2	Tweede iteratie met gecorrigeerde lexicons en validiteitbeoordeling	44
4.3	Vertaling van meetresultaten naar verandering van tone of voice	46
4.4	Vergelijking met de verwachte uitkomsten.....	48
5	Evaluatie door experts.....	49
5.1	Kanttekeningen bij de conceptuele operationalisering.....	49
5.2	Kanttekeningen bij de onderzoeksmethode en technische uitvoering.....	51
5.3	Eenheden van de meetresultaten	51
6	Anatomie, alternatieven en antwoorden.....	53
6.1	Anatomie van het meetinstrument.....	53
6.2	Beantwoording van onderzoekdeelvragen	55
6.2.1	<i>Onderzoekdeelvraag 1 over conceptuele operationalisering</i>	55
6.2.2	<i>Onderzoekdeelvraag 2 over technische operationalisering en uitvoering</i>	58
6.2.3	<i>Onderzoekdeelvraag 3 over gevonden veranderingen van tone of voice</i>	60
6.2.4	<i>Onderzoekdeelvraag 4 over verwachte veranderingen van tone of voice</i>	60
6.2.5	<i>Onderzoekdeelvraag 5 over het vergelijk van gevonden met verwachte veranderingen</i>	61
6.2.6	<i>Onderzoekdeelvraag 6 over de opvatting van deskundigen</i>	61
6.3	Beantwoording van de centrale onderzoeksvraag	61
7	Conclusies en aanbevelingen.....	63

Bijlage A. Verantwoording en leereffecten.....	65
Bijlage B. Index van het onderzoeksdossier.....	67
Bijlage C. Gepland onderzoeksmodel (uit: onderzoeksontwerp).....	70
Bijlage D. Geplande onderzoeksaanpak (uit: onderzoeksontwerp).....	71
Bijlage E. "Readme" bij de Subjectivity Clues lexicon (Wilson 2005)	73
Bijlage F. Instructie aan toezichthouders in annotatiepanel	75
Bijlage G. Her-classificatie van signaalwoorden	76
Literatuur.....	77

1 Inleiding

Dit referaat beschrijft een afstudeeronderzoek naar *tone of voice* in toezichtbrieven van DNB. In dit inleidende hoofdstuk zullen we de context en hoofdlijnen van dat onderzoek schetsen.

Eerst geven we een globale indruk van het toezicht door DNB en kritiek van onder toezicht staande instellingen op de *tone of voice* in toezichtbrieven (paragraaf 1.1). Vervolgens zien we dat die kritiek de eerste aanleiding vormt voor het onderzoek, en mijn belangstelling voor *subjectivity analysis* de tweede aanleiding (paragraaf 1.2). In de probleemstelling zien we de inhoudelijke routekaart van het onderzoek (paragraaf 1.3). We sluiten dit inleidende hoofdstuk af met een indruk op hoofdlijnen over de uitvoering van het onderzoek, zowel hoe die gepland was (paragraaf 1.4) als hoe die feitelijk is verlopen (paragraaf 1.5).

1.1 Setting the stage

DNB is belast met het toezicht op banken, verzekeringsmaatschappijen en pensioenfondsen in Nederland. In het kader van dat toezicht voeren toezichthouders van DNB toezichtonderzoeken uit bij zo'n instelling. De toezichthouder sluit zo'n toezichtonderzoek in de regel af met een toezichtbrief aan de directie van de instelling. Daarin vermeldt en onderbouwt hij¹ de bevindingen uit het onderzoek en de conclusies die DNB daar aan verbindt. Als daar reden toe is beschrijft hij bovendien welke acties de instelling moet ondernemen.

Naast het sturen van een brief bespreekt de toezichthouder de te verwachten uitkomsten van het onderzoek met de instelling. In de regel doet hij dat voordat hij de toezichtbrief verstuurt. Die bespreking is meestal met leden van de directie van de instelling.

Het blijkt nu dat vertegenwoordigers van onder toezicht staande instellingen een verschil ervaren in de *tone of voice* van die twee communicatievormen. Zij ervaren tijdens een overleg met de

¹ Waar ik 'hij' en/of de mannelijke vorm gebruik is ook 'zij' respectievelijk de vrouwelijke vorm bedoeld.

toezichthouder vaak een andere *tone of voice* dan uit een toezichtbrief over hetzelfde onderwerp: "Wat communicatie betreft, ervaren de stakeholders een verschil tussen het gesproken en het geschreven woord [...] Zij hadden naar hun idee een prettig gesprek gevoerd, maar vervolgens valt er een formele, sterk juridisch getinte brief op de deurmat [...] Het grote verschil in tone of voice schept wel eens verwarring" (DNB, 2015, p.12).

DNB neemt die kritiek ter harte en is expliciet van zins om het gesignaleerde verschil "aan te pakken". Die reactie past in een bredere ambitie van DNB om transparant(er) te zijn over het door haar uitgevoerde toezicht (DNB, 2014, p.22).

Een onderzoek naar *tone of voice* in de toezichtbrieven, en de eventuele veranderingen daarin door de tijd, kan belangrijke inzichten bieden voor die aanpak. Voor dit referaat wil ik dan ook een instrument ontwikkelen voor het meten van *tone of voice* in Nederlandstalige geschreven teksten, dat instrument toepassen op toezichtbrieven van DNB en de betrouwbaarheid van die metingen evalueren aan de hand van de kritiek van instellingen over *tone of voice*.

1.2 Aanleidingen

De in paragraaf 1.1 beschreven kritiek van onder toezicht staande instellingen over de *tone of voice* in toezichtbrieven (DNB, 2015, pp.10-12) vormt één van de twee aanleidingen voor dit onderzoek. Uitgaande van die kritiek heb ik twee versmallingen toegepast om uit te komen bij een scherp afgebakend afstudeeronderwerp. De versmallingen betrekking op de keuzes (i) om niet de instellingen als onderwerp van onderzoek te kiezen en (ii) om evenmin het onderzoek te richten op de structurele verschillen tussen mondelinge en schriftelijke communicatie. De afwegingen bij deze keuzes zijn toegelicht in het onderzoeksontwerp dat aan de uitvoering van dit onderzoek ten grondslag ligt. Al met al is het onderwerp van dit referaat de ontwikkeling in de tijd van *tone of voice* in toezichtbrieven.

Een tweede aanleiding is gelegen in mijn belangstelling voor toepassingen van geautomatiseerde tekstanalyse. De analyse van geschreven natuurlijke taal heeft in de afgelopen jaren een vlucht genomen. Die is onder meer aangezwengeld door het commerciële belang van ontsluiting van de schat aan informatie op *social media* en internetfora. In de toegepaste linguïstiek zijn methoden en technieken ontwikkeld om structuur- en betekenis van geschreven natuurlijke taal te onderzoeken; waaronder ook kenmerken die verband houden met *tone of voice*.

Subjectivity analysis, ook wel aangeduid met *opinion mining*, is een verzamelterm voor dergelijke methoden en technieken. *Subjectivity analysis* is erop gericht om de attitude van een spreker of

schrijver ten aanzien van een bepaald onderwerp te identificeren. Die attitude kan betrekking hebben op zijn of haar beoordeling van het onderwerp, zijn of haar affectieve positie ten aanzien van het onderwerp, of zijn of haar bedoelde emotionele communicatie erover.² Het lijkt er dus op dat *subjectivity analysis*, althans in concept, geschikte instrumenten zou moeten kunnen aanreiken voor een analyse van *tone of voice* in toezichtbrieven.

1.3 Probleemstelling

1.3.1 Doelstelling

De onderzoeksdoelstelling luidt als volgt:

Het doel van dit onderzoek is het ontwikkelen van een meetinstrument voor de verandering van *tone of voice* in Nederlandstalige geschreven teksten en het toepassen en evalueren daarvan, door operationalisering van *tone of voice* en toepassing van *subjectivity analysis*-technieken op DNB-toezichtbrieven uit tenminste twee onderscheiden perioden, teneinde DNB inzicht te bieden in de ontwikkeling van *tone of voice* in haar toezichtbrieven.

1.3.2 Vraagstelling

De centrale onderzoeksvraag in het onderzoek luidt:

in hoeverre is het in dit onderzoek te ontwikkelen meetinstrument voor verandering van *tone of voice* praktisch toepasbaar en wat is er in dat geval te zeggen over de validiteit van de meetresultaten?

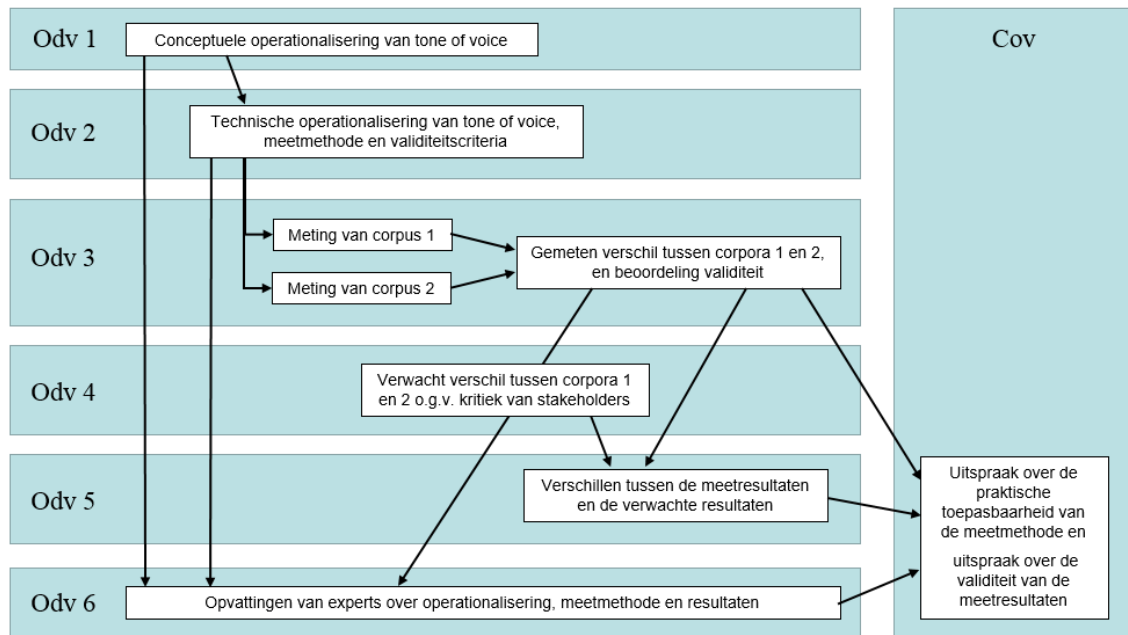
Het antwoord op de centrale onderzoeksvraag wordt opgebouwd uit antwoorden op zes onderzoekdeelvragen (zie ook figuur 1 voor een schematische weergave):

1. Hoe kan *tone of voice* conceptueel worden geoperationaliseerd?
2. Hoe kan *tone of voice* technisch worden geoperationaliseerd en hoe kan het worden gemeten?
3. Wat is de *tone of voice* in toezichtbrieven uit twee onderscheiden corpora (van periode 1 en van periode 2) en welke veranderingen in *tone of voice* blijken er uit de metingen van die beide

² Wikipedia, lemma 'Sentiment analysis'. http://en.wikipedia.org/wiki/Sentiment_analysis (geraadpleegd 02-Aug-2015).

corpora?

4. Wat is de verwachting over verandering in *tone of voice* tussen beide corpora op grond van de opvatting van de stakeholders?
5. Hoe verhouden de gemeten veranderingen tussen beide corpora (Odv 3) zich tot de verwachte veranderingen (Odv 4) tussen beide corpora?
6. Wat is de opvatting van deskundigen over de conceptuele operationalisering (Odv 1), de meetmethode (Odv 2) en de meetresultaten (Odv 3)?



Figuur 1 Onderzoeksdeelvragen (Odv'n) en centrale onderzoeksvraag (Cov)

1.4 Onderzoekopzet

Dit onderzoek aan te merken als een empirisch onderzoek gericht op het ontwikkelen van een praktisch toepasbare methode. De doelstelling van dit onderzoek is immers om een praktisch bruikbaar meetinstrument te ontwikkelen waarvan de toepasbaarheid getoetst wordt in de empirie: meting van *tone of voice* van concrete corpora en vergelijking van de meetresultaten tegen de

opvattingen van lezers van brieven uit die corpora.

Bijlagen C en D bevatten het onderzoeksmodel respectievelijk de onderzoeksaanpak zoals die waren voorzien in het goedgekeurde onderzoeksontwerp dat ten grondslag ligt aan de uitvoering van dit onderzoek (versie 1.0 van 9 maart 2016).

1.5 Verloop van het onderzoek en afwijkingen van het onderzoeksontwerp

Het onderzoek is voor het grootste deel uitgevoerd zoals was voorzien in het goedgekeurde onderzoeksontwerp (versie 1.0 van 9 maart 2016). Op twee punten is daar van afgeweken. Ten eerste was in het onderzoeksontwerp voorzien dat de analyse van *style guides* en interne DNB-rapporten zou leiden tot aanvullende signaalwoorden, bovenop de inventarisatie van signaalwoorden door annotatie in een panel. Dat is niet gelukt. De *style guides* boden daartoe onvoldoende aanknopingspunten. De analyse van *style guides* is wel gebruikt als bron voor de conceptuele operationalisering van *tone of voice* (paragraaf 2.1.1 van dit referaat). De tweede afwijking betreft de verdeling van onderwerpen over de hoofdstukken en paragrafen van dit referaat. Op onderdelen bleek een iets andere indeling logischer dan in het onderzoeksontwerp van voorzien.

In de hoofdstukken 2, 3 en 4 behandelen we de kern van het onderzoek gezien. Achtereenvolgens komen de conceptuele operationalisering van *tone of voice*, de technische operationalisering en uitvoering van het onderzoek, en de meetresultaten en interpretatie daarvan aan bod. Vervolgens zien we in hoofdstuk 5 een kritische beschouwing van het onderzoek door externe deskundigen. Met al die gegevens kunnen we de centrale onderzoeksvraag en de onderzoekdeelvragen beantwoorden, dat doen we in hoofdstuk 6. Het referaat wordt afgesloten met conclusies en aanbevelingen in hoofdstuk 7.

2 Conceptuele operationalisering van tone of voice

In dit hoofdstuk zullen we het begrip *tone of voice* conceptueel operationaliseren.

In de eerste paragraaf stellen we een voorlopige definitie op van wat onder *tone of voice* kan worden verstaan, op basis van enkele internetartikelen, academische publicaties en bedrijfsspecifieke *style guides* (paragraaf 2.1). Vervolgens zoeken we in twee kennisdomeinen naar theoretische aanknopingspunten voor een verdere begripsmatige operationalisering. Eerst gebruiken we de notie van *private state* uit het linguïstische vakgebied *subjectivity analysis* als basis voor een analytisch kader (paragraaf 2.2). Een *private state* is een subjectieve uiting van iemand over iets. In dit onderzoek kunnen we concrete voorkomens van *tone of voice* in geschreven tekst zien als subjectieve uitingen van de schrijver (DNB of de toezichhouder) over "iets met de lezer" (de onder toezicht staande instelling). Voor een uitwerking van dat "iets met de lezer" kijken we wat diverse communicatiemodellen zeggen over boodschappen en onderliggende bedoelingen (paragraaf 2.3). Aan de hand van het *four ears*-model van Schulz von Thun overwegen we dat *tone of voice* met name terugkomt in het relationele aspect van een boodschap. Daarmee kunnen we voorkomens van *tone of voice* begrijpen als subjectieve uitingen van de schrijver (DNB) over zijn relatie tot de lezer (instelling). Aan het eind van het hoofdstuk combineren we de voorlopige definitie, de theoretische noties van *private state* en het *four ears*-model in een conceptuele operationalisering van *tone of voice* (paragraaf 2.4)

2.1 Voorlopige definitie van het begrip tone of voice

Voor een verkenning van het begrip *tone of voice* raadplegen we het internet. In internetartikelen, academische publicaties en bedrijfsspecifieke *style guides* zien we een aantal kenmerken meermaals terugkomen. Daaruit leiden we een eerste indruk af van wat *tone of voice* kennelijk behelst. In het tweede deel van de paragraaf plaatsen we enkele accenten en nuanceringen en komen we tot een voorlopige definitie. Die geldt als startpunt voor het zoeken naar theoretische aanknopingspunten in de rest van het hoofdstuk.

2.1.1 Internetartikelen, academische publicaties en bedrijfsspecifieke style guides

Een zoekvraag in Google met als trefwoord "tone of voice" geeft een eerste indruk van wat in diverse internetartikelen onder *tone of voice* wordt begrepen. De treffers verwijzen overwegend naar websites over marketingcommunicatie, bijvoorbeeld uit het Nederlandse taalgebied.

Omschrijvingen van *tone of voice* luiden:

- Het begrip tone of voice verwijst naar "[...] de stijl waarin een bedrijf of merk communiceert met de doelgroep. [...] Een tone of voice kan formeel zijn of juist informeel, serieus of juist luchtig, direct of indirect, commercieel of informatief, technisch of juist toegankelijk." ³
- Of naar: "[...] het jasje waarin je een tekst giet en de manier waarop je de lezer aanspreekt. Concreet gaat het bijvoorbeeld om woordkeus, zinsopbouw en aanspreekvorm." ⁴
- Of: "Je tone of voice gaat niet over wat je zegt, maar over hoe je het zegt. Het is de taal die je gebruikt, de manier waarop je zinnen in elkaar zet, de woorden die je kiest. Het is de verbale zus van je visuele identiteit (logo, huisstijl). In geschreven tekst heb je geen hulp van non-verbale communicatie. Geschreven tekst heeft alleen toon om meer te zeggen dan de losse woorden die op papier staan. Die moet dus herkenbaar, duidelijk en relevant zijn voor de lezer." ⁵

In Engelstalige publicaties vinden we vergelijkbare omschrijvingen van *tone of voice*:

- "A tone of voice is not what you say, but how you say it. This encompasses not only the words you choose, but their order, rhythm and pace. Rather confusingly, when seen in the world of business and marketing, the phrase 'tone of voice' refers to written – rather than spoken – words. A company's tone of voice will inform all of its written copy, including its website, social media messages, emails and packaging" (Cummings, publicatiejaar onbekend).
- Of: "Written tone of voice is simply the 'personality' of your brand or company as expressed through the written word. Tone of voice governs what you say in writing, and how you say it – the content and style of textual communications, in any setting and in any medium" (Albrighton 2010).

De academische literatuur over *tone of voice* in geschreven communicatie oogt tamelijk beperkt.⁶

³ Marketingtermen.nl, lemma 'tone of voice'. <http://www.marketingtermen.nl/begrip/tone-of-voice> (geraadpleegd 02-Aug-2015).

⁴ <http://www.setcommunicatie.nl/tekstschrijven/schrijftips/de-juiste-tone-of-voice> (geraadpleegd 01-Sept-2015).

⁵ <http://www.fijn.net/tone-of-voice/> (geraadpleegd 21-Mrt-2016)

⁶ In wetenschappelijke literatuur is overigens wel het nodige te vinden over onderzoek naar *tone of voice* in gesproken communicatie (Markel e.a. 1973; Zuckerman e.a. 1982; Mayew en Venkatachalam 2012; Allee en Deangelis 2015).

Het zoekargument "*tone of voice*" levert geen enkel toepasselijk collegiaal getoetst artikel op. Met de zoekwoorden "*tone style written business communication*" komen drie relevante, maar gedateerde, publicaties naar voren.

In haar onderzoek naar stijlverschillen tussen mannelijke en vrouwelijke managers operationaliseert Sterkel een schrijfstijl in 20 linguïstische tekstkenmerken (1988, p.19). Die kenmerken zijn allemaal kwantitatief, zoals bijvoorbeeld het aantal woorden, het percentage passieve woorden en de gemiddelde zinslengte.

Macpherson deed onderzoek naar stijl van teksten (1992, p.111). Hij schetst zes verschillende manieren om naar de stijl van een tekst te kijken: (1) als afwijking van een norm; (2) als een manifestatie van het individuele; (3) met dimensies op inhoud en/of vorm; (4) als een keuze tussen alternatieve manieren om hetzelfde idee tot uitdrukking te brengen; (5) als een product van de context; en (6) in de zin van goed of mooi (literair) schrijven.

Kenman tenslotte heeft voor trainingsdoeleinden de toon van een tekst onderzocht en geeft dat begrip al een lading; hij verwijst met *tone* naar beleefdheid, positiviteit, vriendelijkheid en gerichtheid op de lezer (2007, p.306).

Bedrijfsspecifieke *style guides* vormen - naast omschrijvingen op websites en academische publicaties - een derde bron voor aanknopingspunten over wat *tone of voice* behelst:

- "*Our written communications should be clear, concise, honest and open. The words we use should closely follow and support our principles and values. They need to demonstrate our: professionalism; clarity; respect; accessibility; straightforwardness*" (National Health Service 2010);
- "*Our tone of voice is the way in which we write and speak, what we say and how we say it. Like a person, what we say is dictated by our principles, experiences and aspirations, how we say it is informed by our personality*" (University of Leeds, publicatiejaar onbekend);
- "*Our tone of voice is the way in which our brand is expressed in writing, particularly in relation to our marketing materials. This includes what we say and how we say it*" (University of Salford Manchester, publicatiejaar onbekend).

