

UNIVERSITEIT UTRECHT

MA NEERLANDISTIEK

MASTER SCRIPTIE

Sensatiezoekers

EEN VERGELIJKEND COMPUTATIONEEL ONDERZOEK NAAR HET SENSATIONEEL TAALGEBRUIK IN
DE KOPPEN VAN *DE TELEGRAAF* EN *NRC HANDELSBLAD*

Student:

Lennart VAN WINZUM

4022122

Lennart@Winzum.nl

Begeleider:

dr. Henk PANDER MAAT

Tweede lezer:

dr. Jacqueline EVERS-VERMEUL

31 mei 2021



Universiteit Utrecht

Samenvatting

In dit onderzoek worden de krantenkoppen van *De Telegraaf* en het *NRC Handelsblad* computationeel met elkaar vergeleken door te scannen op twee talige elementen van sensationisme. In dit stuk zijn deze twee elementen intensiverend taalgebruik en emotioneel taalgebruik. Na een kwalitatieve toets, waarbij computationele beoordelingen van verschillende tools worden vergeleken met menselijke beoordelingen van krantentitels, kies ik voor T-Scan voor het constateren van intensiverend taalgebruik, en de *Python* extensie *Pattern* voor het constateren van emotioneel taalgebruik. De verwachting is dat de stereotypes omtrent tabloids en populaire kranten bevestigd zou worden en dat deze koppen dus vaker en in grotere mate intensiverend en emotioneel zijn. Uiteindelijk wordt er geconcludeerd dat het onderscheid tussen populaire en kwaliteitskrant niet zo scherp is als door sommigen misschien gedacht wordt. Zowel bij *NRC Handelsblad* als bij *De Telegraaf* zijn beide elementen van sensationeel taalgebruik aanwezig, alhoewel deze bij *De Telegraaf* iets vaker bij elkaar voorkomen dan bij *NRC Handelsblad*. Desondanks is de belangrijkste conclusie dat er een onverwacht effect meespeelt: dat van onderwerp. Bij verschillende secties worden verschillende resultaten gevonden. Zo valt er in de sectie Financieel nieuws bijvoorbeeld geen significant verschil tussen de kranten te constateren in de hoeveelheid intensiverende titels, terwijl dit bij de secties Binnenland en Buitenland wel het geval is. Ook het gebruik van citaten in de titel hangt samen met een hogere mate van intensiverend en emotioneel taalgebruik, wat ook nog eens krantoverstijgend is. Toekomstig onderzoek zou er goed aan doen om talig sensationisme niet alleen te onderzoeken en te vergelijken op basis van de krant waarin het verschijnt, maar ook op basis van elementen zoals sectie, onderwerp en actor.

Voorwoord

Op het moment van schrijven beginnen we (naar ik hoop) bijna zicht te krijgen op het einde van de coronacrisis. Als ik de kranten lees, waren de jaren 2020 en 2021 een tijd die vol zat met spanning en sensatie. Tegelijkertijd was het voor mij een tijd die, als ik me besef dat ik me gedurende een jaar vooral op mijn studentenkamer heb afgezonderd, erg lang leek te duren. Maar na een lange schrijvensweg is dit interdisciplinaire stuk, waarin zowel mijn technische vaardigheden als mijn Neerlandistieke interesses zijn verweven, eindelijk af.

Ten eerste wil ik graag Henk Pander Maat bedanken. Mijn begeleider die mij met eindeloos veel geduld te woord heeft gestaan, die mijn stukken las, me waardevolle feedback gaf, en mij de juiste kant op wist te sturen wanneer ik weer eens te veel onderwerpen tegelijk aan wilde pakken. Ik heb bijzonder veel geleerd tijdens het uitvoeren van mijn onderzoek en het schrijven van dit stuk en hoop dat ik in de toekomst meer kan doen op deze overlap van taal en digitale analyses.

Ik wil graag mijn nauwe vriendengroep bedanken zonder wie ik de lockdowns niet had overleefd. Stuk voor stuk hebben ze me gesteund, hebben ze interesse getoond voor mijn studieonderwerp dat soms mijlenver van hun eigen vakgebied lag en stonden ze als trouwe honden altijd klaar om een wandeling te maken.

Mijn familie bedank ik graag, juist in de tijden waarin we fysiek zo ver van elkaar af staan is het fijn om via andere wegen nog dicht bij elkaar te zijn.

Ten slotte bedank ik in het bijzonder Marita Vaes, voor haar hulp bij het uitvogelen van de statistieken, lieve grapjes en de zetjes in de rug die ik nodig had bij het afmaken van dit stuk.

Lennart van Winzum

Utrecht, mei 2021

Inhoudsopgave

1	Introductie	5
2	Theoretisch kader	7
2.1	Tabloids	7
2.2	Sensationalisme en taalintensiteit	9
2.3	Intensiteit, sentiment en <i>negativity bias</i>	12
2.4	Het herkennen van emotie in teksten	13
2.5	Krantenkoppen en <i>Relevance Theory</i>	14
3	Onderzoeksvraag en hypotheses	17
4	Methode	19
4.1	Samenstelling onderzoeksmateriaal	19
4.2	Corpus	19
4.3	Onderzoeksmethode	24
4.4	T-Scan	24
4.5	SentiStrength	25
4.6	Pattern	25
4.7	Kwalitatieve toets	26
4.7.1	Conclusie kwalitatieve toets	28
5	Resultaten	29
5.1	Intensiveerders in krantenkoppen	29
5.2	Sentiment in krantenkoppen	33
5.3	Correlatie tussen intensiveerders in koppen en polariteit	38
5.4	Aanhalingstekens	41
6	Discussie	43
6.1	Beantwoording hoofdvraag	43
6.2	De validiteit van tools	45
6.3	Toekomstig onderzoek	45
6.3.1	Vergelijken van gelijke onderwerpen	45
6.3.2	Vergelijking van online en offline vormgeving	46
6.3.3	Analyse op basis van de geografie	46
6.3.4	Het gebruik van de krantenpagina's	47
7	Eindconclusie	49

A	Normaliteit	53
B	Top 10's en valentie	54

1 Introductie

In een onderzoek naar het gebruik van communicatiestrategieën in de headlines van de online media van 28 EU landen concludeerden Orosa, Santorum en Garcia (2017) dat in bijna de helft van de gevallen de aandacht voor de artikelen niet gegenereerd werd door hoge kwaliteit van de inhoud maar door pakkende, provocerende en sensationele headlines die de aandacht opwekken bij de lezer, ongeacht de relevantie van deze titels. De Nederlandse krant *De Telegraaf* was meegenomen in dit onderzoek. In eerste instantie lijkt dit misschien zorgwekkend, gezien een groot deel van de Nederlandse bevolking deze krant gebruikt als primaire bron voor het verzamelen van informatie. Maar in hoeverre is het gebruik van sensationele koppen behouden tot *De Telegraaf*? En in hoeverre is dit behouden tot online media?

Sensationalisme, het fenomeen waarbij niet-neutrale communicatiestrategieën worden toegepast om de aandacht van de lezer te grijpen en te behouden, bestaat al langer dan het internet. Het onderscheid tussen de zogeheten ‘tabloids’ en kwaliteitsjournalistiek dateert bijvoorbeeld uit het begin van de twintigste eeuw (Esser, 1999, pp. 291–292). Hierbij wordt het begrip tabloid gebruikt als pejoratief voor kranten die sensationeel taalgebruik toepassen, en wordt er bij kwaliteitsjournalistiek gedacht aan kranten die dit niet doen. Ook in Nederland is dit onderscheid nog steeds te vinden. Zo wordt *De Telegraaf* door Scholten en Bakker (2006) gerekend tot de populaire kranten, en *NRC Handelsblad* tot de kwaliteitsjournalistiek. In dit onderzoek bekijk ik in hoeverre deze aanname klopt door de koppen van beide kranten computationeel te analyseren op elementen van sensationalisme. Dit definieer ik in dit stuk als een niet-neutrale taalstrategie met twee pijlers, enerzijds intensiverend en anderzijds emotioneel. De onderzoeksvraag die ik heb opgesteld, luidt:

Is er een correlatie tussen het gebruik van intensiveerders en het gemeten sentiment in de koppen, en hoe verschillen populaire en kwaliteitskranten hierin van elkaar?

Als eerste zal ik in het theoretisch kader in hoofdstuk 2 de bestaande literatuur over de relevante fenomenen behandelen: de onstaansgeschiedenis en eigenschappen van tabloids in sectie 2.1, hoe sensationalisme samenhangt met taalintensiteit in sectie 2.2, hoe emotionaliteit samenhangt met sensationalisme in sectie 2.3, het herkennen van emotionele taal in sectie 2.4 en waarom nou specifiek krantenkoppen van belang zijn in sectie 2.5. Vervolgens worden de hoofdvraag en de gevormde deelvragen en hypotheses toegelicht in sectie 3, en wordt de onderzoeksmethode beschreven in sectie 4. Als voorbereidend onderzoek heb ik een kwalitatieve test uitgevoerd waarbij de beoordelingen van computationele tools zijn vergeleken met handmatige beoordelingen, dit is te vinden in sectie 4.7. In sectie 5 behandel ik de resultaten en beantwoord ik de deelvragen, waarna ik in sectie 6 de hoofdvraag beantwoord en inga op inzichten die tijdens

het uitvoeren van het onderzoek aan het licht zijn gekomen. In de conclusie in sectie 7 geef ik mijn uiteindelijke, concluderende antwoord op mijn onderzoeksvraag.

2 Theoretisch kader

2.1 Tabloids

Om dieper in te kunnen gaan op sensationele journalistiek is het belangrijk om eerst te kijken naar het onderliggende proces: tabloidisering. Het woord “tabloid” als benaming voor journalistieke instanties is een relatief nieuwe term die in het begin van de twintigste eeuw zijn intrede deed, en doelt op kranten die nieuws produceren dat vooral gericht is op de verkoopbaarheid van de artikelen. Het heeft een medische achtergrond, waarbij bedoeld werd op het formaat van de pil en het gemak waarbij deze geslikt kan worden. Het gemak dat dit medische begrip met zich meebracht werd gebruikt om Londense dagbladen te typeren, die door hun kleinere formaat gemakkelijk meegenomen konden worden om in het openbaar vervoer gelezen te worden (Esser, 1999, pp. 291–292). De fysieke verandering hing ook samen met verandering van inhoud en vorm. Zo werden de titels groter, foto’s belangrijker en werden berichten soms meer ontspannen en soms sensationeler. Wanneer we tegenwoordig spreken van tabloid wordt er zowel naar het formaat verwezen, als naar de inhoud en vorm (Fang, 1997, p. 104).

De definitie, met een vergelijking met een medicijn dat ‘makkelijk geslikt wordt’, kan opgevat worden als pejoratief. De term tabloidisering wordt dan ook vaak gebruikt om de gang van zaken omtrent dagbladen te bekritisieren. Bon (2016) haalt Sparks (2000) aan en heeft het over een negatief fenomeen dat beschouwd wordt als een ‘crisis voor de democratie’ door de elite en als een onderscheid tussen ‘slechte journalistiek’ versus de ‘goede journalistiek’ van kwaliteitskranten (Bon, 2016, p. 3). Volgens Hauttekeete (2004) is hier bij het gros van de auteurs sprake van een soort ongenoegen over het begrip en is het wetenschappelijke discours dus bij uitstek negatief. Voor de lezers van kranten kan het echter best zo zijn dat sommige elementen van tabloidisering gezien worden als positief omdat ze zich niet ergeren aan een verhoogde focus op onderwerpen als lifestyle. Ook binnen de journalistiek zelf zullen de meningen over tabloidisering variëren van een noodzakelijke modernisering tot een beschadiging van de journalistieke waarde van de krant. Tabloidisering is dus een veelzijdige term, waarover voor- en tegenstanders het nog steeds niet eens zijn (Hauttekeete, 2004, pp. 131–132). Volgens Esser (1999, p. 293) moet tabloidisering gezien worden als een proces op zowel de micro- als de macroschaal. Op de microschaal is het een mediafenomeen waarbij de traditionele kranten en andere mediaformaten veranderen onder de druk van de voorkeur van de lezers en commerciële vereisten. Op de macroschaal is het een sociaal fenomeen waarbij het zowel oorzaak is van, en symbool staat voor, grote veranderingen in de maatschappij zoals minder belang van educatie en een groter belang voor politieke propaganda met een steeds groter wordende politieke kloof als gevolg. Hetzelfde punt maakt

Rowe (2000) na het presenteren van een lijst met veelvoorkomende kenmerken van tabloids. Volgens Rowe bevat tabloidisering, bekeken vanuit een breed perspectief, invloed van sociale, ideologische en economische krachten en structuren op de media-instituten. Omdat deze brede definitie inherent vaag is, is het nuttiger om het begrip nauwer en technisch te bekijken, en het te beschouwen als een eigen genre en een weerspiegeling van het instituut waarin het geproduceerd is, met bestandsdelen die met variërende intensiteiten voorkomen in verschillende mediaformaten (Rowe, 2000, p. 6).

Esser (1999) geeft aan dat er bij tabloidisering sprake is van een ‘besmetting’ van de serieuze media die een ‘tabloid agenda’ aannemen. Daarom is het niet de tabloidpers, maar de kwaliteitspers die onder de loep genomen moet worden. Ook benadrukt hij dat tabloidisering een vaag begrip blijft dat het beste multi-dimensionaal benaderd kan worden, in plaats van het te aarden in één monocultureel analytisch kader (Esser, 1999, pp. 293–294). In dit stuk zal vooral ingegaan worden op taalgebruik en dus de microschaal. De elementen die tot de macroschaal gerekend worden, kunnen eventueel gebruikt worden bij duidingen in de resultaten en discussie.

De precieze definitie van wat een tabloid tot een tabloid maakt, en hoe publicaties tabloidiseren, staat dus nog ter discussie. In zijn stuk „On going tabloid: A preliminary analysis” presenteert Rowe een lijst met typische en overwegend negatieve kenmerken van het fenomeen die na een lezing van een brede selectie met artikelen over tabloids en tabloidisering naar voren komen.

1. decline of orthodox, official or formal institutional politics;
 2. passage (often voyeuristically intrusive) of formerly separate private concerns into the public domain;
 3. sensationalisation and spectacularisation of news, accompanied by the accelerated ephemerality both of media celebrity and news stories;
 4. collapse of the boundaries between news and entertainment (‘infotainment’) and news and advertising (‘advertorial’);
 5. increasing obsession with and fetishisation of celebrity and stardom;
 6. uncontrolled expansion of the media and of media products and uses;
 7. overwhelming concern with rapidly accessible and digestible ‘bites’ of information, frequently coupled with a direct, personalised mode of address;
 8. decline in media ethics and of the ‘serious’, Fourth Estate role of the press, with hypocritical moralising replacing moral seriousness;
 9. drift towards a right-wing, proprietor-sponsored authoritarian populism, and a corresponding failure of the political Left to engage with the life concerns of ‘the people’;
 10. general coarsening of the cultural climate and of individual sensibilities.
- (Rowe, 2000, p. 5)

Rowe geeft deze kenmerken verschillende oorsprongen zouden kunnen hebben. Sommige (2, 3, 4, en 8) komen vooral vanuit de mediaconcerns zelf, andere vooral door invloed van buitenaf (1, 9 en 10) en sommige eerder door een interactie tussen deze twee elementen (5, 6 en 7). Veel definities

van tabloid bevatten een (sub)set van bovenstaande opsomming. Toch is het belangrijk om aan te geven dat de scheiding tussen interne en externe oorzaken niet altijd binair is, dat bovenstaande processen niet altijd lineair verlopen en dat bovenstaande eigenschappen niet exclusief tot tabloids behoren (Rowe, 2000, p. 6). Integendeel, deze nieuwe nieuwswaarden en de manier waarop deze gepresenteerd worden lijkt in de Westelijke wereld invloed te hebben op de presentatie van alle onderwerpen, maar speciaal de serieuze onderwerpen die invloed hebben op de lezer. Door een stijgende tabloidisering is het te verwachten dat deze serieuze onderwerpen ook in steeds grotere mate op een triviale wijze gepresenteerd worden (Bon, 2016, pp. 3–4). Volgens Costera Meijer (2012), die de overlap tussen de twee vormen van journalistiek ook in ogenschouw neemt, is het niet zo dat inspelen op de wensen van het beoogde publiek automatisch leidt tot een verlaging van de nieuwsstandaarden, maar juist een verbreding van de mogelijke narratieven en maakt het journalistiek tot een meer democratisch proces. Kwaliteitsjournalistiek hoeft in een persoonlijke ervaring niet als zodoende ervaren te worden. Hierbij houdt Costera Meijer rekening met drie extra factoren om de kwaliteit te meten: participatie, representatie en presentatie van (door het publiek) waardevol geachte informatie. Bij zowel tabloids- als kwaliteitsjournalistiek is er dus sprake van inspelen op, en interactie met, het publiek (Costera Meijer, 2012, pp. 765–767).

Beide vormen spreken een ander publiek aan en vertegenwoordigen dus ook een ander deel van de maatschappij van de landen waarin de publicaties verschijnen. In dit stuk wordt er gekeken naar de Nederlandse journalistiek en is, in navolging van Schaap en Pleijter (die zich baseren op Scholten en Bakker (2006), gekozen voor *De Telegraaf* als vertegenwoordiging van de tabloid, of ‘populaire kranten’, zoals Scholten en Bakker dat noemen, en *NRC Handelsblad* als vertegenwoordiging van de kwaliteitskranten.

2.2 Sensationalisme en taalintensiteit

De begrippen ‘sensationalisme’ en ‘tabloid’ zijn nauw gerelateerd. Het trekken van de aandacht van een bepaald publiek uit zich vaak in sensationalisme. In de kern doet sensationalisme een beroep op het menselijke instinct door zich te richten op het creëren van emotionele reacties, de empathie van de lezer of nieuwsgierigheid te creëren. Het grijpt de aandacht van de lezer en dramatiseert elementen al dan niet overbodig. Tevens draagt het bij aan het vervagen van de grens tussen tabloids- en kwaliteitsjournalistiek. Het is dan ook niet verbazend dat er overwegend negatief gesproken wordt over het verschijnsel (Kilgo, Harlow, García-Perdomo & Saleverría, 2018, p. 1498-1499). Toch is het begrip volgens sommigen nog onvoldoende gedefinieerd. Volgens Grabe, Zhou en Barnett (2001) is het onderzoek naar sensationalisme gefragmenteerd en bevat het verschillende definities van het concept. Ondanks dat wordt het in het publieke debat veel gebruikt, met een negatieve connotatie. Zij geven aan dat zowel woordenboekdefinities als

definities in verschillende onderzoeken zich vooral richten op het emotionele effect van sensationisme, maar tevens dat empirisch onderzoek naar het effect van sensationisme of de emotionele toestand van de consument van sensationeel nieuws zeldzaam is. Net als bij de scheiding tussen tabloid- en kwaliteitsjournalistiek wordt er aandacht besteed aan het winstmotief: de selectie van sensationele onderwerpen, en het presenteren van onderwerpen op een sensationele manier, kan leiden tot extra aandacht van het publiek (en dus meer omzet). Hierbij melden Grabe et al. (2001) dat er in de meeste onderzoeken echter vooral een focus is op de inhoud - het selecteren van emotionele en sensationele onderwerpen - en er niet ingegaan wordt op de vormelijke elementen van de presentatie. Deze zijn minstens even belangrijk als de selectie van de onderwerpen, aangezien sensationele onderwerpen op een niet-sensationele manier gebracht kunnen worden en vice versa (Grabe et al., 2001, pp. 636-639). Veel van het onderzoek naar sensationisme richt zich op televisiejournisme en de aspecten die daar verband mee houden, maar kijkt niet naar de relatie tussen sensationisme en nieuws in andere media, zoals print en online media. Kilgo et al. bekijken sensationisme in een digitale context en definiëren het begrip op een manier die bovengenoemde aspecten verenigt: *sensationisme is een stijl die een emotie bij de lezer opwekt en die onderwerpen behandelt zoals een tabloid dat doet* (2018, p. 1499).

