

Onderzoek naar de leesbaarheid van online nieuwsartikelen van de vijf populairste Nederlandse nieuwssites

*Research into the readability of online news articles from the five most popular Dutch news
sites*



Radboud Universiteit

Bachelorscriptie

LET-CIWB351

Lara Pap

Eerste beoordelaar: Dr. M. Zwets

Tweede beoordelaar: Dr. W.G. Reijnierse

Thema 12

Laaggeletterdheid en journalistiek

Woorden: 7192

Definitieve versie

8 juni 2020

Samenvatting

Een belangrijk onderdeel van maatschappelijke participatie is het lezen van nieuwsmedia (Wasike, 2018). Nieuwsmedia informeren, dragen bij aan de publieke opinie en voorzien in informatie bij politieke en maatschappelijke ontwikkelingen (Van Keulen, Korthagen, Diederer & Van Boheemen, 2018). De Amerikaanse wetenschapper Wasike (2018) stelde dat het taalniveau van Texaanse lokale kranten voor het merendeel van de samenleving niet begrijpelijk is, daarnaast vond hij een verschil tussen de leesbaarheid van nieuwsrubrieken. Afkortingen worden vaak gebruikt om ruimte te besparen, maar ze zijn vaak moeilijk(er) te begrijpen voor lezers die niet bekend zijn met een onderwerp (Terada, Tokunaga & Tanaka, 2004), en zijn daardoor van invloed op de complexiteit. Hierdoor raken mensen in de war of ontmoedig om verder te lezen (Nolan, 1991).

Er is relatief weinig bekend over het taalniveau van Nederlandse nieuwsmedia. Om te onderzoeken wat het taalniveau in Nederlandse nieuwsartikelen is, is in dit onderzoek een corpusstudie gedaan naar de leesbaarheid van de vijf populairste nieuwssites van Nederland (1a), namelijk RTL Nieuws, NOS, AD, NU.nl en De Telegraaf. Ook is onderzocht of er tussen deze nieuwssites een verschil in leesbaarheid is (1b) en of er een verschil in leesbaarheid per nieuwsrubriek is (1c), namelijk show, binnenland, sport, buitenland en economie. Daarnaast is onderzocht wat het gebruik van afkortingen op de vijf populairste nieuwssites van Nederland is (2a) en of het gebruik van afkortingen per nieuwssite (2b) en per nieuwsrubriek (2c) verschilt.

De gemiddelde leesbaarheidsscore van Nederlandstalige online nieuwsartikelen blijkt niet overeen te komen met taalniveau B1. Artikelen van RTL Nieuws en AD bleken gemiddeld genomen het gemakkelijkste leesbaar, ten opzichte van de artikelen van NOS, NU.nl en De Telegraaf. Daarnaast blijkt de nieuwsrubriek shownieuws het makkelijkst leesbaar, economische en buitenlandse nieuwsartikelen zijn het moeilijkst leesbaar. In het algemeen gebruik van afkortingen, per nieuwssite en per nieuwsrubriek blijkt geen verschil te bestaan.

1. Inleiding

1.1 Aanleiding

Onderzoek naar leesbaarheid houdt wetenschappers al decennialang bezig. Niet ieder persoon is in staat teksten van hetzelfde niveau te lezen. Zo hebben geletterden een ander taalniveau dan laaggeletterden en ongeletterden en is er binnen de groep geletterden ook een verschil in taalniveaus. Kranten spelen een unieke en integrale rol binnen de samenleving (Wasike, 2018). Mensen lezen nieuwsartikelen omdat zij geïnformeerd willen worden over wereldwijde gebeurtenissen en omdat nieuws hen voorziet in extra informatie bij politieke en maatschappelijke ontwikkeling (Van Keulen, Korthagen, Diederer & Van Boheemen, 2018). Hiervoor is het van belang dat een groot deel van de samenleving in staat is teksten te begrijpen. Nieuwsmedia zouden daarom een taalniveau van teksten in nieuwsartikelen moeten aanhouden dat aansluit bij het taalniveau van de gemiddelde Nederlander.

De meest relevante onderzoeken over het taalniveau van nieuws dateren uit de jaren 40 (Wasike, 2018). Daarentegen stammen de meest recente onderzoeken uit de jaren 70 en 80. Omdat het hier om relatief oude onderzoeken gaat heeft Wasike in 2018 onderzoek gedaan naar het taalniveau van artikelen in Texaanse lokale kranten. Echter is er nog geen onderzoek gedaan naar het taalniveau van nieuwsartikelen in Nederland. Dit is opvallend, want wanneer journalisten hun publiek succesvol willen bereiken, moeten zij nieuws voor een zo groot mogelijke groep leesbaar maken stelde Stempel (in Wasike, 2018). Er is bekend dat teksten op taalniveau B1 voor de meerderheid van de samenleving begrijpelijk zijn (Rijksoverheid, z.d.), namelijk 80% (Jansen, 2013). Maar hoe verhoudt het taalniveau van de gemiddelde Nederlander zich tot het taalniveau van nieuwsartikelen in Nederland? Hierover zijn tot nu toe geen wetenschappelijke inzichten beschikbaar en daarom focust dit onderzoek zich op het taalniveau van online nieuwsmedia in Nederland.

1.2 Theoretisch kader

Nieuwsconsumptie

Een belangrijk onderdeel van maatschappelijke participatie is het lezen van nieuwsmedia. Uit onderzoek van Wasike (2018) bleek dat kranten een unieke en integrale rol spelen binnen de samenleving. Nieuws gaat over nieuwe informatie over recente gebeurtenissen, waarbij deze nieuwe gebeurtenissen het onderwerp van een verslag of gesprek zijn (Oxford Shorter Dictionary, 2014). Nieuwsmedia informeren, onderwijzen, vergroten het kritisch denken van

lezers en brengen activiteiten van de overheid op een constructieve manier dichterbij mensen (Babalola, 2002). Echter zijn zowel het onderzoek van Wasike als het onderzoek van Babalola Amerikaanse onderzoeken. Ook Van Keulen, Korthagen, Diederik en Van Boheemen (2018) stellen dat journalistiek verschillende functies heeft. Dit onderzoek is een Nederlands onderzoek. Allereerst gaat het over het informeren van lezers over wereldwijde gebeurtenissen. Daarnaast draagt het lezen van nieuws bij aan de ontwikkeling van de publieke opinie en voorziet het men in extra informatie bij politieke en maatschappelijke ontwikkelingen (Van Keulen et al., 2018).

De inhoud waarover wordt geschreven in nieuwsartikelen is veelal erg complex en heeft invloed op de leesbaarheid van nieuws (Lostutter, 1947). Deze complexiteit neemt steeds verder toe en nieuwsartikelen worden daardoor steeds moeilijker leesbaar (Danielson and Lasorsa: in Dalecki, Lasorsa & Lewis, 2009). Volgens Dalecki et al. (2009) is deze complexiteit toe te schrijven aan verschillende factoren, zoals hoogopgeleide journalisten, contextualisering van nieuwsgebeurtenissen en de complexe wereld waarin verslaggevers opereren.