In de genoemde omschrijvingen en publicaties komt een aantal verwijzingen meermaals voor. Hieruit dringt zich een eerste indruk op van hoe we *tone of voice* zouden kunnen begrijpen. Bij *tone of voice* gaat het kennelijk (i) om een stijl, (ii) in geschreven tekst, (iii) die is afgestemd op - en uitdrukking geeft aan - de *corporate identity* van de schrijver, (iv) waarmee niet-feitelijke

aspecten van een boodschap tot uitdrukking kunnen worden gebracht, en (v) die concreet vorm krijgt door de woordkeus, zinsopbouw, aanspreekvorm en andere linguïstische tekstkenmerken.

2.1.2 Kritische beschouwing en een voorlopige definitie van tone of voice

Voor de toepassing in dit afstudeeronderzoek zijn, uitgaande van de bovengenoemde eerste indruk, enkele accenten en nuancerings relevant. De accenten betreffen: (i) geschreven tekst; (ii) perspectief van de zender; en (iii) de niet-feitelijke aspecten, de nuancering gaat over (iv) het kunnen differentiëren van *tone of voice* binnen de afstemming met de *corporate identity*.

Het eerste accent, en een bevestiging van de eerste indruk, is dat in dit onderzoek wordt gekeken naar *tone of voice* in geschreven tekst. Dit vloeit rechtstreeks voort uit de doelstelling van het onderzoek, het ontwikkelen van een instrument om *tone of voice* in geschreven teksten te kunnen meten.

Ten tweede wordt *tone of voice* beschouwd vanuit het perspectief van de schrijver. Ook dat is in lijn met de bovengenoemde eerste indruk. Merk echter op dat het onderzoek door dit accent wegdrijft van de aanleiding van het onderzoek, die immers was ingegeven door kritiek van ontvangers op de *tone of voice* in toezichtbrieven. Een belangrijke overweging om niet het perspectief van de ontvanger tot onderwerp van het onderzoek te maken is dat er in het geval van DNB al verdiepend onderzoek is (en wordt) gedaan naar de opvattingen van ontvangers. De toegevoegde waarde van nóg een onderzoek zou beperkt zijn.

Het derde accent betreft de niet-feitelijke aspecten van een boodschap. Dit afstudeeronderzoek is gericht op het kunnen meten van hetgeen "tussen de regels door" staat, de bedoelingen, associaties, lading die voorbij gaan aan de feitelijke, objectieve boodschappen in de geschreven tekst. In paragraaf 2.2.3 komt aan bod hoe dit accent op niet-feitelijke aspecten in het bijzonder van belang is in tekstgenres met een beoordelende aard, zoals bijvoorbeeld toezichtbrieven.

Als nuancering op de eerste indruk wordt *tone of voice* in dit onderzoek beschouwd als iets dat per boodschap moet kunnen worden gedifferentieerd, binnen de randvoorwaarde van niet discongruent zijn met de *corporate identity* van de schrijver. De specifieke omstandigheden van de boodschap spelen mee met het kiezen van een bij die boodschap passende *tone of voice*. Met deze nuance wordt ruimte gemaakt voor de realiteit waarin DNB (of een individuele toezichthouder) in verschillende toezichtbrieven een verschillende *tone of voice* moet kunnen hanteren.

Voorlopige definiëren we het begrip *tone of voice* als een schrijfstijl waarmee de schrijver niet-feitelijke en specifiek bij de boodschap passende bedoelingen, associaties, lading tot uitdrukking

kan brengen. Die stijl is consequent toegepast in de tekst waarin die bepaalde boodschap centraal staat en krijgt concreet vorm door de toegepaste linguïstische tekstenkenmerken. En die stijl is tenminste niet discongruent met de *corporate identity* van de schrijver.

In de volgende paragraaf zien we hoe de linguïstische notie van *private state* - een subjectieve uiting van iemand over iets - een goede basis biedt voor een analytisch kader voor *tone of voice*.

2.2 Analytische concepten uit de linguïstiek

2.2.1 Linguïstiek, corpustaalkunde, computationele taalkunde, subjectivity analysis

De linguïstiek is de wetenschappelijke studie van de natuurlijke talen⁷. Het doel van de linguïstiek is om talen te beschrijven zoals ze zijn; dat wil zeggen, zoals mensen ze in werkelijkheid gebruiken. Voor dit afstudeeronderzoek zijn met name twee disciplines in de linguïstiek relevant, te weten de corpustaalkunde en de computationele taalkunde. Corpustaalkunde⁸ is de studie van taalverschijnselen met behulp van verzamelingen van bestaande teksten. Dergelijke tekstverzamelingen worden corpora genoemd. Typische onderzoeken in deze discipline gaan over de frequentie van bepaalde woorden en de contexten waarin ze voorkomen. Dit gebeurt tegenwoordig over het algemeen digitaal, in de computationele taalkunde⁹. Dat is de specialisatie waarin de computationele modellering van taalkundige verschijnselen centraal staat. Typische onderzoeksgebieden binnen de computationele taalkunde zijn het bepalen van de betekenis van ambigue woorden in hun context en het identificeren van het onderwerp in stukken tekst.

De analyse van geschreven natuurlijke taal heeft in de afgelopen jaren een vlucht genomen. Die is onder meer aangezwengeld door het commerciële belang van ontsluiting van meningen op *social media* en internetfora. De behoefte aan kennis van wat potentiële klanten, burgers of kiezers denken, vinden, verwachten, verwerpen is al oud, maar met het wijdverbreide gebruik van internet en andere openbare media is er een nieuwe rijke bron beschikbaar om die kennis aan te ontfen (Wiebe 2005, p.1; Pang en Lee 2008, p.1).

De corpus- en computationele taalkunde bieden methoden en technieken om teksten te onderzoeken op de subjectieve opvattingen en meningen van de schrijvers. In het bijzonder komt

⁷ Wikipedia, lemma 'taalkunde'. <https://nl.wikipedia.org/wiki/Taalkunde>, geraadpleegd 20-Mrt-2016.

⁸ Wikipedia, lemma 'corpustaalkunde'. <https://nl.wikipedia.org/wiki/Corpustaalkunde>, geraadpleegd 20-Mrt-2016.

⁹ Wikipedia, lemma 'computationele taalkunde'. https://nl.wikipedia.org/wiki/Computationele_taalkunde, geraadpleegd 20-Mrt-2016.

dat tot uiting in het vakgebied *subjectivity analysis*.¹⁰ Dat is een verzamelterm voor methoden en technieken die erop gericht zijn om de attitude van een spreker of schrijver ten aanzien van een bepaald onderwerp te identificeren. In de linguïstiek worden de te onderzoeken attitudes, opvattingen en meningen veelal aangeduid met het bredere begrip *private state*.

2.2.2 *Private state als basis voor een analytisch kader voor tone of voice*

Het begrip *private state* verwijst naar een subjectieve uiting van iemand over iets. Een *private state* is een zinsnede die een antwoord bevat op de samengestelde vraag: wie vindt wat waarover? Een *private state* geeft weer dat een actor een bepaalde subjectieve opvatting of attitude, overtuiging, vraag, twijfel, of emotie heeft ten aanzien van een onderwerp (Wiebe 1994, p.235). Wilson definieert *private state* als "*a general term that covers opinions, beliefs, thoughts, feelings, emotions, goals, evaluations, and judgments*" (2005, p.53).

Wiebe onderscheidt drie analytische eigenschappen van een *private state*, te weten de *experiencer*, zijn *attitude* en het *object* waarop de *attitude* van de *experiencer* betrekking heeft (1994, p.235). Zij stelt dat elke *private state* in een tekst kan worden geanalyseerd als een combinatie van deze drie eigenschappen. Wilson voegt daar nog twee eigenschappen aan toe, namelijk de aard van de attitude (de polariteit): positief of negatief, en de mate (intensiteit) ervan: zwak of sterk (2008, p.4).

Door voorkomens van tone of voice in geschreven teksten te beschouwen als *private states*, en daar de bijbehorende analytische eigenschappen aan te relateren, kan een *tone of voice*-voorkomen worden geïdentificeerd en bestudeerd in termen van:

- wie: wie is de *experiencer* die kennelijk een bepaalde *tone of voice* hanteert?;
- vindt wat: welke bedoeling, associatie of lading drukt die *experiencer* uit?; heeft dat een positieve of negatieve connotatie?; is dat een sterke of zwakke uiting?;
- waarover: wat is het *object*?; op welk niet-feitelijke aspect van de boodschap slaat die *attitude* van de *experiencer* nu precies?

Zodoende biedt het concept van de *private state* een bruikbaar kader voor de begripsmatige operationalisering van *tone of voice* in geschreven teksten.

2.2.3 *Subjectiviteit uit de private state en het niet-feitelijke van tone of voice*

In de context van onderzoek van beoordelende teksten (zoals bijvoorbeeld toezichtbrieven) is goed om stil te staan bij de subjectiviteit die essentieel is in het concept van *private state*.

¹⁰ Het onderzoeksterrein van *subjectivity analysis* wordt ook wel aangeduid met *opinion mining*, *sentiment analysis*, *review mining* en *appraisal extraction* (Pang en Lee 2008, p.8).

Toezichtbrieven, auditrapporten, filmrecensies, juryrapporten, klachtmeldingen en andere beoordelende teksten staan vanwege hun functie vol met opvattingen en oordelen van iemand ergens over. Toch kwalificeren die zinsneden niet per se als *private state*. Er is namelijk per definitie alleen sprake van een *private state* als het een subjectieve uiting betreft.

Uitingen die controleerbaar - en daarmee objectief - zijn, zoals dat typisch geldt voor bevindingen in auditrapporten en toezichtbrieven, lijken misschien wel een opvatting of oordeel weer te geven, maar moeten toch worden gezien als feitelijke informatie. Een bevinding geeft feitelijke informatie over hoe een waargenomen toestand zich verhoudt tot een norm. Bij een bevinding is het typisch de verwachting dat een andere auditor of een andere toezichthouder die dezelfde situatie onderzoekt tot hetzelfde bevinding komt. Daarmee is een bevinding typisch niet subjectief en dus ook geen *private state*.

Die voorwaarde van subjectiviteit sluit goed aan bij het *tone of voice*-criterium over de niet-feitelijke aspecten van een boodschap. Het gaat bij *tone of voice* immers om de schrijfstijl waarmee de schrijver niet-feitelijke aspecten van de boodschap tot uitdrukking brengt. Feitelijke informatie, zoals bijvoorbeeld bevindingen, valt niet alleen buiten het *private state*-begrip maar ook buiten de voorlopige definitie van *tone of voice*.

Het is dus goed verdedigbaar om *tone of voice* in geschreven tekst te beschouwen als een grootheid die tot uitdrukking komt in *private states* en omgekeerd om een *private state* te beschouwen als een goede indicator voor het voorkomen van *tone of voice*.

In paragraaf 3.1 komt aan de orde hoé de daadwerkelijke identificatie en analyse van concrete *private states* in z'n werk gaat, aan de hand van methoden en technieken uit de *subjectivity analysis*.

Eerst hebben we nog een verdieping nodig van de waarover-eigenschap, ofwel het *object*, van de *private state*. Op welk niet-feitelijke aspect van een boodschap heeft de *attitude* van de *experiencer* nu eigenlijk betrekking? Daartoe kijken we in de volgende paragraaf wat communicatiemodellen zeggen over boodschappen en onderliggende bedoelingen.

2.3 Theoretische modellen voor communicatie

Communicatiemodellen geven houvast voor het theoretische denken over wat onder communicatie moet worden verstaan. In het *conduit model* wordt communicatie louter gezien als het overbrengen, door een zender, van een boodschap aan een ontvanger. Meer geavanceerde

cooperative models, waarvan in deze paragraaf twee uitwerkingen worden behandeld, besteden daarnaast ook aandacht aan mogelijke verschillen tussen de boodschap en de achterliggende bedoeling. Dergelijke verschillen zijn essentieel voor een beschouwing van *tone of voice*.

2.3.1 *Het conduit (of: transitional) model*

Een basaal denkmodel voor communicatie is het *conduit model*. Pauly (1977) beschrijft communicatie volgens dat model (dat hij overigens aanduidt met de term *transitional model*) als het proces waarin mensen hun ideeën en gevoelens omzetten in een gemeenschappelijk symbolisch ruilmiddel ("*currency*") en vervolgens doorgeven aan anderen. Hagge en Kostelnick (1989) stellen het iets simpeler: in het *conduit model* wordt communicatie eenvoudigweg beschouwd als het overbrengen van in taal verpakte ideeën en gedachten. Rasmussen (1991) schetst een vergelijkbaar beeld: het *conduit model* geeft communicatie weer als een eenvoudig proces van het versturen van betekenis, opgenomen in woorden en gebaren, van de ene persoon naar de andere. Het model betoogt bovendien dat betekenis betrekkelijk eenvoudig kan worden overgebracht en dat misverstanden zelden voorkomen (!).

Pauly erkent het belang van het *conduit model* als richtinggevend paradigma vanaf de jaren '50 in zowel de communicatiewetenschap als in toepassingen daarvan, zoals bijvoorbeeld bedrijfscommunicatie, marketing en reclamepraktijk. Tegelijk wijst hij op een cruciale interne tegenstrijdigheid van het *conduit model*, juist waar het die toepassingsgebieden betreft: "*The model claims to show how to manipulate others' behavior even as it argues that sensible, rational people cannot be manipulated*" (1977, p.15). De belangrijkste kritiek van Pauly op dit model is dat het communicatie uitdrukt in té mechanistische termen en te weinig aandacht heeft voor het symbolische aspect ervan.

Ook Hagge en Kostelnick doen het *conduit model* als te simpel af. Het adagium "opschrijven wat je bedoelt en bedoelen wat je opschrijft", zoals dat volgens hen door communicatiedocenten wordt ontleend aan het *conduit model*, gaat voorbij aan de realiteit van alledag waarin juist grote verschillen kunnen blijken tussen de feitelijke uiting en de bedoeling van de zender (1989, pp.317-318). Zij pleiten ervoor om die verschillen tussen wat een zender schrijft of zegt en wat hij bedoelt sterker te betrekken in de analyse van het boodschappenverkeer.

In de navolgende subparagrafen worden twee modellen behandeld die meer aandacht geven aan het onderscheid tussen "wat er staat" en "wat er werd bedoeld", te weten de in wetenschappelijk werk veel geciteerde *politeness theory* van Brown en Levinson en het meer op de (trainings)praktijk gerichte *four ears*-model van Schulz von Thun.

2.3.2 Brown en Levinson: politeness theory en face

Eén voorbeeld van een communicatiemodel waarin rekening wordt gehouden met verschillen tussen boodschap en bedoeling is de *politeness theory* (Brown & Levinson, 1987). Die theorie behandelt beleefdheid en de functie ervan in sociale interactie. Op basis van een fundamentele beschouwing beschrijft de theorie enkele strategieën die sprekers toepassen om in een dialoog beleefd te zijn, of anders gezegd: de "*face*" van hun gesprekspartner niet te bedreigen. Het dilemma van de zender bestaat enerzijds uit de sociale verplichting om de *face* van de ontvanger niet te bedreigen en anderzijds uit zijn eigen belang om ideeën over te brengen die voor de ontvanger *face-threatening* kunnen zijn (1987, p.58). Het model is net zo goed van toepassing op geschreven teksten als op gesproken dialogen, zoals blijkt uit wetenschappelijke publicaties over toepassing van politeness theory op geschreven tekst (Cherry 1987; Limberg 2009).

Face verwijst naar het positieve beeld dat een persoon (de ontvanger) van zichzelf wil vestigen of handhaven in het sociale verkeer. Die persoon claimt een positieve sociale waardering voor zichzelf. Het wordt als onbeleefd gezien als anderen (bijvoorbeeld de zender) die positieve sociale waardering niet respecteren of zelfs bedreigen. Dit *face*-begrip werd uitgewerkt door de socioloog Erving Goffman (1955).

Brown en Levinson beschrijven het gedrag van twee deelnemers (zender en ontvanger, die gedurende de dialoog van rol kunnen wisselen) in een mondelinge interactie, uitgaande van het *face*-begrip van Goffman. Uit het gedrag van de deelnemers in de interacties concluderen Brown en Levinson dat de ontvanger de bedoelingen van de zender niet zozeer opmaakt uit wat die zender feitelijk zegt, maar uit de subtiële linguïstische details van de uiting en uit begeleidende lichaamstaal (1987, p.57).

Verder onderscheiden Brown en Levinson verschillende strategieën in het gedrag van de zender, kennelijk met de bedoeling om de *face* van de ontvanger niet te bedreigen. Die strategieën variëren van de "*bald on-record*", gewoon botweg zeggen waar het op staat, in het ene uiterste tot "*Don't do any face threatening action*", waarbij de zender de dialoog uit de weg gaat, in het andere uiterste. Tussen die twee uitersten liggen nog strategieën die uitgaan van een positieve beleefdheid en strategieën die uitgaan van een negatieve beleefdheid.

De zender kiest zijn strategie op basis van een inschatting van het effect ervan, en dat wordt bepaald door drie onafhankelijke factoren, te weten (i) de sociale afstand tussen de zender en de ontvanger; (ii) de mate waarin de zender, op basis van zijn machtspositie ten opzichte van de ontvanger, acties kan afdwingen; en (iii) de sociaal bepaalde mogelijkheden voor de ontvanger om zich te onttrekken aan de door de zender opgelegde acties (1987, p.74 e.v.).

De eigenschappen van de relatie tussen de zender en de ontvanger worden ook in het hierna te

behandelen *four ears*-model expliciet uitgewerkt.

2.3.3 Schulz von Thun: *four ears*-model

Een ander voorbeeld van een communicatiemodel waarin nadrukkelijk onderscheid wordt gemaakt tussen de letterlijke boodschap en onderliggende bedoelingen is het *four ears*-model. In 1977 ontwierp de psycholoog Friedemann Schulz von Thun een model voor de analyse van communicatie-uitingen, als hulpmiddel voor in communicatietrainingen (2010, p.5). Het model is het *four ears*-model gaan heten omdat het de aandacht vestigt op vier afzonderlijke aspecten van elke communicatie-uiting, namelijk een referentieel, een appellerend, een expressief en een relationeel aspect.

In een simpel voorbeeld kunnen de afzonderlijk aspecten al worden herkend: twee mensen rijden in een auto en de passagier (zender) zegt tegen de bestuurder (ontvanger): "daar verderop staat het verkeerslicht op groen" (uiting) (2010, pp.17-20).

In het referentiële aspect van een boodschap, ook wel het zakelijke aspect genoemd, komt een beschrijving van feiten tot uitdrukking. In het voorbeeld is dat de feitelijke informatie dat het verkeerslicht, dat zich verderop bevindt, op het moment van spreken op groen staat. Niet meer en niet minder. Hierbij wordt overigens het feit zelf niet ter discussie gesteld. Of de passagier het wel goed gezien heeft, of er verderop ook een ander verkeerslicht staat dat niet op groen staat, of het eigenlijk wel een verkeerslicht is dat op groen staat en vergelijkbare vragen naar de geldigheid van de uitgewisselde feitelijke informatie doen in dit verband niet ter zake.

Het appellerende aspect verwijst naar het door de zender beoogde doel van de uiting, de oproep aan de ontvanger. De passagier zou met de mededeling over het groene verkeerslicht de bedoeling kunnen hebben om de bestuurder aan te sporen iets te versnellen, zodat ze het groene licht nog kunnen halen. Een belangrijke aanname in het *four ears*-model is dat er in elke communicatie-uiting een bedoeling ligt. Die aanname lijkt terecht in de context van bedrijfscommunicatie en trainingen daarvoor, waarvoor Schulz von Thun het model heeft ontwikkeld. In een breder perspectief op communicatie kan echter ook sprake zijn van non-intentionele communicatie, bijvoorbeeld het onbedoeld of onbewust uitzenden van signalen door blozen of knippen met de ogen.

Het expressieve aspect bevat informatie over de zender. Uit de mededeling in het voorbeeld kan worden opgemaakt dat de passagier wakker is, dat hij niet blind is, dat hij mogelijk haast heeft of misschien om een praatje verlegen zit, en - hoe triviaal ook - dat hij er voor gekozen heeft om de kleur van het verkeerslicht te melden. In het expressieve aspect komt de presentatie door de zender van zichzelf tot uitdrukking. De onderliggende aanname in het model is dat een ontvanger

elke boodschap analyseert in termen van wat die boodschap hem zegt over de zender. Dat gaat niet alleen over de persoonlijkheid en emoties van de zender maar ook over zijn keuze voor presentatietechnieken. Geeft de zender zich bloot, probeert hij te imponeren of misschien zijn nadelige kanten te verhullen (Schulz von Thun 2010, pp.37-41)?

Het relationele aspect gaat over de verstandhouding tussen zender en ontvanger, over hoe de zender zichzelf ziet in zijn relatie tot de ontvanger en over hoe de ontvanger dat interpreteert. Zo kan de mededeling in het voorbeeld erop duiden dat de passagier zichzelf beter vindt dan de bestuurder. De zender permitteert zich immers zich te bemoeien met de besturen van de auto, althans de bestuurder kan dat zo opvatten. In het relationele aspect komt bijvoorbeeld tot uiting of de zender zichzelf onder-, boven- of nevenschikt ziet ten opzichte van de ontvanger. En of de zender expliciet sturend is in zijn boodschap of juist meer ruimte laat (2010, pp.51-54).