Het onderscheid tussen vorm en inhoud wordt ook benoemd door Burgers en De Graaf (2013). Zij kiezen ervoor om geen holistische definitie van sensationisme te gebruiken, maar juist om te kijken naar hoe een afzonderlijk element (in dit geval de vorm) kan bijdragen aan sensationisme. Ook zij constateren dat de meeste literatuur zich richt op de vormelijke elementen van televisie, met voorbeelden als zoombewegingen, muziek en interviews met “normale” mensen. In print dragen andere vormelijke elementen bij aan het sensationiseren van een bericht. Bij het bekijken van de stijl van een krant zijn elementen als woordkeuze en de informatiestructuur veelvoorkomend. Hierbij wordt kwaliteitsjournalistiek geassocieerd met een sobere en duidelijke stijl, en is een sensationistische rapportage het tegenovergestelde. Ook wordt woordkeuze gerelateerd aan journalistieke kwaliteit, waarbij tabloidsjournalistiek bestempeld wordt als extra sensationeel. In hoeverre een artikel ervoor kiest om wel of geen neutrale taal te gebruiken kan dus een indicator zijn voor sensationisme (Burgers & De Graaf, 2013, pp. 6-7). Een voorbeeld van niet-neutrale taal zijn intensiveerders. Burgers en De Graaf leggen uit (hierbij gebruikmakend van Hamilton, Hunter en Bagoon (1990) en Hamilton en Stewart (1993)) dat intensiveerders twee dimensies hebben: (1) specificiteit en (2) emotionaliteit. Een specifiek (en intenser) bericht is hierbij meestal concreet en levendig, een minder specifiek (en minder intens) bericht is hierbij meestal abstracter. Het verschil tussen concreet en abstract taalgebruik wordt duidelijk gemaakt in het *Linguistic Category Model* (LCM, Semin & Fiedler,

1988 en 1992). Concreet taalgebruik gebruikt descriptieve werkwoorden (Jan slaat Piet), abstracte taal gebruikt adjectieven (Jan is aggressief). Deze laatste beschrijven niet de besproken situatie maar een kwaliteit van een van de objecten of actoren in de situatie (Burgers & De Graaf, 2013, p. 7). Emotionaliteit is de tweede dimensie van taalintensiteit en is de dimensie die voor deze scriptie het meest van belang is. Emotionaliteit wordt gezien als “the degree of affect expressed in the source’s language” (Burgers & De Graaf, 2013, p. 7). Positieve en negatieve berichten hebben volgens Liebrecht, Hustinx en Van Mulken (2019) verschillende effecten en negatieve berichten hebben vaak een grotere impact (de reden hiervan wordt uitgediept in sectie 2.3). Hiernaast hebben betekenisvolle intensiveerders, waar de polariteit (positief of negatief) van de intensiveerder een herhaling is van de context, een versterkend effect op de kracht van het bericht. Tevens onderstrepen ze het belang van het onderzoeken van de context van het bericht, aangezien de polariteit van een artikel niet goed te berekenen is door het tellen van positieve en negatieve woorden an sich (meer hierover in sectie 2.4)(Liebrecht et al., 2019, pp. 185–186).

De hierboven beschreven benaderingen gaan uit van het effect van intensiveerders op de attitude van de lezer, wat een definitie is uit de sociaal-psychologisch georiënteerde taalwetenschap. Het begrip “taalintensivering” wordt ook behandeld vanuit andere wetenschappelijke domeinen, waarvan de traditionele taalkunde een van de meest prominente is (Renkema, 1997, p. 496). Dit domein ziet de intensiteit niet als een effect op de lezer, maar als een eigenschap in de boodschap zelf en wordt gekenmerkt door een poging tot het categoriseren van intensiveerders in gradaties. Renkema (1997) stelt drie vormen van intensiveerders voor: stillistische (herschrijven door middel van stijlfiguren), semantische (het vervangen van woorden met een specifiekere betekenis, van *mooi* naar *schitterend*) en lexicale (het toevoegen van woorden, bijvoorbeeld het *heel* in *heel gek*). Hiernaast presenteert Renkema 18 categorieën die verdeeld zijn over de bovenstaande drie vormen. In het geval van de semantische vorm wordt er onderscheid gemaakt in werkwoorden (gedaald → *gekelderd*), zelfstandige naamwoorden (schande → *blamage*), bijvoeglijke naamwoorden (groot → *kolossaal*) en bijwoorden (enigszins → *behoorlijk*). Bij de lexicale intensiveerders behoren onder andere de basisversterkers als *erg*, *heel* en *zeer*, kwantificators als *bijna* iedereen, *slechts* weinigen en kwalificators als *puur* suiker en de *onstuimige* zee (Renkema, 1997, pp. 497–498). Pander Maat (2004) baseert zich onder andere op het schema van Renkema en voegt hieraan toe dat er niet alleen sprake hoeft te zijn over *intensiveerders* maar ook van *detensiveerders*. Dit is vooral in negatieve zinnen, waarin de detensiveerder gebruikt kan worden om de zin minder negatief over te laten komen. Denk hierbij aan constructies als *slechts* en *minder dan*: “Wij kochten dit bedrijf voor *minder dan* 5 miljoen euro” (Pander Maat, 2004, p. 213). Volgens het onderzoek van Burgers en De Graaf heeft het gebruik van deze intensiveerders een positief effect op de geobserveerde taalintensiteit en kan dat

een artikel meer nieuwswaardig kan maken (2013, p. 22).

2.3 Intensiteit, sentiment en *negativity bias*

Tom, Pollmann en Goudbeek (2018) beschrijven de asymmetrie in de reacties van facebookgebruikers op de emotionele valentie van berichten: emotionele berichten zorgen voor meer reacties dan niet-emotionele, en gebruikers voelen sneller de behoefte om hun emotionele reactie bekend te maken als het gaat om negatief nieuws (Tom et al., 2018, p. 153). Deze asymmetrie in het menselijke gedrag als het gaat om positieve en negatieve berichten of concepten is al langer bekend in de literatuur. Liebrecht et al. (2019) bekijken de interactie tussen intensiverende taaluitingen en de emotionele polariteit van het bericht (positief of negatief), hierbij onder andere gebruikmakend van de *negativity bias* als kader. Volgens Rozin en Royzman (2001) uit de *negativity bias* zich op vier manieren:

1. **Negative Potency:** Dit principe stelt dat, wanneer er twee gebeurtenissen gegeven worden met een objectief even grote impact, er aan de negatieve gebeurtenis meer impact wordt toegeschreven.
2. **Greater Steepness of Negative Gradients:** De negativiteit van een negatieve gebeurtenis groeit sneller hoe dichterbij de ervaren deze gebeurtenis nadert in tijd en ruimte dan een respectievelijk positieve gebeurtenis groeit. Met andere woorden: een tandartsafpraak wordt negatiever ervaren hoe dichterbij hij komt, dan een aankomend feestje wordt ervaren als positiever.
3. **Negativity Dominance:** Volgens de negativiteitsdominantie worden gebeurtenissen die zowel positieve als negatieve elementen bevatten holistisch als negatiever beoordeeld dan de opsomming van de afzonderlijke elementen in eerste instantie zou vermoeden.
4. **Greater Negative Differentiation:** Negatieve stimuli worden uitgebreider en op meer verschillende manieren gedifferentieerd dan positieve stimuli. Volgens Peeters (1971) is bijvoorbeeld het vocabulaire voor negatieve evaluatieve woorden uitgebreider dan het vocabulaire voor positieve evaluatieve stimuli.

(Rozin & Royzman, 2001, pp. 298–299)

De negativity bias kan worden uitgelegd met behulp van cognitief psychologisch onderzoek dat aantoonde dat negatieve uitingen als sterker worden beoordeeld dan positieve uitingen (Liebrecht et al., 2019, p. 172). Negatieve uitingen krijgen meer aandacht, roepen sneller emotie op en hebben meer invloed op het gedrag van de ontvanger. Dit gedrag heeft twee theoretische oorzaken. Het eerste is een evolutionaire verklaring: negativiteit wordt geassocieerd met angst en gevaar, en positiviteit met veiligheid. Een persoon let automatisch meer op onplezierige informatie dan op plezierige informatie. De tweede verklaring stelt dat negatieve berichten als sterker gezien worden op een sociaal niveau. Volgens wat het Pollyana principe genoemd wordt, verwachten mensen eerder positieve dan negatieve communicatie. Wanneer een bericht onverwervelijk negatief is, wordt deze als intenser gezien dan wanneer een bericht wel te verwerpen is of positief is. Positieve informatie is hier dus de standaard en negatieve informatie wijkt hier vanaf (Liebrecht et al., 2019, p. 172).

De relatie tussen kracht en polariteit wordt door Liebrecht (2015) ook uitgediept in haar proefschrift „Intens krachtig. Stilistische intensiveerders in evaluatieve teksten”. Gedurende de afgelopen vijftig jaar zijn lijsten verschenen van evaluatieve intensiveerders, met hun bijbehorende polariteit en verschillen in intensiteit. Deze lijsten worden samengesteld op basis van de intuïtie van de onderzoekers zelf of door het ondervragen van participanten. Van deze samengestelde lijsten is dankbaar gebruikgemaakt in onderzoeken, maar in de literatuur en onderzoeken naar taalintensiteit is er nauwelijks onderzoek gedaan naar de polariteit van de totale uiting, alleen naar de polariteit van de intensiveerders op zichzelf, maar zoals hierboven al werd aangegeven is polariteit van een boodschap geen simpele cumulatieve kwestie vanwege de negativity bias (Liebrecht, 2015, pp. 200–204).

Hoewel de onderzoeken naar negativiteit in teksten over de Engelse taal gaan, zijn er ook onderzoeken naar de Nederlandse taal verschenen. In Kamoen, Holleman en Van den Bergh (2007) zijn bijvoorbeeld Nederlandstalige contraire woordparen (*makkelijk/moeilijk, boeiend/saai*) gebruikt in een evaluatie van een tekst. Uit dit onderzoek bleek dat de ondervraagden positiever reageerden op een negatieve vraagstelling dan vice versa. Uit latere studies bleek tevens dat respondenten het vaker met negatieve vraagstellingen oneens waren dan met positieve vraagstellingen. Ook voor het Nederlands geldt dat negatieve berichten op een andere manier worden geëvalueerd dan positieve berichten (Liebrecht, 2015, p. 204).

2.4 Het herkennen van emotie in teksten

Volgens Liebrecht et al. (2019) is het onderzoeken van de polariteit van de context waarin intensiveerders staan belangrijk wanneer er onderzoek gedaan wordt naar de relatie tussen polariteit en intensiverend taalgebruik (Liebrecht et al., 2019, pp. 185–186). Trilling en Boumans (2018) bespreken in hun overzicht alle automatische inhoudsanalyses voor Nederlandse teksten die tot in het jaar van het schrijven van het artikel bekend zijn. Hierin wordt vermeld dat Nederlandse sentimentanalyses (automatische onderzoeksmethodes om de emotionele polariteit van een tekst te achterhalen) vaak gebaseerd zijn op verouderde woordenlijsten en niet zo goed gevalideerd worden als hun Engelse varianten (Trilling & Boumans, 2018, p. 10). Veel eerdere onderzoeken naar emotionele valentie in Nederlandse teksten maken dan ook gebruik van vertaaldiensten om de teksten te vertalen naar het Engels, om hier een Engelse onderzoeksmethode op los te laten. Het spreekt voor zich dat dit weer andere problemen met zich meeneemt, de Nederlandse taal verschilt immers van de Engelse en niet elk woord is letterlijk te vertalen. In de afgelopen twee tot drie jaar is er echter veel gebeurd op het gebied van Nederlandse inhoudsanalyses. Eén van de nieuwste ontwikkelingen komt vanaf Google en is een taalanalysemodel genaamd BERT (Bidirectional Encoder Representations from Transformers)

(Devlin, Chang, Lee & Toutanova, 2018). Dit is een model dat getraind is op het herkennen van zinsonderdelen. Hoewel het originele model in eerste instantie getraind is op de Engelse taal zijn er in het laatste jaar nieuwe modellen gebaseerd op de Nederlandse taal: RobBERT (Delobelle, Winters & Berendt, 2020) en BERTje (De Vries, Van Cranenburgh, Bisazza, Caselli, Van Noord & Nissim, 2019). Allebei deze modellen zijn door onderzoekers getraind op het analyseren van Nederlandse teksten. Het herkennen van sentiment gaat met deze modellen veel beter dan met eerdere technieken, zo had RobBERT bij het analyseren van boekreviews een succespercentage van 95.1%.¹

2.5 Krantenkoppen en *Relevance Theory*

Als laatste puzzelstukje ga ik nog nader in op waarom ik de focus op de krantenkop leg ten opzichte van het complete artikel. Het primaire doel van een krantenkop is onderdeel geweest van veel discussie. Het heeft twee communicatiedoelen die met elkaar verbonden zijn. Het eerste doel is dat het de aandacht van de lezer moet trekken en deze moet verleiden om het hele verhaal te lezen (Nir, 1993). Het tweede doel van een goede krantenkop is om een informatierijke samenvatting van het artikel te vormen die tegelijkertijd kort en duidelijk is.

Dor (2003) noemt krantenkoppen *relevance optimizers*. Relevance verwijst naar de *Relevance Theory* gemunt door Sperber en Wilson (1986). Deze communicatietheorie is een poging om de oneindig complexe communicatiesystemen van de mens te reduceren tot een selectie aan variabelen. Het denken van de ontvanger van een bericht wordt hierbij gerepresenteerd door een collectie *assumpties*: kennis en overtuigingen die een persoon heeft, de sterkte hiervan is het niveau tot waarin deze persoon gelooft in deze assumptie. Wanneer een persoon een nieuwe assumptie ontvangt, wordt deze geïnterpreteerd in een *context*: een subset van assumpties die de ontvanger al heeft. Deze elementen worden dan vergeleken in de gedachten van de ontvanger. Als de nieuwe assumptie tegen de oude ingaat, kunnen de oude verzwakt of gewist worden, aan de andere kant kan het de oude versterken als ze elkaar niet tegenspreken. Het cruciale element van deze theorie is waar koppen een grote rol spelen: nieuwe assumpties worden alleen vergeleken met een subset van de assumpties die de ontvanger heeft, omdat het vergelijken met *alle* assumpties een cognitief onmogelijke taak is. Daarom is de context van een ontvanger van een bericht gebaseerd op de assumpties die door de ontvanger gemaakt wordt bij het lezen van het bericht (Dor, 2003, pp. 699–700). Kort gezegd: *Een assumptie is alleen relevant in een context wanneer deze een effect heeft in deze context.*

In deze definitie zijn er drie gevallen waarin een assumptie geen effect kan hebben op de context, en dus irrelevant is. Ten eerste kan een assumptie informatie toevoegen die totaal geen verband

¹Voor meer informatie over RobBERT zie: <https://people.cs.kuleuven.be/pieter.delobelle/robbert/>

heeft met de informatie die al present is in de context. Ten tweede kan een assumptie informatie presenteren die al present is in de context, waardoor deze niets toevoegt, oninformatief is en dus irrelevant is. De laatste mogelijkheid is dat een nieuwe assumptie inconsistent is met de context van eerdere assumpties, maar ook te zwak om deze te beïnvloeden, waardoor het verwerken van deze assumptie niets toevoegt (Sperber & Wilson, 1986, pp. 120–122).

In de praktijk vertrouwen redacteurs bij het schrijven van hun koppen op intuïtie en de vaardigheid die hen aangeleerd is door ervaring. De kunst van het creëren van goede krantenkop is dus niet theoretisch, maar praktisch van aard (Dor, 2003, p. 707). Na een analyse van 134 koppen van nieuwsberichten geeft Dor een lijst van 10 kenmerken die een kop kan hebben, die elk onder de optimalisatie van relevantie vallen:

1. "Een kop moet zo kort mogelijk zijn"
2. "Koppen moeten duidelijk, makkelijk te begrijpen en niet ambigu zijn."
3. "Koppen moeten interesse bij de lezer opwekken"
4. "Koppen moeten nieuwe informatie bevatten"
5. "Koppen moeten *niet* uitgaan van informatie die de lezer niet kent"
6. "Koppen moeten namen en concepten met een hoge 'nieuws waarde' voor de lezers bevatten"
7. "Koppen moeten *geen* namen en bevatten met een lage 'nieuws waarde' voor de lezers"
8. "Koppen moeten het verhaal verbinden met eerder bekende feiten en gebeurtenissen"
9. "Koppen moeten het verhaal verbinden met eerder verwachtingen en assumpties"
10. "Koppen moeten het verhaal op een gepaste manier framen" De kop moet dus overeenkomen met de informatie die het artikel zelf bevat.
(Dor, 2003, pp. 708–716)

Kuiken, Schuth, Spitters en Marx (2017) beschrijven hoe de manier waarop mensen nieuws consumeren, getransformeerd is van het scannen van koppen op papier naar het online lezen van individuele artikelen, eventueel buiten de originele publicatie doordat deze aangeraden of gedeeld wordt door kennissen. Hierbij is het verdienmodel van publicaties veranderd: waar artikelen eerst naast elkaar gepubliceerd werden in één krant verschijnen ze nu vaker individueel online. Het doel van een krantenkop is inmiddels niet langer om een correcte samenvatting te geven van het bijbehorende artikel, maar om de lezer aan te trekken op de link te klikken om het artikel te openen (Chen, Conroy & Rubin, 2015). Kuiken et al. (2017) onderzoeken de specifieke karakteristieken van een effectieve kop van nieuwsartikelen in een digitale omgeving en analyseren hierbij data van Blendle, een Nederlandse online dienst die artikelen van alle grote Nederlandse kranten en tijdschriften aanbiedt. Gebruikers kunnen individuele artikelen kopen. Deze kunnen geselecteerd worden door de gebruikers of aangeraden worden door Blendle in een dagelijkse nieuwsbrief. Deze nieuwsbrief bevat een inleiding en een headline die geschreven wordt door de redactie van Blendle (Kuiken et al., 2017, p. 1301). Om deze reden kunnen Kuiken

et al. (2017) drie factoren analyseren: de originele headline, de herschreven headline en de zogeheten *click-through-rate* (CTR). Hiermee wil Kuiken et al. drie subvragen beantwoorden: hoe worden headlines herschreven? Welke eigenschappen hebben een significante invloed op de performance van die headline? En is het mogelijk om de performance van een headline te voorspellen (Kuiken et al., 2017, p. 1301)? De gebruikte data bestond uit 1828 unieke items die uit 202 nieuwsbrieven gehaald waren. Kuiken et al. hebben gebruikt gemaakt van zelfgeschreven Python-scripts en de library genaamd *Pattern* gemaakt door De Smedt en Daelemans (2012b) om een lijst met eigenschappen uit elke krantenkop te extraheren. Hiernaast gebruiken ze de *Dutch language subjectivity lexicon* van Jijkoun en Hofmann (2008)(Kuiken et al., 2017, pp. 1304–1305).