Allereerst is de contextualisering van nieuws in opkomst. Dit verwijst naar een schrijfstijl die interpretatiever is dan enkel beschrijvende feiten (Wasike, 2018). Hierbij worden gebeurtenissen in beredenerende hoeken en analytischer neergezet met het doel de gebeurtenis uit te leggen, te verhelpen, te onderzoeken of om mensen te inspireren (Fink & Schudson, 2014). Nieuws moet uitgebreider worden beschreven en niet alleen feitelijk. Dit maakt dat teksten langer zijn, waardoor ze moeilijker te lezen zijn (Fink & Schudson, 2014). Ten tweede leven journalisten in het huidige tijdperk in een complexere wereld dan vroeger. Een groot deel van de journalisten, 69%, vindt dat rapportagetechnieken van nieuws deze complexiteit moet weerspiegelen (Willnat & Weaver, 2014). Ten derde wordt de complexiteit van nieuwsmedia versterkt doordat journalisten moeten concurreren met nieuwe mediatypen, zoals bloggers en burgerjournalisten, die zelf verslagen maken van nieuwsgebeurtenissen en functies van de reguliere journalistiek op zich nemen (Wasike, 2018).

Een andere ontwikkeling is dat de nieuwsconsumptie van online nieuwsmedia toeneemt. Waar men vroeger traditionele nieuwsmedia consumeerde, gebruikt de meerderheid van de Nederlanders nu online kanalen voor nieuwsmedia (Commissariaat voor de Media, 2019). Dit hangt samen met digitalisering, 90% van de Nederlanders van twaalf jaar of ouder beschikten in 2019 over een smartphone met internet. Daarnaast heeft 45,9% van de 75-plussers een smartphone, terwijl dit in 2012 nog maar 2% was. Nieuws speelt een belangrijke rol in het leven van de gemiddelde Nederlander. 41% van de jongeren consumeert nieuws voornamelijk online, waarbij zij gebruik maken van kwalitatieve nieuwsmedia (Commissariaat voor de Media,

2019). De meest gebruikte Nederlandse online nieuwsmedia zijn NU.nl, NOS, AD, De Telegraaf en RTL Nieuws (Commissariaat voor de Media, 2019).

De complexiteit van nieuwsmedia heeft effect op het taalniveau van journalisten in online nieuwsmedia, waardoor teksten moeilijk(er) leesbaar zijn (Christoffels, Baay, Bijlsma & Levels, 2016). Teksten moeten wel toegankelijk zijn voor de lezers en daarom worden geschreven op het juiste taalniveau (Heij & Visser, 2006). Er is niet bekend op welk taalniveau Nederlandse nieuwsmedia hun nieuwsartikelen schrijven, maar uit onderzoek van Heij en Visser (2006) bleek dat publiekscommunicatie vooral plaats vindt op taalniveau C1, waarbij een klein deel plaats vindt op taalniveau B2. Tabel 1 geeft de verdeling van taalniveaus in publiekscommunicatie weer.

Tabel 1. Verdeling van het gebruik van taalniveaus in publiekscommunicatie in percentages.

Taalniveau	A1	A2	B1	B2	C1	C2
Verdeling in percentages	0%	0%	5%	15%	75%	5%

Taalniveau

Terwijl nieuwsartikelen steeds complexer lijken te worden, is het wel van belang dat deze artikelen begrijpelijk kunnen worden gelezen. Hiervoor is het van belang dat het niveau van een nieuwsartikel overeen komt met het algemeen taalniveau van lezers (Heij & Visser, 2006). Bij taalniveau wordt vaak verwezen naar het Common European Framework of Reference of Languages (CEFR) dat onderscheid maakt tussen verschillende niveaus van taalvaardigheid (Kraf, Lentz & Pander Maat, 2011). De CEFR-schaal onderscheidt A1, A2, B1, B2, C1 en C2 niveau. Hierbij verwijst A naar beginnend, B naar onafhankelijk en C naar vaardig niveau. De CEFR-schaal is een Europees referentiekader, daarnaast bestaat er ook een Nederlands referentiekader.

Het Nederlandse referentiekader wordt opgedeeld in taalniveau 1F, 2F, 3F en 4F (Bonset, 2009; Evers-Vermeul & Holtermann, 2013). Deze taalniveaus worden gekoppeld aan Nederlandse opleidingsniveaus. Taalniveau 1F staat gelijk aan eind primair onderwijs; taalniveau 2F staat gelijk aan einde vmbo; taalniveau 3F staat gelijk aan einde mbo-4 en havo; taalniveau 4F staat gelijk aan einde vwo (Evers-Vermeul & Holtermann, 2013).

Daarnaast kunnen de CERF-schaal en het Nederlandse referentiekader met elkaar vergeleken worden. Hierbij komen Europees taalniveau A1 en A2 overeen met Nederlands taalniveau 1F; Europees taalniveau B1 met Nederlands taalniveau 2F; Europees taalniveau B2

met Nederlands taalniveau 3F; Europees taalniveaus C1 en C2 met Europees taalniveau 4F (Litjens, Kuijpers, Wijers & Jonker, 2010). Met deze en bovenstaande informatie kan de CERF-schaal gekoppeld worden aan Nederlandse opleidingsniveaus. Dit wordt weergegeven in tabel 2.

Tabel 2. CERF-schaal gekoppeld aan de Nederlandse opleidingsniveaus.

CERF-schaal	A1	A2	B1	B2	C1	C2
Nederlands referentiekader		1F	2F	3F	4F	
Opleidingsniveaus		Einde primair basisonderwijs	Einde vmbo	Einde mbo-4 en havo	Einde vwo	

De meerderheid van de Nederlanders beschikt over het taalniveau B1 (Rijksoverheid, z.d.). Rijksoverheid (z.d.) adviseert teksten op dit niveau te schrijven, omdat dit voor de meerderheid van de bevolking te begrijpen is. Wanneer mensen taalniveau B1 beheersen, zijn zij in staat om zonder voorbereiding een gesprek te voeren over dagelijkse onderwerpen, zoals hobby's en interessegebieden (Kwakernaak, 2010). Taalniveau B1 kent een aantal tekstkenmerken: 1) De gemiddelde zinslengte ligt rond de tien woorden per zin; 2) De tekststructuur wordt aangegeven met een duidelijk frequente signaal- en verbindingswoorden; 3) Er is één hoofdgedachte per zin en er worden niet teveel gegevens in één keer aangeboden; 4) Er worden voornamelijk actieve zinnen gebruikt (Bohenn, Jansen, Kuijpers, Schot & Stockmann, 2007).

Iemand met taalniveau B1 kan teksten op taalniveau B1, A2 en A1 lezen, terwijl teksten op B2, C1 en C2 niet voldoende begrijpelijk kunnen worden gelezen (Heij & Visser, 2006). De taalniveaus C1, B2, B1, A2 en A1 worden respectievelijk opgedeeld in 15%, 25%, 40%, 15% en 5% van de Nederlandse bevolking (Jansen, 2013; Heij & Visser, 2006). Dit betekent dat 80% van de Nederlanders in staat is om teksten te lezen op taalniveau B1 (Van Kleij-Van Kleffens & Smit, 2016; Jansen, 2013).