2.3.4 *Het relationele aspect uit het four ears-model als kader voor tone of voice*

Zowel de *politeness theory* als het *four ears*-model bieden een kader voor de analyse van de niet-feitelijke aspecten van boodschappen en de bedoelingen, associaties en lading die door *tone of voice* tot uitdrukking worden gebracht. In dit afstudeeronderzoek is het *four ears*-model gekozen om dat kenmerk van *tone of voice* mee te operationaliseren. Die keuze is ingegeven door de omstandigheid dat het *four ears*-model wordt gebruikt in de schrijfvaardigheidstrainingen binnen DNB.

De niet-feitelijke inhoud van de boodschap, de subjectieve uitingen, bedoelingen, associaties en lading die door *tone of voice* tot uitdrukking worden gebracht komen bij uitstek terug in het relationele aspect uit het *four ears*-model. Dus in de manier waarop de zender zichzelf positioneert in zijn relatie met de ontvanger.

De positie die de zender kiest in zijn relatie met de ontvanger kan worden uitgedrukt in twee dimensies, te weten (i) waarderend of afkeurend en (ii) sturend of vrijlatend (Schulz von Thun 2010, pp.51-54). In de trainingen binnen DNB wordt ook de sociale afstand tussen de zender en ontvanger als dimensie onderkend, vergelijkbaar met één van de factoren waarmee een zender zijn beleefdheidsstrategie uit de *politeness theory* kiest.

Voor dit onderzoek wordt het niet-feitelijke in een uiting, dat door *tone of voice* tot uitdrukking wordt gebracht, begripsmatig geanalyseerd in termen van waardering: is het de bedoeling van de schrijver om waardering uit te drukken of afkeuring? en in termen van sturing: is het de bedoeling van de schrijver om directief (sturend) te zijn of juist indicatief en vrijheid te laten?

In paragraaf 3.1 komt aan de orde hoe *tone of voice* technisch kan worden geoperationaliseerd in de genoemde dimensies waardering en sturing.

In de volgende paragraaf, de laatste van dit hoofdstuk, vatten we de theoretische inzichten uit dit hoofdstuk samen in een begripsmatige operationalisering van *tone of voice*.

2.4 Synthese en een conceptuele operationalisering van tone of voice

In paragraaf 2.1 hebben we kennis gemaakt met het begrip *tone of voice* en het voorlopig gedefinieerd als een schrijfstijl waarmee de schrijver niet-feitelijke en specifiek bij de boodschap passende bedoelingen, associaties, lading tot uitdrukking kan brengen. Die stijl is consequent toegepast in de tekst waarin die bepaalde boodschap centraal staat en krijgt concreet vorm door de toegepaste linguïstische tekstkenmerken. En die stijl is tenminste niet discongruent met de *corporate identity* van de schrijver.

In paragraaf 2.2 hebben we aan de hand van linguïstische theorie geconstateerd dat *tone of voice* in geschreven tekst kan worden beschouwd als een grootheid die tot uitdrukking komt in *private states* en omgekeerd dat een *private state* te beschouwen is als een goede indicator voor het voorkomen van *tone of voice*. De analytische eigenschappen van een *private state* (de *experienter*, zijn *attitude*, het *object*, de polariteit van de *attitude* en de intensiteit van de *attitude*) kunnen ook worden toegepast bij de identificatie en analyse van *tone of voice*-voorkomens.

Kijkend naar theoretische communicatiemodellen hebben we in paragraaf 2.3 gezien dat het niet-feitelijke uit de voorlopige definitie van *tone of voice*, evenals het subjectieve uit het concept van *private state*, kan worden uitgewerkt aan de hand van het relationele aspect van een boodschap zoals uiteengezet in het *four ears*-model van Schulz von Thun. *Tone of voice* heeft betrekking op twee relationele dimensies te weten (i) waardering: is het de bedoeling van de schrijver om waardering uit te drukken of afkeuring? en (ii) sturing: is het de bedoeling van de schrijver om directief te zijn of juist indicatief?

Op basis van deze theoretische noties is de volgende conceptuele operationalisering van *tone of voice* af te leiden.

Definitie

Voor dit afstudeeronderzoek wordt *tone of voice* gedefinieerd als een schrijfstijl waarmee de schrijver niet-feitelijke en specifiek bij de boodschap passende bedoelingen, associaties, lading

tot uitdrukking kan brengen. Die stijl is consequent toegepast in de tekst waarin die bepaalde boodschap centraal staat. De schrijfstijl die op *tone of voice* duidt is tenminste niet discongruent met de *corporate identity* van DNB. Die *identity* van DNB wordt gekenmerkt door de waarden stabiel, daadkrachtig, zichtbaar en toonaangevend.

Private state als analytisch kader

In een geschreven tekst is *tone of voice* herkenbaar aan bepaalde combinaties van linguïstische kenmerken. Die combinaties vormen *private states*. Dat wil zeggen er is herkenbaar sprake van een *experiencer*, zijn *attitude*, het *object*, de polariteit van de *attitude* en de intensiteit van de *attitude*.

Private states die op *tone of voice* duiden hebben als *object* de relatie tussen de *experiencer* en degene tot wie hij zich richt. In de context van dit onderzoek, waarin toezichtbrieven van DNB op *tone of voice* worden onderzocht, wordt de rol van de *experiencer* typisch door DNB of door een toezichthouder ingenomen. De *attitude* van dergelijke *private states* is een uitdrukking op het continuüm tussen waardering en afkeuring (van die relatie) of een uitdrukking op het continuüm tussen sturend en vrij latend.

Meetmethode

De meting van *tone of voice* in een toezichtbrief (of corpus van toezichtbrieven) gebeurt door het tellen van de *private states* - in die brief - die duiden op *attitudes* in vier categorieën:

- (i) Een afkeurende, geringschattende houding van de DNB briefschrijver ten aanzien van de relatie tussen DNB en de onder toezicht staande instelling;
- (ii) Een waarderende houding van de DNB briefschrijver ten aanzien van de relatie tussen DNB en de onder toezicht staande instelling;
- (iii) Een sturende, directieve, bovengeschatte positie van de DNB briefschrijver ten opzichte van de onder toezicht staande instelling;
- (iv) Een vrij-latende, indicatieve, gelijkwaardige positie van de DNB briefschrijver ten opzichte van de onder toezicht staande instelling.

Deze vier categorieën zullen in het vervolg worden aangeduid als indicatoren van *tone of voice*.

Op de dimensie "afkeuren↔waarderen" geldt dat hoe meer *private states* worden geteld waar afkeuring uit blijkt des te scherper de *tone of voice* is en hoe meer *private states* worden geteld waar waardering uit spreekt des te milder de *tone of voice* is. Voor de dimensie "sturen↔vrij laten" geldt dat hoe meer sturende-*private states* er worden geteld des te scherper de *tone of voice* is en hoe meer vrij latende-*private states* des te milder is de *tone of voice* is.

Er is sprake van een verandering in *tone of voice* tussen toezichtbrieven uit twee onderscheiden perioden als de meetresultaten op de vier indicatoren veranderen. Met vier indicatoren die elk óf hoger óf lager kunnen zijn zijn er 16 verschillende combinaties mogelijk. In tabel 1 zijn de veranderingen in *tone of voice* tussen twee corpora samengevat als functie van veranderingen in de aantallen *private states* op de indicatoren A (afkeurend), W (waarderend), S (sturend) en V (vrij latend) tussen diezelfde corpora:

Verandering in <i>tone of voice</i>	A		W		S		V	
Is scherper naarmate	(hoger	en/of	lager)	en/of	(hoger	en/of	lager)	
	(hoger	en/of	lager)	en	(hoger	en	hoger)	
	(hoger	en/of	lager)	en	(lager	en	lager)	
	(hoger	en	hoger)	en	(hoger	en/of	lager)	
	(lager	en	lager)	en	(hoger	en/of	lager)	
Is milder naarmate	(lager	en/of	hoger)	en/of	(lager	en/of	hoger)	
	(lager	en/of	hoger)	en	(hoger	en	hoger)	
	(lager	en/of	hoger)	en	(lager	en	lager)	
	(hoger	en	hoger)	en	(lager	en/of	hoger)	
	(lager	en	lager)	en	(lager	en/of	hoger)	
Is niet bepaald als	(hoger	en/of	lager)	en	(lager	en/of	hoger)	
	(lager	en/of	hoger)	en	(hoger	en/of	lager)	
	(hoger	en	hoger)	en/of	(hoger	en	hoger)	
	(hoger	en	hoger)	en/of	(lager	en	lager)	
	(lager	en	lager)	en/of	(hoger	en	hoger)	
	(lager	en	lager)	en/of	(lager	en	lager)	

Tabel 1. Beslissingstabel voor verandering in *tone of voice*

Evaluatie

De evaluatie van het meetinstrument gebeurt deels door de gevonden *tone of voice*-veranderingen in de corpora te vergelijken met verwachtingen vooraf over *tone of voice*-veranderingen. Die verwachting vooraf is dat de *tone of voice* in toezichtbrieven door de jaren heen scherper is geworden. Die verwachting is afgeleid van opvattingen van vertegenwoordigers van onder toezicht staande instellingen. In opdracht van DNB wordt periodiek een onafhankelijk stakeholderanalyse uitgevoerd onder banken, pensioenfondsen, verzekeraars en trustkantoren. De meest recente stakeholderanalyse is uitgevoerd in 2014. Die wijst uit dat vertegenwoordigers van onder toezicht staande instellingen een verschil ervaren tussen de *tone of voice* van de toezichtbrief en de *tone of voice* van het corresponderende toezichtgesprek over hetzelfde

onderwerp (DNB, 2015, p.12). Deze uitkomst is voor dit onderzoek bewerkt tot de verwachting dat de *tone of voice* in toezichtbrieven door de jaren heen scherper is geworden.

De evaluatie, op deze manier, is beperkt omdat de vergelijking van de gevonden *tone of voice*-veranderingen met deze verwachtingen onzuiverheden kent. Een eerste bron van onzuiverheid is de wisseling van perspectief in het vergelijk: de meetresultaten in dit onderzoek gaan uit van een conceptuele operationalisering van *tone of voice* vanuit het schrijverperspectief. De vooraf gestelde verwachtingen over verandering van *tone of voice* - waarmee de gevonden meetresultaten worden vergeleken - zijn gebaseerd op het perspectief van de lezers van toezichtbrieven. Het effect van die perspectiefwisseling is niet onderzocht noch zijn daarover aannames gemaakt. Ten tweede is de vooraf opgestelde verwachting slechts een afgeleide van de uitkomsten van de laatste stakeholderanalyse. De observaties uit de stakeholderanalyse hebben namelijk betrekking op een - door de brieflezers gepercipieerd - verschil tussen *tone of voice* in een brief en *tone of voice* in het gesprek over hetzelfde onderwerp. Daarbij komt bovendien dat stakeholders niet direct bevraagd zijn over *tone of voice*, maar dat de bevinding in termen van *tone of voice* zich als conclusie opdrong aan de onderzoekers. Ten derde zijn de uitkomsten van de stakeholderanalyse niet precies gekaderd in de tijd; het is niet duidelijk op welke periode de opmerkingen over *tone of voice* in brief en gesprek betrekking hebben.

Deze onzuiverheden maken dat de vergelijking tussen de in dit onderzoek gevonden meetresultaten en de vooraf gestelde verwachtingen maar beperkt kan gelden als betrouwbare basis voor de evaluatie van het meetinstrument.¹¹ Aanvullende evaluaties moeten worden gezocht in de bespreking met de experts en in een kritische beschouwing van de praktijkervaringen met het ontwikkelde meetinstrument tijdens de daadwerkelijke uitvoering van het onderzoek.

In het volgende hoofdstuk, over de onderzoeksmethode, komt aan bod hoe concreet uitvoering wordt gegeven aan deze conceptuele operationalisering van *tone of voice*.

¹¹ Voor conclusies over de oorzaken van *tone of voice*-veranderingen in toezichtbrieven is dit vergelijk sowieso niet geschikt. Maar die vraag, naar de oorzaken van *tone of voice*-veranderingen, ligt ook niet in de kern van dit onderzoek.

3 Onderzoeksmethode en technische uitvoering

In dit hoofdstuk beschrijven we de activiteiten en overwegingen bij de technische uitvoering van het onderzoek.

In de doelstelling van dit onderzoek (paragraaf 1.3.1) wordt verwezen naar de "toepassing van *subjectivity analysis*-technieken". Vandaar dat we in dit hoofdstuk eerst in gaan op enkele algemene uitvoeringsaspecten van *subjectivity analysis*-onderzoeken, aan de hand van wetenschappelijke publicaties uit dat onderzoeksterrein (paragraaf 3.1).

Vervolgens zien we hoe daadwerkelijk uitvoering wordt gegeven aan dit specifieke onderzoek. Dat begint met het prepareren van de te onderzoeken toezichtbrieven en het groeperen daarvan in drie corpora (paragraaf 3.2). De inhoudelijk meest uitdagende werkstroom is het samenstellen van lexicons met signaalwoorden die duiden op *tone of voice* (paragraaf 3.3). Corresponderend met de conceptuele operationalisering van *tone of voice* worden vier lexicons samengesteld met signaalwoorden voor afkeurende, waarderende, sturende en vrij latende *private states*. De daadwerkelijke doorzoekingen van de drie corpora op de aanwezigheid van signaalwoorden uit die lexicons voeren we uit met het computerprogramma WordSmith (paragraaf 3.4). Alvorens de resultaten uit die doorzoekingen te interpreteren in termen van een eventuele *tone of voice*-verandering beoordelen we de validiteit ervan (paragraaf 3.5). De interpretatie van de gevalideerde meetresultaten en de vertaling naar verandering van *tone of voice* komt ten slotte aan het eind van het hoofdstuk aan bod (paragraaf 3.6).

In het onderzoeksontwerp, dat aan de uitvoering van dit onderzoek ten grondslag ligt, is de uitvoering van het onderzoek op hoofdlijnen geschetst. Het onderzoeksmodel daaruit is nogmaals opgenomen in bijlage C bij dit referaat, de geplande onderzoeksaanpak is nogmaals opgenomen in bijlage D.

3.1 Nogmaals: subjectivity analysis

3.1.1 Hoofdstappen van subjectivity analysis

In *subjectivity analysis* staat de *private state* centraal als aangrijpingspunt voor onderzoek. Om concrete *private states* in bestaande corpora te kunnen identificeren en construeren moeten in de *subjectivity analysis* twee hoofdstappen worden uitgevoerd (Wiebe e.a. 2006, p.3), te weten het verzamelen van signaalwoorden en het doorzoeken van teksten op die signaalwoorden.

De eerste hoofdstap is het verzamelen van woorden die een lading van subjectiviteit hebben. Dergelijke woorden, woordvormen en zinsneden zullen we hierna aanduiden met "signaalwoorden", de verzameling van signaalwoorden noemen we "lexicon". In de volgende subparagraaf wordt het samenstellen van een lexicon nader toegelicht.

In de tweede hoofdstap worden de te onderzoeken corpora geanalyseerd op het voorkomen van de signaalwoorden uit het lexicon. Hiervoor wordt gebruik gemaakt van computerprogramma's. Zo'n analyse leidt typisch tot output in de vorm van beschrijvende statistieken zoals het aantal keren dat een signaalwoord voorkomt, en tot concordanties, dat zijn de zinsneden waarin een signaalwoord voorkomt (dus het signaalwoord in z'n directe context).

3.1.2 Lexicon: bestand van signaalwoorden

De essentie van een *subjectivity analysis* zit in de eerste hoofdstap: het samenstellen of verkrijgen van een lexicon met passende signaalwoorden. Er zijn ruwweg drie manieren om een lexicon te verkrijgen: (i) handmatig signaalwoorden aanwijzen in bestaande teksten; (ii) geautomatiseerd signaalwoorden laten aanwijzen in bestaande teksten; (iii) gebruik maken van een al bestaande lexicon.

Handmatig samenstellen

Het handmatig samenstellen van een lexicon gebeurt door het doorzoeken van bestaande teksten op signaalwoorden. Dit proces heet annoteren. Bij annoteren wordt een verzameling van bestaande teksten (de zogenaamde trainingscorpus) geanalyseerd door een panel van beoordelaars (*annotators*). De *annotators* identificeren en classificeren - ieder voor zich - die woorden die zij associëren met subjectiviteit. Om individuele interpretatieverschillen te reduceren wordt elke tekst door meerdere *annotators* beoordeeld waarna de verschillen in overleg tussen de *annotators* worden overbrugd of uitgesloten. Veelal wordt in twee of drie substapjes geannoteerd: eerst wordt van woorden bepaald of ze subjectiviteit dragen of niet, dan worden de woorden die subjectiviteit uitdrukken geclassificeerd naar de aard van het sentiment (de polariteit) dat ermee

wordt uitgedrukt: positief of negatief, en soms ook naar de mate (intensiteit) van dat sentiment: weinig of veel.

In wetenschappelijk onderzoek wordt geëxperimenteerd met variaties op deze algemene aanpak voor handmatige annotatie, waarbij de betrouwbaarheid van de toegepaste annotatiemethode veel aandacht krijgt. Een gangbare indicator is het aantal annotatieverschillen tussen twee *annotators* over dezelfde tekst: hoe eensgezinder de *annotators*, des te betrouwbaarder de methode (Wilson & Wiebe 2005; Wilson, Wiebe & Hofman 2005; Al-Subaihin e.a. 2011; Shin e.a. 2012;).

Een voorbeeld van een handmatig samengestelde lexicon is de zogenaamde *Subjectivity Clues Lexicon* (Wilson, Wiebe & Hofman 2005). Figuur 2 toont een aantal regels uit dat lexicon.

```
type=strongsubj len=1 word1=amusing pos1=anypos stemmed1=y priorpolarity=positive
type=strongsubj len=1 word1=amusingly pos1=anypos stemmed1=y priorpolarity=positive
type=weaksubj len=1 word1=analytical pos1=adj stemmed1=n priorpolarity=neutral
type=strongsubj len=1 word1=anarchism pos1=noun stemmed1=n priorpolarity=negative
type=strongsubj len=1 word1=anarchist pos1=noun stemmed1=n priorpolarity=negative
type=strongsubj len=1 word1=anarchistic pos1=adj stemmed1=n priorpolarity=negative
type=strongsubj len=1 word1=anarchy pos1=noun stemmed1=n priorpolarity=negative
type=weaksubj len=1 word1=anemic pos1=adj stemmed1=n priorpolarity=negative
type=weaksubj len=1 word1=angel pos1=noun stemmed1=n priorpolarity=positive
type=strongsubj len=1 word1=angelic pos1=adj stemmed1=n priorpolarity=positive
type=strongsubj len=1 word1=anger pos1=adj stemmed1=n priorpolarity=negative
type=strongsubj len=1 word1=anger pos1=noun stemmed1=n priorpolarity=negative
type=strongsubj len=1 word1=anger pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=angrily pos1=anypos stemmed1=y priorpolarity=negative
```

Figuur 2 Fragment uit de *Subjectivity Clues* lexicon (Wilson e.a. 2005)

De *Subjectivity Clues* lexicon telt circa 8.200 Engelstalige signaalwoorden, elk signaalwoord voorzien van enkele linguïstische kenmerken. Elke regel in het bestand is opgezet volgens een vast stramien, om bewerking met computerprogrammatuur mogelijk te maken. Het signaalwoord zelf wordt voorafgegaan door het kenmerk "word1=". In de figuur zijn de regels van de signaalwoorden *amusingly*, *analytical* en *anarchism* gemarkeerd. Het kenmerk dat wordt aangeduid met "type=" (aan het begin van elke regel) geeft aan of het signaalwoord is aangemerkt als een sterke of een zwakke uitdrukking van subjectiviteit. De polariteit van het signaalwoord is weergegeven bij het kenmerk "priorpolarity=" (aan het eind van elke regel). In bijlage D is de "Readme" opgenomen die bij deze lexicon hoort.

Het signaalwoord *amusingly* draagt volgens deze lexicon dus een sterke subjectiviteit en een positieve polariteit, het signaalwoord *anarchism* is eveneens sterk maar dan met een negatieve polariteit, en het signaalwoord *analytical* is aangemerkt als zwaksubjectief met een neutrale polariteit. Het signaalwoord *orthodoxy* (niet in de figuur) heeft eveneens een neutrale polariteit

maar is als sterk subjectief geannoteerd.

Geautomatiseerd samenstellen

De samenstelling van een lexicon op basis van handmatige annotatie is in opzet inzichtelijk en met betrekkelijk eenvoudige middelen uitvoerbaar, maar wel erg arbeidsintensief. Sinds de jaren '90 worden alternatieven ontwikkeld basis van kunstmatige intelligentie (Sebastiani 2002). Deze ontwikkelingen liggen in het domein van de computationele taalkunde. Die alternatieven maken de uitvoering van annotaties veel efficiënter, maar vergen zeer specialistische kennis voor het ontwerpen en programmeren van de software.

In hoofdzaak zijn tenminste twee typen te onderscheiden (Sebastiani 2002, pp.8-9). Het eerste type bouwt voort op de op het principe van kennissystemen en wordt ook wel aangeduid met de term *knowledge engineering* (KE). Bij kennissystemen wordt de inhoudelijke kennis van domeindeskundigen - bijvoorbeeld een *annotator*, dus van iemand die heel goed is in het herkennen van woorden die subjectiviteit uitdrukken - geëxpliciteerd, uiteengehaald en in de computer gecodeerd. Om in het voorbeeld te blijven: de *annotator* wordt door de KE-analist uitgehooard over hoe hij subjectiviteit herkent. De KE-analist zet die kennis om in elementaire beslisregels in het kennissysteem. Het idee is dat het kennissysteem alle beslisregels bevat die de domeindeskundigen impliciet en/of onbewust toepassen. Na testen van het kennissysteem kan dat dan de taken overnemen van die domeindeskundigen.