Kuiken et al. (2017) tonen aan dat de herschreven headlines significant verschillen op alle onderzochte eigenschappen wanneer ze vergeleken worden met het origineel. Geen van de nieuwe headlines was hetzelfde als het origineel (Kuiken et al., 2017, pp. 1311–1312). De herschreven headlines bevatten meer van de eigenschappen die Kuiken et al. (2017) bekijken, maar niet alle elementen toonden een significante verhoging van de effectiviteit. Deze prestatie werd door Kuiken et al. berekend met behulp van de CTR en met het vergelijken van de computationeel verzamelde karakteristieken van de headlines werden de significante verschillen berekend. De significant meetellende eigenschappen waren: het gebruik van korte woorden, de afwezigheid van vragen, de afwezigheid van citaten, het gebruiken van signaalwoorden en het gebruiken van voornaamwoorden. Deze elementen dragen volgens Kuiken et al. significant bij aan een hogere CTR, maar de performance van een headline is niet te voorspellen door een machine-learning model dat getraind is op deze elementen. Er spelen dus meer elementen mee dan enkel de door Kuiken et al. geteste factoren.

3 Onderzoeksvraag en hypotheses

In dit stuk zal het onderscheid tussen tabloid- en kwaliteitsjournalistiek onder de loep genomen worden door te kijken naar het gebruik van intensiveerders in krantenkoppen ten opzichte van de polariteit van het bijbehorende krantenartikel. Binnen dit onderzoek zullen de artikelen bestudeerd worden van twee Nederlandse kranten, *NRC Handelsblad* en *De Telegraaf*. Hoe dit corpus precies geconstrueerd wordt, zal toegelicht worden in 4.2. De hoofdvraag die ik ga beantwoorden is:

Is er een correlatie tussen het gebruik van intensiveerders en het gemeten sentiment in de koppen, en hoe verschillen populaire en kwaliteitskranten hierin van elkaar?

Om deze hoofdvraag te helpen beantwoorden stel ik de volgende deelvragen:

1. *Is er een verschil in het gebruik van intensiveerders in de krantenkoppen tussen de onderzochte kranten?*
2. *Is er een verschil in het gebruik van emotie in de krantenkoppen tussen de onderzochte kranten?*
3. *Is er een correlatie tussen het gebruik van intensiveerders en het gebruikte sentiment?*

Hierbij definieer ik een aantal hypotheses. Zoals uiteengezet in deel 2 heeft tabloidjournalistiek een lange ontstaansgeschiedenis en is het innig verbonden met sensationisme (Bon, 2016, p. 3). Intensiveerders zijn talige een manier waarop een boodschap versterkt kan worden, wat de geobserveerde nieuwswaardigheid kan verhogen (Burgers & De Graaf, 2013). Hoewel het sensationeler brengen van nieuws geobserveerd wordt bij alle vormen van journalistiek, komen de elementen in grotere schaal voor bij tabloids (Rowe, 2000, p. 6). De eerste hypothese is daarom:

- **H1:** Een tabloid (*De Telegraaf*) bevat gemiddeld een hoger percentage intensiverende krantenkoppen dan een kwaliteitskrant (*NRC*)

Hiernaast verwacht ik, dat de grotere schaal van intensiveerders bij tabloids er ook voor zorgt dat de titels gemiddeld intensiever zijn dan intensiveerders bij kwaliteitskranten, omdat in titels van tabloids bijvoorbeeld meerdere intensiveerders gebruikt worden.

- **H2:** Een tabloid (*De Telegraaf*) heeft intensiverende titels met een gemiddeld hoger intensiviteitsniveau dan een kwaliteitskrant (*NRC*)

Onder de door Rowe samengestelde lijst van eigenschappen van tabloidisering staan elementen die zich richten op het persoonlijke, de vervaging van de scheiding tussen privé en publiek domein,

en een verruwing van het publiek debat. Het is te verwachten dat deze eigenschappen ook leiden tot een nieuwsweergave die minder objectief overkomt en dus een met een hoger percentage in geïllustreerde emotie. Hoe de emotie computationeel gemeten wordt is toegelicht in sectie 4.6.

- **H3:** Een tabloid (*De Telegraaf*) bevat een hoger percentage emotie in de koppen dan kwaliteitskrant (*NRC*)

Er is een asymmetrie in de reactie van mensen op positieve en negatieve berichtgeving, waarbij mensen sneller geneigd zijn om tot actie over te gaan bij het lezen van een negatief bericht. Daarom is het te verwachten dat tabloids (met een neiging naar polarisatie en speculatie) hier sneller gebruik van maken dan kwaliteitskranten, wat leidt tot artikelen met een negatievere polariteit die door lezers sneller onthouden en met anderen gedeeld worden:

- **H4:** Een tabloid (*De Telegraaf*) bevat krantenkoppen met een gemiddeld negatievere polariteit dan een kwaliteitskrant (*NRC*)

De laatste twee hypothesen zijn gericht op het onderzoeken van een correlatie tussen intensiveerders en een verhoogd sentiment. Als het zo is dat koppen niet meer een correcte samenvatting van het artikel vormen, maar juist de lezer moeten intrigeren, verwacht ik daar elementen van sensationeel taalgebruik in. Aangezien zowel intensiverend als sentimenteel taalgebruik vormen van niet-neutraal taalgebruik zijn, kan het zo zijn dat er een correlatie is tussen het gebruik van deze twee vormen in sensationele koppen.

- **H5:** Het gebruik van intensiveerders in de kop correleert met een verhoogde polariteit (positief of negatief) in de kop

Ten slotte is het vanwege de eerder genoemde negativity bias de verwachting dat het gebruik van intensiveerders niet alleen samenhangt met verhoogde polariteit in zijn algemeen (dus positief of negatief), maar specifiek met een negatieve polariteit.

- **H6:** Het gebruik van intensiveerders in de kop hangt samen met een negatieve polariteit in de kop

4 Methode

4.1 Samenstelling onderzoeksmateriaal

In dit onderzoek onderzoek ik de headlines en bijbehorende artikelen van twee Nederlandse kranten, waarbij één gezien wordt als meer kwalitatief (*NRC Handelsblad*) en één als meer sensationeel (*De Telegraaf*), zoals gedefinieerd door Scholten en Bakker (2006). Het samenstellen van dit corpus heeft verschillende stappen gehad. Met de programmeertaal Python heb ik een script geschreven dat in de online database *Nexis Uni* kijkt, die voor studenten toegankelijk is („Nexis Uni - Utrecht University Library - Universiteit Utrecht”, 2019). Dit script zoekt op basis van dag naar de artikelen die per krant verschenen zijn en slaat deze informatie lokaal op. Omdat er niet in de database gezocht kan worden met een lege opdracht, zoekt het script naar alle artikelen die die dag verschenen zijn die het lidwoord ‘de’ bevatten. Dit betekent dat een artikel niet wordt meegenomen als dat specifieke lidwoord die dag niet in de titel of het artikel verscheen. De kans dat een heel krantenartikel op geen enkele plek het lidwoord ‘de’ gebruikt leek echter verwaarloosbaar klein.²

4.2 Corpus

Het door mij geschreven Python script heeft niet alleen de krantenkop en het artikel opgehaald uit *Nexis Uni*, maar ook andere informatie wanneer dat beschikbaar was. Bij de complete lijst bevat elk artikel de volgende automatisch meegenomen eigenschappen:

- De bron (*NRC Handelsblad*, *De Telegraaf*)
- Verschijningsdatum van het artikel (jaar-maand-dag)
- Titel van het artikel
- Inhoud van het artikel (complete artikeltekst)
- Druk van het document (bijvoorbeeld: 3 september 2019, gehele oplage)
- Sectie van de krant met paginanummer (bijvoorbeeld: REGIO, blz. 3)
- Lengte van het artikel
- *Dateline* (locatie van waaruit de redacteur/correspondent rapporteert)
- Onderwerp (automatisch gegenereerd door Nexus Uni)

Bij artikelen van *De Telegraaf* heeft *Nexis Uni* geen dateline aangegeven, het script zoekt daarom in de eerste 150 tekens naar een liggend streepje, dat in de meeste gevallen voorgegaan wordt door de schrijver en de locatie, en deelt de voorliggende tekst onder dateline in. Voor *NRC* geldt dat er in sommige gevallen themawoorden gebruikt worden boven de titel om de context aan te geven van het onderwerp van het artikel. Een voorbeeld van deze themawoorden is te zien in Figuur 1.

²De scripts zijn bruikbaar voor mensen met een *Nexis*-account en kunnen per mail verstuurd worden.

BUZZFEED

'Trump vroeg Cohen in Congres te liegen over contact met Moskou'

AMSTERDAM. Trumps toenmalige advocaat Michael Cohen heeft op verzoek van Trump in het Congres gelogen over diens zakelijke belangen in Rusland ten tijde van de presidentsverkiezingen. Dat schrijft de Amerikaanse nieuwssite *Buzzfeed* op basis van twee anonieme bronnen rond het onderzoek. Cohen verklaarde in 2017 dat een project voor een Trump Tower in Rusland was stilgelegd vóór de eerste Republikeinse voorverkiezing in januari 2016. Eind november bleek dat dit niet klopte. Cohen heeft zich tot ongeveer juni 2016 ingezet bij de Russische overheid voor die 'Tower'. Ook zou Trump een plan van Cohen hebben gesteund om tijdens de presidentiële campagne de Russische president Poetin te bezoeken om de onderhandelingen te versnellen. 'Regel het', zou Trump tegen Cohen hebben gezegd, aldus *Buzzfeed*. (NRC)

Figuur 1: Themawoorden in het rood boven de titel in NRC Handelsblad

In sommige gevallen zijn deze woorden rechtstreeks overgenomen en door Nexis Uni voor de titel geplaatst. Dit is een probleem, aangezien deze woorden niet behoren tot de krantenkop. In andere gevallen zijn verschillende varianten van dezelfde krantenkop neergezet in de krantenbank, gescheiden door een puntkomma. Om de titels die rechtstreeks uit de krantenbank worden overgenomen op te schonen zijn een paar regels vastgesteld:

1. Titels die een puntkomma (;) bevatten met meerdere variaties van die titel worden gesplitst op het eerste teken. De eerste variatie wordt meegenomen.
2. Themawoorden die van de titel gescheiden zijn met behulp van een regeleinde worden verwijderd.
3. Themawoorden zonder hoofdletters die voor de kop staan worden verwijderd.
4. Mocht er sprake zijn van een procentteken (%) in de sectie ECONOMIE, dan wordt het deel van de kop na dit teken beschouwd als de krantenkop, in de economische sectie van *NRC Handelsblad* worden de koppen in sommige gevallen namelijk voorgedaan door de beursprestaties van het onderwerp, die aangegeven worden met een stijging in procenten.
5. Dubbele koppen en artikelen worden enkelvoudig gemaakt en inhoudsloze koppen worden verwijderd. Hiertoe behoren ook de elementen die automatisch meegenomen zijn uit de krantenbank maar feitelijk geen artikelen bevatten: *Inhoud*, *quote*, *colofon*, *No Headline In Original*.
6. Titels waarin aanhalingstekens (" " of ' ') gevonden worden krijgen de eigenschap 'citaat' mee. We gaan ervan uit dat er in deze koppen een actor in het verhaal aan het woord komt of dat een uitdrukking die deze actor gebruikt wordt aangehaald.

Deze werkwijze haalt een groot deel van de ambiguïteit weg die bij het invoegen van de titels in de krantenbank is aangebracht. In het geval van de themawoorden bij *NRC Handelsblad* worden deze vanaf 21 maart 2019 gescheiden van de titel met een regeleinde, als deze themawoorden al meegenomen worden. Dit is namelijk niet altijd het geval. Om onzekerheid tegen te gaan is ervoor gekozen om de dagen tot 21 maart 2019 buiten de uiteindelijke selectie te laten. In totaal heeft het script gezocht naar alle artikelen van de te onderzoeken kranten die in het jaar 2019 verschenen zijn. Deze zijn volgens de hiervoor vermelde regels opgeschoond. Uiteindelijk resulteerde dat in een totale hoeveelheid van 44726 artikelen van *De Telegraaf* en 26333 artikelen van *NRC Handelsblad* gepubliceerd in het jaar 2019, te zien in Figuur 2.

Aantal artikelen per jaar		Gemiddeld aantal woorden	
	2019	outlet	2019
De Telegraaf	44.726 62,94%	De Telegraaf	344,6
NRC Handelsblad	26.333 37,06%	NRC Handelsblad	592,5
TOTAAL	71.059 100,00%		

Figuur 2: Overzicht van het totale aantal artikelen uit 2019

De totale selectie aan artikelen bevat ook niet-journalistieke secties als advertenties, roddels en tv-gidsen. Zoals aangegeven in sectie 2.1 hebben de begrippen *tabloid* en *tabloidisering* te maken met zowel de inhoud als de vorm. Omdat dit onderzoek betrekking heeft op de stijl, hoe berichten gebracht worden, is het dus belangrijk om secties met elkaar te vergelijken die ook daadwerkelijk vergelijkbaar zijn, en niet inhoudelijk te veel van elkaar verschillen. In Figuur 3 zijn de secties van beide kranten te vinden die meegenomen worden in dit onderzoek, met de gegevens van de overgebleven segmenten vermeld per krant in Figuur 4.

Aantal artikelen per sectie per krant 21/03/2019 - 31/12/2019

	De Telegraaf		NRC Handelsblad	
	Aantal artikelen	Percentage	Aantal artikelen	Percentage
BINNENLAND	8.043	44,71%	953	5,30%
BUITENLAND	1.571	8,73%	995	5,53%
ECONOMIE			2.463	13,69%
FINANCIEEL	3.964	22,04%		
TOTAAL	13.578	75,48%	4.411	24,52%

Gegevens artikelen

	Aantal artikelen	Gemiddelde lengte titel	Gemiddelde lengte artikel
NRC Handelsblad	4.411	6,2	599
De Telegraaf	13.578	4,4	350
TOTAAL	17.989	4,9	413

Figuur 3: Overzicht van de mee te nemen secties van 21 maart tot en met 31 december 2019

Er is hier gekozen voor een nauwe selectie van secties die enigszins overeenkomen: binnenlands nieuws, buitenlands nieuws en economisch nieuws. Regionaal nieuws in het geval van *De Telegraaf* en nieuws uit bijlages die alleen in bepaalde steden meegeleverd worden (Amsterdam, Rotterdam) in het geval van *NRC Handelsblad* zijn ook buiten beschouwing gelaten. In *De Telegraaf* zijn onder de sectie *Binnenland* ongeveer evenveel artikelen verschenen als bij *NRC Handelsblad*. Dit komt omdat *NRC Handelsblad* ook een Sectie “In Het Nieuws” bevat, waaronder zowel binnen- als buitenlands nieuws behandeld kan worden. Omdat hierin in de verwerking geen onderscheid gemaakt kan worden is ervoor gekozen deze sectie buiten beschouwing te laten. Zoals Bon (2016) ook al opmerkt bevat *NRC Handelsblad* gemiddeld artikelen met meer woorden dan *De Telegraaf* (Bon, 2016, pp. 20–21), 599 bij *NRC Handelsblad* versus 347 woorden per artikel bij *De Telegraaf*. Ditzelfde onderscheid is terug te zien in de lengte van de titels, gemiddeld 4,3 woorden per titel bij *De Telegraaf* versus 6,2 woorden per titel bij *NRC Handelsblad*. Ten slotte valt het op dat het totaal aantal artikelen in het overzicht (17989) iets minder is dan het totaal aantal artikelen in het onderzoek (18009). Het overzicht is gegenereerd in het programma Tableau waarbij de optie om duplicaten te negeren aanstond. In het corpus zitten echter verschillende artikelen met dezelfde kop. In het onderzoek zelf is het corpus handmatig gecontroleerd om dit soort fouten te voorkomen. Omdat het om een miniem verschil gaat tussen grote aantallen heb ik besloten om dit overzicht te laten staan, omdat het verschil in verhoudingen tussen de kranten wel belangrijk is om aan te geven.

Aantal artikelen per sectie per krant
21/03/2019 - 31/12/2019

	De Telegraaf		NRC Handelsblad	
	Aantal artikelen	Percentage	Aantal artikelen	Percentage
ACHTERPAGINA			551	1,56%
ADVERTENTIE	3	0,01%		
AMSTERDAM			315	0,89%
AUTOVISIE	239	0,68%		
BOEKEN			551	1,56%
CARRIÈRE			123	0,35%
COVER			27	0,08%
CULTUREEL SUPPLEMENT			336	0,95%
CULTUUR	869	2,47%	560	1,59%
DE GIDS			2.545	7,22%
FILM			532	1,51%
HET BLAD			112	0,32%
HET GROTE VERHAAL			3	0,01%
IN HET NIEUWS			2.501	7,10%
JOURNAAL	754	2,14%		
LEVEN			1.388	3,94%
LIFESTYLE	2.019	5,73%		
MEDIA			181	0,51%
OPINIE			908	2,58%
OPINIE DEBAT			411	1,17%
PRIVE	570	1,62%		
RADIOTV	671	1,90%		
REDACTIESERVICE	498	1,41%		
REGIO	5.565	15,79%		
REPORTAGE	331	0,94%		
ROTTERDAM			266	0,75%
SPECIAL			19	0,05%
SPORT	5.969	16,94%	946	2,68%
THEMA	879	2,49%		
VOORPAGINA			1.392	3,95%
VROUW	251	0,71%		
WATUZEGT	1.711	4,86%		
WEEKEND			441	1,25%
WETENSCHAP			535	1,52%
ZOMER	313	0,89%		
TOTAAL	20.642	58,57%	14.599	41,43%

Gegevens artikelen

	Aantal artikelen	Gemiddelde lengte titel	Gemiddelde lengte artikel
NRC Handelsblad	14.599	5,7	599
De Telegraaf	20.642	4,1	340
TOTAAL	35.241	4,8	450

Figuur 4: Overzicht van de buitengesloten secties

4.3 Onderzoeksmethode

Het onderzoek bestaat uit een kwantitatief en een kwalitatief deel. Als eerste zijn de krantenkoppen geanalyseerd op aanwezige intensiveerders, hierbij gebruikmakend van T-Scan (Pander Maat, Kraf, van den Bosch, Dekker, Gompel, Kleijn, Sanders & Sloot, 2014).³ Hierna werd de emotionele polariteit van de artikelen geanalyseerd. Als derde ging ik over op de analyse van de data. Hierin maakte ik gebruik van t-tests en loglineaire analyses.