Zoals eerder genoemd is bekend dat nieuwsmedia op te hoog taalniveau schrijven, maar op welk taalniveau Nederlandse nieuwsmedia schrijven is niet eerder onderzocht. Daarentegen is wel bekend dat publiekscommunicatie over het algemeen op taalniveau C1 wordt geschreven (Van Kleij-Van Kleffens & Smit, 2016). Dit zou betekenen dat, indien dit ook van toepassing

is op Nederlandse nieuwsartikelen, 85% van de Nederlanders nieuwartikelen niet begrijpelijk kan lezen.

Leesbaarheid

Kraf et al. (2011) stellen dat de leesbaarheid van een tekst een afgeleide van taalniveau is. Volgens Richards, Platt en Platt (in Zamanian & Heydari, 2012) wordt leesbaarheid gedefinieerd als ‘hoe makkelijk geschreven materialen kunnen worden gelezen en begrepen’. Zij stellen dat dit afhangt van verschillende factoren, waaronder gemiddelde zinslengte, aantal woorden en de complexiteit van taal die wordt gebruikt.

Om de leesbaarheid van een tekst te bepalen kan men gebruik maken van leesbaarheidsformules. Volgens Fitzgerald en Watkins (in Burke en Greenberg, 2010) worden leesbaarheidsformules vaak aanbevolen aan journalisten om hen te helpen bij het inschatten van de moeilijkheid van een door hen geschreven tekst. Een leesbaarheidsformule is een wiskundige vergelijking die het niveau van leesvaardigheid voorspelt (Redish, 2000). Dit wordt gedaan op basis van syntactische en semantische moeilijkheid (Fry, 2002). De lengte van een zin wordt gebruikt om de syntactische moeilijkheid te meten en lettergrepen worden gebruikt om de semantische moeilijkheid te meten. Om te beoordelen op welk taalniveau nieuwsartikelen geschreven zijn, kan gebruik worden gemaakt van de Flesch Reading Ease Scale. De Flesch Reading Ease Scale geeft teksten een score tussen 0 en 100. Des te hoger de score, des te makkelijker het materiaal te lezen is. Een lagere score geeft aan dat (delen in) de tekst moeilijk(er) te lezen zijn. De formule die hiervoor wordt gebruikt is (Flesch, 1948): $206,835 - (1,015 \times ASL) - (84,6 \times ASW)$. Hierbij staat ASL voor de gemiddelde zinslengte en ASW voor de gemiddelde woordlengte.

De Flesch Reading Ease Scale formule houdt rekening met Engelse lettergrepen en is niet toepasbaar op Nederlandse teksten. W.H. Douma heeft daarom eind jaren 50 een onderzoek gedaan naar de verschillen tussen Engelse en Nederlandse teksten om te onderzoeken hoe de formule kan worden toegepast op Nederlandse teksten (Renkema, 1982). Uit zijn onderzoek is gebleken dat Nederlandse zinnen en woorden over het algemeen 11% langer zijn dan in het Engels. Op basis van deze bevinding heeft hij de Flesch Reading Ease Scale zodanig aangepast dat deze toepasbaar is op Nederlandse teksten. Deze aanpassingen berusten op twee hypothesen: 1) De Nederlandse vertalingen hebben dezelfde leesbaarheid als de bijbehorende Engelse teksten en 2) Woordlengte en zinslengte moeten in het Nederlands en Engels hetzelfde worden gewaardeerd (Renkema, 1982). De formule die Douma hierbij heeft ontwikkeld luidt:

$206,83 - (0,93 \times SL) - (0,77 \times WL)$ (Janzen & Lentz, 2008). Aangezien dit onderzoek zich richt op Nederlandse teksten wordt de Flesch-Douma formule gebruikt.

Hoe de scores uit deze formule geïnterpreteerd moeten worden, staat weergegeven in tabel 3. Tevens kan hierbij de koppeling worden gemaakt tussen de leesbaarheidsscores en CERF-schaal, waaruit blijkt dat de leesbaarheidsscores tussen 60 en 70 overeen komen met vmbo, wat overeen komt met taalniveau B1.

Tabel 3. Leesbaarheidsscores volgens de Flesch-Douma formule.

Leesbaarheidsscore	Leesbaarheid	Opleidingsniveau
90 – 100	Zeer gemakkelijk	Groep 6
80 – 90	Gemakkelijk	Groep 7
70 – 80	Vrij gemakkelijk	Groep 8
60 – 70	Standaard	Vmbo
50 – 60	Vrij moeilijk	MBO, onderbouw havo/vwo
30 – 50	Moeilijk	Bovenbouw havo/vwo
0 – 30	Zeer moeilijk	Universiteit

Ook bleek uit onderzoek van Wasike (2018) dat de leesbaarheid verschilt per nieuwsrubriek. Hij onderzocht of verschillende nieuwsrubrieken, namelijk hard nieuws, sport nieuws, shownieuws, economisch nieuws, gezondheidsnieuws en onderwijsnieuws, verschilden aan de hand van de Flesch Reading Ease Score. Uit dit onderzoek bleek dat alle rubrieken op moeilijk niveau geschreven werden. Hard nieuws en economisch nieuws bleken het moeilijkst leesbaar te zijn. Sport nieuws was in dit onderzoek het makkelijkste leesbaar. Opvallend is dat hard nieuws relatief moeilijk te lezen is, terwijl dit het meest geconsumeerde nieuws type is (Wasike, 2018). De kanttekening die hierbij geplaatst moet worden is dat dit een Amerikaans onderzoek is. Het is tot op heden volstrekt onbekend hoe de leesbaarheid van nieuwsrubrieken op Nederlandse nieuwssites zich tot elkaar verhouden, daarom zal dit onderzoek hierop ingaan.

Bovenstaande vorm van testen is wereldwijd de meest algemeen gevalideerde leesbaarheidsformule (Kincaid, Fishburne, Rogers & Chissom, 1975). Onderzoeken naar de betrouwbaarheid van deze formule laten zien dat de formule een test-hertest en interbeoordelaarsbetrouwbaarheid van +0.90 heeft, waarbij -1 een sterk negatief verband is en +1 een sterk positief verband (Thomas, Hartley & Kincaid, 1975). Burke en Greenberg (2010)

stelden daarentegen dat leesbaarheidsformules de neiging hebben om tekstuele complexiteit te onderschatten. Overal het algemeen gebruiken leesbaarheidsformules een aantal oppervlakkige tekstkenmerken, zoals zinslengte, woordlengte en/of lettergrepen, en doen daarop vergaande uitspraken over begrijpelijkheid (Kraf, Lentz & Pander Maat, 2011). Zo spelen lezerskenmerken, zoals woordenschat, geen rol in leesbaarheidsformules. Dit terwijl voor globaal begrip van een tekst de lezer 90% van de woorden uit een tekst moet kennen (Bossers, Van der Werf & Van Veen, 2012). Voor gedetailleerd begrip van een tekst moet de lezer 95% van de woorden uit een tekst kennen. Teksten met teveel onbekende woorden zijn onbegrijpelijk en ontoegankelijk.