Het andere type oplossing gaat uit van het principe van zelflerende computernetwerken. Bij deze zogenaamde *machine learning* (ML) techniek wordt de domeindeskundige ingezet als trainer van in computertaal geprogrammeerde algoritmes. Die algoritmes zijn erop gericht om patronen te herkennen in teksten vervolgens die teksten te classificeren aan de hand van criteria die de domeindeskundige vooraf heeft opgegeven. In een aantal trainingsiteraties - telkens met andere teksten - beoordeelt de deskundige de output van het algoritme en geeft zijn correcties aan het algoritme terug. Dat is zodanig opgezet dat het zijn onderscheidend vermogen kan verbeteren met de feedback van de deskundige. Na verloop van tijd heeft het algoritme zichzelf zodanig "deskundig" gemaakt dat het idealiter de taken van de deskundige kan overnemen.

Bestaande lexicons

Bestaande lexicons voor *subjectivity analysis*, zoals bijvoorbeeld de *Subjectivity Clues* lexicon, zijn - hoewel niet talrijk - goed te vinden via referenties in academische publicaties. Het betreft overwegend onderzoeksmateriaal (corpora en lexicons) uit analyses van Engelstalige teksten.

Voor Nederlandstalig tekstonderzoek zijn weliswaar bestaande corpora en lexicons beschikbaar,

maar die zijn niet goed bruikbaar voor *subjectivity analysis*. Hoeksema heeft een lexicon samengesteld van meer dan 800 negatief-polaire Nederlandstalige uitdrukkingen (Hoeksema 2013). Hoewel hij geen mededeling doet over de toepasbaarheid van het lexicon in onderzoek van geschreven danwel gesproken taal, lijken de uitdrukkingen zelf vooral ontleend aan spreektaal of aan populaire schrijftaal. Voor onderzoeksdoeleinden stelt de Nederlandse Taalunie gratis corpora en lexicons beschikbaar (waaronder Referentiebestand Nederlands; SoNaR; Cornetto) maar deze bronnen bevatten geen informatie over polariteit of intensiteit van signaalwoorden.¹²

Na deze bespiegeling op enkele algemene uitvoeringsaspecten van *subjectivity analysis*-onderzoeken, gaan we in de resterende paragrafen van dit hoofdstuk in op de daadwerkelijke uitvoering van dit specifieke onderzoek. Dat begint, in de volgende paragraaf, met het prepareren van de te onderzoeken toezichtbrieven en het groeperen daarvan in drie corpora.

3.2 Corpora van toezichtbrieven prepareren

De teksten die in dit onderzoek op *tone of voice* worden onderzocht zijn toezichtbrieven. Een toezichtbrief wordt in dit verband gedefinieerd als een brief van DNB in haar capaciteit als toezichthoudende instantie, aan een bepaalde onder toezicht staande instelling, over de uitkomsten van een toezichtonderzoek dat betrekking had op die instelling. In een toezichtbrief zijn bevindingen uit het onderzoek vermeld en onderbouwd, de conclusies die DNB daar aan verbindt zijn erin vermeld en zonodig zijn de eventuele acties beschreven die de instelling moet ondernemen.

Toezichtbrieven worden primair opgesteld door de toezichthouder voor de betreffende instelling en gereviewed door een tweede toezichthouder of door het hoofd van de toezichtafdeling, ze worden in de regel ondertekend door de toezichthouder en een leidinggevende.

Voor dit onderzoek zijn toezichtbrieven onderzocht die zijn verstuurd in de jaren 2006, 2010 en 2014. De keuze van die jaren is ingegeven door de overweging om (i) een redelijk tijdsverschil te hebben tussen twee jaargangen; (ii) brieven van vóór de cultuurverandering van toezicht (ingezet in 2010, zie DNB 2010) te kunnen vergelijken met brieven van erna; en (iii) de beschikbaarheid van ongeveer evenveel brieven in elke jaargang.

¹² http://tst-centrale.org/nl/producten/0?category_id=0&page=shop.browse&limit=20&limitstart=0

Binnen die jaargangen 2006, 2010 en 2014 is de selectie van brieven gebaseerd op twee criteria. Ten eerste moest het gaan om brieven die de weerslag zijn van een afgerond toezichtonderzoek. Het is goed om hierbij te bedenken dat DNB ook andersoortige brieven stuurt. Bijvoorbeeld aankondigen van onderzoek of brieven met een verzoek om informatie op te leveren. De tweede voorwaarde is dat er ook een toezichtgesprek heeft plaatsgevonden tussen toezichthouder en in stelling over de uitkomsten van het onderzoek. Dit criterium hangt direct samen met de oorspronkelijke aanleiding voor het onderzoek; namelijk dat brieflezers een verschil ervaren in *tone of voice* van de brief en *tone of voice* tijdens een corresponderend gesprek met de toezichthouder over hetzelfde onderwerp.

Toezichtbrieven zijn strikt vertrouwelijk, zoals alle communicatie tussen een onder toezicht staande instelling en DNB. Voor dit onderzoek zijn de brieven geanonimiseerd zodat er geen verwijzing meer is af te leiden naar de instelling. Concreet betekent dit dat minimaal de instellingsnaam, persoonsnamen, alsmede verwijzingen naar specifieke producten, systemen, projecten, dossiers, documenten, datums en tijdvakken zijn weggenomen. Ten behoeve van de verwerking in het programma WordSmith zijn de brieven opgeslagen als .txt-bestand.

Uit elk van de genoemde jaren zijn circa 15 toezichtbrieven geselecteerd en per jaargang in één corpus ondergebracht.

Die drie corpora met toezichtbrieven worden doorzocht op het voorkomen van signaalwoorden. In de volgende paragraaf zullen we zien hoe het verzamelen van signaalwoorden en het samenstellen van lexicons in z'n werk gaat.

3.3 Lexicons van signaalwoorden samenstellen

De vier lexicons, met signaalwoorden voor respectievelijk afkeurende, waarderende, sturende en vrij latende *private states*, zijn handmatig samengesteld. Ze bevatten signaalwoorden waarvan verwacht wordt dat die concreet uitdrukking geven aan het relationele aspect van een boodschap zoals uiteengezet in het *four ears*-model van Schulz von Thun.

Circa 30 toezichthouders hebben signaalwoorden geselecteerd uit een trainingscorpus van toezichtbrieven die geen deel uitmaken van de drie te onderzoeken corpora. Deze toezichtbrieven waren vooraf geanonimiseerd en opgeknipt in fragmenten. Elk fragment is voorgelegd aan twee of drie toezichthouders met het verzoek om die woorden en woordcombinaties aan te strepen die uitdrukking geven aan één van de indicatoren afkeurend, waarderend, sturend (directief) of vrij

latend (indicatief).

Die annotatie is uitgevoerd tijdens een al geplande interne bijeenkomst van de toezichthouders. In het programma van die bijeenkomst, over professionele communicatie, kon een half uur worden ingeruimd om de signaalwoorden aan te strepen en te classificeren. Voorafgaand aan de annotatie is een korte toelichting gegeven op het *four ears*-model en het onderscheid tussen de vier indicatoren. Bijlage F bevat de instructie aan de *annotators*.

Helaas was er geen gelegenheid voor de toezichthouders om *interannotator*-verschillen onderling te bespreken. De aangestreepte signaalwoorden én de daarbij opgegeven classificatie zijn daarom in eerste instantie ongewijzigd overgenomen en ondergebracht in vier lexicons: één met afkeurende signaalwoorden, één met waarderende woorden, één lexicon voor sturing en één met signaalwoorden die uitdrukking geven aan vrij laten. Daarbij zijn de lexicons bovendien uitgebreid met andere woordvormen van de erin opgenomen geldige signaalwoorden. Zo is bijvoorbeeld naast "u dient" ook "dient u" in het betreffende lexicon opgenomen. Ten behoeve van de verwerking in het programma WordSmith zijn de lexicons opgeslagen als .txt-bestand.

De samenstelling van de lexicons zal een wezenlijk effect blijken te op de meetresultaten en daarmee op de uitkomsten van het onderzoek. Meer hierover in paragraaf 4.1, over de validatie van de eerste meetresultaten, en in paragraaf 5.2, over de kritiek van experts op de omvang van de lexicons.

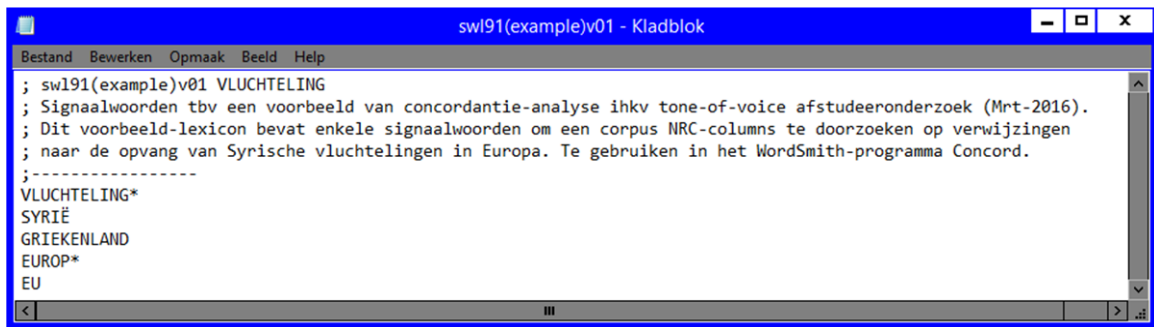
Als we voor elk van de vier indicatoren van *tone of voice* aparte lexicons met signaalwoorden hebben kunnen we de drie corpora met toezichtbrieven daadwerkelijk doorzoeken op het voorkomen van die signaalwoorden. Dat doen we met het programma WordSmith. In de volgende paragraaf zien we hoe dat gaat.

3.4 Analyses uitvoeren in WordSmith

Teksten - bijeengebracht in corpora - kunnen met behulp van het programma WordSmith Concord worden doorzocht op het voorkomen van signaalwoorden uit één of meerdere lexicons. Dit programma telt per corpus/lexicon-combinatie het aantal keren dat de signaalwoorden uit het lexicon voorkomen in de teksten van het corpus, en geeft bovendien per gevonden voorkomen de letterlijke directe context van het signaalwoord.

Een voorbeeld ter verduidelijking: stel we willen een corpus van 18 columns uit het NRC van

najaar 2015 onderzoeken op verwijzingen naar de opvang van Syrische vluchtelingen in Europa. Daartoe hebben we een lexicon samengesteld met de signaalwoorden "vluchteling*", "Syrië", "Europ*" en "EU", zie figuur 3. Merk op dat WordSmith een asterisk als *wildcard* interpreteert; aan de hand van het signaalwoord "vluchteling*" worden ook woorden zoals "vluchtelingen", "vluchtelingencrisis" en "vluchtelingendebat" als treffer aangemerkt.



Figuur 3: Lexiconbestand swl91(example)v01

In het programma WordSmith Concord koppelen we het lexicon (technisch vormgegeven door een .txt-bestand) aan de corpus van de 18 columnteksten.

De doorzoeking door WordSmith Concord levert onder meer de volgende uitkomsten op. In het tabblad "plot" (zie figuur 4) is op de bovenste regel vermeld dat de gehele corpus (alle 18 columns samen) 8.145 woorden telt, waarvan 51 gelden als treffer van signaalwoorden uit het lexicon, ofwel een overall trefferfrequentie van 6,26 treffers per 1000 woorden tekst.

N	File	Words	Hits	per 1,000
1	Overall	8.145	51	6,26
2	VLUCHTELING tc NRCcol	712	1	1,40
3	VLUCHTELING tc NRCcol	714	3	4,20
4	VLUCHTELING tc NRCcol	715	1	1,40

Figuur 4: Tabblad "plot" uit de WordSmith-output

De overall trefferfrequentie van 6,26 treffers per 1000 woorden tekst geeft enige indicatie van de aandacht die de columnisten in het najaar van 2015 hebben besteed aan de opvang van Syrische

vluchtelingen in Europa. Als we zo'n zelfde doorzoeking maken van NRC-columns uit bijvoorbeeld het najaar van 2008 dan kunnen we door vergelijking van de trefferfrequenties een indruk krijgen van de verandering in aandacht voor dit onderwerp tussen beide periodes.

In het onderzoek naar verandering in *tone of voice* in toezichtbrieven hebben we precies ook die aanpak gevolgd: door trefferfrequenties van *tone of voice*-signaalwoorden in brieven uit verschillende jaren te vergelijken kunnen we een verandering van *tone of voice* onderkennen. Althans, voor zover de gevonden treffers allemaal geldige verwijzingen zijn naar *tone of voice*. De validiteitstoetsen ter beoordeling daarvan komen in de volgende paragraaf aan de orde.

3.5 Validiteit toetsen en zo nodig trefferfrequenties corrigeren

Alvorens de trefferfrequenties van verschillende corpora te interpreteren in termen van een eventuele *tone of voice*-verandering moet worden beoordeeld in hoeverre de onderliggende treffers geldige verwijzingen zijn naar *tone of voice*.

3.5.1 Signaalwoordvaliditeit en treffervaliditeit

Begripsmatig kan onderscheid worden gemaakt tussen ongeldige treffers als gevolg van een signaalwoord dat (bij nader inzien) systematisch geen goede aanduiding blijkt te zijn van *tone of voice* en ongeldige treffers omdat de betekenis van zo'n treffer in zijn specifieke context geen goede aanduiding is van *tone of voice*. Het eerste type duiden we hier aan met signaalwoordongeldigheid en het tweede met trefferongeldigheid.

Signaalwoordongeldigheid is aan de orde als de voorkomens van een bepaald signaalwoord systematisch niet voldoen aan de criteria uit de betreffende operationalisering van *tone of voice*. Ondanks de zorg en beste bedoelingen waarmee de signaalwoorden voor die operationalisering zijn verzameld kunnen daarin toch systematische fouten optreden. Voor de analyse op basis van het *four ears*-model is elk signaalwoord geïdentificeerd in een annotatie met onervaren *annotators* waarbij bovendien geen gelegenheid was voor bespreking van *interannotator*-verschillen. Bij de alternatieve analyse (op basis van de schrijfrichtlijnen) zijn de signaalwoorden vertaald uit een Engelstalige lexicon. Het is in beide gevallen niet op voorhand uit te sluiten dat signaalwoorden in de te onderzoeken corpora systematisch een "vals alarm" afgeven.

Als nu blijkt dat een signaalwoord in de te onderzoeken corpora systematisch niet verwijst naar wat in de betreffende *tone of voice*-operationalisering was bedoeld, dan geeft de trefferfrequentie

van dat signaalwoord een verkeerd beeld van de *tone of voice* volgens die operationalisering. Het betreffende signaalwoord wordt dan als ongeldig aangemerkt en moet uit het lexicon worden verwijderd. Alle drie de corpora worden dan opnieuw geanalyseerd met het aangepaste lexicon.

De trefferongeldigheid gaat over situaties waarin een bepaald voorkomen van een signaalwoord in z'n specifieke context niet voldoet aan de criteria uit de operationalisering van *tone of voice*. Zo kan bijvoorbeeld de aanwezigheid van een negatie in het directe zinsverband van de treffer de betekenis ervan zodanig verdraaien dat die niet past binnen de criteria. Of, bij de eerste analyse op basis van het *four ears*-model, uit het zinsverband van de treffer blijkt dat de *private state* betrekking heeft op een andere *experiencer* dan DNB waardoor die *private state* niet meer voldoet aan de criteria.

Als blijkt dat treffers van eenzelfde signaalwoord incidenteel, in specifieke zinsverbanden, ongeldig zijn, en in andere zinsverbanden wel geldig, dan moeten alleen die treffers als ongeldig worden aangemerkt en in mindering worden gebracht van de trefferfrequentie. Het signaalwoord kan dan als geldig signaalwoord in het lexicon worden gehandhaafd.

3.5.2 Twee kwalitatieve validiteitscriteria

Stel dat van een bepaald signaalwoord 100 treffers zijn gevonden. Uit de interpretaties van elk van de zinsverbanden blijkt dat alle 100 treffers recht doen aan de van toepassing zijnde operationalisering van *tone of voice*. Stel verder dat van een tweede signaalwoord ook 100 treffers zijn gevonden waarvan er 23 niet geldig zijn. En dat een derde signaalwoord ook 100 treffers heeft waarvan er 58 ongeldig zijn. En stel ten slotte dat bij een vierde signaalwoord met 100 treffers blijkt dat 90 treffers niet voldoen. Wat zeggen deze geldigheidstoetsen nu over de geldigheid van de signaalwoorden of over die van hun treffers?

Alleen voor het eerste signaalwoord kan worden geconcludeerd dat er geen sprake is van signaalwoordongeldigheid noch van trefferongeldigheid. Op zich zou voor elk van de drie andere signaalwoorden kunnen worden betoogd dat er sprake is van signaalwoordongeldigheid of alleen van trefferongeldigheid. Zonder aanvullende criteria is het begripsmatige onderscheid tussen signaalwoordongeldigheid (systematisch ongeldig) en trefferongeldigheid (incidenteel ongeldig) niet te maken.

In dit onderzoek is het begripsmatige onderscheid tussen signaalwoordongeldigheid en trefferongeldigheid uitgewerkt in kwalitatieve criteria. We zullen een signaalwoord als ongeldig aanmerken als (i) een groot deel van zijn treffers niet aansluit op de bedoeling van de betreffende *tone of voice*-operationalisering, en (ii) er daarvoor hooguit een beperkt aantal algemene oorzaken is aan te wijzen. We zullen concluderen dat er sprake is van trefferongeldigheid als (i) een goed

deel van de andere treffers van hetzelfde signaalwoord wel geldig is, en (ii) er niet veel andere ongeldige treffers zijn die om een vergelijkbare oorzaak als ongeldig zijn aangemerkt. Aan deze criteria liggen twee validiteitindicatoren ten grondslag. De eerste indicator is het aantal ongeldige treffers van een bepaald signaalwoord ten opzichte van het totaal aantal treffers van datzelfde signaalwoord. De tweede indicator is het aantal ongeldige treffers van een bepaald signaalwoord dat om dezelfde reden ongeldig is verklaard ten opzichte van het totaal aantal ongeldige treffers van dat signaalwoord. De beide indicatoren en de criteria voor ongeldigheid van signaalwoord of trefter zijn weergegeven in tabel 2.

<i>Type ongeldigheid</i>	Aantal ongeldige treffers van het signaalwoord t.o.v. totaal aantal treffers van dat signaalwoord		Aantal ongeldige treffers van het signaalwoord als gevolg van dezelfde oorzaak t.o.v. het totaal aantal ongeldige treffers van dat signaalwoord
Signaalwoord-ongeldigheid als	veel	en	veel
Trefterongeldigheid als	weinig	en	weinig
Ad hoc te bezien als	veel	en	weinig
	weinig	en	veel

Tabel 2. Beslissingstabel voor het onderscheid tussen signaalwoord- en trefterongeldigheid

3.5.3 Concordantieanalyse om de validiteit te beoordelen

De beoordeling van de geldigheid van een trefter is gebaseerd op de analyse van de betekenis van de trefter in zijn directe zinsverband. Die betekenis moet recht doen aan de operationalisering van *tone of voice* zoals afgeleid in paragraaf 2.4. Dus uit het zinsverband moet blijken dat (i) de trefter betrekking heeft op de relatie tussen DNB en de onder toezicht staande instelling (het *object* uit de *private state*), (ii) wijst op een uiting van DNB (de *experiencer*) die (iii) subjectief is in één van de vier indicatoren afkeurend, waarderend, sturend en vrij latend (de *attitude*). Zowel voor de treffervaliditeit als voor de signaalwoordvaliditeit wordt de betekenis van de trefter in zijn directe zinsverband bezien. Zo'n direct zinsverband van een trefter wordt concordantie genoemd.

Voortbouwend op het voorbeeld uit de vorige paragraaf kan de uitvoering van een concordantieanalyse worden uitgelegd. In het tabblad "concordance" van de WordSmith-output (zie figuur 5) staan de 51 treffers uit het voorbeeld weergegeven in concordantie. De trefter zelf staat in het midden van de regel, links en rechts geflankeerd door de circa 10 à 12 voorafgaande en opvolgende woorden uit de lopende tekst zoals die in de corpus is opgenomen.



Bij de derde concordantie is dat minder duidelijk. Daar gaat het over tienduizenden die kennelijk vluchtten voor de Russische bezetter. Hoe dat in verband staat met de opvang van Syrische vluchtelingen moet nader worden gezien. Door op die regel te dubbelklikken kunnen we in het tabblad "source text" de oorspronkelijke tekst zien. Daaruit blijkt dat het tekstfragment betrekking heeft op de Hongaarse opstand in 1956 (figuur 6). Deze treffer is dus niet geldig als verwijzing naar de opvang van Syrische vluchtelingen.



pagina 39 van 79

in de New York Times over de opvang van vluchtelingen in Oost Europese landen. Hier zouden we nadere criteria willen bedenken om uit te maken of deze treffers nu wel of niet valide zijn voor het tekstonderzoek naar opvang van Syrische vluchtelingen in Europa.

3.5.4 *Meerdere iteraties van doorzoeken, valideren en corrigeren*

Door per treffer de concordantie te analyseren kan worden beoordeeld of die geldig is en bij ongeldigheid wat de oorzaak voor die ongeldigheid is. Aan de hand van de beslissingstabel 2 kan wordt bepaald of er sprake is van trefferongeldigheid of van signaalwoordongeldigheid.