Voor het computationele onderzoek maak ik gebruik van twee tools: T-Scan en Pattern, waarvan de werking in sectie 4.4 en sectie 4.6 wordt uitgelegd. Hiernaast heb ik een kwalitatieve proef uitgevoerd om de werking van deze tools toe te lichten en vast te stellen óf, en zo ja wát, er ontbreekt aan deze programma's.

4.4 T-Scan

T-Scan is een tool om Nederlandse teksten automatisch te analyseren en kan naar ongeveer driehonderd elementen zoeken op tekst-, zins- en woordniveau (Pander Maat et al., 2014, p. 53). Met de door Renkema (1997) aangegeven tweedeling in benadering (traditioneel taalkundige versus sociaal-psychologische) in het achterhoofd kunnen we ook vaststellen welke benadering hier gebruikt wordt: de taalkundige. Deze tool onderzoekt namelijk de eigenschappen van een bericht, en niet hoe deze door ontvangers van dat bericht wordt opgenomen. In dit onderzoek wordt T-Scan gebruikt om de intensiveerders in de krantenkoppen aan te geven. Hierbij put T-Scan uit een woordenlijst van ongeveer 3700 uitdrukkingen (Pander Maat, Kraf & Dekker, 2020, p. 62). Deze zijn verzameld van RBN-Lijsten met nomina, adjectieven en werkwoorden (*Referentiebestand Nederlands*⁴), het Woordenboek van Nederlandse Intensiveringen van de website onderwoorden.nl en sterke bijwoorden uit een verzameling bijwoorden van de SoNaR-frequentielijst⁵. Van elk woord is vervolgens ook nagegaan of of ze afleidingen hebben die ook versterkend werken, bijvoorbeeld de corresponderende werkwoorden en bijwoorden van een sterk nomen (overheersing > overheersen, overheersend) (Pander Maat et al., 2020, p. 97). Deze afleidingen zijn alleen meegenomen in gevallen dat ze ook versterkend werkten. Intensiveerders worden in de handleiding gedefinieerd als:

- *Sterke* woorden of uitdrukkingen die verwijzen naar een bijzondere graad van en bepaalde eigenschap (*fenomenaar*)
- *Versterkende* woorden die de interpretatie versterken van de uiting waar ze in staan (*hogelijk*)

(Pander Maat et al., 2020, p. 96)

³<https://webservices-1st.science.ru.nl/portal/>

⁴http://tst-centrale.org/images/stories/producten/documentatie/rbn_documentatie_nl.pdf

⁵<https://taalmaterialen.ivdnt.org/download/tstc-sonar-corpus/>

Deze definities komen overeen met de door Renkema uiteengezette varianten van intensiverders, zoals eerder gemeld in sectie 2.2, respectievelijk de *semantische* (vervangen van woorden met een specifiekere betekenis) en de *lexicale* (het toevoegen van woorden). Hoewel de door Renkema genoemde *stillistische* (het gebruik van stijlfiguren) variant in eerste instantie niet expliciet in de handleiding lijkt te staan, wordt er wél melding van gemaakt. T-Scan houdt rekening met 190 combinaties die een intense variant zijn van een enkelvoudige uitdrukking. Enkele voorbeelden van dit soort uitdrukkingen zijn *geen bal* (niets), *voor geen cent* (niet) en *stinkende best* (best) (Pander Maat et al., 2020, p. 107). Na het analyseren beoordeelt T-Scan de tekst met een score van nul tot duizend die berekend wordt op basis van de verhouding wel- en niet-intensiverende woorden in de tekst. Een tekst zonder intensiverende woorden krijgt dus de score nul, een tekst met alleen maar intensiverende woorden de score duizend.

4.5 SentiStrength

SentiStrength (Thelwall, Buckley, Paltoglou, Cai & Kappas, 2010) is een algoritme dat emotionele valentie uit teksten kan halen.⁶ Hoewel het in eerste instantie gemaakt is voor korte informele Engelse stukken tekst is het ook toepasbaar op andere soorten tekst, en is het met behulp van door andere samengestelde lijsten ook toepasbaar op andere talen. SentiStrength is uniek ten opzichte van andere sentimentanalysetools in dat het twee waarden teruggeeft in plaats van één: een negatieve én een positieve waarde. Dit is belangrijk omdat veel teksten niet maar één polariteit bevatten, maar zowel positieve als negatieve elementen, wat het meten van één polariteit dan natuurlijk bemoeilijkt. Net als T-Scan is SentiStrength in de kern gebaseerd op woordenlijsten, in het geval van de laatste op een woordenlijst met woorden die sentiment uitdrukken met verschillende wegingen (Thelwall et al., 2010, p. 2545). SentiStrength komt zelf alleen met de Engelse woordlijsten, maar er zijn ook woordlijsten voor andere talen beschikbaar. In het geval van dit onderzoek wordt gebruikgemaakt van Nederlandse woordenlijsten met bijbehorende wegingen.⁷ Na het analyseren van een stuk tekst geeft SentiStrength twee waarden terug op een ordinale schaal: een positieve schaal van 1 (niet positief) tot 5 (heel positief) en een negatieve schaal van -1 (niet negatief) tot -5 (extreem negatief).

4.6 Pattern

Pattern (De Smedt & Daelemans, 2012b) is een pakket voor Python dat geïnstalleerd kan worden om nieuwe functies toe te voegen. Bij deze functies behoren onder andere *Web Mining*, netwerkanalyse, *Machine Learning* en *Natural Language Processing* (NLP) Die laatste is voor dit onderzoek het meest van belang, De NLP-module bevat namelijk ook een optie voor het

⁶<http://sentistrength.wlv.ac.uk/>

⁷<https://github.com/felipebravom/StaticTwitterSent/tree/master/extra/SentiStrength/dutch>

analyseren van sentiment. Dit is gebaseerd op ongeveer 4000 Nederlandse lemma's met verschillende wegingen en getest op het analyseren van Nederlandse boek- en CD-recensies. Hier behaalde het een correctheidspercentage van ongeveer 70% (De Smedt & Daelemans, 2012a, p. 3571). Het algoritme geeft twee waarden terug: een polariteitsschaal van -1 tot +1, die het sentiment weergeeft, en een subjectiviteitsschaal van 0 tot 1. Bij de eerste schaal is -1 zeer negatief en +1 zeer positief. Bij de tweede schaal is 0 heel objectief en +1 zeer subjectief.

4.7 Kwalitatieve toets

Om te zien in hoeverre de hierboven vermelde tools overeenkomen met een menselijke beoordelaar is ervoor gekozen om een steekproef te vergelijken met een menselijke beoordelaar. Met behulp van een Python-script is een bestand gegenereerd met een puur willekeurige greep van 200 titels per krant. Deze zijn vervolgens handmatig beoordeeld met drie vijfpuntsschalen: een van intensiteit (hoe intensiverend komt de kop over), een van negativiteit (hoe negatief komt de kop over) en een van positiviteit (hoe positief komt de kop over). Voor elke schaal geldt dat bij 1 het fenomeen niet voorkomt en bij 5 het fenomeen extreem voorkomt. Er is gekozen voor het gescheiden schalen voor wat betreft de emotionele polariteit - dus de aan- of afwezigheid van positiviteit en negativiteit worden apart van elkaar beoordeeld. De reden hiervoor is dat SentiStrength twee aparte schalen levert, en er gekeken moest worden naar hoe en of deze afzonderlijk correlatie zouden vertonen met handmatige beoordelingen. Om de uiteindelijke beoordeling van het artikel te berekenen worden resultaten van de negatieve vijfpuntsschaal per artikel afgetrokken van de resultaten van de positieve vijfpuntsschaal per artikel. Dit levert als eindresultaat een negenpuntschaal op die loopt van -4 tot +4, waarbij -4 correspondeert met zeer negatief, en +4 correspondeert met zeer positief. Ten slotte zijn negen van de 400 koppen buiten het onderzoek gehouden, Dit waren dubbele en foutieve koppen genaamd *Inhoud*, *In het kort*, *Dit weekend op DFT.NL* en *No Headline In Original*. De overige koppen zijn vervolgens beoordeeld door T-Scan, SentiStrength en Pattern.

Om te bepalen of slechts 1 beoordelaar het gehele corpus betrouwbaar kon coderen, is allereerst gecheckt of dit betrouwbaar kan. Een set van 50 krantenkoppen is door 2 beoordelaars beoordeeld. Uit de interbeoordelaarsbetrouwbaarheidsanalyse van de drie schalen bleek dat de ICC-waarden voldoende waren om dit te kunnen doen: 0.868 voor de intensiverende schaal, 0.788 voor de positieve schaal, 0.927 voor de negatieve schaal.

De voorspellingen van deze tools over de krantenkoppen in het sample zijn vergeleken met de uitkomsten van de handmatige beoordelingen. Op basis van deze analyse vertonen zowel Pattern ($r_s = .231, p < .001$) als SentiStrength ($r_s = .247, p < .001$) een relatief lage correlatie met de handmatige beoordelingen. Ook wanneer we, in navolging van de handmatige analyses hierboven,

de handmatig als neutraal beoordeelde koppen weglaten zien we soortgelijke correlaties bij Pattern ($r_s = .266, p < .001$) en SentiStrength ($r_s = .259, p < .001$). Natuurlijk gaat het hierbij om een sample waarbij het menselijk oordeel van tevoren bekend is, wat in het echte onderzoek, met duizenden koppen, een stuk moeilijker kan. Wat wél kan is om bij de uitkomsten van beide beoordelingen de neutrale beoordelingen (dus rond de nul) en vooral te kijken naar de uitschieters aan beide kanten, bijvoorbeeld in navolging van De Smedt en Daelemans (2012a), die voor Pattern een regel opstellen die alle beoordelingen tussen -0.10 en +0.10 weglaat. Om de prestaties van beide tools met deze benadering met elkaar te vergelijken is er gekozen voor het weglaten van alle beoordelingen van SentiStrength die, wanneer de positieve en negatieve assen bij elkaar opgeteld worden, uitkomen op nul. Wanneer we hierbij de resultaten van SentiStrength ($N=122$) vergelijken met Pattern ($N=126$) zien we dat Pattern een iets sterkere correlatie vertoont met de handmatige beoordelingen ($r_s = .415, p < .001$) dan SentiStrength ($r_s = .353, p < .001$). Bij het nóg groter maken van de weggelaten sectie bij SentiStrength, dus bij het weglaten van de cumulatieve oordelen van -1 tot en met +1, is er geen significante correlatie meer vast te stellen ($N = 50, r_s = .256, p < .07$). Wanneer Pattern bekeken wordt met het weglaten van oordelen tussen -0.25 en +0.25 vertoont het een nog sterkere correlatie ($N = 77, r_s = .47, p < .001$), en dit geldt ook bij een marge van -0.30 tot +0.30 ($N = 58, r_s = .49, p < .001$).

Met deze resultaten in beschouwing genomen worden blijkt dus dat, bij het gebruik van tools om sentiment in kranten koppen te analyseren, er significantere conclusies getrokken kunnen worden wanneer er een subset van het totale corpus in beschouwing genomen wordt in plaats van het corpus als geheel. In het geval van de laatste twee selecties gaat het dus selecties van 19,69% en 14,83%. Bij een corpus van ongeveer 18.000 artikelen blijven er dan nog rond 3.600 over (20% van 18.000), wat nog steeds een aanzienlijke selectie is.

De voorspellingen van T-Scan zijn vergeleken met de uitkomsten van de handmatige beoordelingen. Deze vertonen een significante correlatie ($N = 391, r_s = .49, p < .001$). De op woorden gebaseerde benadering van T-Scan lijkt dus overeen te komen met handmatige beoordelingen zonder dat hiervoor een voorselectie nodig bij is, zoals wel het geval bij de analyses van het sentiment hierboven. Deze overeenkomst zien we ook terug bij het vergelijken van de verschillen in de beoordelingen van T-Scan tussen verschillende kranten. Koppen van *De Telegraaf* worden gemiddeld als intensiever beoordeeld dan koppen van *NRC Handelsblad*, volgens een t-test is dit verschil significant ($p < .001$).

	N	M	SD
NRC Handelsblad	199	15.13	46.63
De Telegraaf	192	25.62	69.95

Tabel 1: Gemiddelde beoordeling van de intensivering van de koppen volgens T-Scan

4.7.1 Conclusie kwalitatieve toets

Voor het grootschalige corpusonderzoek is uiteindelijk gekozen voor twee tools die het meest overeenkomen met handmatige beoordelingen. Voor het analyseren van sentiment in de koppen heb ik gebruikgemaakt van Pattern, in combinatie met een marge waarin beoordelingen tussen -0,25 en +0,25 weggelaten worden. Zoals aangetoond vertonen de resultaten dan een sterkere correlatie met de handmatig beoordeelde titels en zullen deze dus dichterbij de opvattingen van de lezers van de titels in het grote corpus liggen. Voor het beoordelen van intensiteit in de koppen heb ik gebruikgemaakt van T-Scan, zonder het toepassen van een marge. De resulterende data is getest op normaliteit; de skewness en kurtosis is te vinden in Tabel 1 in de appendix.

5 Resultaten

In de hieropvolgende subsecties worden de deelvragen zowel kwantitatief als kwalitatief benaderd. Voor het kwalitatief benaderen zijn er ranglijsten gegenereerd waarin de tien meest extreem beoordeelde titels staan op basis van de gemeten positiviteit, negativiteit en intensiteit per sectie per krant. Op deze manier kan er wat inzicht gegeven worden in waar de tools naar gezocht hebben. Deze lijsten zijn terug te vinden in appendix B.

5.1 Intensiveerders in krantenkoppen

De eerste twee hypothesen zijn gevormd om de eerste deelvraag te helpen beantwoorden en hebben betrekking op het gebruik van intensiveerders in de koppen. Beide hypothesen gaan ervan uit dat een tabloid - in dit geval *De Telegraaf* - meer gebruik zal maken van sensationeel taalgebruik en niet-neutrale taal en dus meer intensiveerders in de koppen zal bevatten dan een kwaliteitskrant (*NRC Handelsblad*). De eerste hypothese verwacht dat een tabloid een hoger percentage intensiveerders in de koppen heeft dan een kwaliteitskrant, en dus ook een groter percentage koppen bevat waarin intensiverende taal wordt gebruikt. Aangezien sensationisering een proces zou zijn van zowel inhoud als vorm is het belangrijk om waar mogelijk de inhoud gelijk te trekken als er gekeken gaat worden naar de vorm. In dit geval is het dus gunstig om niet alleen de verschillen te bekijken tussen de kranten maar ook per sectie tussen de kranten, dus de secties Binnenland, Buitenland en Financieel. Om te kijken naar de verhouding tussen wel en niet intensiverende koppen is er een binair construct gemaakt waarbij alle koppen die door T-Scan zijn beoordeeld als intensiverend een 1 hebben gekregen, en alle andere koppen een 0 hebben gekregen. Er is vervolgens een log-lineaire analyse uitgevoerd. Bij een log-lineaire analyse wordt de relatie tussen meer dan twee categoriale variabelen onderzocht door de te kijken naar hoe de variantie in de geobserveerde frequenties binnen de categorieën het beste verklaard kunnen worden. In dit geval wordt er gekeken naar de secties, de soort krant en het intensiverend construct als variabelen. Uit deze analyse blijkt een significante driewegsinteractie ($X^2(2) = 24.081, p < .001$). Dit betekent dat, bekeken per sectie, de verhouding in wel en niet intensiverende titels verschilt per krant.

Om dit verschil verder uit te diepen en te verkennen is er een contingentietabel gemaakt per sectie en het totaal, te vinden in Tabel 2. Over het totaal genomen was de interactie significant: $X^2(1) = 6.926$, $p = .008$. Wanneer er gekeken wordt naar het beeld per sectie is te zien dat enkel de Binnenland ($X^2(1) = 19.532$, $p < .001$) en Financieel ($X^2(1) = 4.785$, $p = .029$) significante effecten vertonen. Bij de sectie Buitenland is er geen effect ($X^2(1) = 0.057$, $p = .812$). Wanneer er gekeken wordt naar de verhouding van intensiverende en neutrale titels is te zien dat, over het geheel genomen, in de titels van *De Telegraaf* een groter percentage intensiverende titels te vinden is: 12.6% bij *De Telegraaf* tegenover 11.1% bij *NRC*. Ditzelfde verschil is terug te zien in de sectie Binnenland, hier heeft *De Telegraaf* een intensiverend aandeel van 14.1% ten opzichte van een intensiverend aandeel van 8.9% bij *NRC Handelsblad*. De sectie Financieel gaat tegen deze trend in: 9.2% intensiverende titels bij *De Telegraaf* en 10.9% intensiverende titels bij *NRC Handelsblad*. Wat betreft het beantwoorden van de eerste hypothese is er dus een gemengd beeld te zien. De eerste hypothese, die verwacht dat een tabloid gemiddeld een hoger percentage intensiverende titels heeft kan alleen voor de sectie Binnenland worden **aangenomen**. De sectie Financieel laat een omgekeerd beeld zien, de hypothese kan voor deze sectie **niet worden bevestigd**.

Sectie			Neutraal	Intensiverend	Totaal
Binnenland	De Telegraaf	Aantal	6922	1135	8057
		Percentage	85.9%	14.1%	100 %
	NRC Handelsblad	Aantal	869	85	953
		Percentage	91.1%	8.9%	100 %
Buitenland	De Telegraaf	Aantal	1368	208	1571
		Percentage	86.8%	13.2%	100 %
	NRC Handelsblad	Aantal	860	135	995
		Percentage	86.4%	13.6%	100 %
Financieel	De Telegraaf	Aantal	3602	367	3969
		Percentage	90.8%	9.2%	100 %
	NRC Handelsblad	Aantal	2194	269	2463
		Percentage	89.1%	10.9%	100 %
Totaal	De Telegraaf	Aantal	11887	1710	13597
		Percentage	87.4%	12.6%	100 %
	NRC Handelsblad	Aantal	3923	489	4412
		Percentage	88.9%	11.1%	100 %

Tabel 2: Contingentietabel aandeel intensiverende titels in secties van De Telegraaf en NRC Handelsblad

Ik constateer een een gemengd beeld, de verdeling van intensiverende en niet-intensiverende titels lijkt te verschillen afhankelijk van de sectie. Om het verschil tussen de twee kranten beter te begrijpen worden in Tabel 3 de meest intensiverende titels van eenzelfde sectie met elkaar vergeleken.