Er zijn naast zinslengte, woordlengte en lettergrepen meer tekstkenmerken die invloed hebben op de complexiteit van een tekst. De tekstkenmerken kunnen worden onderverdeeld in kenmerken op lexicaal, syntactisch en discours niveau (Vandeghinste & Bulté, 2019). Op lexicaal niveau moeten schrijvers woorden gebruiken die voor de lezer bekend zijn of makkelijk te begrijpen zijn. Indien er moeilijke woorden gebruikt worden, moeten deze worden uitgelegd. Schrijvers moeten consistent blijven in woordkeuze en afwisseling tussen woorden die hetzelfde betekenen wordt daarom afgeraden. Ook moeten metaforen, afkortingen, leenwoorden, lange getallen en percentages worden vermeden. Op syntactisch niveau moet men gebruik maken van actieve zinnen, waarbij de lezer direct moet worden aangesproken. Op discours niveau moeten schrijvers informatie op een eenvoudige manier structureren en de tekst op onderwerp sorteren. Naast bovenstaande factoren zijn er nog meer factoren die de complexiteit van een tekst beïnvloeden. In dit onderzoek, bestaande uit vijf deelonderzoeken, wordt gekeken naar de factoren tangconstructies, buitenlandse woorden, informatiedichtheid, metaforen en afkortingen (Vandeghinste & Bulté, 2019; Inclusion Europe, 2009; Kraf en Pander Maat, 2009).

De eerste factor is tangconstructies. Uit onderzoek is gebleken dat tangconstructies een negatieve invloed hebben op de begrijpelijkheid van een tekst (Jansen & Lentz, 2008; Kraf & Pander Maat, 2009). Renkema (1989a) stelde dat een tangconstructie een constructie is waarin zinsdelen die bij elkaar horen uit elkaar zijn geplaatst. In de literatuur wordt daarom geadviseerd tangconstructies te vermijden (Renkema, 1989b).

Ook het gebruik van buitenlandse woorden is een factor die de leesbaarheid van een tekst moeilijk maakt (Inclusion Europe, 2009). Er bestaat een lineaire relatie tussen het percentage van vocabulaire die men kent en tekstbegrip (Schmitt, Jiang & Grabe, 2011). Eerdere studies lieten zien dat lezers 90% tot 95% van de woorden in een tekst moeten kennen (Hu & Nation: in Schmitt et al., 2011).

Ten derde is informatiedichtheid een beperking, omdat informatiedichte teksten moeilijk(er) leesbaar zijn. Informatiedichtheid heeft betrekking op de hoeveelheid nieuwe informatie die een tekst bevat (Pander Maat, 2009). Deze wordt bepaald door de verhouding tussen het aantal verschillende woorden en het totaal aantal woorden. Een tekst met meer verschillende woorden wordt langzamer gelezen en minder goed onthouden (Kintsch, Kozminsky, Streby, McKoon & Keenan, 1975).

De aanwezigheid van metaforen is ook een factor die een tekst moeilijk(er) leesbaar zou kunnen maken (Inclusion Europe, 2009). Metaforen zijn verschijnselen die worden geconceptualiseerd in termen van een ander verschijnsel (Jansen, Van Nistelrooij, Olislagers, Van Sambeek & De Stadler, 2009). Een voorbeeld hiervan is: ‘Hij woonde in een kamer als een zwijnenstal’. Het begrip van een metafoor is afhankelijk van de complexiteit ervan (Van Mulken, Van Hooft & Nederstigt, 2014). Uit ditzelfde onderzoek is gebleken dat onbekende metaforen minder goed worden begrepen dan minder onbekende metaforen.

Afkortingen

Dit deelonderzoek focust zich het gebruik van afkortingen in teksten, omdat het gebruik van afkortingen een tekst moeilijk leesbaar maken bleek uit onderzoek van Vandeghinste en Bulté (2019) en Inclusion Europe (2009). Een afkorting is een korte beschrijving van een oorspronkelijk langer(e) woord(verzameling) (Zhang, Li, Wang, Sun & Meng in Fumani, 2016), bijvoorbeeld d.w.z. is de afkorting van ‘dat wil zeggen’. Er bestaan verschillende soorten afkortingen, namelijk ‘echte’ afkortingen, symbolen, initiaalwoorden, letterwoorden en verkortingen. Een ‘echte’ afkorting is een verkorte vorm die wordt uitgesproken, zoals het woord of de woordgroep waarvoor de afkorting staat, zoals d.w.z. voor dat wil zeggen (Team Taaladvies, z.d.). Een symbool is een notatie van een wetenschappelijk begrip, een eenheid, een grootheid of een evaluatie, zoals cm voor centimeter (Team Taaladvies, z.d.; Genootschap OnzeTaal, z.d.). Een initiaalwoord is een verkorte vorm die is samengesteld uit de beginletters van een woord en letter voor letter wordt uitgesproken, zoals EU voor Europese Unie (Team Taaladvies, z.d.; Genootschap OnzeTaal, z.d.). Een letterwoord is een verkorte vorm die uit de beginletters van een woord is samengesteld, maar als een gewoon woord wordt uitgesproken, zoals aids voor acquired immune deficiency syndrome (Team Taaladvies, z.d.; Genootschap OnzeTaal, z.d.). Een verkorting is een verkorte vorm die uit een of meer lettergrepen is samengesteld, zoals info voor informatie (Team Taaladvies, z.d.; Genootschap OnzeTaal, z.d.).

Het gebruik van afkortingen is geen nieuw verschijnsel en worden gebruikt om communicatie te versnellen (Fumani, 2016). Ondanks dit voordeel bestaat er altijd een kans dat

afkortingen lezers in de war brengen of ontmoedigen verder te lezen (Nolan, 1991). Afkortingen worden vaak gebruikt in krantenkoppen, bijschriften en andere soorten teksten om ruimte te besparen, maar ze zijn vaak moeilijk(er) te begrijpen door lezers die niet bekend zijn met de context waarin de afkorting wordt gebruikt (Terada, Tokunaga & Tanaka, 2004). Hierdoor vormen afkortingen een obstakel in de tekstverwerking.

In sommige teksten is het gebruik van universele afkortingen, zoals FBI, vereist, maar hierbij moeten journalisten een alfabet soep vermijden. Om deze reden wordt geadviseerd geen afkortingen te gebruiken die lezers niet snel zouden herkennen (Nolan, 1991). Nolan (1991) heeft een studie uitgevoerd naar het gebruik van afkortingen. In dit onderzoek werden aan twintig universitaire studenten 45 afkortingen voorgelegd die veelvuldig in de campusnieuwsbrief gebruikt werden. Hieruit kwam naar voren dat slechts zes van de 45 afkortingen begrijpelijk waren voor alle twintig studenten. Deze uitkomsten roepen de vraag op wat dit betekent voor de begrijpelijkheid van afkortingen in teksten voor de gemiddelde lezer van een nieuwsartikel.