In het geval van trefferongeldigheid moet de ongeldige treffer in mindering worden gebracht van de trefferfrequentie. In het geval van signaalwoordongeldigheid moet het signaalwoord uit het lexicon worden verwijderd en een nieuwe doorzoeking worden gedaan (in WordSmith) met het aangepaste lexicon. De validiteit van de uitkomsten van die nieuwe doorzoeking wordt dan weer op dezelfde manier beoordeeld, zo nodig weer gevolgd door aanpassing van het lexicon en in dat geval ook door een nieuwe doorzoeking. Dit proces herhaalt zich net zolang totdat er geen sprake meer is van signaalwoordongeldigheid. De uitkomsten van de laatste doorzoeking worden op treffervaliditeit beoordeeld en zo nodig worden de trefferfrequenties gecorrigeerd.

Al met al resulteert dan een per doorzoeking trefferfrequentie die vrij is van systematische ruis doordat verkeerde signaalwoorden zijn gebruikt en die vrij is van incidentele ruis doordat de treffer nét niet aangaf wat werd verwacht.

Aan het eind van ons onderzoek kunnen we dergelijke gecorrigeerde trefferfrequenties gebruiken om tot een uitspraak te komen over een eventuele verandering in *tone of voice* - volgens de van toepassing zijnde operationalisering - in de onderzochte corpora. Hoe we uitspraak afleiden zien we in de volgende paragraaf, de laatste in dit hoofdstuk over de onderzoeksmethode en technische uitvoering.

3.6 **Interpretatie van de meetresultaten**

In de analyse met WordSmith wordt elk van de drie te onderzoeken corpora doorzocht (mogelijk in meerdere iteraties) op het voorkomen van signaalwoorden uit elk van de lexicons afkeurend, waarderend, sturend, vrij latend. Voor elk van die vier indicatoren worden de gecorrigeerde trefferfrequenties uit de verschillende corpora met elkaar vergeleken. De eventuele veranderingen, per indicator, van gecorrigeerde trefferfrequenties worden vertaald in veranderingen van *tone of*

voice aan de hand van de beslisregels in tabel 1 in paragraaf 2.4.

Zo zou bijvoorbeeld kunnen worden geconcludeerd dat de *tone of voice* in het 2010-corpus scherper is dan die in het 2006-corpus als - *ceteris paribus* - in het 2010-corpus een hoger aantal afkeurende *private states* is geteld dan in het 2006-corpus. Of indien in het 2010-corpus een lager aantal waarderende *private states* is geteld dan in het 2006-corpus. Of bij een combinatie van deze twee uitkomsten. Ook als in het 2010-corpus een hoger aantal sturende *private states* is geteld dan in het 2006-corpus is de *tone of voice* scherper geworden. Of als er minder vrij latende *private states* zijn geteld. Vergelijkbare, maar precies omgekeerde, relaties zijn af te leiden voor situaties waarin de *tone of voice* milder is in het ene corpus ten opzicht van het andere corpus.

In het volgende hoofdstuk zien we tot welke concrete uitkomsten de onderzoeken hebben geleid, tot welke uitspraken over verandering van *tone of voice* ze aanleiding geven, en hoe die uitspraken zich verhouden tot de vooraf gestelde verwachtingen.

4 Meetresultaten

In dit hoofdstuk presenteren we de meetresultaten uit de daadwerkelijke doorzoekingen van de drie corpora van toezichtbrieven.

De doorzoekingen zijn uitgevoerd zoals beschreven in het vorige hoofdstuk. Er waren twee iteraties nodig om tot gevalideerde trefferfrequenties te komen. We behandelen eerst - voor beide iteraties - de ruwe uitkomsten en correcties daarop (paragrafen 4.1 en 4.2). Daarna vertalen we de gevalideerde meetuitkomsten uit de tweede iteratie naar veranderingen in *tone of voice* (paragraaf 4.3). We grijpen daarvoor terug op de beslistabel uit de conceptuele operationalisering in hoofdstuk 2. Aan het eind van dit hoofdstuk vergelijken we de aldus gevonden veranderingen in *tone of voice* met de verwachtingen die daarover vooraf waren gesteld, en we bespreken de onzuiverheden bij die vergelijking (paragraaf 4.4).

4.1 Eerste iteratie met initiële lexicons en validiteitbeoordeling

De eerste doorzoeking, op basis van de door het panel samengestelde initiële lexicons, resulteerde in 488 treffers over alledrie de corpora heen (tabel 3).

indicator ---->		AFKEUREND		WAARDEREND		STUREND		VRIJ LATEND		
initieel lexicon ---->		swl03(pnlA)02		swl03(pnlW)02		swl03(pnlS)02		swl03(pnlV)02		
metric ---->		[hits] [p.1000w]		[hits] [p.1000w]		[hits] [p.1000w]		[hits] [p.1000w]		
corpus	# words									
toezichtbrieven uit 2006	26.252	45	1,714	1	0,038	69	2,628	0	0,000	115
toezichtbrieven uit 2010	29.079	120	4,127	0	0,000	68	2,338	0	0,000	188
toezichtbrieven uit 2014	23.896	125	5,231	6	0,251	53	2,218	1	0,042	185
totaal aantal treffers ---->		290		7		190		1		488

Tabel 3. Samenvatting van de eerste iteratie

Is 488 treffers over een totaal van circa 79.000 woorden nu veel of weinig? Om deze uitkomst te kunnen interpreteren is het goed om stil te staan bij enkele statistieken van de corpora. De 79.000 woorden zijn voorkomens van 6.319 unieke woorden. Gemiddeld komt elk uniek woord dus $79.000/6.319 = 12,5$ keer voor. Maar die verdeling is zeker niet gelijkmatig. Het meest gebruikte unieke woord van die 6.319 unieke woorden is het woord "de". Dat woord komt 5.219 keer voor. De 10 meest gebruikte unieke woorden komen elk gemiddeld 2.160 keer voor en de 100 meest gebruikte woorden komen elk gemiddeld 433 keer voor. Dat komt overeen met circa 43.300 van de 79.000 woorden in de drie corpora; ofwel circa 55%. Dus circa 55% van de woorden in de drie corpora zijn voorkomens van slechts 100 van de 6.319 unieke woorden.

Aan de andere kant van de verdeling van unieke woorden over de 79.000 woorden in de drie corpora ligt een staart van 4.725 unieke woorden die 5 keer of minder voorkomen en 2576 unieke woorden die precies één keer voorkomen in de drie corpora.

Het unieke woord "DNB" telt 884 voorkomens en staat op de 13e plaats in de rangorde van meest gebruikt unieke woorden. Het unieke woord "dient" (uit de signaalwoorden "instelling dient" en "dient instelling") komt 189 keer voor en staat op de 46ste plaats. Het unieke woord "acht" (uit de signaalwoorden "DNB acht" en "acht DNB") staat met 66 voorkomens op de 143ste plaats.

Kijkend naar de verdeling van treffers per corpus lijkt er in ordegrootte geen uitschieter te zijn. In deze eerste iteratie met de initiële lexicons zijn in het 2006-corpus 115 treffers gevonden, 188 in het 2010-corpus en 185 in het 2014-corpus.

Ter beoordeling van de signaalwoordvaliditeit zijn de treffers steekproefsgewijs¹³ onderzocht in een concordantieanalyse. De analyse gaf op drie manieren aanleiding tot correctie van die initiële lexicons. Ten eerste bleken veel woorden die tijdens de annotatie waren aangemerkt als afkeurende woorden in de onderzochte corpora geen betrekking te hebben op de relatie tussen DNB en de instelling maar op een inhoudelijke bevinding.¹⁴ Signaalwoorden zoals "onvoldoende", "overtreding", "vraagtekens", "opmerkelijk," en "tekortkoming" hebben in de drie corpora bij elkaar 174 treffers maar die hebben in bijna alle gevallen betrekking op een inhoudelijke kwestie. De briefschrijvende toezichthouder vond dan die kwestie onvoldoende, niet zijn relatie met de instelling. Zoals toegelicht in paragraaf 2.2.3 worden dergelijke woorden, in het kader van dit onderzoek, beschouwd als feitelijke informatie en niet als uiting van subjectiviteit, waardoor ze geen geldige signaalwoorden zijn voor *tone of voice*. Deze begripsmatige fout kwam 29 keer voor in het 2006-corpus, 74 keer in het 2010-corpus en 71 keer

¹³ Steekproef met omvang $n=216$, op basis van populatie 488, met 95% betrouwbaarheid, 5% steekproefrisico en (vooraf onbekende) spreiding van 50%. Systematische trekking met interval van 2 uit de treffers per signaalwoord.

¹⁴ De uitsluiting van bevindingen, zoals uiteengezet in paragraaf 2.2.3, van de conceptuele operationalisering heeft een wezenlijk effect gehad op de resultaten. Dat leidde bij de bespreking met de experts tot discussie, zie paragraaf 5.1.

in het 2014-corpus. De correctie bestond eruit dat de betreffende signaalwoorden uit de lexicons zijn verwijderd.

Ten tweede bleken de 30 treffers van het signaalwoord "binnen * * na" in alle gevallen te gaan om een reactietermijn bij een verzoek aan de instelling. De *tone of voice* van zo'n verzoek wordt echter primair bepaald door het onderwerp en werkwoord in de betreffende zin (bijvoorbeeld "Wij verzoeken u" of "[instelling] dient" of "Graag ontvangen wij"). Daarvoor zijn andere signaalwoorden in het lexicon opgenomen (en die blijken ook treffers te hebben), met als gevolg dat de treffers voor "binnen * * na" dubbeltellingen zijn. Deze fout trad 15 keer op in het 2006-corpus, 8 keer in het 2010-corpus en 7 keer in het 2014-corpus. De correctie bestond eruit dat het signaalwoord "binnen * * na" uit het lexicon is verwijderd.

Ten derde bleek het conceptuele onderscheid tussen de indicatoren "afkeurend" (afkomstig van de dimensie "afkeurend ↔ waarderend") en "sturend" (uit de dimensie "sturend ↔ vrij latend") lastig te kunnen worden gemaakt bij het annoteren. Woordcombinaties zoals "DNB acht", "[instelling] dient" en "mag worden verwacht" werden wel aangemerkt als signaalwoorden die iets zeggen over de relatie tussen DNB en de brieflezer, maar vervolgens geclassificeerd als afkeurend en ook als sturend. De correctie bestond eruit dat de overgebleven geldige signaalwoorden opnieuw zijn verdeeld over de indicatoren afkeurend, waarderend, sturend (directief) en vrij latend (indicatief). Voor die indeling zijn per indicator de omschrijvingen aangehouden zoals die in het *four ears*-model zijn genoemd (Schulz von Thun 2010, pp.51-52). Bijlage G vat samen hoe die her-classificatie is uitgepakt.

De lexicons zijn aangepast voor de genoemde vormen van signaalwoordongeldigheid en met de gecorrigeerde lexicons is een tweede iteratie van 12 doorzoekingen uitgevoerd.

4.2 Tweede iteratie met gecorrigeerde lexicons en validiteitbeoordeling

De tweede serie doorzoekingen, op basis van de gecorrigeerde lexicons, resulteerde in 284 treffers over alledrie de corpora heen (tabel 4).

Het meten van tone of voice in Nederlandstalige geschreven tekst

indicator ---->		AFKEUREND			WAARDEREND			STUREND			VRJ LATEND		
gecorrigeerd lexicon ---->		swl03(pnlA)03			swl03(pnlW)03			swl03(pnlS)03			swl03(pnlV)03		
metric ---->		[hits]	[corr.]	[p.1000w]	[hits]	[corr.]	[p.1000w]	[hits]	[corr.]	[p.1000w]	[hits]	[corr.]	[p.1000w]
corpus	# words												
toezichtbrieven uit 2006	26.252	17	-2	0,5714	21	2	0,8761	34	-7	1,0285	0	0	0,0000
toezichtbrieven uit 2010	29.079	30	-2	0,9629	20	2	0,7566	54	-3	1,7538	0	0	0,0000
toezichtbrieven uit 2014	23.896	21	-2	0,7951	17	2	0,7951	69	-3	2,7620	1	0	0,0418
		68			58			157			1		
													284

Tabel 4. Samenvatting van de tweede iteratie

De beoordeling van de signaalwoordvaliditeit gaf geen aanleiding tot verder correctie van de lexicons. Ter beoordeling van de treffervaliditeit zijn alle treffers onderzocht in een concordantieanalyse. Daar waar de treffers onterecht bleken te zijn meegeteld in de aantallen treffers zijn die aantallen gecorrigeerd (zie telkens de kolom [corr] in tabel 4). Het betrof in hoofdzaak incidentele correcties in de volgende categorieën.

De meeste incidentele correcties zijn aangebracht naar aanleiding van het gebruik van het signaalwoord "dien* te worden" uit het sturende lexicon. Daar bleken 9 treffers niet te verwijzen naar een directief vanuit DNB, maar merendeels naar observaties door de toezichthouder. Alle 9 onterechte treffers, 6 in het 2006-corpus, 2 in het 2010-corpus en 1 in het 2014-corpus, zijn in mindering gebracht op het getelde aantal.

Twee voorkomens van het sturende "verwacht DNB" uit het 2014-corpus zijn overgeheveld naar de waarderende treffers, vanwege de vriendelijke bejegening in de zinsnede "Wat verwacht DNB van u?" Eén treffer, uit het 2010-corpus, is uit de telling weggelaten vanwege een technisch manco: in de zinsnede "... het kredietrisico groter is dan verwacht. DNB vraagt u ..." heeft WordSmith kennelijk geen aandacht besteed aan de zinbeëindiging midden in de treffer.

Vier treffers van het signaalwoord "DNB acht" uit het afkeurende lexicon zijn vanwege hun positieve context overgeheveld naar het aantal waarderende treffers: "DNB acht dit een goede ontwikkeling" en "DNB acht dit zorgvuldig en doordacht" uit het 2006-corpus en "DNB acht dit een plausible toelichting" en "DNB acht het een goede zaak dat ..." uit het 2010-corpus.

Twee treffers van het afkeurende signaalwoord "overtreding", beide uit het 2014-corpus, zijn niet meegeteld omdat hun zinsnede betrekking heeft op feitelijke bevindingen: "overtreding van [interne] limieten" en "geen sprake van overtreding [...]" van het EU wapenembargo.

Al met al zijn 13 treffers uit de telling weggelaten en 6 treffers overgeheveld naar het aantal waarderende treffers. Na aanpassing van de aantallen treffers (zie telkens de kolom [corr] in tabel 4) zijn per corpus/lexicon-combinatie trefferfrequenties opnieuw berekend.

4.3 Vertaling van meetresultaten naar verandering van tone of voice

De trefferfrequenties in tabel 4 zijn gevalideerd en vrij van systematische en incidentele ruis. Ze vormen de basis om af te leiden hoe *tone of voice* verandert. Daartoe bepalen we eerst voor elk tweetal corpora en voor elk van de vier indicatoren (afkeurend, waarderend, sturend, vrij latend) of de meetwaarde van de indicator in het jongste corpus hoger of lager is dan in het oudste corpus. Vervolgens leiden we aan de hand van de beslissingsregels in tabel 1 af hoe de veranderingen in de indicatoren zich vertalen in een verandering in *tone of voice* tussen het jongste en het oudste corpus.

Uit tabel 4 lezen is af te lezen dat de trefferfrequentie voor de indicator afkeurend (A) in het 2006-corpus 0.5714 bedraagt en in het 2010-corpus 0.9629. De trefferfrequentie voor afkeurend (A) is tussen 2006 en 2010 dus hoger geworden. Dat is weergegeven in de blauwe cel in tabel 5. De indicator waarderend (W) heeft in deze corpora trefferfrequenties van 0.8761 respectievelijk 0.7566, een verlaging tussen 2006 en 2010, zoals vermeld in de gele cel. Uit tabel 4 is af te lezen dat de indicator sturend (S) is verhoogd van 1.0285 in het 2006-corpus tot 1.7583 in het 2010-corpus, leidend tot de aanduiding "hoger" in de lila cel. De indicator vrij latend (V) tenslotte heeft in beide corpora dezelfde trefferfrequentie: 0.0000. Daarvoor is "gelijk" vermeld in de groene cel.

Voor de aldus samengestelde combinatie van veranderingen op de vier indicatoren (tussen 2006 en 2010) is een vertaling naar verandering van *tone of voice* af te leiden aan de hand van de combinaties in tabel 1. De bovenste rij van tabel 1 past het beste bij de gevonden veranderingen in de bovenste regel van tabel 5. Op die rij in tabel 1 is af te lezen dat bij die combinatie van veranderingen op de vier indicatoren hoort dat de *tone of voice* scherper is geworden tussen beide jaren. Dat is vermeld in de oranje cel.

Op vergelijkbare wijze is voor de twee andere tijdsperiodes 2010→2014 en 2006→2014 bepaald hoe de *tone of voice* is veranderd, zie de rijen 2 en 3 in tabel 5.

Van corpus naar corpus	Verandering per indicator				Verandering in <i>tone of voice</i>
	A	W	S	V	
2006 → 2010	(hoger)	en (lager)	en (hoger)	of (gelijk)	Scherper geworden
2010 → 2014	(lager)	en hoger)	en (hoger)	en hoger)	Milder geworden
2006 → 2014	(hoger)	en lager)	en (hoger)	en hoger)	Scherper geworden

Tabel 5. Vertaling van meetresultaten naar verandering in *tone of voice*

Het blijkt dat de *tone of voice* tussen 2006 en 2010 scherper is geworden, althans afgemeten aan de gevalideerde trefferfrequenties voor de geselecteerde signaalwoorden. Die verscherping wordt in relatief gelijke mate gedreven door de toename van de frequenties voor "afkeurend" (toename met 69%) en voor "sturend" (toename met 71%). De relatieve afname van de frequentie voor "waarderend" (afname met 14%) draagt in dezelfde richting bij aan de verscherping van *tone of voice*, maar met minder gewicht. Op de indicator "vrij latend" is geen verandering gemeten.

Vervolgens is de verscherping iets teruggebogen want tussen 2010 en 2014 werd de *tone of voice* milder. Althans, die conclusie moet worden getrokken op basis van veranderingen per indicator en de beslisregels in tabel 1. Kijkend naar de cijfermatige veranderingen per indicator is echter het nodige af te dingen op die conclusie. Het blijkt dat de frequentie voor "afkeurend" met 17% is afgenomen, dat rechtvaardigt het milder worden van de *tone of voice*. De frequentie voor "waarderend" is licht toegenomen met 5%, dat is eveneens een begrijpelijke drijfveer voor het milder worden. De frequentie voor "sturend" is echter met 57% toegenomen en dat zou op zichzelf een verscherping van *tone of voice* impliceren. De frequentie voor "vrij latend" tenslotte is van 0 met één treffer toegenomen tot 1. Een procentuele verandering is op die indicator niet uit te drukken. De toename van deze frequentie met één treffer geeft weliswaar steun aan het milder worden van *tone of voice*, maar het cijfermatige gewicht van deze indicator is zeer gering.

Dus kijkend naar procentuele veranderingen zien we twee indicatoren die met veranderingen van 17% respectievelijk 5% steun geven aan de conclusie dat de *tone of voice* milder is geworden en één indicator die met een verandering van 57% daar juist een contra-indicatie voor geeft. De vraag in hoeverre de contra-indicatie zwaarder zou moeten wegen dan de (som van de) twee pro-indicaties blijft hier onbeantwoord en leidt tot een aanknopingspunt voor verder onderzoek. Het punt dat alvast moet worden gemaakt is dat de beslisregels in tabel 1 voorbij gaan aan de cijfermatige verhoudingen tussen gevalideerde meetresultaten. Deze kritiek zullen we terugzien in de beantwoording van onderzoekdeelvraag 1, in paragraaf 6.2.1.

Over de hele periode van 2006 tot en met 2014 bezien is de *tone of voice* scherper geworden. De verscherping over de hele periode wordt vooral gedreven door de toename van de frequentie voor "sturend" (toename met 169%). In mindere mate draagt ook de toename van de frequentie voor "afkeurend" bij (toename met 39%) en de afname van de frequentie voor "waarderend" (afname met 9%). De toename van frequentie voor "vrij latend" met één treffer geldt als contra-indicatie maar heeft numeriek beduidend minder gewicht dan het opgetelde effect van de drie pro-indicaties.

4.4 Vergelijking met de verwachte uitkomsten

De verwachting vooraf ten aanzien van verandering van *tone of voice* in toezichtbrieven was dat die door de jaren heen scherper is geworden (zie paragraaf 2.4).

De op de beslisregels in tabel 1 gebaseerde veranderingen in *tone of voice* komen ten dele overeen met die verwachtingen. Voor de periode van 2006 tot 2010 geven de meetresultaten een verscherping van *tone of voice* aan, dat is dus in lijn met de verwachting. Kijkend naar de periode tussen 2010 en 2014 geven de gevonden resultaten aan dat de *tone of voice* milder is geworden, terwijl een verscherping was verwacht.¹⁵ Over de achtjaarsperiode tussen 2006 en 2014 is in het onderzoek een verscherping gevonden en die werd ook verwacht.

De verwachtingen vooraf zijn niet gespecificeerd in de *tone of voice*-indicatoren zoals hier onderzocht. Dus het relateren van de meetresultaten op die indicatoren aan de kritiek van de stakeholders is speculatief. Dat zo zijnde kan worden bedacht dat de kritiek van de stakeholders mogelijk samenhangt met de gemeten toename van sturend woordgebruik in beide periodes.