NRC Handelsblad		De Telegraaf	
Titel	valentie	Titel	valentie
'Hoe meer risico, des te groter de euforie'	571,43	Steeds krachtiger	1000,00
Nieuwe KPN-baas loyaal tot op het bot	571,43	Chantage	1000,00
Hypotheekrentes dalen heel snel	500,00	Lofzang	1000,00
Veel meer meldingen van medicijntekorten	500,00	Noodkreten	1000,00
De Russen wacht nog meer onheil	500,00	Lievelingen	1000,00
Albert Heijn neemt nog meer Marqt-winkels over	428,57	Jammerklacht	1000,00
Snel: pure spaarder eerlijker belasten	400,00	Jarenlang extra speelplezier	666,67
Chinezen zijn écht heel anders	400,00	'Veel meer overstappen'	666,67
Brillengigant wil nog meer winkels	400,00	'Woningen onbetaalbaar'	500,00
Het geld is er, nu nog meer curlers	375,00	Pijnlijk duidelijk	500,00

Tabel 3: Top 10 Financieel - Intensief van NRC Handelsblad en De Telegraaf

Hierin vallen drie dingen op: het gebruik van aanhalingstekens, het verschil in titellengte en het hoogteverschil van de titelbeoordelingen. De aanhalingstekens worden besproken in discussiepunt 5.4. De laatste twee genoemde eigenschappen lijken in zeker opzicht samenhang te hebben. Bij de top van *De Telegraaf* zien we zeer korte titels, van vaak maar één of twee woorden. Bij diezelfde toppen zien we ook de hoogste beoordelingen van intensiviteit. Dit heeft waarschijnlijk meer te maken met de criteria die gebruikt worden door de tool: de verhouding van het aantal wel- en niet-intensieve woorden, op een maximum van een score van 1000. Een korte zin waar twee intensiverende woorden in staan zal dus als intensiever beoordeeld worden dan een langere zin waar ook twee intensiverende woorden in staan. In het discussiepunt 6.2 zal een laatste blik geworpen worden op wat dit betekent voor toekomstig onderzoek.

Hoe zit het met de gemiddelde hoogte van de intensiteit in de intensiverende titels? Hoewel Intensiteit niet normaal verdeeld is (zie de Tabel 1 in de appendix), is er toch voor gekozen om een t-test uit te voeren, aangezien deze bij het vergelijken van grote aantallen robuust is tegen schendingen van normaliteit (Bartlett, 2017). Het resultaat van deze test, en de testen per sectie, zijn te vinden in Tabel 4. Over het geheel bekeken zijn de intensiverende titels van *De Telegraaf* gemiddeld significant intensiever dan de intensiverende titels van *NRC Handelsblad*. Hetzelfde patroon is tevens terug te vinden bij elke onderzochte sectie.

Dit betekent dus ook dat zelfs in de sectie Financieel, waar *NRC Handelsblad* een groter aandeel intensiverende koppen had, de intensiverende koppen van *De Telegraaf* gemiddeld toch intensiever waren. Dit was ook al terug te zien in Tabel 3. De hypothese die verwacht dat een tabloid intensiverende titels bevat die gemiddeld intensiever zijn dan een kwaliteitskrant dat heeft, kan dus voor alle secties worden **aangenomen**.

Secties	Kranten				t	df
	Telegraaf		NRC			
	M	SD	M	SD		
Binnenland	284.554	181.715	187.823	77.994	9.641**	164
Buitenland	298.091	201.705	173.838	72.076	8.121**	279
Financieel	239.527	138.218	184.645	82.724	6.234**	611
Totaal	276.537	176.984	182.214	79.120	16.908**	1819

*p < 0.05, **p < 0.01

Tabel 4: t-testen intensiverders in intensiverende titels totaal en per sectie (indien van toepassing zijn er gecorrigeerde vrijheidsgraden)

Ter beantwoording van de eerste deelvraag kan er dus gezegd worden dat er zeker een verschil is in het gebruik van intensiverders in de krantenkoppen tussen tabloids en kwaliteitskranten. Hoewel niet elke sectie hetzelfde beeld laat zien qua verdeling van wel en niet-intensiverende titels, is het wél zo dat titels van een tabloid gemiddeld intensiverender zijn dan de titels van een kwaliteitskrant.

5.2 Sentiment in krantenkoppen

De tweede deelvraag is beantwoord met behulp van twee hypothesen. De eerste hypothese heeft betrekking tot het aantal emotionele koppen per krant. Er wordt verwacht dat een tabloid vaker gebruik zal maken van emotionele taal, en dus een groter percentage koppen zal bevatten waarin emotionele taal wordt gebruikt. Om dit te kunnen beantwoorden is gebruikgemaakt van de marge die gedefinieerd is in sectie 4.7. Hierbij is een construct gemaakt waarbij de koppen met een Pattern-sentiment-score onder de -0.25 en boven de 0.25 zijn beoordeeld als 'emotioneel' en de koppen tussen die marge als neutraal. Er is een log-lineaire analyse uitgevoerd om de verdeling van emotionele en niet-emotionele koppen te analyseren per sectie van de twee kranten. Uit deze analyse blijkt geen significant driewegseffect (Secties x Krant x Emotioneel). Omdat er geen drieweginteractie blijkt, wordt er bij het beantwoorden van de twee hypothesen geen onderscheid gemaakt tussen de secties van de kranten, maar enkel op de krant zelf.

De *Krant x Sentiment*- interactie was significant ($X^2(1) = 5.044, p = .025$). In Tabel 5 zijn de respectievelijke verhoudingen te zien. Bij *NRC Handelsblad* werden 843 (17.9%) van de koppen geteld als emotioneel en bij *De Telegraaf* werden er 2353 (16.9%) van de koppen geteld als emotioneel. Dit lijkt er in eerste instantie op te wijzen dat er een verschil is in het gebruik van emotie in de koppen wanneer deze twee kranten vergeleken worden. Tevens lijkt dit verschil het omgekeerde te zijn van wat er verwacht wordt: juist de kwalitatieve krant heeft een groter percentage emotioneel beoordeelde titels, terwijl er verwacht zou worden dat een tabloid meer gebruik maakt van niet-neutrale taal.

		Neutraal	Emotioneel	Totaal
De Telegraaf	Aantal	11294	2303	13597
	Percentage	83.1%	16.9%	100 %
NRC Handelsblad	Aantal	3622	790	4412
	Percentage	82.1%	17.9%	100%
Totaal	Aantal	14916	3093	18009
	Percentage	82.7%	17.3%	100%

Tabel 5: Contingentietabel Krant x Emotionaliteit

De hypothese die stelt dat een tabloid een groter aandeel emotionele titels in de koppen bevat kan worden bestempeld als **verworpen**. Wanneer er onderscheid gemaakt wordt op het soort krant is te zien dat er bij *NRC Handelsblad* een hoger aandeel emotionele titels te vinden is (17.9%) dan bij *De Telegraaf* (16.9%).

Naast het onderscheid tussen tabloid en kwaliteitskrant was ook een andere significante interactie vast te stellen: *Sectie x Sentiment* ($X^2(2) = 13.269$, $p = .001$). Een significant onderscheid in verdeling van emotioneel en neutraal is dus ook te maken op basis van sectie in plaats van op basis van krant. Kijkend naar de contingentietabel zien we dat er in de sectie Buitenland ondergemiddeld veel emotionele titels te constateren zijn (15.0%). Verder bevatte de sectie Financieel niet significant meer emotionele titels (17.3%), en bevatte de sectie Binnenland net wat meer emotionele titels (17.7%). Blijkbaar bevatten de secties dus onderwerpen die in verschillende hoeveelheden emotioneel behandeld worden.

		Neutraal	Emotioneel	Totaal
Binnenland	Aantal	7414	1597	9011
	Percentage	82.3%	17.7%	100 %
Buitenland	Aantal	2180	386	2566
	Percentage	85.0%	15.0%	100%
Financieel	Aantal	5322	1110	6432
	Percentage	82.7%	17.3%	100%
Totaal	Aantal	14916	3093	18009
	Percentage	82.7%	17.3%	100%

Tabel 6: Contingentietabel Sectie x Emotionaliteit

De tweede hypothese hield zich bezig met het gemiddeld gemeten sentiment in de koppen. Een asymmetrie in de gepresenteerde emotie zou zorgen voor een gemiddeld negatievere berichtgeving, die mensen sneller aanzet tot actie, het verwerken en delen van het bericht. De verwachting was dat tabloids daar sneller gebruik van zouden maken dan kwaliteitskranten, en dus een gemiddeld negatievere polariteit in de koppen zouden hebben dan kwaliteitskranten. Om het verschil binnen de emotionele titels verder uit te diepen is er een construct gemaakt waarbij de emotionele titels gescheiden zijn op polariteit: positieve titels kregen een 1 en negatieve titels een 0. In combinatie met de variabelen Krant en Sectie is hiermee een loglineaire analyse uitgevoerd. Hieruit bleek dat er in dit geval wél een significant driewegseffect vast te stellen was: ($X^2(2) = 12.792$, $p = .002$). Het soort sectie speelt dus mee bij het correct verklaren van de verdeling in positieve en negatieve titels, in tegendeel tot de verdeling in emotioneel en niet-emotioneel, en zal dus meegenomen worden. De verdeling van de emotionele titels is vervolgens in een kruistabel gezet, te vinden in Tabel 7. De bijbehorende Chi-kwadraatstoetsen gaven aan dat over het totaal genomen er geen significante relatie waar te nemen is ($X^2(1) = 1.605$, $p = .205$), maar dat deze alleen binnen de secties afspeelen. In de sectie Binnenland heeft *NRC Handelsblad* een hoger percentage negatieve titels (60.4%) ten opzichte van *De Telegraaf* (51.5%) ($X^2(1) = 5.494$, $p = .019$). In de sectie Buitenland heeft *De Telegraaf* juist een hoger percentage negatieve titels (59.0%) ten opzichte van de kwaliteitskrant (46.1%) ($X^2(1) = 6.192$, $p = .013$). Bij de sectie Financieel heeft *NRC Handelsblad* een hoger percentage negatief beoordeelde titels (53.1%) ten opzichte van de tabloid (46.9%) ($X^2(1) = 3.990$, $p = .046$).

			Negatief	Positief	Totaal
Binnenland	De Telegraaf	Aantal	721	679	1400
		Percentage	51.5%	48.5%	100 %
	NRC Handelsblad	Aantal	119	78	197
		Percentage	60.4%	39.6%	100 %
Buitenland	De Telegraaf	Aantal	138	96	234
		Percentage	59.0%	41.0%	100 %
	NRC Handelsblad	Aantal	70	82	152
		Percentage	46.1%	53.9%	100 %
Financieel	De Telegraaf	Aantal	314	355	669
		Percentage	46.9%	53.1%	100 %
	NRC Handelsblad	Aantal	234	207	441
		Percentage	53.1 %	46.9%	100 %

Tabel 7: Contingentietabel aandeel emotionele titels in secties van De Telegraaf en NRC Handelsblad

Bij het behandelen van de hypothese is er onderscheid gemaakt op positief en negatief sentiment. Om meer inzicht te krijgen in hoe de gebruikte tool onderscheid maakt in sentiment is in Tabel 8 de top 10 positief beoordeelde titels van de sectie Binnenland te vinden van beide kranten.

NRC Handelsblad		De Telegraaf	
Titel	valentie	Titel	valentie
Ooit vrolijke' SP zoekt de weg omhoog	.90	Grenzeloos gezellig	1.00
Soldaat wil nu echt net zo veel loon als schoonmaker	.88	Steengoede bevalling	1.00
Zonnebaars: geliefd voor in de vijver, nu verboden in EU	.80	'Zorg niet zo briljant'	1.00
Dit examen was makkelijk te lezen'	.75	Alles voor perfect plaatje	1.00
Tot in Apeldoornen Utrecht weten ze: in de Wielewaalkunje makkelijk kraken	.75	Prachtig uithangbord voor Nederland	1.00
Mooie zomer? De lol is erwel een beetje vanaf	.70	Prins: 'Niemand is perfect...'	1.00
De dokter kreeg altijd 't beste stuk vlees'	.68	Zonder twijfel de allerbeste	1.00
Wachtlijsten voor speciaal onderwijs'	.60	'Het is vooral heel gezellig'	1.00
'Zo top is ons onderwijs niet meer'	.60	Stikstofleed als de perfecte vuilnisdag	1.00
Afkeer van de regio? Amsterdam vindt vooral zichzelf heel bijzonder	.60	'Heel leuk dat ze hier komen'	.96

Tabel 8: Top 10 Binnenland - Positief van NRC Handelsblad en De Telegraaf

Wanneer de sentimentsranglijsten beschouwd worden valt iets interessants op; in de positief beoordeelde titels zitten namelijk nog steeds negatieve titels: *'Zorg niet zo briljant'*, *Prins: 'Niemand is perfect...'*, *'Zo top is ons onderwijs niet meer'*. Dit lijkt er op te wijzen dat het zoeken op woordniveau ervoor zorgt dat zowel semantische betekenis als negatie niet altijd goed meegenomen wordt. In deze voorbeelden worden de woorden 'perfect', 'briljant' en 'top' door Pattern blijkbaar als positief meegenomen, terwijl er in de complete kop een negatie is met 'niet' of 'niemand'. Dit gegeven plaatst dus wel een kanttekening bij de werkwijze: als het voor de tool moeilijk is om te bepalen wanneer een titel positief is, zegt het gemeten verschil bij de positieve titels in de sectie Binnenland dan nog wat? Als het zo is dat de selectie positieve titels nog steeds

negatieve titels bevat, en Pattern blijkbaar moeite heeft met het maken van onderscheid tussen deze twee uitersten, is het goed om een alternatieve metriek op te stellen. In het geval van deze scriptie is dat een construct waarbij van alle titels alleen de absolute waardes meegenomen wordt. Op deze manier wordt er alleen gekeken naar de hoogtes van het sentiment dat in de titels gemeten werd. Toch zullen de positief en negatief beoordeelde titels in het onderzoek meegenomen worden, aangezien er in de kwalitatieve toets in sectie 4.7 aangetoond is dat deze beoordelingen genoeg overeenkomen met menselijke beoordelingen wanneer de gebruikte marge van +0.25 tot -0.25 in acht genomen wordt.

Er wordt binnen de secties door de kranten verschillend omgegaan met positieve en negatieve koppen, vandaar dat het goed zou kunnen zijn om deze afzonderlijk van elkaar te bekijken. Er zijn t-testen uitgevoerd van het gemiddelde in positieve titels, negatieve titels en titels bij elkaar. Hiernaast is er ook een t-test uitgevoerd op de absolute waardes van de titels. Deze t-testen zijn hiernaast ook allemaal uitgevoerd op de artikelen van de secties Binnenland, Buitenland en Financieel, om te zien of hier verschil in zat. Deze zijn allemaal te vinden in Tabel 9.

Emotie	Secties	Kranten				t	df
		Telegraaf		NRC			
		M	SD	M	SD		
Positief	Binnenland	.495	.168	.454	.140	2.379*	104
	Buitenland	.459	.150	.465	.151	-.306	176
	Financieel	.484	.161	.481	.160	.185	560
	Totaal	.488	.165	.472	.154	1.690	1495
Negatief	Binnenland	-.519	.149	-.502	.150	-1.163	838
	Buitenland	-.531	.159	-.514	.162	-.683	206
	Financieel	-.489	.158	-.487	.152	-.159	546
	Totaal	-.512	.153	-.496	.153	-1.917	1594
Beide	Binnenland	-.027	.531	-.123	.491	2.547*	264
	Buitenland	-.124	.512	.014	.514	-2.602*	384
	Financieel	.027	.511	-.032	.508	1.907	1108
	Totaal	-.021	.525	-.046	.507	1.179	1411
Absoluut	Binnenland	.507	.159	.483	.148	2.035*	1595
	Buitenland	.501	.159	.488	.158	.784	384
	Financieel	.486	.159	.484	.156	.210	1108
	Totaal	.500	.159	.485	.154	2.447*	3091

*p < 0.05, **p < 0.01

Tabel 9: t-testen sentiment in titels totaal en per sectie (indien van toepassing zijn er gecorrigeerde vrijheidsgraden)

Zoals te zien heeft *NRC Handelsblad* in de sectie Binnenland een gemiddeld negatievere polariteit in de koppen. Dit geldt zowel voor de positief beoordeelde koppen als de positieve en negatieve koppen bij elkaar. Daarentegen heeft *De Telegraaf* bij de sectie Buitenland een gemiddeld negatievere scores. De hypothese die verwacht dat *De Telegraaf* gemiddeld negatievere sentimenten in de koppen heeft is bij dezen dan ook **verworpen**. Het is afhankelijk van de sectie waarin de koppen verschijnen óf er een verschil is en welke kant dit verschil op helt. Wat ten

slotte nog opvalt is dat, wanneer de absolute beoordelingen van de titels bekeken worden, de titels van *De Telegraaf* gemiddeld wél hoger scoren dan de titels van *NRC Handelsblad*. Dit zegt ons dus dat, hoewel de titels van de *NRC* een groter percentage emotionele titels bevatten, en deze in sommige secties gemiddeld negatiever zijn dan die van *De Telegraaf*, de emotionele titels van *De Telegraaf* gemiddeld wel emotioneler zijn dan die van *NRC Handelsblad*.

5.3 Correlatie tussen intensiveerders in koppen en polariteit

Om de derde deelvraag te beantwoorden zijn er twee hypothesen gevormd. De eerste heeft betrekking tot de correlatie tussen intensiveerders en het gemeten sentiment met de verwachting dat het gebruik van intensiveerders in de kop correleert met een verhoogde emotionaliteit. Omdat er hier niet per se onderscheid gemaakt hoeft te worden in positieve en negatieve emoties is er een construct gemaakt waarin de absolute waarde van het gemeten sentiment buiten de eerder vastgestelde marge meegenomen wordt. Koppen die eerder negatief waren, en dus een beoordeling onder nul hadden, zijn dan positief. De andere hypothese had ook betrekking tot de polariteit, en verwachtte dat het gebruik van intensiveerders vooral samenhangt met een negatieve polariteit in de kop. Voor deze hypothese zijn twee constructen gemaakt waarin óf alleen alle positief beoordeelde koppen boven de marge meegenomen worden, óf alle negatief beoordeelde koppen onder de marge meegenomen worden. Omdat intensiteit niet normaal verdeeld is kan er geen Pearson Product-Moment uitgerekend worden (Akoglu, 2018, p. 92). Daarom is ervoor gekozen om in plaats daarvan te kijken naar de Spearman Rho rangcorrelatiecoëfficiënt. Er is voor gekozen om alleen de significante correlaties boven de .1 in oogschouw te nemen, deze duiden nog een kleine correlatie aan (Akoglu, 2018, p. 92). Het resultaat van deze berekeningen is te zien in Tabel 10.

Sectie	Absoluut (n)	Sentiment	
		Positief (n)	Negatief (n)
Alle	.027 (3093)	.044 (1497)	-.016 (1596)
De Telegraaf	.042* (2303)	.068* (1130)	-.017 (1173)
NRC Handelsblad	-.044 (790)	-.057 (803)	.011 (423)
Binnenland	.034 (1597)	.083* (757)	.009 (840)
Buitenland	.114* (386)	.107 (178)	-.075 (208)
Financieel	-.030 (1110)	-.044 (562)	-.003 (548)

*p < 0.05, **p < 0.01

Tabel 10: Spearman Rho coëfficiënten van sentimenteel en intensiverend taalgebruik. Coëfficiënten boven de .1 zijn groen gemaakt.