Wetenschappelijke relevantie

Wasike (2018) stelde dat krantleesbaarheid een vergeten gebied is binnen de journalistiek. Daarentegen laat onderzoek zien dat kranten wel degelijk te moeilijk zijn in vergelijking met het algemeen taalniveau B1. In het onderzoek van Wasike (2018) werd gesteld dat nieuwsartikelen op een hoog taalniveau geschreven zijn en dat het taalniveau per rubriek verschilt. Terwijl eerder werd benoemd dat het lezen van nieuwsmedia een belangrijk onderdeel is van maatschappelijke participatie (Wasike, 2018), omdat nieuwsmedia informeren, bijdragen aan de publieke opinie en voorzien in informatie bij politieke en maatschappelijke ontwikkelingen (Van Keulen et al., 2018). Het is daarom van belang dat iedereen toegang heeft tot nieuwsartikelen. In Nederland is niet eerder wetenschappelijk onderzoek gedaan naar de leesbaarheid van nieuwsartikelen. Dit onderzoek verrijkt de kennis over het taalniveau en het gebruik van afkortingen die van invloed zijn op de complexiteit van nieuwsartikelen in vergelijking met de vijf grootste nieuwssites van Nederland.

1.3 Onderzoeksvragen

De volgende onderzoeksvragen zijn opgesteld:

1a. Wat is de leesbaarheid van de vijf populairste nieuwssites van Nederland?

1b. Is er verschil in leesbaarheid per nieuwssite?

1c. Is er verschil in leesbaarheid per nieuwsrubriek?

2a. Wat is het gebruik van afkortingen op de vijf populairste nieuwssites van Nederland?

2b. Is er verschil in het gebruik van afkortingen per nieuwssite?

2c. Is er verschil in het gebruik van afkortingen per nieuwsrubriek?

2. Methode

Dit onderzoek is een corpusstudie naar de leesbaarheid en het gebruik van afkortingen in nieuwsartikelen van de vijf populairste Nederlandse nieuwssites. In dit hoofdstuk zijn het materiaal, de procedure en de statistische toetsing besproken.

2.1 Materiaal

Het soort corpus dat in deze studie is onderzocht zijn nieuwsartikelen van de nieuwssites NU.nl, AD, NOS, De Telegraaf en RTL Nieuws. Er is voor deze nieuwssites gekozen omdat deze de populairste online nieuwsmedia van Nederland zijn. De nieuwsartikelen per nieuwssite zijn onderverdeeld in vijf nieuwsrubrieken, namelijk binnenland, buitenland, economie, sport en show. Hier is voor gekozen omdat alle vijf de nieuwssites deze rubrieken hanteren, zoals te zien in tabel 4.

Tabel 4. Nieuwssites en hun nieuwsrubrieken

Nieuwssite	Nieuwsrubrieken
Nu.nl	Economie, sport, technologie, media en cultuur, binnenland, buitenland, politiek, klimaat.
NOS.nl	Binnenland, buitenland, regionaal nieuws, politiek, economie, koningshuis, technologie, cultuur en media, sport.
AD.nl	Binnenland, buitenland, politiek, economie, gezond, sport, show, wetenschap, auto, technologie.
Telegraaf.nl	Binnenland, buitenland, sport, entertainment, financieel, vrouw, lifestyle.
RTLnieuws.nl	Economie, sport, entertainment, technologie, lifestyle, Nederland, buitenland, politiek.

In totaal bestond het corpus uit 500 nieuwsartikelen, waarvan 100 artikelen per nieuwssite. Van de 100 nieuwsartikelen per nieuwssite zijn er 20 per nieuwsrubriek gecodeerd. De nieuwsartikelen in het corpus zijn afkomstig uit het jaar 2019 en zijn minimaal 200 woorden

lang. Dit woordenaantal was exclusief titels van de nieuwsartikelen. Enkel de inhoudelijke tekst en tussenkoppen werden meegerekend. De steekproeftrekking is random afgenomen. De methode hierbij was een gestratificeerde willekeurige steekproef. Hierbij is de steekproef opgedeeld in nuttige categorieën en binnen deze categorieën is een willekeurige steekproef getrokken. Hier is voor gekozen, omdat uit iedere nieuwsrubriek evenveel nieuwsartikelen gecodeerd moesten worden.

Dit onderzoek maakt onderdeel uit van een groter geheel, bestaande uit vijf onderzoeken naar verschillende factoren die invloed hebben op de leesbaarheid van een nieuwsartikel. Voor deze deelonderzoeken is gebruik gemaakt van hetzelfde corpus. Hierbij is voor het bepalen van de leesbaarheidsscore gebruik gemaakt van alle 500 artikelen, waarbij iedere onderzoeker 100 artikelen heeft gebruikt om zijn/haar eigen factor te onderzoeken. In dit onderzoek is de factor *afkortingen* geanalyseerd.

2.2 Procedure

Nadat alle 500 artikelen zijn uitgezocht moest worden gecodeerd op leesbaarheid en afkortingen. De Flesch-Douma formule is gebruikt om de leesbaarheid van de nieuwsartikelen (1a), per nieuwssite (1b) en per nieuwrubriek (1c) te meten. Bij alle drie de onderzoeksvragen is sprake van één afhankelijke variabele, namelijk leesbaarheid. Bij onderzoeksvraag 1b is de onafhankelijke variabele nieuwssite en bij onderzoeksvraag 1c is de onafhankelijke variabele nieuwsrubriek. De definitie van leesbaarheid die werd aangehouden is: ‘hoe makkelijk geschreven materialen kunnen worden gelezen en begrepen’ (Richards, Platt en Platt: in Zamanian & Heydari, 2012). Bij de variabele leesbaarheid is sprake van ordinaal meetniveau. Het meetniveau van de variabelen nieuwssite en nieuwrubriek is nominaal.

Om het corpus te coderen op leesbaarheid is gebruik gemaakt van de website <https://jeroenvandergun.nl/leesbaarheidstest>. Deze rekenmachine is ontwikkeld door Jeroen van der Gun, postdoctoraal onderzoeker aan de TU Delft. Zijn leesbaarheidsrekenmachine is, net als de Flesch-Douma formule, gebaseerd op het aantal zinnen, aantal woorden en het aantal lettergrepen in een tekst (Van der Gun, 2020). De lettergrepen in deze leesbaarheidsrekenmachine worden berekend door het aantal klinkers en medeklinkers in een tekst te analyseren. Het script van deze leesbaarheidsformule is gebaseerd op de Flesch-Douma formule. Vervolgens komt uit deze leesbaarheidsrekenmachine een score tussen de 0 en 100. Des te hoger de score, des te makkelijker de tekst te lezen is. Deze score wordt door de

leesbaarheidsrekenmachine gekoppeld aan een opleidingsniveau van het Nederlandse scholensysteem.

In deze leesbaarheidsrekenmachine worden getallen die niet voluit geschreven zijn niet meegenomen, aldus J. van der Gun (persoonlijke communicatie, 12 mei 2020). Het is onduidelijk of de Flesch-Douma formule deze zelf wel meeneemt. Een andere beperking van deze leesbaarheidsrekenmachine is dat trema's ook niet worden meegenomen, want de leesbaarheidsrekenmachine telt hierbij een lettergreep te weinig volgens J. van der Gun (persoonlijke communicatie, 12 mei 2020). Echter kunnen de resultaten wel worden gebruikt, omdat voor alle nieuwsartikelen dezelfde leesbaarheidsrekenmachine is gebruikt. Dit zorgt ervoor dat de scores wel onderling met elkaar vergeleken kunnen worden en op basis daarvan de complexiteit ten opzichte van elkaar bepaald kan worden. Daarnaast is er voor deze leesbaarheidsrekenmachine gekozen, omdat in het kader van het tijdsbestek het niet haalbaar was om handmatig op leesbaarheid te coderen en de leesbaarheidsrekenmachine alle artikelen hetzelfde beoordeelde, waardoor geen interbeoordelaarsbetrouwbaarheid nodig was.