Bovendien kent het vergelijk tussen de in dit onderzoek gevonden meetresultaten en de vooraf gestelde verwachtingen onzuiverheden (zie paragraaf 2.4) die maken dat zo'n vergelijk maar beperkt kan gelden als betrouwbare basis voor de evaluatie van het meetinstrument.

In de hoofdstukken 2, 3 en 4 hebben we de kern van het onderzoek gezien. Achtereenvolgens kwamen de conceptuele operationalisering van *tone of voice*, de technische operationalisering en uitvoering van het onderzoek, en de meetresultaten en interpretatie daarvan aan bod. Daarmee is materiaal verzameld om vijf van de zes onderzoekdeelvragen (zoals opgesomd in hoofdstuk 1) te kunnen beantwoorden. De zesde onderzoekdeelvraag betreft de kritische beschouwing van het onderzoek door externe deskundigen. Dat is het onderwerp van het volgende hoofdstuk.

¹⁵ Merk hierbij op dat de gevonden verzachting (milder worden) van de *tone of voice* is gebaseerd op de beslisregels in tabel 1. Zoals in paragraaf 4.3 is toegelicht zou een meer cijfermatige analyse tot een verscherping van *tone of voice* kunnen concluderen.

5 Evaluatie door experts

Na uitvoering van het onderzoek was ik in de gelegenheid om mijn onderzoek te bespreken met Michiel Boswinkel en Tigran Spaan, beide deskundig op het terrein van schriftelijke communicatie.

Michiel Boswinkel werkt meer dan 20 jaar als trainer van "schrijvende professionals en hun afdelingen"¹⁶, de laatste jaren in dienst van DNB. Michiel is tevens de begeleider van dit onderzoek vanuit DNB.

Tigran Spaan is oprichter en directeur van GridLine¹⁷, een bureau dat zich specialiseert in computersoftware om van Nederlandse teksten onder meer trefwoorden te bepalen, zoekvragen efficiënt uit te voeren en begrijpelijkheid te meten en bevorderen.

In het gesprek met hen heb ik gevraagd om hun commentaar in het bijzonder op de conceptuele operationalisering van *tone of voice*, de onderzoeksmethode en de uitdrukking van de meetresultaten. In dit hoofdstuk geef ik hun opmerkingen in mijn eigen woorden weer.

5.1 Kanttekeningen bij de conceptuele operationalisering

Een wezenlijk onderdeel van de conceptuele operationalisering is dat daarin consequent het perspectief van de briefschrijver als maatgevend is gekozen. In hoofdstuk 2 is *tone of voice* is een schrijfstijl waarin de schrijver kiest óf en hoe hij subjectieve uitingen onder woorden wil brengen.

Maar, in weerwil van de eerste indrukken over *tone of voice* in paragraaf 2.1, bij alle communicatie is het oogpunt van de lezer minstens zo belangrijk als dat van de schrijver. *Tone of voice is in the eye of the beholder*, het gaat erom hoe de boodschap op de lezer overkomt, hoe de

¹⁶ <http://www.focusconferences.nl/sprekers/michiel-boswinkel/>

¹⁷ <http://www.gridline.nl/>

lezer de *tone of voice* ervaart. De focus op het schrijverperspectief - hoewel goed gemotiveerd en consequent gevolgd - leidt tot een aanzienlijke beperking. Deze kanttekening komt terug in de beantwoording van onderzoekdeelvraag 1, in paragraaf 6.2.1.

Een ander wezenlijk onderdeel van de conceptuele operationalisering gaat over de uitsluiting van bevindingen (zie paragraaf 2.2.3). De gedachtegang is dat een bevinding feitelijke informatie geeft over hoe een waargenomen toestand zich verhoudt tot een norm. Een andere auditor of een andere toezichthouder die dezelfde situatie onderzoekt zou typisch tot hetzelfde bevinding moeten komen. Daarmee is een bevinding niet subjectief en dus ook geen *private state*.

Daar is de nuance bij te maken dat er meerdere manieren zijn om een bevinding onder woorden te brengen. Ook situaties die een toezichthouder als onvoldoende beschouwt kunnen op een "vriendelijke" of "onvriendelijke" manier worden beschreven. Zo beschouwd zou het interessant zijn om bevindingen juist wél in het bereik van het onderzoek te nemen. Ook deze gedachte komt terug in de beantwoording van onderzoekdeelvraag 1, in paragraaf 6.2.1.

Een derde kanttekening gaat over de toepassing van het *four ears*-model en in het bijzonder de keuze voor de dimensies "afkeurend ↔ waarderend" en "sturend ↔ vrij latend" (zie paragraaf 2.3.4). Die keuze is ontleend aan een (bewerkte) publicatie door Schulz von Thun zelf.

In de DNB-praktijk en -schrijftrainingen wordt in plaats van de dimensie "afkeurend ↔ waarderend" vaak de afstandsdimensie gebruikt om het relationele aspect van een boodschap in uit te drukken: "dichterbij/informeler ↔ afstandelijker/formeler". Signaalwoorden zoals "(DNB) acht" en "(u) dient" zijn, samen met bijvoorbeeld nominalisering in plaats van actieve werkwoorden¹⁸, typische indicatoren voor grote afstand. Deze aanvulling komt terug in paragraaf 6.1, bij enkele alternatieve invullingen voor onderdelen van de conceptuele operationalisering.

Een laatste opmerking betreft het verband tussen *tone of voice* en de begrijpelijkheid van de tekst. Aan dat verband wordt in de conceptuele operationalisering geen aandacht besteed, terwijl begrijpelijkheid een belangrijke factor is voor de effectiviteit van de schriftelijke communicatie.

¹⁸ Bijvoorbeeld: *Het indienen van de door u samengestelde rapportages geschiedt door het uploaden van...* in plaats van *De door u samengestelde rapportages kunt u indienen door ze te uploaden ...*

5.2 Kanttekeningen bij de onderzoeksmethode en technische uitvoering

De toepasbaarheid van de meetmethode heeft een achilleshiel, namelijk de controle op de geldigheid van afzonderlijke treffers. Op hoofdlijn zullen de gekozen signaalwoorden best een goede indicatieve waarde hebben, maar het is ook voorzienbaar dat een deel van de signaalwoorden een vals alarm afgeeft. Dit omdat die woorden, afhankelijk van de context waarin ze worden gebruikt, zowel een indicatie van afkeuring kunnen zijn als een indicatie van goedkeuring. Er is - vermoedelijk een handmatige - aanvullende toets nodig op de toon van de context waarin het betreffende signaalwoord is gebruikt. Deze opmerking krijgt veel gewicht bij de beantwoording van onderzoekdeelvraag 2, in paragraaf 6.2.2.

Verder zal de gekozen aanpak zeker nog kunnen worden verbeterd door van de signaalwoorden veel meer gerelateerde woordvormen in de lexicons op te nemen dan in dit onderzoek is gebeurd. Het laten aanstrepen van signaalwoorden in een annotatiepanel is prima maar zou - meer dan hier is gebeurd - moeten worden gevolgd door een systematische analyse van de andere woordvormen die bij de aangestreepte woorden horen. Ook dit punt komt terug bij de beantwoording van onderzoekdeelvraag 2, in paragraaf 6.2.2.

De derde kanttekening bij de gekozen aanpak is dat die alleen maar kijkt naar treffers op woordniveau. *Tone of voice* zit echter vaak ook in taalconstructies op zinsniveau en in de structuur van de tekst. Die voorkomens van *tone of voice* zijn in deze analyse buiten beeld gebleven. Deze aanvulling komt terug in paragraaf 6.1, bij enkele alternatieve invullingen voor onderdelen van de technische operationalisering

5.3 Eenheden van de meetresultaten

In het onderzoek zijn alle meetresultaten uitgedrukt als aantal treffers per 1000 woorden tekst. In professionele analyses is een andere uitdrukkingseenheid gebruikelijker namelijk de TF IDF-metric. TF IDF is de afkorting voor *term frequency inverse document frequency*. De TF IDF kijkt niet alleen naar de frequentie waarin een signaalwoord voorkomt maar weegt die frequentie ook met het belang van het signaalwoord in de tekst. De TF IDF corrigeert voor het feit dat sommige woorden nu eenmaal vaak voorkomen maar minder vaak in een specifieke betekenis. De TF IDF geeft daarmee een zuiverder beeld dan de trefferfrequentie in termen van aantal treffers per 1000 woorden tekst.

Met de kanttekeningen van de beide experts is de verzameling van benodigde gegevens voor het onderzoek compleet. Met alle gegevens kunnen we nu overgaan tot beantwoording van de centrale onderzoeksvraag en de onderzoekdeelvragen. Dat doen we in het volgende hoofdstuk.

6 Anatomie, alternatieven en antwoorden

In dit hoofdstuk keren we terug naar het doel van het onderzoek, te weten het ontwikkelen van een meetinstrument voor de verandering van *tone of voice* in Nederlandstalige geschreven teksten. We zullen een uitspraak doen over de mate waarin dat doel is bereikt door het beantwoorden van de onderzoekdeelvragen en van de centrale onderzoeksvraag uit hoofdstuk 1. Dat doen we op basis van de tijdens het onderzoek opgedane inzichten en onderzoeksgegevens, die we daartoe eerst kritisch zullen beschouwen.

Om te beginnen vatten we nog eens samen wat we in het onderzoek hebben gedaan. We analyseren het ontwikkelde meetinstrument als een raamwerk van afzonderlijke componenten en bezien per component welke specifieke invulling daaraan is gegeven (paragraaf 6.1). Daarbij benoemen we wenselijke verbeteringen en/of eventuele denkbare alternatieve invullingen. Vervolgens gebruiken we die beschouwing om te komen tot een beantwoording van de onderzoekdeelvragen (paragraaf 6.2) en van de centrale onderzoeksvraag (paragraaf 6.3).

6.1 Anatomie van het meetinstrument

Het in dit onderzoek ontwikkelde meetinstrument is te beschouwen als samenstel van bouwstenen. Elk van die bouwstenen heeft een concrete invulling gekregen zodanig dat het geheel een voor dit onderzoek passend instrument werd. Voor de meeste bouwstenen zijn echter ook alternatieve invullingen denkbaar en op enkele punten zijn verbeteringen wenselijk. In tabel 6 zijn de bouwstenen van het meetinstrument opgesomd met de invullingen zoals ze in dit onderzoek zijn gekozen met denkbare alternatieve invullingen dan wel wenselijke verbeteringen.

Bouwsteen	Specifieke invulling voor dit onderzoek	Verbeteringen (I) en/of denkbare alternatieven (A)
Conceptuele operationalisering		
Analyseniveau	Woordniveau	(I,A) Zinsniveau; of (I,A) Tekstniveau
Perspectief	Schrijverperspectief	(A) Lezerperspectief
Linguïstisch analysekader	<i>Private state</i>	
<i>Object</i> uit de private state	De relatie tussen toezichthouder en instelling (waardoor bevindingen buiten de definitie vallen)	(A) <i>Object</i> ruimer definiëren
Inhoudelijke indicatoren	O.b.v. <i>four ears</i> -model: dimensies "afkeurend ↔ waarderend" en "sturend ↔ vrij latend"	(A) Andere dimensies o.b.v. <i>four ears</i> -model; of (A) Dimensies o.b.v. een ander communicatie-model
Meetmethode	Tellen van <i>private states</i>	
Vertaling van meetwaarden naar <i>tone of voice</i> -veranderingen	Binair: scherper of milder al naar gelang hogere of lagere meetwaarden op de indicatoren	(I) Meer "tussenstapjes" (hogere granulariteit in kwalitatieve vertalingen); of (I) Cijfermatige uitdrukkingen
Evaluatie van gevonden <i>tone of voice</i> -veranderingen	Vergelijk met vooraf gestelde verwachtingen o.b.v. stakeholderanalyse	(I) Achteraf voorleggen aan panel van briefschrijvers; en/of (I) Achteraf voorleggen aan panel van brieflezers
Technische operationalisering		
Lexicons	Specifiek voor dit onderzoeksdoel samengesteld in een annotatiepanel	(A) Gebruikmaking van bestaande lexicons;
Validiteitscriteria	Onderscheid naar signaalwoordvaliditeit en treffervaliditeit o.b.v. systematische dan wel incidentele fouten	(I) Verdere "aanscherping" van criteria bijvoorbeeld door gebruikmaking van statistische significanties
Uitvoering van de doorzoekingen en interpretatie van de meetresultaten		

Bouwsteen	Specifieke invulling voor dit onderzoek	Verbeteringen (I) en/of denkbare alternatieven (A)
Doorzoekingen	Met behulp van WordSmith	(A) Met behulp van andere programma's, op woordniveau bijvoorbeeld de AntSuite
Validiteitcontroles	Handmatig met concordantieanalyse	

Tabel 6. Anatomie van het meetinstrument

In de volgende paragrafen wordt elk van de onderzoekdeelvragen beantwoord aan de hand van in dit onderzoek opgedane inzichten, inclusief het in de vorige paragraaf geschetste overzicht van denkbare alternatieven en verbeteringen per component.

6.2 Beantwoording van onderzoekdeelvragen

6.2.1 Onderzoekdeelvraag 1 over conceptuele operationalisering

Onderzoekdeelvraag 1 luidt: Hoe kan *tone of voice* conceptueel worden geoperationaliseerd? De in dit onderzoek gekozen begripsmatige operationalisering van *tone of voice* is uitgewerkt in hoofdstuk 2. Daar is, gefundeerd door theoretische noties van *private state* en het *four ears*-model, *tone of voice* gedefinieerd als een schrijfstijl waarmee de schrijver niet-feitelijke en specifiek bij de boodschap passende bedoelingen, associaties, lading tot uitdrukking kan brengen. *Tone of voice* kan volgens die operationalisering worden gemeten door *private states* te tellen die een uitdrukking zijn op de dimensie "afkeurend ↔ waarderend" of op de dimensie "sturend ↔ vrij latend". De aantallen getelde *private states* in twee jaargangen kunnen via beslisregels in tabel 1 worden terugvertaald naar een verandering van *tone of voice* tussen die jaargangen.

De hele conceptuele operationalisering overziende is er voor drie onderwerpen aanleiding om stil te staan bij de gekozen invulling en verbeteringen of alternatieve invullingen te overwegen, te weten het schrijverperspectief, de uitsluiting van bevindingen, en de vertaling van meetresultaten en de evaluatie van de *tone of voice*-veranderingen.

Schrijverperspectief

Het hier gekozen schrijverperspectief werd ingegeven door de bespiegelingen vanuit marketingcommunicatie: *tone of voice* als vehikel om de identiteit van een merk of een bedrijf

neer te zetten (paragraaf 2.1.1). Dat is, zoals ook bleek uit de bespreking met de deskundigen, een keuze gebleken waarvan de consequenties niet makkelijk intuïtief te volgen zijn. *Tone of voice* is immers iets dat bij uitstek manifest wordt in de taxatie van de lezer van de geschreven tekst.

Om een conceptuele operationalisering te ontwikkelen die uitgaat van het lezerperspectief zou kunnen worden aangesloten bij intern onderzoek door DNB, onder meer in 2012. Ten aanzien van de *tone of voice* gaven respondenten aan dat DNB haar schriftelijke communicatie zou kunnen verbeteren door het positieve meer te benadrukken en het negatieve meer te vermijden.

Uitgaande van de *private state*-notie als analytisch kader is mogelijk een alternatieve conceptuele operationalisering van *tone of voice* af te leiden die recht doet aan deze praktische aanbevelingen en daarmee lezerperspectief centraal stelt. *Tone of voice* wordt dan gedefinieerd als een schrijfstijl waarin eenvoudigweg zoveel mogelijk het positieve is benadrukt en het negatieve is vermeden.

Ook voor deze operationalisering is de notie van *private state* bruikbaar om voorkomens van *tone of voice* te identificeren en te analyseren. *Private states* in het lezerperspectief hebben, net als in het schrijverperspectief, DNB of de toezichthouder als *experiencer*. De verschillen met *private states* in het schrijverperspectief zitten in de *attitude*, in het *object* en in de *intensity*.

In het lezerperspectief is de *attitude* een subjectieve uiting - door de *experiencer* - op het continuüm tussen negatief en positief. De lezer meet *tone of voice* immers alleen maar af aan één criterium: de verhouding tussen het negatieve en het positieve. (Merk op dat de *attitude* in het schrijverperspectief wordt uitgedrukt in twee dimensies, namelijk (i) de verhouding tussen afkeurend en waarderend en (ii) de verhouding tussen sturend en vrij latend.)

Het *object* in het lezerperspectief is niet nader bepaald. Het gaat bij *tone of voice* in het lezerperspectief immers om een stijl waarin eenvoudigweg zoveel mogelijk het positieve is benadrukt en het negatieve is vermeden. Waar dat te benadrukken positieve dan wel te vermijden negatieve betrekking op heeft is open. (In het schrijverperspectief is het *object* wel specifiek bepaald, namelijk op de relatie tussen de toezichthouder en de instelling.)

Uitsluiting van bevindingen (het *object* uit de *private state*)

Een andere kanttekening van de deskundigen betreft de uitsluiting van bevindingen, als zijnde feitelijke informatie, die een forse inperking van de conceptuele operationalisering is gebleken. In aantallen treffers zijn op basis daarvan van die uitsluiting ongeveer 35% van de aanvankelijk gevonden treffers buiten beschouwing gelaten. Bovendien is er begripsmatig veel voor te zeggen om bevindingen juist wél te onderkennen als potentiële dragers van *tone of voice* omdat er meerdere manieren zijn om een bevinding onder woorden te brengen en elke manier z'n eigen

effect op de *tone of voice* kan hebben.

De uitsluiting van bevindingen kreeg analytisch vorm door het *object* van de *private state* af te bakenen op de relatie tussen de toezichthouder en de instelling. Om de bevindingen in de conceptuele operationalisering te kunnen zien is het dus nodig om dat het *object* te verruimen tot de relatie tussen toezichthouder en instelling en/of een door de toezichthouder aangetroffen situatie bij de instelling. Een andere manier om dit vorm te geven is door het *object* in het geheel niet te bepalen. Dan wordt *tone of voice* gezien als subjectieve uiting van de toezichthouder (de *experiencer*) over iets (het *object*) dat verder niet bepaald is. Bevindingen, en de subjectief gekozen bewoordingen daarvan, komen dan in het bereik van *tone of voice*. Een derde denkrichting is het geheel loslaten van de notie van *private state* als analytische eenheid. Bijvoorbeeld door simpelweg alle positief geladen en alle negatief geladen woorden en woordcombinaties te tellen.

Vertaling van meetresultaten (beslisregels in tabel 1)

De in dit onderzoek gekozen vertaling van meetresultaten naar *tone of voice*-veranderingen, zoals weergegeven in tabel 1, is voor verfijning vatbaar, zoals al bleek bij de interpretatie van de meetresultaten in paragraaf 4.3. Met een verfijning zou de beschikbare numerieke informatie in de meetresultaten moeten worden benut om tot rijkere uitspraken over *tone of voice*-veranderingen te kunnen komen. In dit onderzoek zijn de gevalideerde meetresultaten, uitgedrukt in trefferfrequenties per *tone of voice*-indicator en per corpus, in twee reductieslagen ontdaan van waardevolle granulariteit.

In de eerste slag is een koppel van corresponderende trefferfrequenties uit twee corpora gereduceerd tot een verhoging of een verlaging op de betreffende indicator. Daarmee wordt informatie die op te maken is uit het verschil tussen beide frequenties buiten beschouwing gelaten. Bijvoorbeeld, kijkend naar tabel 4, de trefferfrequentie voor "vrij latend" in het 2006-corpus bedraagt 0 en die in het 2014-corpus bedraagt 1. Dat leidt tot een tussenuitkomst "verhoging" op de indicator "vrij latend". De trefferfrequentie voor "sturend" bedraagt (na correctie voor trefferongeldigheid) 27 in het 2006-corpus en 66 in het 2014-corpus. Ook dat leidt tot een tussenuitkomst "verhoging". De informatie die is af te leiden uit de omvang van de verschillen (respectievelijk $1 - 0 = 1$ en $66 - 27 = 39$) is in die tussenuitkomsten niet meer zichtbaar.

In de tweede reductieslag zijn combinaties van aldus bepaalde veranderingen op vier indicatoren gereduceerd tot een verscherping of verzachting (milder worden) van *tone of voice*, zonder enige weging van het relatieve belang van afzonderlijke indicatoren. Ook weer kijkend naar de meetresultaten in tabel 4: over de drie corpora heen heeft de indicator "vrij latend" een trefferfrequentie van 1. De trefferfrequenties voor de andere indicatoren bedragen 68

("afkeurend"), 58 ("waarderend") en 157 ("sturend"). Kijkend naar die scores lijkt het redelijk om te bezien of de frequenties voor "vrij latend" nu met hetzelfde gewicht moeten meetellen in de bepaling van *tone of voice*-verandering als het gewicht dat bij elk van de andere indicatoren wordt toegepast.

Mogelijk dat de toepasbaarheid van het instrument kan worden versterkt met een meer granulaire opzet van beslisregels waarmee ook de numerieke informatie van de meetresultaten op de indicatoren wordt vertaald in veranderingen van *tone of voice*.