Uit Tabel 10 blijkt dat er in de meeste gevallen geen correlatie vast te stellen is tussen het gebruik van intensiveerders in de titels en het gebruik van emotioneel taalgebruik in de titels. Alleen bij het absolute construct bij de sectie Buitenland lag de coëfficiënt netaan boven de .1. Zelfs als dit meegenomen wordt duidt dit op een zeer zwakke relatie (Akoglu, 2018, p. 92). De hypothesen die betrekking hebben tussen vermeende correlaties tussen emotioneel en intensiverend taalgebruik in de titels kunnen allebei worden **verworpen**.

Toch kan er ook op een andere manier gekeken worden naar de data. Tot nu toe is er voor het onderzoeken van deze deelvraag gekeken naar het correleren van de hoogtes van beide maatstaven. Een andere manier van kijken is het gebruik maken van de eerder gemaakte binaire constructies: of een titel emotioneel is of niet (dus binnen of buiten de vastgestelde marge), en of

een titel intensiverend is of niet (dus een uitslag boven nul heeft gehaald of niet). Met deze gegevens kan vervolgens een driewegs loglineaire analyse gemaakt worden over de relatie Krant x Intensiverend x Emotioneel. Ik constateer hierbij een significante driewegsinteractie ($X^2(1) = 18.310, p < .001$). Kijkend naar de contingentietabel 11 merk ik een verschil op tussen *De Telegraaf* en *NRC Handelsblad* waarbij de intensiverende titels van de tabloid vaker ook gecategoriseerd worden als emotioneel (29.5%) dan bij de kwaliteitskrant het geval is (22.1%).

Sectie			Neutraal	Emotioneel	Totaal
De Telegraaf	Neutraal	Aantal	10089	1798	11887
		Percentage	84.9%	15.1%	100 %
	Intensiverend	Aantal	1205	505	1710
		Percentage	70.5%	29.5%	100 %
NRC Handelsblad	Neutraal	Aantal	3241	682	3923
		Percentage	82.6%	17.4%	100 %
	Intensiverend	Aantal	381	108	489
		Percentage	77.9%	22.1%	100 %
Totaal	Neutraal	Aantal	13330	2480	15810
		Percentage	84.3%	15.7%	100 %
	Intensiverend	Aantal	1586	613	2199
		Percentage	72.1%	27.9%	100 %

Tabel 11: Contingentietabel overlap emotioneel en intensiverende titels van De Telegraaf en NRC Handelsblad

Hoe deze zinnen er dan uitzien is terug te vinden in Tabel 12, waar een paar willekeurige titels uit de respectievelijke secties neergezet zijn, met hun respectievelijke scores. Dat intensiveerders en sentimenteel taalgebruik niet snel met elkaar in titels gemeten worden, kan snel verklaard worden door de werking van de tools te bekijken, bijvoorbeeld op de woorden die gemeten worden.⁸ Beide lijsten hebben een lengte van rond de 4000 woorden of frases. In de toppen met de als uiterst intensiverend beoordeelde titels, terug te vinden in de Appendix, is te zien dat de meest intensiverende titels kort zijn. Zo bevat de als extreem intensief beoordeelde titels van *NRC Handelsblad* bevatten in veel gevallen slechts één woord: *Flapuit*, *Chantage*, *Lievelingen*, *Jammerklacht*, *Lofzang*. De kans dat dat ene woord aanwezig is in zowel de zoeklijst van de intensiveerders als de zoeklijst van het sentiment is natuurlijk vrij klein. Bij de titel *Weerzinwekkend* van *De Telegraaf* in Tabel 12 is dit bijvoorbeeld toevallig het geval.

⁸lijst met sentiment afkomstig van <https://github.com/clips/pattern/blob/master/pattern/text/nl/nl-sentiment.xml>, lijst met intensiveerders afkomstig van ontwikkelaars van T-Scan

Krant		N.I.	W.I.
De Telegraaf	N.E.	Blinde Marianne kon zelf stem uitbrengen (0/0)	'Allemaal onzin' (0/1000)
		Gehele ov ligt dinsdag stil (0/0)	Teddybeer echte ijzervreter (0/666.67)
		Overval op hotel in Rotterdam (0/0)	Reünie van radicalen (0/333.33)
		Groningse student op rantsoen (0/0)	Volop Nederlandse wijn (0/333.33)
		Vlamingen wild van Veluwe (0/0)	Ravage door fiets op spoor (0/200)
	W.E.	Transitie lukt alleen met betrouwbare overheid (0.80/0)	Weerzinwekkend (-0.90/1000)
		India dregt vergeefs naar koffiemiljardair (-0.78/0)	'Dan ben je niet helemaal lekker' (0.60/666.67)
		Heeft de havo zijn beste tijd gehad? (0.68/0)	Steengoede bevalling (1.00/500)
		Boetegeld illegaal gokken verdubbeld (-0.60/0)	Draaimolen pretpark draait op volle toeren (0.30/500)
		Feest met een staartje (0.50/0)	Onbewaakt is levensgevaarlijk (-0.60/333.33)
NRC Handelsblad	N.E.	'Durf als docent kleur te bekennen' (0/0)	Geterroriseerd door wapengeweld (0/333.33)
		'De samenleving is te complex om als normaal persoon nog te begrijpen' (0/0)	'WW-export' blijft een splijtzwam (0/250)
		Toezicht op AIVD stemt vaker in met tap of hack (0/0)	De wanhoop onder boeren is door heel India voelbaar (0/222.22)
		De brommerrijder verdedigt zichzelf wel (0/0)	Peildatum brengt onheil voor bioboeren (0/200)
		Grensregio's gebaat bij meer Europese samenwerking (0/0)	Zo'n oude koelkast vréét stroom (0/200)
	W.E.	Hoera, slecht economisch nieuws! (-0.88/0)	Gruwelijk, maar genoeg bewijs voor genocide? (-1.00/400)
		'Het Westen koos voor de makkelijke weg in Congo' (0.75/0)	Gevangenis Epstein 'mensonterend' (-0.70/333.33)
		'De dokter kreeg altijd 't beste stuk vlees' (0.68/0)	Onderhandelen voor extra zetel (0.4/250)
		Zonder afscheid slaat verdriet naar binnen (-0.65/0)	Soldaat wil nu echt net zo veel loon als schoonmaker (0.88/200)
		Lula mogelijk voorlopig vrij door uitspraak Hof (0.48/0)	De perverse prikkels van de Amerikaanse autoleningen (-0.80/142.86)

Tabel 12: Wel- en niet-intensieve (W.I./N.I.) en Wel- en niet-emotionele (W.E./N.E.) Titels van NRC Handelsblad en De Telegraaf (sentimentscore/intensiviteitsscore), sentimentcores binnen de gestelde marge van +0.25 en -0.25 zijn ook beoordeeld met 0

Om de derde deelvraag, het vermeende verband tussen intensiverend en emotioneel taalgebruik, te beantwoorden kan er geconstateerd worden dat er met de gebruikte onderzoeksmethodes dus nagenoeg geen correlatie te vinden is tussen het gemeten intensief taalgebruik en het gemeten sentiment in een titel, zowel bij tabloid als bij kwaliteitskrant. Wel kan er vastgesteld worden dat intensiverende titels bij een tabloid als *De Telegraaf* iets vaker als emotioneel beoordeeld wordt dan het geval is bij een kwaliteitskrant als *NRC Handelsblad*.

5.4 Aanhalingstekens

Tijdens het bekijken van de toppen van de titels begonnen aanhalingstekens op te vallen als een nieuw element dat misschien mee zou tellen bij het wel of niet intensief en/of emotioneel zijn van de titels. Zowel bij de tabloid als de kwaliteitskrant zijn deze aanwezig. Een aanhalingsteken wordt ingezet om aan te geven dat hier een mens aan het woord is (geweest), het is niet vreemd om te verwachten dat de gebruikte taal minder neutraal is en dat er in dit geval meer intensief taalgebruik te constateren is. Om deze reden is er ook gekeken of dit taalgebruik een extra dimensie toevoegt ten opzichte van de secties. Met behulp van een construct zijn alle titels waarin twee aanhalingstekens - enkel (') of dubbel (") - voorkwamen gemarkeerd als citaat. Hierna is er een loglineaire analyse uitgevoerd met 4 dimensies: secties, krant, aanhalingstekens en intensiverend taalgebruik. Uit deze analyse blijkt dat de vierwegsinteractie niet significant was ($X^2(2) = 1.162$, $p = .559$). Dit betekent dus dat er geen verschil te constateren is in in hoe de twee verschillende kranten intensiverend taalgebruik toepassen bij aanhalingstekens wanneer er onderscheid gemaakt wordt in secties. Wel zijn er twee driewegsinteracties geconstateerd. De eerste is de interactie die hierboven al behandeld is: Sectie over soort krant over intensiverend taalgebruik. De andere driewegsinteractie die significant is, is die van sectie op soort krant op aanhalingstekens: $X^2(2) = 8.981$, $p = .011$. Er is een contingentietabel gemaakt met deze factoren: Tabel 13.

			Niet-Citaat	Citaat	Totaal
Binnenland	De Telegraaf	Aantal	6353	1704	8057
		Percentage	78.9%	21.1%	100 %
	NRC Handelsblad	Aantal	781	173	954
		Percentage	81.9%	18.1%	100 %
Buitenland	De Telegraaf	Aantal	1372	199	1571
		Percentage	87.3%	12.7%	100 %
	NRC Handelsblad	Aantal	841	154	995
		Percentage	84.5%	15.5%	100 %
Financieel	De Telegraaf	Aantal	3344	625	3969
		Percentage	84.3%	15.7%	100 %
	NRC Handelsblad	Aantal	2116	347	441
		Percentage	85.9%	14.1%	100 %
Totaal	De Telegraaf	Aantal	11069	2528	13597
		Percentage	81.4%	18.6%	100 %
	NRC Handelsblad	Aantal	3738	674	4412
		Percentage	84.7%	15.3%	100 %

Tabel 13: Contingentietabel aandeel citerende titels in secties van De Telegraaf en NRC Handelsblad

De twee kranten laten in verschillende verhoudingen mensen aan het woord in de titel. Bij de sectie Binnenland bevat *De Telegraaf* een significant ($X^2(1) = 4.702, p = .030$) groter aandeel citaten in de titels: 21.1% ten opzichte van 18.1%. Bij de sectie Buitenland bevat *NRC Handelsblad* juist weer significant ($X^2(1) = 4.055, p = .044$) meer citaten: 15.5% ten opzichte van 12.7%. In de sectie Financieel was geen significant verschil op te merken: $X^2(1) = 3.259, p = .071$. Wat hiernaast opvalt is het tweewegseffect tussen intensief taalgebruik en aanhalingstekens: deze is significant ($X^2(1) = 40.781, p < .001$). Zoals te zien in de contingentietabel 14 was maar liefst 15.9% van de citerende koppen intens ten opzichte van 11.4% van de niet-citerende koppen.

		Niet-intens	Intens	Totaal
Geen citaat	Aantal	13116	1691	14807
	Percentage	88.6 %	11.4%	100 %
Citaat	Aantal	2694	508	3202
	Percentage	84.1%	15.9%	100%
Totaal	Aantal	15810	2199	18009
	Percentage	87.8%	12.2%	100%

Tabel 14: Contingentietabel citaten x intensiverend taalgebruik

Ook op het gebied van emotie is er gekeken naar de invloed van citaten. Om te zien hoe en of deze bijdragen aan de gemeten emotie zijn er twee vierwegs loglineaire analyses uitgevoerd: sectie op krant op citaat op emotioneel (of de titel emotioneel is of niet), en sectie op krant op citaat op emotionaliteit (of de titel positief of negatief is). Geen van beide analyses had een significante vierwegsrelatie: respectievelijk $X^2(2) = .221, p = .895$ en $X^2(2) = 1.628, p = .443$. Hiernaast was citaat wel significant gerelateerd aan emotie ($X^2(1) = 38.925, p < .001$) maar niet aan positieve of negatieve emotionaliteit ($X^2(1) = .957, p = .328$). Zoals te zien in de contingentietabel (Tabel 15) bevatten de citerende titels in hogere mate emotionaliteit (21.1%) dan niet-citerende titels (16.3%).

		Neutraal	Emotioneel	Totaal
Geen citaat	Aantal	12390	2417	14807
	Percentage	83.7 %	16.3%	100 %
Citaat	Aantal	2526	676	3202
	Percentage	78.9%	21.1%	100%
Totaal	Aantal	14916	3093	18009
	Percentage	82.8%	17.2%	100%

Tabel 15: Contingentietabel citaten x emotioneel taalgebruik

Het is dus te zien dat, in de titels waar mensen aan het woord komen, vaker intensieve en emotionele taal wordt gebruikt. Deze principes overstijgen de soort krant en het citaat.

6 Discussie

6.1 Beantwoording hoofdvraag

De gestelde hoofdvraag luidt: *Is er een correlatie tussen het gebruik van intensiverders in de koppen en de gemeten polariteit van de titels, en hoe verschillen tabloids en kwaliteitskranten hierin van elkaar?* Op basis van de hiervoor besproken resultaten concludeer ik dat er met de in dit stuk besproken werkwijze **geen correlatie** tussen gemeten intensiverders en gemeten sentiment in de titels is. **Wel** heb ik een paar interessante inzichten opgedaan die hieronder gedeeld worden.

Zoals verwacht heeft de tabloid, in dit geval *De Telegraaf* een hoger percentage intensiverende titels, die ook gemiddeld intensiever zijn. Het intensiverende aspect van sensationisme, besproken in sectie 2.2 lijkt hierbij dus inderdaad een significant verschil tussen de kranten te veroorzaken. Hiernaast was emotie de andere factor van de definitie van sensationisme die in dit onderzoek is geformuleerd. Zoals aangegeven in sectie 2.3 zouden we op basis van de Negativity Bias verwachten dat emotionele titels vaker bij populaire kranten te vinden zijn en dat deze ook vaker negatief zijn. In tegenstelling tot deze verwachtingen bevatte *NRC Handelsblad* een hoger percentage emotionele titels, die in de sectie Binnenland ook nog eens negatiever waren. Het enige punt waarop de tabloid wel hoger scoorde, was in het absolute construct. Hoewel *NRC Handelsblad* dus een groter percentage titels hadden die gemiddeld iets negatiever scoorden, waren de emotionele titels van *De Telegraaf* wel iets emotioneler dan de emotionele titels van *NRC Handelsblad*.

Naast bovenstaande resultaten was er nog een andere factor die constant meespeelde. Bij de intensivering speelde de sectie mee bij de verdeling van neutraal en intensiverende titels, aangezien de driewegsinteractie ook de tweewegsinteractie van sectie op intensivering bevatte. Bij sentiment speelde de sectie niet mee bij de verdeling emotioneel en niet-emotioneel in de driewegsinteractie, maar had sectie wel een op zichzelf staande significante tweewegsinteractie met de emotionaliteit. Hiernaast had sentiment een driewegsinteractie bij de verdeling van positief en negatieve titels, waar sectie ook een rol bij speelde. Dit hoeft niet per se onverwacht te zijn. Bij economisch nieuws zou je toch sneller berichten over objecten, bedrijven en producten verwachten dan bij de andere secties. Dit is nieuws waarbij je minder snel intensiverend of emotioneel nieuws verwacht aan te treffen dan bij persoonlijk nieuws. Deze verwachting lijkt ook bevestigd te worden door de verdeling van citaten per sectie. Hieruit bleek dat de sectie financieel bij beide kranten een lager aandeel citaten in de titels had dan de in de andere secties. Tegelijkertijd was er ook een sterk verband te constateren tussen citaat en intensief taalgebruik en citaat en sentimenteel taalgebruik. Wanneer de gebruikte niet-neutrale taal binnen citaten bekeken werd en vergeleken per krant dan was hierin geen verschil te herkennen: de niet-neutrale taal die gebruikt wordt binnen citaten overstijgt dus het onderscheid tussen tabloid en kwaliteitskrant.

Hierbij komen we op het onderscheid dat in de theorie gemaakt wordt tussen tabloids en kwaliteitsjournalistiek. In dit stuk is sensationisme gedefinieerd als taal die bij de lezer emotie opwekt en de lezer aanspreekt op een bepaalde, opzweepende toon. Wanneer er door deze lens gekeken wordt naar de beoordeling van de titels is hier echter wel wat op af te dingen. De verschillen in intensiteit en het gemeten sentiment bij de titels waren zowel in het geval van intensiteit - zoals theoretisch verwacht werd bij de tabloid hoger - als bij het gemeten sentiment aanwezig. Maar de verschillen waren vrij klein en sloegen in het geval van het gemeten sentiment in sommige secties zelfs de kant op van de kwaliteitskrant. Kijkend naar het taalgebruik lijkt een hard onderscheid tussen populaire krant en kwaliteitskrant zoals betoogd door Scholten en Bakker (2006) niet op te gaan. Integendeel, het lijkt erop dat er bij *NRC Handelsblad* eveneens sensationeel taalgebruik te constateren is, volgens de onderzoeksdefinitie die in dit stuk gehanteerd is. In dit opzicht lijkt er dus sprake te zijn van de 'tabloidisering' die al aangegeven werd door Esser (1999) en Rowe (2000), waarbij de kwaliteitskrant en populaire krant beiden in min of meerdere mate sensationele tabloidstrategieën lijken toe te passen om op deze manier lezers te trekken.

We zouden ons hierbij kunnen afvragen of het bevorderlijk is om in de toekomst bij voorbaat artikelen te scheiden op de krant waarin het verschenen is. In dit stuk is het immers voorgekomen dat er bij een bepaalde sectie (Financieel nieuws) niet of nauwelijks verschil te constateren was terwijl dat in andere secties (Binnenland en Buitenland) juist weer wel het geval was. De invloed van een citaat oversteeg zelfs het onderscheid tussen de twee kranten. In dit opzicht doemt er dus een alternatieve manier om waarschijnlijke sensationisme in krantenkoppen vast te stellen: een die niet let op de krant waarin het verschijnt, maar die let op het onderwerp en of dat onderwerp vooral te maken heeft met mensen (subjecten) of dingen (objecten).

Zoals in sectie 2.5 al duidelijk werd gemaakt, zijn krantenkoppen een belangrijk onderdeel van een krantenartikel. Als de zogeheten *relevance optimizers* zouden de koppen berichten moeten zijn die bij de lezer de juiste context moeten oproepen. Volgens Dor (2003) zouden deze berichten zowel interesse moeten opwekken bij de lezer, als het verhaal op een gepaste manier framen. We kunnen ons afvragen in hoeverre het gebruik van niet-neutrale taal een verhaal nog gepast framet. Zoals uiteengezet in het theoretisch kader kan het gebruik van sensationisme een emotie bij de lezer opwekken, en kan emotie een lezer ertoe zetten om actie te ondernemen. In het geval van een krantenkop is die actie vaak het lezen van het artikel zelf, waarna de lezer kan bepalen of de inhoud van het artikel de gebruikte sensatie rechtvaardigt. Hoewel het stereotype van de populaire krant versus de kwaliteitskrant anders doet vermoeden, zijn het zowel *De Telegraaf* als *NRC Handelsblad* die zich schuldig maken aan deze strategieën. Geen van beide kranten betitelt alle artikelen sensationeel.