De nieuwsartikelen van NU.nl, AD, NOS en RTL Nieuws zijn geraadpleegd via de website Archive.org. Deze website heeft alle nieuwsartikelen van afgelopen jaren opgeslagen op één grote database. De nieuwsartikelen van De Telegraaf waren niet te raadplegen via de website Archive.org, maar De Telegraaf heeft een eigen database. Daarom zijn nieuwsartikelen van De Telegraaf via het eigen archief geraadpleegd. Vervolgens zijn de teksten van de nieuwsartikelen, behalve de titel, gekopieerd en rechtstreeks geplakt in de leesbaarheidsrekenmachine.

Om onderzoeksvraag 2a, 2b en 2c te beantwoorden is een analyse uitgevoerd naar het gebruik van afkortingen in de nieuwsartikelen (2a), per nieuwssite (2b) en per nieuwsrubriek (2c). Bij alle drie de onderzoeksvragen zijn er twee afhankelijke variabelen, namelijk aanwezigheid afkortingen en hoeveelheid afkortingen. De onafhankelijke variabele bij onderzoeksvraag 2b is nieuwssite en bij onderzoeksvraag 2c is de onafhankelijke variabele nieuwsrubriek. De definitie van afkortingen die in dit onderzoek wordt aangehouden is: 'Een afkorting is een korte beschrijving van een oorspronkelijk langer(e) woord(verzameling)' (Zhang, Li, Wang, Sun & Meng in Fumani, 2016). Bij de afhankelijke variabele aanwezigheid afkortingen is sprake van nominaal meetniveau. Bij de hoeveelheid afkortingen in een tekst is sprake van ratio meetniveau. Ook hierbij zijn de onafhankelijke variabelen nieuwssite en nieuwsrubriek van nominaal meetniveau.

Tijdens het coderen zijn de volgende aannames gedaan over afkortingen (Terada, Tokunaga & Tanaka, 2004): 1) Afkortingen in de nieuwsartikelen bevatten minder tekens dan

hun oorspronkelijke woordvormen, zoals bijv. in plaats van bijvoorbeeld; 2) De volgorde van tekens die in een afkorting worden gebruikt zijn een deelverzameling van de tekens die in de oorspronkelijke woordvorm in dezelfde volgorde voorkomen, zoals i.p.v. voor in plaats van; 3) Afkortingen zijn geen echte woorden. Daarnaast bestaan er verschillende soorten afkortingen, namelijk ‘echte’ afkortingen, symbolen, initiaalwoorden, letterwoorden en verkortingen. Tijdens het coderen is gezocht naar deze vormen in de tekst. Hierbij is geen onderscheid gemaakt tussen het soort afkorting.

Op basis van bovenstaande aannames worden afkortingen die in de Nederlandse samenleving zijn opgenomen, zoals aids en Brexit, niet meegerekend als afkorting. Hier is voor gekozen, omdat de oorspronkelijke woordvorm van bijvoorbeeld aids: ‘Acquired Immune Deficiency Syndrome’ minder bekend is dan de afkorting. Deze woorden worden inmiddels als ‘echte’ woorden beschouwd. Afkortingen van instanties, zoals RIVM en FBI, worden wel meegenomen als afkortingen in dit onderzoek. Hier is voor gekozen, omdat de betekenis van deze instanties zeer relevant is voor de context van het nieuwsartikel.

Om te bepalen wanneer er sprake is van een afkorting in een nieuwsartikel zijn er een aantal criteria opgesteld. Allereerst is er een databestand opgesteld met een link naar de nieuwsartikelen die op afkortingen gecodeerd moesten worden. De nieuwsartikelen zijn inhoudelijk doorgelezen om afkortingen in de teksten te identificeren. Woorden in de tekst werden geïdentificeerd als afkorting wanneer er sprake was van punten tussen de letters, zoals n.v.t., of wanneer letters met hoofdletters geschreven waren, zoals RIVM. Bij twijfel over of iets een afkorting was, is dit opgezocht in een woordenboek. Wanneer er afkortingen aanwezig waren, werd een 1 genoteerd, bij afwezigheid van afkortingen noteerde men een 0. Indien er sprake was van een afkorting werd ook de hoeveelheid afkortingen in het artikel genoteerd. Ook bij het coderen op afkortingen zijn de titels van nieuwsartikelen niet meegenomen. In bijlage A staat een uitgebreid codeboek.

Van de 100 nieuwsartikelen is 25% onafhankelijk gecodeerd door een tweede codeur op aanwezigheid van afkortingen en hoeveelheid afkortingen. Dit is gedaan om de betrouwbaarheid van de analyses te waarborgen. De interbeoordelaarsbetrouwbaarheid is gemeten aan de hand van Cohen’s Kappa (K) en Pearson’s r . De interbeoordelaarsbetrouwbaarheid van de variabele Aanwezigheid afkortingen was adequaat: $K = .78, p < .001$. De interbeoordelaarsbetrouwbaarheid van de variabele Hoeveelheid afkortingen was goed: $r(25) = .89, p < .001$.

2.3 Statistische toetsing

Om aan te tonen of de leesbaarheid van de vijf nieuwssites verschilt ten opzichte van het algemeen taalniveau B1 is naar descriptieve statistieken gekeken. Daarnaast is er getoetst of de leesbaarheid per platform verschilt door middel van de eenwegvariantie-analyse met een Bonferroni-gecorrigeerde Ad Hoc toets. Om te toetsen of leesbaarheid per nieuwsrubriek verschilt is ook de eenwegvariantie toets, met een Bonferroni-gecorrigeerde Ad Hoc toets, gebruikt. Om te analyseren wat het gebruik van afkortingen op de vijf populairste nieuwssites van Nederland is, zijn descriptieve statistieken gebruikt. Om te toetsen of er een samenhang bestaat tussen het aantal afkortingen in een tekst en de soort nieuwsrubriek is de Chi-kwadraat toets uitgevoerd. Tot slot werd er een Chi-kwadraat toets uitgevoerd om de mogelijke samenhang tussen het gebruik van afkortingen in een tekst en de nieuwssite te toetsen.