Er is op basis van dit onderzoek geen principiële reden aan te wijzen om af te zien van zo'n meer granulaire vertaling. De enige reden waarom die verfijning hier niet is aangebracht is pragmatisch; binnen de beschikbare (doorloop)tijd was geen ruimte om hier meer aandacht aan te besteden. Overigens heeft het onderzoek net zo min heldere aanknopingspunten aan het licht gebracht voor een inhoudelijke ontwikkelingsrichting van zo'n verfijning. Om die reden verdient het aanbeveling dat zo'n verfijning in een apart vervolgonderzoek wordt uitgewerkt of wellicht aan bestaande literatuur kan worden ontleend.

6.2.2 Onderzoekdeelvraag 2 over technische operationalisering en uitvoering

Onderzoekdeelvraag 2 luidt: Hoe kan *tone of voice* technisch worden geoperationaliseerd en hoe kan het worden gemeten? De in dit onderzoek gevolgde technische uitwerking is in hoofdstuk 3 uiteengezet. Door in een annotatiepanel signaalwoorden te verzamelen waarmee de - in de conceptuele operationalisering gedefinieerde - *private states* kunnen worden gedetecteerd in de te onderzoeken teksten. Die doorzoeking zelf kan met behulp van het programma WordSmith waarna de uitkomsten moeten worden gevalideerd om "valse alarmen" uit te sluiten van de meting. In de technische operationalisering zijn drie kwetsbare plekken te onderscheiden, te weten het verzamelen van signaalwoorden, het vooraf vaststellen van validiteitscriteria en het uitvoeren van validiteitcontroles aan de hand van die criteria.

Verzamelen van signaalwoorden

De eerste kwetsbaarheid betreft het verzamelen van signaalwoorden. Dat is inhoudelijk de meest uitdagende stap gebleken omdat daarin de concrete vertaling zit van het begripsmatige. Verkeerde, te weinig of te veel concrete signaalwoorden geven verkeerd beeld van de *tone of voice* van de tekst. In dit onderzoek zijn de signaalwoorden verzameld door een annotatiepanel. De instructies aan het waren heel specifiek toegesneden op de conceptuele operationalisering en het trainingscorpus bestond uit vergelijkbare teksten (toezichtbrieven) als de te onderzoeken corpora. De verzamelde signaalwoorden waren dan ook zeer specifiek voor dit onderzoek.

Eén van de kanttekeningen door de deskundigen onderstreept die kwetsbaarheid, namelijk het te zuinig aanvullen van de lexicons met woordvormen die weliswaar niet zijn aangestreept in de annotatie, maar er toch bij horen.

Een alternatief voor het specifiek samenstellen van lexicons is het gebruik maken van bestaande lexicons die passen bij de operationalisering. Bijvoorbeeld zou in het verlengde van de hierboven beschreven alternatieve operationalisering vanuit een lezerperspectief de in paragraaf 3.1.2 beschreven Engelstalige *Subjectivity Clues* lexicon kunnen worden vertaald en toegepast voor de zoekingen.

Mogelijk dat de toepasbaarheid van het instrument kan worden versterkt door (ook) gebruik te maken van zoekingen met signaalwoorden uit bestaande lexicons.

Validiteitscriteria

Een andere kwetsbaarheid is het uitwerking van criteria voor de validiteitsbeoordeling van treffers. In dit onderzoek heeft die uitwerking vorm gekregen door twee criteria waarmee het onderscheid tussen systematische en incidentele fouten inzichtelijk is gemaakt, maar dat onderscheid is nog steeds kwalitatief. De norm die in beide criteria wordt aangelegd is uitgedrukt in termen van "veel" en "weinig".

Mogelijk dat de toepasbaarheid van het instrument kan worden versterkt indien de normwaarden voor de validiteitscriteria kwantitatief kunnen worden uitgedrukt, bijvoorbeeld aan de hand van statistische significantieniveaus.

Validiteitscontroles

De derde kwetsbaarheid is, vanwege het arbeidsintensieve karakter ervan, de controle op de geldigheid van alle gevonden treffers. Die controle was cruciaal om de meetresultaten te corrigeren voor signaalwoord- en trefferongeldigheid. Die controles zijn handmatig uitgevoerd met een concordantieanalyse per treffer. Een alternatief voor die handmatige beoordeling dient zich niet direct aan.

Dit is vooralsnog dan ook een belangrijk structureel nadeel van het instrument, zeker indien de aantallen treffers fors hoger liggen door gebruikmaking van "ruimere" lexicons.

Overigens is de handmatige uitvoering van de validiteitscontroles ook een bron voor mogelijke onderzoekerbias, zeker als die controles door één onderzoeker worden uitgevoerd zoals in dit onderzoek het geval was. Voor een deel kan dat nadeel worden gereduceerd door een systematische verslaglegging en verantwoording daarover in het onderzoeksrapport. Dat is de

aanpak die in dit onderzoek is gekozen. Als alternatief of in aanvulling daarop zou kunnen worden overwogen om ook de concordantie-analyses te laten uitvoeren in een panel zodat de onderzoekerssubjectiviteit wordt gereduceerd.

6.2.3 Onderzoekdeelvraag 3 over gevonden veranderingen van tone of voice

Onderzoekdeelvraag 3 luidt: Wat is de *tone of voice* in toezichtbrieven uit twee onderscheiden corpora (van periode 1 en van periode 2) en welke veranderingen in *tone of voice* blijken er uit de metingen van die beide corpora? De in dit onderzoek gevonden meetresultaten zijn beschreven in hoofdstuk 4. Het blijkt dat de *tone of voice* tussen 2006 en 2010 scherper is geworden, althans afgemeten aan de gevalideerde trefferfrequenties voor de geselecteerde signaalwoorden. Die verscherping is tussen 2010 en 2014 iets teruggebogen want in die periode werd de *tone of voice* milder. Over de hele periode van 2006 tot en met 2014 bezien is de *tone of voice* scherper geworden.

Overigens moet deze beantwoording van onderzoekdeelvraag 3 worden gezien in samenhang met de in de vorige sub-paragrafen toegelichte kwetsbaarheden van de metingen.

6.2.4 Onderzoekdeelvraag 4 over verwachte veranderingen van tone of voice

Onderzoekdeelvraag 4 luidt: Wat is de verwachting over verandering in *tone of voice* tussen beide corpora op grond van de opvatting van de stakeholders? In hoofdstuk 2 over de conceptuele operationalisering is de vooraf gestelde verwachting toegelicht: die was dat de *tone of voice* in toezichtbrieven door de jaren heen scherper is geworden.

Die vooraf gestelde verwachtingen kunnen maar beperkt gelden als betrouwbare basis voor de evaluatie van het meetinstrument vanwege onzuiverheden in het vergelijk. Die onzuiverheden kunnen in beginsel worden weggelaten of althans gereduceerd indien de verwachtingen geformuleerd hadden kunnen worden in termen die tot een zinnig vergelijk hadden geleid. Dat was in dit onderzoek praktisch niet mogelijk. Een andere manier om die onzuiverheden uit de weg te gaan is door de onderzoeksresultaten achteraf voor te leggen aan een panel van brieflezers en/of briefschrijvers in plaats van ze te vergelijken met vooraf bepaalde verwachtingen.

De toepasbaarheid van het instrument kan worden versterkt door de referentie-uitkomsten (verwachtingen vooraf of vrije beoordeling achteraf) beter vergelijkbaar te maken met de uitkomsten uit de doorzoekingen.

6.2.5 Onderzoekdeelvraag 5 over het vergelijk van gevonden met verwachte veranderingen

Onderzoekdeelvraag 5 luidt: Hoe verhouden de gemeten veranderingen tussen beide corpora (Odv 3) zich tot de verwachte veranderingen (Odv 4) tussen beide corpora? In hoofdstuk 4 zijn de in dit onderzoek gevonden veranderingen van *tone of voice* vergeleken met de vooraf gestelde verwachtingen. De in dit onderzoek gevonden veranderingen in *tone of voice* komen ten dele overeen met die verwachtingen. Voor de periode van 2006 tot 2010 geven de meetresultaten een verscherping van *tone of voice* aan, dat is dus in lijn met de verwachting. Kijkend naar de periode tussen 2010 en 2014 geven de gevonden resultaten aan dat de *tone of voice* milder is geworden, terwijl een verscherping was verwacht. Over de achtjaarsperiode tussen 2006 en 2014 is in het onderzoek een verscherping gevonden en die werd ook verwacht.

Overigens kent het vergelijk tussen de in dit onderzoek gevonden meetresultaten en de vooraf gestelde verwachtingen onzuiverheden (zie paragraaf 2.4) die maken dat zo'n vergelijk maar beperkt kan gelden als betrouwbare basis voor de evaluatie van het meetinstrument.

6.2.6 Onderzoekdeelvraag 6 over de opvatting van deskundigen

Onderzoekdeelvraag 6 luidt: Wat is de opvatting van deskundigen over de conceptuele operationalisering (Odv 1), de meetmethode (Odv 2) en de meetresultaten (Odv 3)? In hoofdstuk 5 zijn de kanttekeningen verwoord die twee deskundigen hebben geplaatst bij dit onderzoek. Alle kanttekeningen bieden aanknopingspunten voor verbetering van het meetinstrument, de belangrijkste opmerkingen zijn behandeld in voorgaande sub-paragrafen. Uit de bespreking met beide deskundigen komt niet het beeld naar voren dat de in dit onderzoek ontwikkelde meetmethode als onhaalbaar moet worden afgeschreven, noch dat hij zonder meer algemeen toepasbaar is.

6.3 Beantwoording van de centrale onderzoeksvraag

De centrale onderzoeksvraag in het onderzoek luidt: in hoeverre is het in dit onderzoek te ontwikkelen meetinstrument voor verandering van *tone of voice* praktisch toepasbaar en wat is er in dat geval te zeggen over de validiteit van de meetresultaten?

Op basis van de in de vorige paragraaf afgeleide antwoorden op de onderzoekdeelvragen kan het in dit onderzoek ontwikkelde meetinstrument worden beschouwd als een samenstel van bouwblokken, elk met een specifieke, voor dit onderzoek passende, invulling. Elk bouwblok kan echter ook op andere manieren worden ingevuld, al naar gelang een specifieke toepassing dat

verlangt.

Voor het samenstel van invullingen dat voor dit onderzoek werd gekozen geldt het volgende:

- het geheel werkt, dat wil zeggen er is een werkende aaneenschakeling mogelijk van inhoudelijk consistente en technisch op elkaar aansluitende componenten, vanaf conceptuele operationalisering tot aan met meetresultaten onderbouwde uitspraken over verandering van *tone of voice*;
- de evaluatie aan de hand van referentie-uitkomsten is niet goed uitgevoerd kunnen worden omdat de voor dit onderzoek beschikbare referentie-uitkomsten slechts beperkt vergelijkbaar zijn met de in dit onderzoek gevonden meetresultaten. Dit euvel lijkt in een andere toepassing in beginsel oplosbaar te zijn;
- de validiteit van de resultaten van de doorzoekingen is vastgesteld op het laagste niveau van afzonderlijke treffers;
- de controles op de validiteit zijn cruciaal maar, vanwege het handmatige en dus arbeidsintensieve karakter ervan, ook een wezenlijke tekortkoming van dit instrument als het over praktische toepasbaarheid gaat. Een efficiënt alternatief voor die controles dient zich niet direct aan en vergt nader onderzoek;
- er zijn in dit onderzoek enkele andere verbeterpunten geïdentificeerd die in een andere toepassing in beginsel oplosbaar lijken te zijn.

7 Conclusies en aanbevelingen

In dit onderzoek is een meetinstrument ontwikkeld om *tone of voice* te meten in Nederlandse teksten. Dat instrument is toegepast op toezichtbrieven van DNB.

De uitkomsten van de doorzoekingen met dit meetinstrument bieden DNB slechts een beperkt inzicht in de ontwikkeling van *tone of voice* in haar toezichtbrieven. Dat komt omdat de voor dit onderzoek beschikbare referentie-uitkomsten (vooraf gestelde verwachtingen over de ontwikkeling van *tone of voice*) slechts beperkt vergelijkbaar zijn met de in dit onderzoek gevonden meetresultaten. Dat euvel lijkt overigens goed oplosbaar te zijn. Bijvoorbeeld door, in een aanvullend onderzoek, referentie-uitkomsten te bepalen die qua perspectief en qua tijdsperiode explicieter zijn en beter aansluiten op de conceptuele operationalisering die ten grond lag aan de gevonden meetresultaten. Voor zover de gevonden uitkomsten vergelijkbaar zijn met de referentie-uitkomsten komen de in dit onderzoek gevonden veranderingen in *tone of voice* ten dele overeen met de vooraf gestelde verwachtingen.

Los van de moeilijkheden met de vergelijkbaarheid lijken de uitkomsten te worden vertekend doordat de verzamelingen van signaalwoorden die op *tone of voice* duiden te zuinig zijn samengesteld. Ook dit euvel lijkt in een aanvullend onderzoek of nieuwe analyse goed oplosbaar te zijn, bijvoorbeeld door systematisch gerelateerde woordvormen van aangestreepte signaalwoorden in de lexicons op te nemen of door uit te gaan van passende bestaande lexicons. Het is op voorhand niet te zeggen in welke richting de uitkomsten zullen bewegen indien de doorzoekingen worden uitgevoerd met rijkere lexicons.

Tenslotte zijn de beslisregels voor de vertaling van gevalideerde meetresultaten in *tone of voice*-veranderingen binair opgezet. Het gevolg is dat die regels geen efficiënt gebruik maken van de cijfermatige informatie in de meetresultaten. Met een verfijning zou de beschikbare numerieke informatie in de meetresultaten kunnen worden benut om tot rijkere uitspraken over *tone of voice*-veranderingen te kunnen komen. Het verdient dan ook aanbeveling om juist op dat punt, van de validiteitcontroles, nader onderzoek te doen.

In het algemeen kan het meetinstrument worden beschouwd als een samenstel van bouwblokken, elk met een specifieke, voor dit onderzoek passende, invulling. Elk bouwblok kan echter ook op andere manieren worden ingevuld, al naar gelang een specifieke toepassing dat verlangt.

Een wezenlijk nadeel van het meetinstrument is het arbeidsintensieve karakter van de handmatige validiteitcontroles van afzonderlijke treffers. Omdat de controle cruciaal zijn staat dit punt de praktische toepasbaarheid van het meetinstrument in de weg. Een efficiënt alternatief voor die validiteitcontroles dient zich niet direct aan. Het verdient dan ook aanbeveling om juist op dat punt, van de validiteitcontroles, nader onderzoek te doen.

De toepassing van het meetinstrument heeft ook enkele andere verbeterpunten aan het licht gebracht. Die lijken in een nieuwe of andere toepassing in beginsel goed oplosbaar te zijn.

Al met al biedt dit meetinstrument, mits verder doorontwikkeld voor wat betreft de gesignaleerde kwetsbaarheden, in een aantal richtingen toepassingsmogelijkheden voor operational auditors. Allereerst kunnen auditors het instrument inzetten zoals is gedaan in dit onderzoek: voor het onderzoeken van auditrapporten op *tone of voice*. Maar de toepassing is zeker niet beperkt tot *tone of voice*. Ook andere kwaliteitskenmerken van auditrapporten kunnen worden onderzocht, zoals bijvoorbeeld het gebruik van vaktechnische auditterminologie of juist het gebruik van typische vocabulaire uit de business. Uiteraard moeten de lexicons van signaalwoorden wel specifiek op de toepassing worden samengesteld. Als laatste toepassingsrichting wordt hier genoemd de snelle doorzoeking van documenten die in audits als bronmateriaal moeten worden onderzocht. Betrekkelijk eenvoudige hulpmiddelen zoals WordSmith of vergelijkbare programma's kunnen worden ingezet om - mits passende lexicons beschikbaar zijn - snel een indruk te krijgen van de relevantie van die documenten.

Bijlage A. Verantwoording en leereffecten

Verantwoording

Ik heb de werkzaamheden die in dit referaat zijn beschreven daadwerkelijk uitgevoerd en ik heb ze zelf uitgevoerd. Buiten de begeleiding vanuit ESAA heb ik hulp gehad van een DNB-collega bij het vinden van toezichtbrieven in het DNB-archief, van DNB-collega's bij het samenstellen van de lexicons, van een DNB-collega bij de inhoudelijke begeleiding vanuit DNB, van twee deskundigen (waarvan één de DNB-begeleider) bij het bespreken van de onderzoeksresultaten, van een DNB-collega bij het meedenken over het schrijven van dit referaat zonder de vertrouwelijkheid van toezichtinformatie te schenden, en van een meelezer die de eerste versie van dit referaat kritisch heeft becommentarieerd.

De gevonden meetresultaten heb ik rechtstreeks overgenomen uit de output van het WordSmith-programma en nadien aangepast ten behoeve van de validiteitcontroles. De aanpassingen in dat verband heb ik samengevat beschreven in hoofdstuk 4. Andere dan die aanpassingen heb ik niet aangebracht.

De conclusies heb ik getrokken en verwoord zonder dat ik me daarover heb hoeven of willen verantwoorden anders dan de consultatie binnen DNB over eventuele gevoeligheden met betrekking tot de toezichtvertrouwelijkheid.

Afgezien van enkele als zodanig aangemerkte interne DNB-rapporten heb ik alle gebruikte literatuur expliciet vermeld en opgenomen in de literatuurlijst.

Leereffecten

De leereffecten die inhoudelijk betrekking hebben op de vraagstelling van dit specifieke onderzoek zijn verwoord in de hoofdstukken 6 en 7.

Daarnaast zijn enkele observaties relevant die wat persoonlijker van aard zijn.

Het onderwerp van onderzoek en de toe te passen methoden en technieken waren voor mij nieuw. Ik had geen ervaring met of kennis van taalkundige analyses. De voorbereiding, in het onderzoeksontwerp, en de daadwerkelijke uitvoering van het onderzoek zelf waren daarom in hoge mate experimenteel, althans voor mij. Achteraf gezien zou ik, gegeven de onderwerpkeuze, drie dingen anders doen. Als eerste zou ik de scheiding tussen ontwerp en uitvoering losser toepassen dan misschien wenselijk is in andere afstudeersituaties. Tijdens het ontwerp wist ik namelijk nog volstrekt niet waarnaar ik op zoek zou moeten gaan of hoe ik dat zou kunnen aanpakken. Dat inzicht kwam pas met het verkennende "rommelen" met teksten. En dat had net zo goed kunnen worden aangemerkt als deel van de uitvoering. Ten tweede zou ik mezelf nóg nadrukkelijker vooraf realiseren dat de beschikbare tijd misschien wel te kort is om daarin tot eenzelfde diepgang te kunnen komen als zou kunnen bij een onderwerp of onderzoeksaanpak dat dichter bij huis ligt. Of andersom: de doelstelling terugbrengen omdat een goed deel van de energie en doorlooptijd naar verwachting niet erg efficiënt kan worden besteed in een onderzoek over een vreemd onderwerp. Ten derde zou ik het onderzoek toepassen op passende maar publiekelijk beschikbare teksten. De aanleiding voor het onderzoek dreef me naar de doorzoeking van toezichtbrieven van DNB maar ik heb relatief veel energie gestopt in het anonimiseren daarvan. De essentie van het onderzoek - de toepasselijkheid van een meetinstrument - vergt echter niet noodzakelijkerwijs dat daarvoor vertrouwelijke DNB-brieven worden doorzocht. Dat had misschien net zo goed gekund op basis van openbare rapporten van de Algemene Rekenkamer of van inspectiediensten.

De zonzijde van het vorige punt is de kans om me te kunnen storten op een onderwerp waarvoor ik vermoedelijk niet zoveel aandacht zou hebben kunnen opbrengen als het niet was omwille van de verplichting in het kader van een afstudeeronderzoek. Want het experimentele maakte het onderzoek voor mij juist ook erg leuk om te doen.

Bijlage B. Index van het onderzoeksdossier

Dit referaat vormt het eindverslag over het afstudeeronderzoek. Tijdens het onderzoek zijn allerhande tussen- en bijproducten tot stand gekomen. Die zijn, conform de ESAA-richtlijnen, opgenomen in een onderzoeksdossier. Deze bijlage geeft de index van dat onderzoeksdossier.

Onderzoeksvoorstel

Versie	Datum	Status
1.0	08-Dec-2014	Ter bespreking met afstudeercoördinator Peter Bos.

Onderzoeksontwerp

Versie	Datum	Status
1.0	09-Mrt-2016	Goedgekeurd door de begeleider en akkoord van ESAA. Tov versie 0.2.4 is de doelstelling verlegd van DNB-specifiek naar een doel met meer algemene relevantie.
0.2.4	03-Mrt-2016	Ter bespreking met de begeleider. Tov versie 0.2.3 zijn onzekerheden in de technische uitwerking (na verkennende studie) gereduceerd.
0.2.3	09-Aug-2015	Work in progress, input voor bespreking met de begeleider. Tov versie 0.2.2 is de scope versmald en verlegd. Tevens verdeling in een verkennend en een beschrijvend onderzoeksdeel.
0.2.2	17-Jun-2015	Work in progress, input voor eerste bespreking met de begeleider.
0.2.1	08-Jun-2015	Work in progress, met verzoek om commentaar voorgelegd aan DNB.