6.2 De validiteit van tools

In dit onderzoek is gebruikgemaakt van twee tools waarvoor een kwalitatieve toets is uitgevoerd: Pattern en T-Scan. Nu zijn wij hier als eindgebruiker afhankelijk van de definities van de ontwerpers en hoewel er in dit stuk door middel van de kwalitatieve toets verantwoord is waarom de computationele tools overeenkomen met de menselijke beoordelingen, kunnen de definities in het gereedschap een gevolg hebben. Zowel T-scan als Pattern maken gebruik van woordlijsten. Hierbij gebruikt Pattern een manier waar woorden in het lexicon op verschillende manieren gewogen zijn, en T-scan een definitie waarbij de intensiteit van een zin berekend wordt als verhouding van wel- en niet-intensiverende woorden. Zoals eerder opgemerkt heeft dat bij T-scan tot gevolg dat zinnen met weinig woorden als extreem intens beoordeeld kunnen worden. Om te zien in hoeverre de zinslengte meespeelt bij de beoordelingen is er een ANOVA uitgevoerd met de zinslengte als covariaat. Hieruit bleek dat zowel de zinslengte als het soort krant bijdragen aan het verschil in gemiddelde beoordeling tussen de groepen ($F(2) = 991.216, p < .001$). Wanneer vervolgens gekeken wordt naar de *Estimated Marginal Means*, waarbij de gemiddeldes aangepast worden aan andere variabelen, in dit geval de zinslengte, dan ontstaat er iets opvallends. De titels van *De Telegraaf* hebben een gemiddelde intensiteit van 251.699 en de titels van *NRC Handelsblad* hebben een gemiddelde intensiteit van maar liefst 269.074. Het omgekeerde dus dan de bevindingen die eerder vermeld waren. Dit betekent dat, wanneer titels van gelijke lengte met elkaar vergeleken worden, de titels van *NRC* gemiddeld iets intensiever zijn. Betekent dit dan dat de eerder genoemde resultaten niets waard zijn? Niet per se, aan gezien ook bekend is dat titels van *De Telegraaf* gemiddeld korter zijn dan die van *NRC Handelsblad*, we gaan ervan uit dat dit een bewuste keuze van de redactie is. Vergelijken van titels met dezelfde lengte correspondeert dan niet met de werkelijkheid. Het is belangrijk om in een onderzoek altijd eerst aan te tonen in hoeverre een tool kwalitatief overeenkomt met de menselijke beoordeling, zoals ik in mijn stuk gedaan heb. Hoewel er in dit stuk 3 tools zijn onderzocht (Pattern, SentiStrength en T-Scan), en er twee zijn gebruikt, zijn er natuurlijk ontelbaar veel meer te vinden. De suggesties voor nieuw onderzoek in de volgende sectie (6.3) zouden dus gelezen worden met de optie om de voorgestelde dimensies te meten met andere tools en technieken dan behandeld in dit onderzoek, bijvoorbeeld met de machine learning technieken benoemd in sectie 2.4.

6.3 Toekomstig onderzoek

6.3.1 Vergelijken van gelijke onderwerpen

In dit stuk is aangetoond dat het vergelijken van stukken van verschillende secties kan leiden tot verschillende resultaten, waarbij economisch nieuws bijvoorbeeld iets neutraler nieuws kan

bevatten. In toekomstig onderzoek zou hier dieper op in te gaan en alleen het nieuws rond één of meerdere onderwerpen op te zoeken en te analyseren. Zo zouden bijvoorbeeld artikelen rondom controversiële onderwerpen (Brexit, verkiezingen) tussen soorten kranten vergeleken kunnen worden, of vergeleken worden met minder controversiële onderwerpen. Hiernaast zou er ingegrepen kunnen worden op het inzicht in discussiepunt 5.4 door te kijken naar citaten en de mensen die geciteerd worden. Uit dit onderzoek bleek al dat citaten vaker emotioneel en intensiverend zijn, het zou interessant kunnen zijn om te zien of bepaalde groepen, mensen of organisaties die geciteerd worden ook intensiverender en emotioneler behandeld worden.

6.3.2 Vergelijking van online en offline vormgeving

De strategie om data te verzamelen, beschreven in sectie 4.2, was niet mijn eerste strategie. In eerste instantie heb ik een script geschreven dat op de websites van de publicaties keek en de publieke archieven doorzocht. Op deze manier was het mogelijk om artikelen te verzamelen zoals ze online verschenen zijn. Met behulp van deze informatie heb ik in de online database van Nexis Uni gezocht in de hoop dat het de artikelen zou opleveren zoals ze op papier zijn verschenen. In veel gevallen was er echter sprake van een mismatch, waarbij Nexis Uni niet de gewenste resultaten gaf. Dit kwam dan vaak omdat er online andere woorden gebruikt werden dan offline, bijvoorbeeld door redactionele aanpassingen achteraf. Toch zou deze aanpak een interessante dataset kunnen opleveren, namelijk een waarbij de verschijningsvorm van hetzelfde artikel in verschillende media (online/offline) kan worden vergeleken.

6.3.3 Analyse op basis van de geografie

Met behulp van een simpele tekstverwerkingstechniek (opzoeken in woordlijsten) zouden de vermelde stads-, provincie- en landsnamen per artikel vastgelegd kunnen worden. Deze locaties zouden geprojecteerd kunnen worden op een land- of wereldkaart om zo de geografische dekking per krant te kunnen zien, en of er bijvoorbeeld verschil is tussen tabloid- en kwaliteitsjournalistiek. Hiernaast is het zo ook makkelijk om te zien of en waar er gaten zijn in de dekking. Natuurlijk is het zo dat Nederlandse kranten bovengemiddeld nieuws verstrekken uit Nederland, maar is er een verschil te zien tussen randstedelijke gebieden en de daarbuiten gelegen provincies? En is er op internationaal gebied iets aan te merken zoals overrapportage van nieuws uit de Europese Unie en de Verenigde Staten ten opzichte van Aziatische of Afrikaanse gebieden? Deze gegevens zouden ook weer gecombineerd kunnen worden met intensiverende factoren en polariteit, rapporteert de ene krant met meer emotie vanuit een bepaalde regio dan een andere?

6.3.4 Het gebruik van de krantenpagina's

Het uiteindelijk gebruikte script heeft ook de mogelijkheid om een pdf-document (tekst plus foto's) van de krantenpagina waarin het artikel is gepubliceerd te downloaden. Een voorbeeld van zo'n bestand is te zien in Figuur 5. Wat gelijk opvalt is dat de gebruikte foto's op de pagina's korrelig gemaakt zijn, waarschijnlijk om te besparen op opslagruimte enerzijds en de privacy van de geportretteerden anderzijds. Toch zouden deze bestanden gebruikt kunnen worden in een vervolgonderzoek naar krantenartikelen. Met behulp van deze bestanden zou het bijvoorbeeld mogelijk zijn om de locatie van een artikel ten opzichte van andere artikels op dezelfde pagina uit te rekenen, wat een extra dimensie oplevert. Zo zou je kunnen onderzoeken of intensiverende headlines vaker linksboven de pagina verschijnen, waar we in het Nederlands ook beginnen met lezen, of juist vaker in het midden, waar het oog eerder op valt. Hoewel de foto's korrelig gemaakt zijn, kunnen eigenschappen kleur, grootte en locatie van de foto ten opzichte van het artikel ook gebruikt worden als extra dimensie. In combinatie met de door mij gebruikte factoren als emotionele polariteit en gebruik intensiveerders zou er dan gedacht kunnen worden als het gebruik van kleur in foto's ten opzichte van de emotionele polariteit. In het geval van de door mij onderzochte kranten waren er alleen pagina's beschikbaar van het *NRC Handelsblad*. Als men onderzoek wil doen naar het onderscheid tussen tabloids en kwaliteitsjournalistiek op dit gebied moeten er dus nog pagina's verzameld worden van een tabloid.

Christiaan Weijts

Klaas Dijkhoff aan d'ns overkant

Dat zeven Vlaanderen al als het endend woelt, hoort en doordert bij de Nederlandse... Klaas Dijkhoff is een man die niet van de PVV... Hij is een man die niet van de PVV...

Klaas Dijkhoff kan sympathie overkomen, genot, maar ook firaal uit de bocht vliegen

Wat is dat toch met Dijkhoff? Hij kan sympathie overkomen, intelligent, gevat, maar ook firaal uit de bocht vliegen... Klaas Dijkhoff is een man die niet van de PVV...

Thuisdras Oud-minister Gerrit Zalm wordt commissaris bij Danske Bank

AMSTERDAM. Oud-minister van Financiën Gerrit Zalm (60) is door het bestuur van Danske Bank...

MOORD OP VOLGTOE MOEDER OM verloopd overlegger voor vier insulneuroden

AMSTERDAM. Het Openbaar Ministerie (OM) vervolgt Bahadur A. (22) voor het ernstig delict van insuline bij zelfmoord...

CULTUREEL ERFGOED

Commissie buigt zich over de Erfgoedwet na veiling Rubens

Onder leiding van Alexander Pechold gaat een commissie bekijken of cultureel erfgoed wel voldoende beschermd is.

Door onze redacteur Arian Ribbens AMSTERDAM. Een onafhankelijke commissie onder leiding van oud-minister Alexander Pechold gaat op dinsdag van het ministerie van Onderwijs, Cultuur en Wetenschap onderzoeken of cultureel erfgoed in particuliere handen wel goed beschermd is.

De ophef over de recente veiling door prins Christina van Kamevorenvervoering uit de Koninklijke Verzamelingen is mede de aanleiding voor het onderzoek, bevestigd een woordvoerder van de minister.

De Commissie Pechold gaat onderzoeken of de Erfgoedwet, van kracht sinds juli 2016, wel voldoet. Ook gaat zij kijken of de aanwijzingscriteria voor de lijst met beschermd erfgoed in particulier best nog adequaat zijn.

Prinses Christina liet artikelen over oude tekeningen velen waaronder een Rubens

„In de discussie over identiteit en nationaliteit heeft brand het gevoel dat het belangrijk is om erfgoed te behouden voor toekomstige generaties.“

In januari liet prinses Christina twaalf oude tekeningen, waaronder een Rubens, en drie kerstkaarten met andere werken uit de kunstwereld van vijf Kamevoren. Een kerstkaart met tafeldecoratie werd teruggekocht, omdat de afbeelding, als vroege van NRC, de onomatopoeie had „Jerkend“

RANTSOENEN GEHALVEERD

Noord-Korea alarmeert VN over nijpende voedseltekorten

AMSTERDAM. Noord-Korea kampt met een voedseltekort. De rampen van vorig jaar zijn nog steeds aanwezig, aldus de VN-voedselorganisatie.

Makki Hartono redde de avond. Knap voor zijn overvaldijkt hij niet 'VTV' in een zaal met vooraf geselecteerde, in het dagelijks leven werkzaam in de sociale advocatuur, dus deze is beschikbaar was welkom. „Jij bent nu nog een sociale advocaat in mijn schiet, niet ik; je moet maar hetzelfde als Marianne van Boven en Martin Koning. Toen ze het had overgenomen, wil mijn vrouw ik vermoed je al te je niet doet.“

Tong bij de politiebureau hoort dat het nog steeds met de advocatuur zou gaan. Dan hoort Hartono mischien de mist' voor het theater en hadden wij, en 'n overkant, en een topambulant bij.

Christiaan Weijts schreef op deze plek iedere vrijdag een column.



Klimaatmeisje Brussel in verlegenheid

BRUSSEL. „Jullie gedragen je als een intellie van vervele vanden.“ Klimaatmeisje Greta Thunberg sprak donderdagavond in de volgspraak zaal in een gebouw van de Europese Commissie, waar zij een handout werd toegestuurd door voorzitter Ursula von der Leyen.

Thunberg sprak in een zaal van de Europese Commissie, waar zij een handout werd toegestuurd door voorzitter Ursula von der Leyen.

Thunberg sprak in een zaal van de Europese Commissie, waar zij een handout werd toegestuurd door voorzitter Ursula von der Leyen.

Thunberg sprak in een zaal van de Europese Commissie, waar zij een handout werd toegestuurd door voorzitter Ursula von der Leyen.

Thunberg sprak in een zaal van de Europese Commissie, waar zij een handout werd toegestuurd door voorzitter Ursula von der Leyen.

Verklaringen uit de onderwereld 'Bij coke horen harde problemen'

Kroongetuige Nabil B. is in tegenstelling tot zijn vertoolde broer geen cruchulighe, bijgt uit zijn verklaringen.

Door onze redacteur Jan Meewis AMSTERDAM. Kroongetuige Nabil B. heeft in acht verken getuigt in vijf verhoorverhoor van de rijk als het gesprek kon op zijn eigen tal in de drughand. „Jij softdrugs heb je zelf probleem“, vertelt Nabil geregeld rijk 's vlogover.

Zijn verhoorverhoor was verbaasd als Nabil vertelt dat hij pas een jaar of achtien, negentien was toen hij in de drughand stapte. Nabil B. is een tegengestelde tot zijn broer geen onschuldige paant.

Nabil heeft nauwelijks gekocht voor een leven in de onderwereld, zo bijgt uit zijn verklaringen. Hij is via hooggeplaatste contacten uit de co-canonoffits in Utrecht bij liquidaties betrokken geraakt. Werkloosheid, haat, handel en de handel in handdrugs. Nabil heeft het allemaal gedaan, vertelt hij. Voord de handel in cocaine is wetsgeword maar daar staan ook grote risico's vlogover.

van woning waar de schutters na hun dood konden schuilen.

Hij een tweede liquidatie leverde het ook de staat waarden de schutters na afloop konden vluchten. Bij die vlogover, op 12 januari 2017, werd de verhoorde man doodgeschoten. Het ging om een poging die Nabil goed kende uit zijn jeugd. Halim C. Mar zijn familie, ook actief in het criminele milieu, is bij de goe-berisping.

Nabil B. was betrokken bij twee gedingde liquidaties en één mislukte moord

De harde problemen die boven bij de handdrughand heeft Nabil wel een „beetje onderkant“, vertelt hij. Uiteindelijk komt Nabil in een straatje terecht die voor een normaal mens nauwelijks voor te stellen is. In 2016 en 2017 is hij betrokken geweest bij twee gedingde liquidaties en één mislukte moordaanval. Nabil staat dat het hier ging om een „vriende-diner“ waarvan hij niet is bereid.

De harde problemen die boven bij de handdrughand heeft Nabil wel een „beetje onderkant“, vertelt hij. Uiteindelijk komt Nabil in een straatje terecht die voor een normaal mens nauwelijks voor te stellen is. In 2016 en 2017 is hij betrokken geweest bij twee gedingde liquidaties en één mislukte moordaanval. Nabil staat dat het hier ging om een „vriende-diner“ waarvan hij niet is bereid.

De harde problemen die boven bij de handdrughand heeft Nabil wel een „beetje onderkant“, vertelt hij. Uiteindelijk komt Nabil in een straatje terecht die voor een normaal mens nauwelijks voor te stellen is. In 2016 en 2017 is hij betrokken geweest bij twee gedingde liquidaties en één mislukte moordaanval. Nabil staat dat het hier ging om een „vriende-diner“ waarvan hij niet is bereid.

Uit alles blijkt dat alle mensen die betrokken zijn bij de eerste liquidatie dat van uitdruk niet mogen weten. Als twee hand- schatters „af begint te schieten voorde Nabil in zijn rijt als getuige is betrokken, contact en even punitik. „Ik zit met doe je nou de mensen doen actie waar ik bij heb“, doeltling. „Krijgt Nabil dan te horen dat hij betrokken is bij de eerste liquidatie?“, vraagt de journalist. Nabil antwoordt dat hij niet weet. Hij zegt dat hij niet weet. Hij zegt dat hij niet weet.

Uit alles blijkt dat alle mensen die betrokken zijn bij de eerste liquidatie dat van uitdruk niet mogen weten. Als twee hand- schatters „af begint te schieten voorde Nabil in zijn rijt als getuige is betrokken, contact en even punitik. „Ik zit met doe je nou de mensen doen actie waar ik bij heb“, doeltling. „Krijgt Nabil dan te horen dat hij betrokken is bij de eerste liquidatie?“, vraagt de journalist. Nabil antwoordt dat hij niet weet. Hij zegt dat hij niet weet. Hij zegt dat hij niet weet.

7 Eindconclusie

Talige constructen computationeel onderzoeken vereist een interdisciplinaire aanpak. In dit stuk is een definitie van sensationisme in koppen gebaseerd op intensiverend en emotioneel taalgebruik geformuleerd. Er is een corpus gegenereerd met behulp van een Python-script. Digitale tools om intensiteit en sentiment vast te stellen in tekst zijn kwalitatief getoetst, waarbij er in het geval van de sentimentsanalyse een marge is aangebracht om de resultaten overeen te laten komen met de handmatige beoordeling. Deze tools hebben vervolgens een subset van de artikelen geanalyseerd. Met de in dit stuk gemunte definitie van sensationeel taalgebruik komt er toch een meer genuanceerd beeld naar boven dan een tweedeling tussen populaire krant en kwaliteitskrant waarbij de eerste vooral sensationele taal gebruikt en de laatste niet. Het gebruik van sensationeel taalgebruik lijkt eerder onderhevig aan andere factoren die in sommige gevallen het onderscheid van de soort krant overstijgen. Het genre van het stuk bijvoorbeeld, en of er expliciet een actor aanwezig is of niet.

Toekomstig onderzoek zou er goed aan doen juist deze verschillen in het achterhoofd te houden door stukken te vergelijken die bij voorbaat al in dezelfde genres vallen, of over dezelfde onderwerpen gaan. Op deze manier wordt de invloed van de verscheidenheid aan onderwerpen en genres die verschillende kranten behandelen uit de weg gegaan. Hiernaast is er in dit onderzoek gebruikgemaakt van twee tools die intensiteit en sentiment in titels kunnen meten, in toekomstig onderzoek zou er ook voor gekozen kunnen worden om hetzelfde onderzoek uit te voeren met andere tools, om te kijken of dit potentieel een ander resultaat tot gevolg kan hebben.

Na het lezen van dit onderzoek zal het duidelijk zijn, sensatie is geen binair verschijnsel waarvan de aan- of afwezigheid makkelijk aan te wijzen is. Het is deze eigenschap die het moeilijk kwantificeerbaar maakt en die het op grote schaal computationeel onderzoeken bemoeilijkt. Vooralsnog is het nog altijd de menselijke intuïtie die met kop en schouders boven de beoordeling van computers uitsteekt als het op het herkennen van sensatie aankomt. Juist daarom is het belangrijk om de bevooroordeelde mening die misschien kan bestaan bij bepaalde kranten te laten varen. Want we zien: sensatie is niet alleen van alle tijden, maar ook van alle kranten.