3. Resultaten

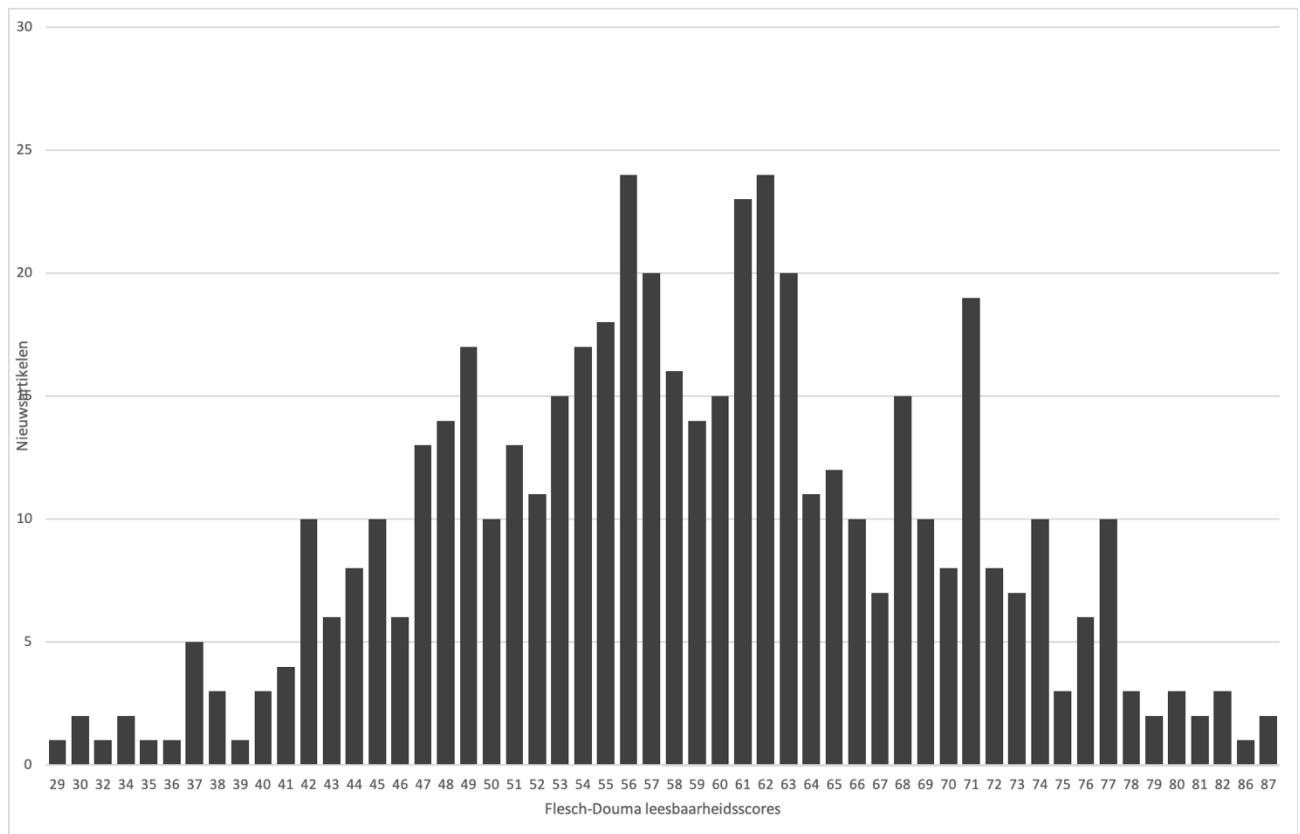
In dit hoofdstuk staan per onderzoeksvraag de resultaten uit de corpusstudie weergegeven.

1a. Wat is de leesbaarheid van de vijf populairste nieuwssites van Nederland?

Er is door middel van descriptieve statistieken gekeken naar de leesbaarheid van de populairste nieuwssites van Nederland. Hieruit is gebleken dat de nieuwsartikelen over het algemeen redelijk moeilijk leesbaar zijn, namelijk een gemiddelde leesbaarheidsscore van 58.58. Een leesbaarheidsscore tussen de 60.00 en 70.00 duidt op een normale leesbaarheid volgens de Flesch-Douma leesbaarheidsformule. De standaarddeviatie is 10.72, met een minimale leesbaarheidsscore van 29.00 een maximum leesbaarheidsscore van 87.00. Het gemiddelde, de standaarddeviatie, het minimum en het maximum staan weergegeven in tabel 5. De verdeling van de Flesch-Douma leesbaarheidsscores is weergegeven in grafiek 1.

Tabel 5. Gemiddelde, standaardafwijking, minimum en maximum van de Flesch-Douma leesbaarheidsscores over de gemiddelde leesbaarheid van alle nieuwsartikelen.

	Gemiddelde (<i>M</i>)	Standaarddeviatie (<i>SD</i>)	Minimum	Maximu m	Aantal (<i>n</i>)
Totaal nieuwsartikelen	58.58	10.72	29.00	87.00	500



Grafiek 1. Verdeling Flesch-Douma leesbaarheidsscores.

1b. Is er een verschil in leesbaarheid per nieuwssite?

Uit een eenwegvariantie-analyse, met Bonferroni-gecorrigeerde post hoc toetsen, van Leesbaarheid op Nieuwssite bleek een significant hoofdeffect van Leesbaarheid ($F(4,495) = 7,59, p < .001$). De Leesbaarheid van RTL Nieuws ($M = 63.15, SD = 10.16$) bleek hoger dan die van NOS ($p = .008$, Bonferroni-correctie; $M = 58.17, SD = 9.87$), NU.nl ($p < .001$, Bonferroni-correctie; $M = 56.59, SD = 11.35$) en De Telegraaf ($p < .001$, Bonferroni-correctie; $M = 55.80, SD = 10.03$). De Leesbaarheid van de overige nieuwssites verschilden niet significant van elkaar (p 's $> .077$, Bonferroni-correctie). De gemiddelden en bijbehorende standaarddeviatie zijn hieronder weergegeven in tabel 6.

Tabel 6. Gemiddelden en standaardafwijkingen van de Flesch-Douma leesbaarheidsscores tussen Nieuwssites (0 = zeer moeilijk leesbaar, 100 = zeer makkelijk leesbaar).

	Gemiddelden (M)	Standaarddeviatie (SD)	Aantal (n)
RTL Nieuws	63.15	10.16	100

NOS	58.17	9.87	100
AD	59.20	10.74	100
NU.nl	56.59	11.35	100
De Telegraaf	55.80	10.03	100

1c. Is er een verschil in leesbaarheid per nieuwsrubriek?

Uit een eenwegvariatie-analyse, met Bonferroni-gecorrigeerde post hoc toetsen, van Leesbaarheid op Nieuwsrubriek bleek een significant hoofdeffect van Leesbaarheid ($F(4,495) = 30.43, p < .001$). De Leesbaarheid van Show ($M = 65.77, SD = 10.47$) bleek hoger dan die van Binnenland ($p < .001$, Bonferroni-correctie; $M = 58.50, SD = 8.96$), Buitenland ($p < .001$, Bonferroni-correctie; $M = 53.30, SD = 10.06$), Economie ($p < .001$, Bonferroni-correctie; $M = 53.68, SD = 8.45$) en Sport ($p = .027$, Bonferroni-correctie; $M = 61.66, SD = 10.47$). De Leesbaarheid van Binnenland ($M = 58.50, SD = 8.96$) bleek hoger dan die van Buitenland ($p = .002$, Bonferroni-correctie; $M = 53.30, SD = 10.06$) en Economie ($p = .004$, Bonferroni-correctie; $M = 53.68, SD = 8.45$). De Leesbaarheid van Sport is hoger dan die van Buitenland ($p < .001$, Bonferroni-correctie; $M = 53.30, SD = 10.06$) en Economie ($p < .001$, Bonferroni-correctie; $M = 53.68, SD = 8.45$). De Leesbaarheid van de overige rubrieken verschilde niet significant van elkaar (p 's $> .208$, Bonferroni-correctie). De gemiddelden en bijbehorende standaarddeviatie zijn hieronder weergegeven in tabel 7.