Communicatie met de begeleider

Zie de versiehistorie op pagina

Datum	Aard en inhoud
17-Apr-2016	E-mail van begeleider met feedback op versie 0.38 van het slotexamendocument.
31-Mrt-2016	E-mail van begeleider met feedback op versie 0.17 van het slotexamendocument.
12-Mrt-2016	E-mail van begeleider. Bevestiging van akkoord door ESAA op

	onderzoeksontwerp v.1.0.
11-Mrt-2016	E-mail van begeleider. Bevestiging van goedkeuring van onderzoeksontwerp v.1.0.
08-Mrt-2016	Telefonisch overleg. Bespreking van onderzoeksontwerp v.0.2.4.
08-Jan-2016	E-mail aan begeleider. Melding van hervatting scriptiewerk.
05-Sept-2015	E-mail aan begeleider. Melding van oponthoud van scriptiewerk i.v.m. familie-omstandigheden.
18-Aug-2015	Telefonisch overleg. Bespreking van onderzoeksontwerp v.0.2.3.
17-Jun-2015	Persoonlijk overleg. Kennismaking, bespreking van onderzoeksontwerp v.0.2.2.

Communicatie met ESAA/OCC

Datum	Aard en inhoud
14-Apr-2016	Telefonisch overleg. Bespreking van de bezorging van de fysiek in te leveren exemplaren van referaat.
14-Mrt-2016	Telefonisch overleg. Bespreking van de exacte status van het in te leveren referaat op het moment van inlevering.
08-Jan-2016	Telefonisch overleg. Bespreking van hervatting scriptiewerk.
02-Oct-2015	Telefonisch overleg. Bespreking van vertraging i.v.m. familie-omstandigheden en herziening planning (uiterlijke inleverdatum 30-Apr-2016).
05-Sept-2015	E-mail aan OCC. Melding van oponthoud van scriptiewerk i.v.m. familie-omstandigheden.
18-Dec-2014	E-mail van OCC. Toewijzing scriptiebegeleider.

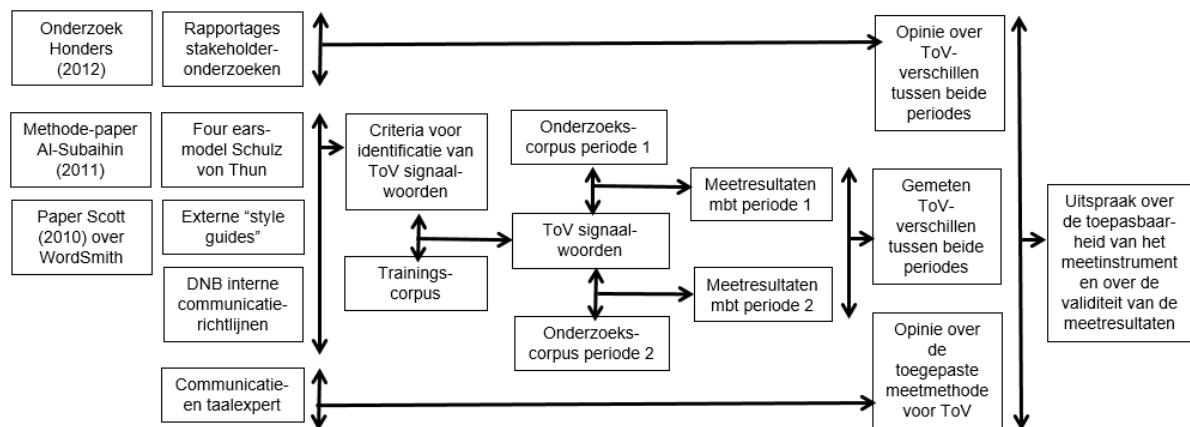
Belangrijke keuzes en afwegingen

Datum	Aard en inhoud
16-Apr-2016	Besluit om, vanwege de tijdsdruk, de opzet met twee alternatieve analyses los te laten en het referaat te beperken tot de ene aanvankelijk voorziene analyse o.b.v. <i>four ears</i> -model. Met name de validatie van de meetresultaten (meer dan 2.600 treffers) uit die alternatieve analyse bleek veel meer tijd te kosten dan beschikbaar was binnen de planning.
08-Apr-2016	Uitbreiding van het referaat met een aanvullende operationalisering en analyse o.b.v. schrijfrichtlijnen. Het werk daarvoor was deels al gedaan en de aanvullende analyse zou een rijker beeld geven van <i>tone of voice</i> .
Aug-2015	Uitgaande van de inhoudelijke aanleiding voor het onderzoek (kritiek van stakeholders over verschil tussen de <i>tone of voice</i> in toezichtbrieven en de <i>tone of voice</i> in corresponderende toezichtgesprekken) heb ik twee versmallingen toegepast om uit te komen bij een scherp afgebakend afstudeeronderwerp.

De versmallingen betrekking op de keuzes (i) om niet de instellingen als onderwerp van onderzoek te kiezen en (ii) om evenmin het onderzoek te richten op de structurele verschillen tussen mondelinge en schriftelijke communicatie. De afwegingen bij deze keuzes zijn toegelicht in versie 0.2.3 van het onderzoeksontwerp.

Bijlage C. Gepland onderzoeksmodel (uit: onderzoeksontwerp)

In het onderzoeksontwerp dat ten grondslag ligt aan de uitvoering van dit onderzoek is het geplande onderzoeksmodel weergegeven (onderzoeksontwerp, versie 1.0, p.6):

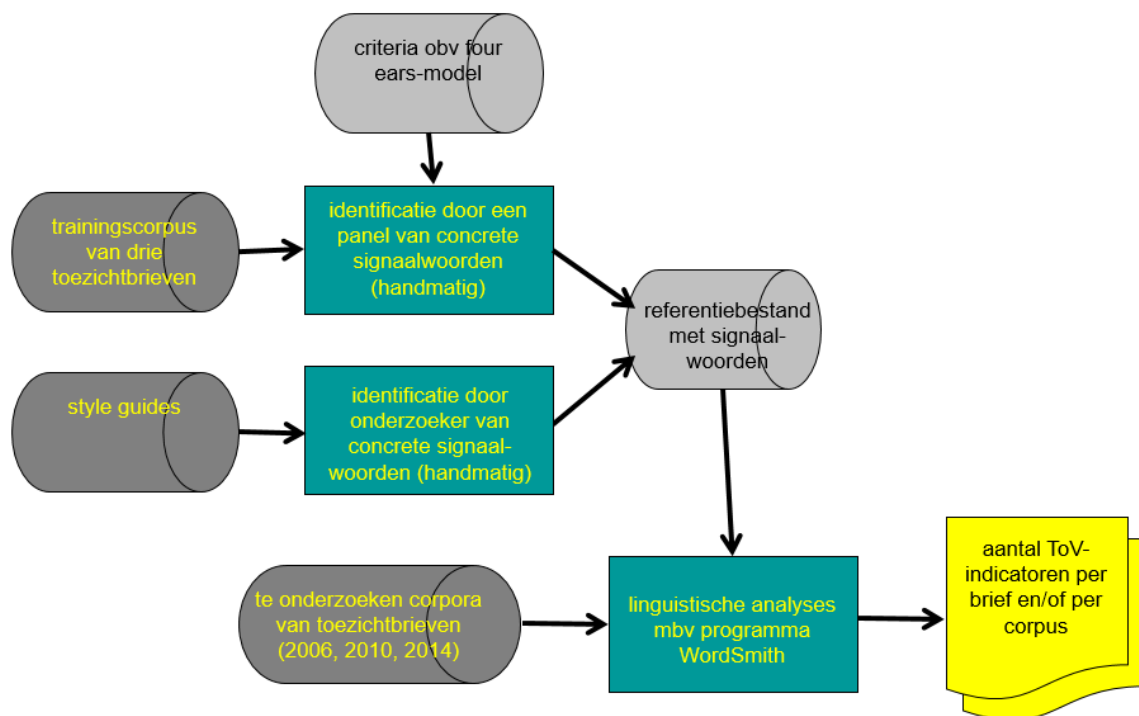


Figuur 1 Onderzoeksmodel

Bijlage D. Geplande onderzoeksanpak (uit: onderzoeksontwerp)

In het onderzoeksontwerp dat ten grondslag ligt aan de uitvoering van dit onderzoek is de geplande onderzoeksanpak beschreven (onderzoeksontwerp, versie 1.0, pp.11-12):

"In figuur 3 is op hoofdlijnen de aanpak geschetst dit in afstudeeronderzoek wordt gevolgd.



Figuur 3 Aanpak op hoofdlijnen

Aan de hand van het four ears-model van Schulz von Thun worden criteria geformuleerd waarmee signaalwoorden kunnen worden geselecteerd uit drie concrete toezichtbrieven in een trainingscorpus. Het gaat om het kunnen selecteren van woorden en woordcombinatie die iets

zeggen over de manier waarop DNB (de briefschrijver) zich positioneert in de relatie met de instelling (de brieflezer). De positie die de briefschrijver kiest kan worden uitgedrukt in twee dimensies, te weten (i) sturend of vrij latend, en (ii) geringschattend of waarderend (Schulz von Thun, 2010, pp. 51-54).

Geïnspireerd door een methode die is beschreven door Al-Subaihin (2011) gebeurt de feitelijke identificatie van signaalwoorden uit dat trainingscorpus gebeurt in een panel van toezichthouders. De (geanonimiseerde) toezichtbrieven uit het trainingscorpus worden, opgeknipt in fragmenten, voorgelegd aan panelleden met het verzoek om die woorden en woordcombinaties aan te strepen die voldoen aan de eveneens aangeboden criteria. De door de panelleden aangestreepte woorden en woordcombinaties worden opgenomen in een referentiebestand met signaalwoorden.

Een tweede bron van signaalwoorden wordt gevormd door style guides en interne DNB-rapporten. Die worden door de onderzoeker geanalyseerd op aanknopingspunten voor tone of voice. Dergelijke aanknopingspunten worden vertaald in signaalwoorden en toegevoegd aan het referentiebestand.

Het referentiebestand wordt vervolgens uitgebreid met andere woordvormen van de opgenomen signaalwoorden. Zo wordt bijvoorbeeld naast "u dient" ook "dient u" in het referentiebestand opgenomen.

Het referentiebestand wordt ter kritische beschouwing voorgelegd aan adviseurs en trainers van de afdeling Communicatie van DNB.

De te onderzoeken corpora worden geanalyseerd op het voorkomen van signaalwoorden uit het aldus samengestelde referentiebestand. Daarvoor wordt het programma WordSmith gebruikt. Dat programma geeft weer hoe vaak een bepaald signaalwoord voorkomt per brief en/of per corpus, in absolute aantallen en relatief ten opzichte van het totaal aantal woorden. Ook kunnen andere "beschrijvende statistieken" over het gebruik van signaalwoorden in de corpora worden ontleend aan de WordSmith-output.

Per onderzochte corpus wordt het gebruik van signaalwoorden bepaald (ter beantwoording van de onderzoeksvragen 2 en 3) en vervolgens worden eventuele verschillen tussen de corpora geanalyseerd (onderzoeksvraag 4). Ten slotte worden die verschillen geïnterpreteerd in termen die bruikbaar zijn voor de analyse van oorzaken (onderzoeksvraag 5)."

Bijlage E. "Readme" bij de Subjectivity Clues lexicon (Wilson 2005)

The subjectivity clues in the file subjclueslen1-HLTEMNLP05.tff were used in the work presented in: Theresa Wilson, Janyce Wiebe and Paul Hoffmann (2005). *Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis*. Proceedings of HLT/EMNLP 2005, Vancouver, Canada.

I. Where the subjectivity clues are from

The clues in this file were collected from a number of sources.

- Some were culled from manually developed resources.
- Others were identified automatically using both annotated and unannotated data.
- A majority of the clues were collected as part of the work reported in (Riloff and Wiebe, 2003).

II. Format of the clues

Each line in the file contains one subjectivity clue. Below is an example:

type=strongsubj len=1 word1=abuse pos1=verb stemmed1=y priorpolarity=negative

a. type - either strongsubj or weaksubj. A clue that is subjective in most context is considered strongly subjective (strongsubj), and those that may only have certain subjective usages are considered weakly subjective (weaksubj).

b. len - length of the clue in words. All clues in this file are single words.

c. word1 - token or stem of the clue.

d. pos1 - part of speech of the clue, may be anypos (any part of speech).

e. stemmed1 - y (yes) or n (no). Is the clue word1 stemmed? If stemmed1=y, this means that the clue should match all unstemmed variants of the word with the corresponding part of speech. For example, "abuse", above, will match "abuses" (verb), "abused" (verb), "abusing" (verb), but not "abuse" (noun) or "abuses" (noun).

f. priorpolarity - positive, negative, both, neutral. The prior polarity of the clue. Out of context, does the clue seem to evoke something positive or something negative.

Riloff and Wiebe (2003). Learning extraction patterns for subjective expressions. EMNLP-2003.

Theresa Wilson
11/16/2005

Bijlage F. Instructie aan toezichthouders in annotatiepanel

MINI-ENQUÊTE

Het tekstfragment op de achterkant van dit A4-tje komt uit een real-life (maar bewerkte) toezichtbrief¹⁾, te weten een concept terugkoppeling uit een kredietonderzoek bij de (gefigeerdde) Rommeldamsche Duitsebank, afgekort RDB.

Ik ben op zoek naar woorden en/of woordcombinaties in toezichtbrieven die iets zeggen over de manier waarop DNB (de briefschrijver) zich positioneert in de relatie met deze instelling, de RDB (de brieflezer).

Die positie die de briefschrijver kiest in de relatie tot de brieflezer kan worden uitgedrukt in twee dimensies, namelijk (i) sturend of vrij latend, en (ii) geringschattend of waardierend.

Het verzoek is om:

- (i) het fragment aandachtig te lezen;
- (ii) die woorden/woordcombinaties te onderstrepen waarin – naar jouw mening – een uiting te lezen is van een positie die DNB kiest in de relatie met de instelling;
- (iii) in de kantlijn per woord/woordcombinatie te coderen welke uiting jij erin leest:
 - SG: DNB is hier sturend en geringschattend i.r.t. RDB
 - SW: DNB is hier sturend en waardierend i.r.t. RDB
 - VG: DNB is hier vrij latend en geringschattend i.r.t. RDB
 - VW: DNB is hier vrij latend en waardierend i.r.t. RDB
 - S: DNB is hier alleen sturend i.r.t. RDB
 - V: DNB is hier alleen vrij latend i.r.t. RDB
 - G: DNB is hier alleen geringschattend i.r.t. RDB
 - W: DNB is hier alleen waardierend i.r.t. RDB
- (iv) het A4-tje met je aantekeningen in te leveren bij Jeroen-Paul.

Achtergrond:

De geïnventariseerde woorden en woordcombinaties gebruik ik in mijn afstudeeronderzoek naar de ontwikkeling van tone-of-voice in toezichtbrieven.

De aanleiding voor het afstudeeronderzoek is gelegen in kritiek van instellingen op de tone-of-voice in toezichtbrieven. Die instellingen vinden de tone-of-voice in de toezichtbrieven te scherp en/of niet corresponderend met de tone-of-voice in toezichtgesprekken over hetzelfde onderwerp.

Ik onderzoek tone-of-voice in toezichtbrieven aan de hand van het model van Schulz von Thun. Hij onderscheidt vier aspecten aan elke boodschap, waarvan het relationele aspect er één is.

Hartelijk dank voor je medewerking!
Jeroen-Paul

¹⁾ In verband met de vertrouwelijkheid is het fragment bewerkt: namen zijn gefingeerd of weggelaten; we doen alsof het kredietonderzoek is uitgevoerd bij de Rommeldamsche Duitsebank, afgekort RDB. Verder zijn datums en periodes veranderd; we doen alsof het onderzoek in 2004 is uitgevoerd. En bedragen en andere specifieke omstandigheden zijn veranderd of weggelaten. Afgezien van deze bewerkingen is de lopende tekst uit de real-life brief gehandhaafd.

De signaalwoorden uit de geschoonde lexicons zijn als volgt opnieuw geclassificeerd:

[illegible]

Literatuur

Al-Subaihini, A.A.; H.S. Al-Khalifa; A.S. Al-Salman (2011): *A Proposed Sentiment Analysis Tool for Modern Arabic Using Human-Based Computing*.

Albrighton, T. (2010): *How to define your brand's tone of voice*.
<http://www.abccopywriting.com/2010/08/31/tone-of-voice-brand>, geraadpleegd 21-Feb-2016.

Allee, K.D. en M.D. Deangelis (2015): *The Structure of Voluntary Disclosure Narratives: Evidence from Tone Dispersion*. Journal of Accounting Research, volume 53, number 2, May 2015, pp.241-274

Bernstein, D. (1995): *Corporate Tone Of Voice*. Design Management Journal, Volume 6, Issue 1, winter 1995, pp.17-21

Brown, P. en S.C. Levinson (1987): *Politeness: Some universals in language usage*. Cambridge, Cambridge University Press.

Cherry, R.D. (1987): *Politeness in Written Persuasion*. Journal of Pragmatics, volume 12, 1988, pp.63-81.

Cummings, H. (publicatiejaar niet bekend): *Finding Your Brand's Voice - How to Shape a Tone of Voice*. <https://www.distilled.net/tone-of-voice/>, geraadpleegd 21-Feb-2016.

DNB, De Nederlandsche Bank (2010): *Van analyse naar actie - Plan van aanpak cultuurverandering toezicht DNB*.

DNB, De Nederlandsche Bank (2012): *Onderzoek naar de schriftelijke communicatie van De Nederlandsche Bank*. Intern rapport, niet publiekelijk beschikbaar.

DNB, De Nederlandsche Bank (2014): *Visie DNB toezicht 2014-2018*.

- DNB, De Nederlandsche Bank (2015): *Stakeholders over het toezicht*. DNB Magazine, #2 2015.
- Hagge, J. en Ch. Kostelnick (1989): *Linguistic Politeness in Professional Prose*. Written Communication, volume 6, number 3, July 1989, pp.312-339.
- Hoeksema, J. (2013): *De negatief-polaire uitdrukkingen van het Nederlands - Inleiding en lexicon*. Rijksuniversiteit Groningen, versie april 2013
- Kenman, L.F. (2007): *Tone and Style: Developing a Neglected Segment of Business Communication*. Business Communication Quarterly, volume 70, number 3, September 2007, pp.305-309
- Limberg, H. (2009): *Impoliteness and threat responses*. Journal of Pragmatics, volume 41, 2009, pp.1376–1394.
- Macpherson, N.A. (1992): *Style: Theories and practical application*. Language Sciences, Volume 14, Issues 1–2, January–April 1992, Pages 109-127
- Markel, N.N.; M.F. Bein en J.A. Phillis (1973): *The Relationship Between Words And Tone-Of-Voice*. Language and Speech, volume 16, number 1, January 1973, pp.15-21
- Mayew, W.J. en M. Venkatachalam (2012): *The Power of Voice: Managerial Affective States and Future Firm Performance*. The Journal of Finance, volume 67, number 1, February 2012, pp.1-43
- National Health Service (2010): *NHS tone of voice: words and written communications*. www.nhsidentity.nhs.uk/generic-content/generic-guideline-content/tone-of-voice
- Pauly, J. (1997): *The Case for a New Model of Business Communication*. The Journal of Business Communication, volume 14, issue 4, pp.11-23.
- Schulz von Thun, F. (2010): *Hoe bedoelt u? Een psychologische analyse van menselijke communicatie*. Groningen, Noordhoff Uitgevers. Derde herziene druk, bewerkt door Maruschka Gijzen.
- Scott, M. (2012): *WordSmith Tools, version 6.0*. Liverpool, Lexical Analysis Software Ltd.
- Sebastiani, F. (2002): *Machine Learning in Automated Text Categorization*. Computing Surveys (CSUR), volume 34, issue 1, March 2002, pp. 1-47

- Shin, H., M. Kim, Y. Jo, H. Jang en A. Cattle (2012): *Annotation Scheme for Constructing Sentiment Corpus in Korean*. 26 th Pacific Asia Conference on Language, Information and Computation pages 181-190
- Sterkel, K.S. (1988): *The Relationship Between Gender And Writing Style In Business Communication*. Journal of Business Communication, volume 25, issue 4, 1988, pp.17-38.
- Rasmussen, R.V. (1991): *A Communication Model Based on the Conduit Metaphor*. Management Communication Quarterly, volume 4, number 3, February 1991, pp.363-374.
- University of Leeds (publicatiejaar onbekend): *Write. Tone of voice guidelines*.
http://www.leeds.ac.uk/comms/tov/tone_of_voice.pdf.
- University of Salford, Manchester (publicatiejaar onbekend): *Tone of voice and style guide - How to express our brand in writing*.
http://www.salford.ac.uk/__data/assets/pdf_file/0010/327997/UOS-style-guide-final.pdf
- Wiebe, J.M. (1994): *Tracking point of view in narrative*. Computational Linguistics, volume 20, number 2, 1994, pp. 233–287.
- Wiebe, J.M. en R. Mihalcea (2006): *Word Sense and Subjectivity*. Proceedings of the Conference on Computational Linguistics / Association for Computational Linguistics (COLING/ACL), 2006.
- Wilson, T. en J.M. Wiebe (2005): *Annotating attributions and private states*. Proceedings of the Workshop on Frontiers in Corpus Annotations II: Pie in the Sky, pp.53-60. June 2005, Ann Arbor, Michigan. Association for Computational Linguistics.
- Wilson, T.; J.M. Wiebe en P. Hoffmann (2005). *Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis*. Proceedings of HLT/EMNLP 2005, Vancouver, Canada.
- Zuckerman, M. e.a. (1982): *Face and Tone of Voice in the Communication of Deception*. Journal of Personality and Social Psychology, volume 43, number 2, 1982, pp.347-357