Bibliografie

- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish journal of emergency medicine*, 18(3), 91–93.
- Bartlett, J. (2017). *The t-test and robustness to non-normality – The Stats Geek*. Verkregen 1 maart 2021, van <https://thestatsgeek.com/2013/09/28/the-t-test-and-robustness-to-non-normality/>
- Bon, E. V. (2016). *Sensationalism and Humanitarian Content in Online Dutch Newspaper News and User Comments: A study of the text of refugee (crisis) related news articles and the comments posted to them on Facebook*. (masterscriptie). Universiteit van Amsterdam.
- Burgers, C. & De Graaf, A. (2013). Language intensity as a sensationalistic news feature: The influence of style on sensationalism perceptions and effects. *Communications - the European Journal of Communication Research*, 38(2), 167–188.
- Chen, Y., Conroy, N. K. & Rubin, V. L. (2015). News in an online world: The need for an “automatic crap detector”. *Proceedings of the Association for Information Science and Technology*, 52(1), 1–4.
- Costera Meijer, I. (2012). Valuable journalism: The search for quality from the vantage point of the user. *Journalism*, 14(6), 754–770.
- De Smedt, T. & Daelemans, W. (2012a). "Vreselijk mooi!"(terribly beautiful): A Subjectivity Lexicon for Dutch Adjectives. *LREC*, 3568–3572.
- De Smedt, T. & Daelemans, W. (2012b). Pattern for Python. *Journal of Machine Learning Research*, 13(66), 2063–2067.
- Delobelle, P., Winters, T. & Berendt, B. (2020). RobBERT: a dutch RoBERTa-based language model. *arXiv preprint arXiv:2001.06286*.
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- de Vries, W., van Cranenburgh, A., Bisazza, A., Caselli, T., van Noord, G. & Nissim, M. (2019). Bertje: A dutch bert model. *arXiv preprint arXiv:1912.09582*.
- Dor, D. (2003). On newspaper headlines as relevance optimizers. *Journal of pragmatics*, 35(5), 695–721.
- Esser, F. (1999). ‘Tabloidization’ of News: A comparative analysis of Anglo-American and German press journalism. *European Journal of Communication*, 14(3), 291–324.
- Fang, I. (1997). *A history of mass communication: Six information revolutions*. Taylor & Francis.
- George, D. & Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference*. Routledge.

- Grabe, M. E., Zhou, S. & Barnett, B. (2001). Explicating Sensationalism in Television News: Content and the Bells and Whistles of Form. *Journal of Broadcasting & Electronic Media*, 45(4), 635–655.
- Hamilton, M. A., Hunter, J. E. & Burgoon, M. (1990). An Empirical Test of an Axiomatic Model of the Relationship Between Language Intensity and Persuasion. *Journal of Language and Social Psychology*, 9(4), 235–255.
- Hamilton, M. A. & Stewart, B. L. (1993). Extending an information processing model of language intensity effects. *Communication Quarterly*, 41(2), 231–246.
- Hauttekeete, L. (2004). *De tabloidisering van kranten: mythe of feit?: de ontwikkeling van een meetinstrument en een onderzoek naar de tabloidisering van Vlaamse kranten* (proefschrift). Ghent University.
- Jijkoun, V. & Hofmann, K. (2008). Task-based evaluation report: Building a Dutch subjectivity lexicon. *ILPS-ISLA*, 2.
- Kamoen, N., Holleman, B. C. & Van den Bergh, H. (2007). Hoe gemakkelijk is een niet moeilijke tekst?: Een meta-analyse naar het effect van vraagformulering bij het meten van tekstwaardering. *Tijdschrift voor taalbeheersing*, 29(4), 314–330.
- Kilgo, D. K., Harlow, S., García-Perdomo, V. & Salaverría, R. (2018). A new sensation? An international exploration of sensationalism and social media recommendations in online news publications. *Journalism*, 19(11), 1497–1516.
- Kuiken, J., Schuth, A., Spitters, M. & Marx, M. (2017). Effective Headlines of Newspaper Articles in a Digital Environment. *Digital Journalism*, 5(10), 1300–1314.
- Liebrecht, C. (2015). *Intens krachtig. Stilistische intensiveerders in evaluatieve teksten* (proefschrift). Radboud Universiteit Nijmegen.
- Liebrecht, C., Hustinx, L. & van Mulken, M. (2019). The Relative Power of Negativity: The Influence of Language Intensity on Perceived Strength. *Journal of Language and Social Psychology*, 38(2), 170–193.
- Nexis Uni - Utrecht University Library - Universiteit Utrecht. (2019). Verkregen 7 januari 2020, van <https://www.uu.nl/en/university-library/help-in-searching/search-engines-explained/nexis-uni>
- Nir, R. (1993). A discourse analysis of news headlines. *Hebrew Linguistics*, 37, 23–31.
- Orosa, B. G., Santorun, S. G. & García, X. L. (2017). Use of clickbait in the online news media of the 28 EU member countries. *Revista Latina de Comunicación Social*, 72, 1261–1277.
- Pander Maat, H. (2004). Wervend taalgebruik in persberichten – werkt het? Hoe journalisten omgaan met persberichten in de luchtvaartsector. *Tijdschrift voor taalbeheersing*, 26(3), 207–223.

- Pander Maat, H., Kraf, R. & Dekker, N. (2020). *Handleiding T-Scan*. (Handleiding). UiL-OTS. Universiteit Utrecht.
- Pander Maat, H., Kraf, R., van den Bosch, A., Dekker, N., Gompel, M. v., Kleijn, S. d., Sanders, T. & Sloot, K. (2014). T-Scan: a new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal*, 4, 53–74.
- Peeters, G. (1971). The positive-negative asymmetry: On cognitive consistency and positivity bias. *European Journal of Social Psychology*, 1(4), 455–474.
- Renkema, J. (1997). Geintensiveerd taalgebruik: een analyseschema. In H. van den Bergh, D. Janssen, N. Bertens & M. Damen (Red.), *Taalgebruik Ontrafeld* (pp. 495–504). Foris.
- Rowe, D. (2000). On going tabloid: A preliminary analysis. *Metro Magazine: Media & Education Magazine*, (121/122), 78.
- Rozin, P. & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review*, 5(4), 296–320.
- Schaap, G. & Pleijter, A. (2012). Het sensatiegehalte van voorpaginafoto's: Een inhoudsanalyse van populaire en kwaliteitskranten in Nederland. *Tijdschrift voor Communicatiewetenschap*, 40(1), 71–86.
- Scholten, O. & Bakker, P. (2006). *Communicatiekaart van Nederland: overzicht van media en communicatie*. Alphen aan den Rijn: Kluwer.
- Semin, G. R. & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of personality and Social Psychology*, 54(4), 558–568.
- Semin, G. R. & Fiedler, K. (1992). The inferential properties of interpersonal verbs. In G. R. Semin & K. Fiedler (Red.), *Language, interaction and social cognition*. Sage Publications, Inc.
- Sparks, C. (2000). Introduction: The panic over tabloid news. In C. Sparks & J. Tulloch (Red.). Rowman & Littlefield Lanham, MD.
- Sperber, D. & Wilson, D. (1986). *Relevance: Communication and cognition*. Harvard University Press Cambridge, MA.
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D. & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American society for information science and technology*, 61(12), 2544–2558.
- Tom, D., Pollmann, M. & Goudbeek, M. (2018). De emotionele respons van lezers op affectieve headlines op sociale media. *Tijdschrift voor Taalbeheersing*, 40(1), 137–155.
- Trilling, D. & Boumans, J. (2018). Automatische inhoudsanalyse van Nederlandstalige data. *Tijdschrift voor Communicatiewetenschap*, 46(1), 5–24.

A Normaliteit

		Skewness	Kurtosis
Intensiteit	Telegraaf	4.691	29.185
	NRC	3.596	14.787
	Binnenland	4.498	26.757
	Buitenland	4.945	33.262
	Financieel	5.074	38.582
Intensiteit zonder 0	Telegraaf	2.640	7.694
	NRC	1.772	4.229
	Binnenland	2.613	7.513
	Buitenland	2.803	8.996
	Financieel	3.245	15.910
Sentiment	Telegraaf	0.063	-1.601
	NRC	0.130	-1.592
	Binnenland	0.132	-1.601
	Buitenland	0.126	-1.611
	Financieel	-0.007	-1.587

Tabel 1: Skewness en Kurtosis - voor Skewness en Kurtosis wordt een marge van $-2/ + 2$ aangehouden (George & Mallery, 2019)

B Top 10's en valentie

Titel	valentie
Grenzeloos gezellig	1.00
Steengoede bevalling	1.00
'Zorg niet zo briljant'	1.00
Alles voor perfect plaatje	1.00
Prachtig uithangbord voor Nederland	1.00
Prins: 'Niemand is perfect...'	1.00
Zonder twijfel de allerbeste	1.00
'Het is vooral heel gezellig'	1.00
Stikstofleed als de perfecte vuilnisdag	1.00
'Heel leuk dat ze hier komen'	.96

Tabel 2: Top 10 Binnenland Telegraaf - Positief

Titel	valentie
'Ideale schoonzoon' wint in Wenen	.90
'Poetin is welkom op G7 in 2020'	.80
Cultuur houdt relatie met Rusland nog beetje warm	.70
'Muur moet zwart, puntig en mooi'	.70
'Was enorme klap geweest als Trump dit had laten lopen'	.70
Corruptie Z-Afrika enorm	.70
Enorme wapenvondst bij Italiaanse rechtsextremisten	.70
Ferme taal over Brexit	.70
Hoop voor legendarische spion	.70
Nomaden blijven leefstijl trouw	.70

Tabel 3: Top 10 Buitenland Telegraaf - Positief

Titel	valentie
'Opkoopprogramma ECB is briljant'	1.00
Aderlating Beyond Meat na prachtige rit omhoog	1.00
Goudlokje en de perfecte storm	1.00
Schilderij Basquiat blijkt allerbeste investering	1.00
Beleggers laten Beyond Meat uitstekend smaken	.90
'72 uur werken, prima!'	.88
'AkzoNobel heeft het juist erg goed gedaan'	.85
'Doorbraak slimme apparaten'	.80
'Nederland aantrekkelijk'	.80
'Nederland zien ze als betrouwbaar'	.80

Tabel 4: Top 10 Financieel Telegraaf - Positief

Titel	valentie
Ooit vrolijke' SP zoekt de weg omhoog	.90
Soldaat wil nu echt net zo veel loon als schoonmaker	.88
Zonnebaars: geliefd voor in de vijver, nu verboden in EU	.80
Dit examen was makkelijk te lezen'	.75
Tot in Apeldoornen Utrecht weten ze: in de Wielewaalkunje makkelijk kraken	.75
Mooie zomer? De lol is erwel een beetje vanaf	.70
De dokter kreeg altijd 't beste stuk vlees'	.68
Wachtlijsten voor speciaal onderwijs'	.60
'Zo top is ons onderwijs niet meer'	.60
Afkeer van de regio? Amsterdam vindt vooral zichzelf heel bijzonder	.60

Tabel 5: Top 10 Binnenland NRC - Positief

Titel	valentie
Dit keer stemt de herdenking van '1989' minder vrolijk	.90
Slim stemmen' probleem voor Poetin	.80
De Rohingya waren welkom in Cox's Bazar, maar nu niet meer	.80
De schorskever sloopt de geliefde Duitse wouden	.80
Het Westen koos voor de makkelijke weg in Congo'	.75
Duitse militairen zweren trouw voor de Rijksdag	.70
Enorme vangst van neushoornhoorns	.70
Hij klonk als zijn vader, retorisch begaafd	.70
Voor de vrouw is er enorme vooruitgang in Saoedi-Arabië	.70
Een dienstmeisje kun je best opsluiten in Libanon	.68

Tabel 6: Top 10 Buitenland NRC - Positief

Titel	valentie
Ik wil als super sportvrouw herinnerd worden'	.90
Groots regeren zonder plan	.90
Wonderbaarlijke redder van Chrysler	.90
Mijn carrière is nu al geslaagd'	.80
5 miljoen subsidie voor slimme laadpalen	.80
Een waardevol luxemerk bepaalt zelf zijn prijs	.80
Geld sparen in een pensioenfonds, hoe slim is dat nog?	.80
Geslaagd voor examen van Fed	.80
Het is de minst geliefde vacature in Den Haag	.80
Het meest waardevol is de mening van vrienden	.80

Tabel 7: Top 10 Financieel NRC - Positief

Titel	valentie
Lugubere eretitel	-.80
Lugubere vondsten	-.80
Politie denkt aan tragisch ongeval	-.80
Iraanse regiem kolossaal duivels	-.85
'Zijn we helemaal besodemieterd?!'	-.88
Barbaars vermaak verholen	-.90
Weerzinwekkend	-.90
'Afschuwelijke daden, geen strafbare dader'	-1.00
'We voelen ons erg belazerd'	-1.00
Gruwelijke kopschopvideo's duiken vaker op	-1.00

Tabel 8: Top 10 Binnenland Telegraaf - Negatief

Titel	valentie
Republikeinen boos op Trump	-.70
Robert Mueller is boos	-.70
Spanjaarden boos over misbruikzaak	-.70
Trump boos om bewijs	-.70
Queen krijgt het nog lastig	-.75
Pelgrims massaal naar Spanje	-.88
Hypocriet	-.90
Tories 'gemakzuchtig'	-.90
Vulgair	-.90
België rouwt na gruwelijke moord	-1.00

Tabel 9: Top 10 Buitenland Telegraaf - Negatief

Titel	valentie
Lastige woningen	-.75
Energiebedrijf vaak financieel erg zwak	-.77
India dregt vergeefs naar koffiemiljardair	-.78
Project Omaha was 'echt verschrikkelijk'	-.80
Pijnlijk!	-.81
We gaan massaal naar het wereldrestaurant	-.88
Boskalis: 'windenergie afgrijselijke markt'	-.90
Oncomfortabel toilet moet pauzes verkorten	-.90
Slaapverwekkend Philips	-.90
UBO Abnormaal	-.90

Tabel 10: Top 10 Financieel Telegraaf - Negatief

Titel	valentie
Curaçao pakt drugshandel slecht aan'	-.70
Burgemeester boos op NCTV omwindmolenactivisme	-.70
De broers, het geld en hun demente moeder	-.70
Dronken, baldadig en boos - met aansteker	-.70
Een demente oom die z'n testament wijzigde	-.70
Huurstichting was 'beetje dom' in Molukse buurt	-.70
Politici wilden luisteren, maar zagen alleen nog boze burgers	-.70
SGP'ers zijn boos over misstanden	-.70
Verhalen drugsschade valt slecht bij Raad voor de Rechtspraak	-.70
Emoties in de Tweede Kamer? Staatsrechtelijk lastig	-.75

Tabel 11: Top 10 Binnenland NRC - Negatief

Titel	valentie
Te ziek om naar de galg te vliegen	-.68
China boos om 'inmenging'	-.70
De cholera, de dodenen nog meer slecht nieuws uit Jemen	-.70
Duitse coalitie heeft alleen slechte opties	-.70
Een slechte zomer voor nucleaire akkoorden	-.70
Gevangenis Epstein 'mensonterend'	-.70
Schimmig machtsspel in elite Syrië	-.70
Schimmig machtsspel in Kazachstan	-.70
Trump duwt Democraten indefensief met racistische uitval	-.70
Gruwelijk, maargenoeg bewijs voor genocide?	-1.00

Tabel 12: Top 10 Buitenland NRC - Negatief

Titel	valentie
Meer zorg thuis klinkt goed, is lastig	-0.75
Lastige automarkt hindert leverancier	-0.75
Sancties VS lastig te ontlopen voor Nederlandse bedrijven	-0.75
Italië maakt het de eurozone opnieuw lastig	-0.75
1,5 miljoen huizen van het gas af, dat is best lastig	-0.77
Paddepoel verwarmen wordt best lastig	-0.77
De perverse prikkels van de Amerikaanse autoleningen	-0.80
Een meedogenloze spiegel	-0.80
Hoera, slecht economisch nieuws!	-0.88
Aan spreekkoren heeft zij een godsgrouwelijke hekel	-1,00

Tabel 13: Top 10 Financieel NRC - Negatief

Titel	valentie
'Allemaal onzin'	1000,00
Victorie kraaien	1000,00
'Onaanvaardbaar paardenmiddel'	1000,00
Gretig griezelen	1000,00
Afstraffing	1000,00
Puinhoop	1000,00
Bemoezucht	1000,00
Geklungel	1000,00
Wantoestanden	1000,00
Buiten-beentjes	1000,00

Tabel 14: Top 10 Binnenland Telegraaf - Intensief

Titel	valentie
Echte helden	1000,00
Flapuit	1000,00
Nachtmerrie	1000,00
Uittocht	1000,00
Tragedie	1000,00
Voorbeeldig	1000,00
Verkwisting	1000,00
Kopzorgen	1000,00
Noodklok	1000,00
Laster	1000,00

Tabel 15: Top 10 Buitenland Telegraaf - Intensief

Titel	valentie
Steeds krachtiger	1000,00
Chantage	1000,00
Lofzang	1000,00
Noodkreten	1000,00
Lievelingen	1000,00
Jammerklacht	1000,00
Jarenlang extra speelplezier	666,67
'Veel meer overstappen'	666,67
'Woningen onbetaalbaar'	500,00
Pijnlijk duidelijk	500,00

Tabel 16: Top 10 Financieel Telegraaf - Intensief

Titel	valentie
Studie Studenten: steeds meer studieschuld	500,00
Levensindekliniek Recordhoogte euthanasieverzoeken	500,00
CBS Steeds meer jongeren met overgewicht	400,00
Bruggen Extra controlein Zaandam	333,33
Amsterdam Kraakpanden sneller ontruimd	333,33
Appingedam Weer aardbeving met kracht van 1.9	333,33
Een militair met maar liefst 25 kentekenplaten; De zitting	285,71
Xtc 'gaat te hard', 'keta' kan verlammen	285,71
Radicale aanjagers in islamonderwijs'	250,00
Groningers zijn moedeloos, machteloos; Aardbeving Groningers zijn moedeloos, machteloos ; Aardbeving	250,00

Tabel 17: Top 10 Binnenland NRC - Intensief

Titel	valentie
Ze islinks, direct en wars van gekissebis	428,57
Gruwelijk, maargenoeg bewijs voor genocide?	400,00
Een heel landgaat te gronde aan de grillen van één man	363,64
Alsnog levenslang Karadzic	333,33
Geterroriseerd door wapengeweld	333,33
Levenslang voordrugsbaas Guzmán	333,33
Gevangenis Epstein 'mensonterend'	333,33
Harderoptredentegen haat enrechts-extremisme	333,33
Kritiek opPoolse tuchtrechtspraak	333,33
Vuurproef Zelensky bijPoetin	333,33

Tabel 18: Top 10 Buitenland NRC - Intensief

Titel	valentie
'Hoe meer risico, des te groterde euforie'	571,43
Nieuwe KPN-baas loyaal tot op het bot	571,43
Hypotheekrentes dalen heel snel	500,00
Veel meer meldingen van medicijntekorten	500,00
De Russen wacht nog meer onheil	500,00
Albert Heijn neemt nog meer Marqt-winkels over	428,57
Snel: pure spaarder eerlijker belastingen	400,00
Chinezen zijn écht heel anders	400,00
Brillengigant wil nog meer winkels	400,00
Het geld is er, nu nog meer curlers	375,00

Tabel 19: Top 10 Financieel NRC - Intensief