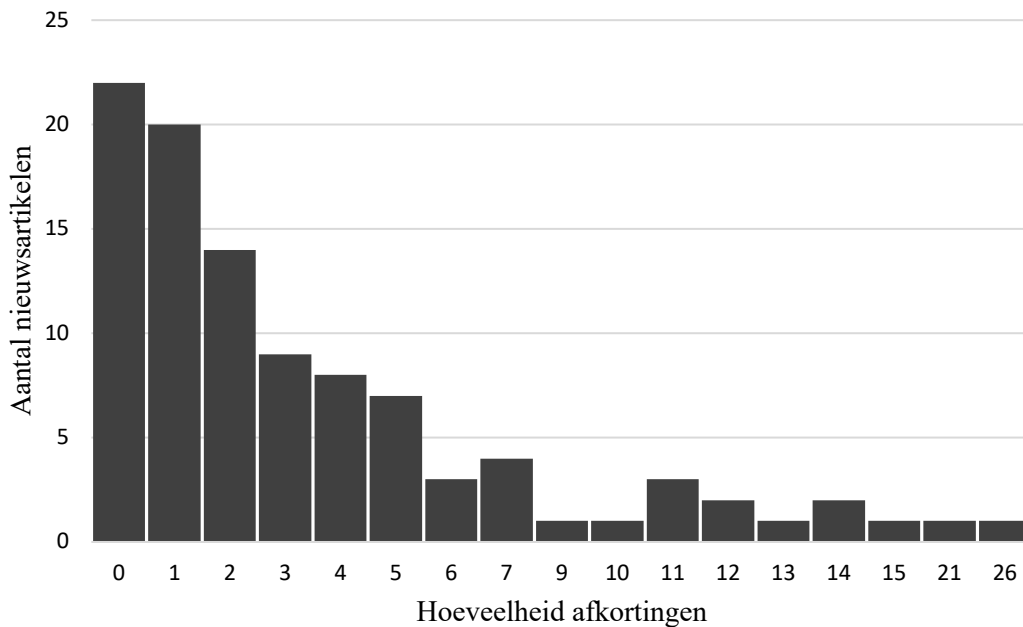
Tabel 7. De gemiddelden en standaardafwijkingen van de Flesch-Douma leesbaarheidsscores tussen Nieuwsrubrieken (0 = zeer moeilijk leesbaar, 100 = zeer makkelijk leesbaar).

	Gemiddelden (M)	Standaarddeviatie (SD)	Aantal (n)
Show	65.77	10.47	100
Binnenland	58.50	8.96	100
Sport	61.66	10.11	100
Buitenland	53.30	10.06	100
Economie	53.68	8.45	100

2a. Wat is het gebruik van afkortingen op de vijf populairste nieuwssites van Nederland?

Er is door middel van descriptieve statistieken gekeken naar het gebruik van afkortingen op de vijf populairste nieuwssites van Nederland. Hieruit is gebleken dat in de 100 geanalyseerde

nieuwsartikelen 78 artikelen afkortingen gebruiken ($M = .78$, $SD = .42$). In de 100 geanalyseerde artikelen werden gemiddeld 3.67 afkortingen per nieuwsartikel gebruikt ($M = 3.67$, $SD = 4.66$). Grafiek 2 de verdeling weer van het gebruik van afkortingen in nieuwsartikelen, dus in twintig nieuwsartikelen stond één afkorting.



Grafiek 2. Hoeveelheid afkortingen ten opzichte van het aantal nieuwsartikelen.

2b. Is er een verschil in het gebruik van afkortingen per nieuwssite?

Uit de χ^2 -toets tussen Aanwezigheid afkortingen en Nieuwssite bleek geen verband te bestaan ($\chi^2(4) = 5.01$, $p = .286$). Tabel 8 geeft de aanwezigheid of afwezigheid van afkortingen per nieuwssite weer.

Tabel 8. De aanwezigheid of afwezigheid van afkortingen per nieuwssite weergegeven in hoeveelheden met bijbehorende percentages (tussen haakjes).

	Afkorting aanwezig	Afkorting afwezig	Aantal (<i>n</i>)
RTL Nieuws	13 (65%)	7 (35%)	20
NOS	16 (80%)	4 (20%)	20
AD	17 (85%)	3 (15%)	20
Nu.nl	14 (70%)	6 (30%)	20
De Telegraaf	18 (90%)	2 (10%)	20

Uit een eenwegvariatie-analyse, met Bonferroni-gecorrigeerde post hoc toetsen, van Hoeveelheid afkortingen op Nieuwssite bleek geen significant hoofdeffect ($F(4,95) = 2.18, p = .077$). De gemiddelden en bijbehorende standaarddeviatie zijn hieronder weergegeven in tabel 9.

Tabel 9. De gemiddelden en standaardafwijkingen van de hoeveelheid afkortingen tussen Nieuwsrubrieken.

	Gemiddelden (M)	Standaarddeviatie (SD)	Aantal (n)
RTL Nieuws	4.65	6.69	20
NOS	2.40	2.52	20
AD	3.45	4.92	20
NU.nl	2.15	2.28	20
De Telegraaf	5.70	4.82	20

2c. Is er een verschil in het gebruik van afkortingen per nieuwsrubriek?

Uit de χ^2 -toets tussen Aanwezigheid afkortingen en Nieuwsrubriek bleek geen verband te bestaan ($\chi^2(4) = 5.59, p = .232$). Tabel 10 geeft de aanwezigheid of afwezigheid van afkortingen per nieuwsrubriek weer.

Tabel 10. De aanwezigheid of afwezigheid van afkortingen per nieuwsrubriek weergegeven in hoeveelheden met bijbehorende percentages (tussen haakjes).

	Afkorting aanwezig	Afkorting afwezig	Aantal (n)
Show	15 (75%)	5 (25%)	20
Binnenland	19 (95%)	1 (5%)	20
Sport	16 (80%)	4 (20%)	20
Buitenland	13 (65%)	7 (35%)	20
Economie	15 (75%)	5 (25%)	20

Uit een eenwegvariatie-analyse, met Bonferroni-gecorrigeerde post hoc toetsen, van Hoeveelheid afkortingen op Nieuwsrubriek bleek geen significant hoofdeffect ($F(4,95) = 1.17, p = .330$). Tabel 11 geeft de gemiddelden en standaarddeviatie weer.

Tabel 11. De gemiddeldes en standaardafwijkingen van de hoeveelheid afkortingen tussen Nieuwsrubrieken.

	Gemiddelden (<i>M</i>)	Standaarddeviatie (<i>SD</i>)	Aantal (<i>n</i>)
Show	3.05	3.87	20
Binnenland	5.15	4.08	20
Sport	3.10	3.87	20
Buitenland	2.50	4.61	20
Economie	4.55	6.32	20

4. Conclusie en discussie

3.1 Conclusie

Met behulp van de resultaten kan beantwoord worden wat de leesbaarheid van de vijf populairste nieuwssites van Nederland (1a), namelijk RTL Nieuws, NOS, AD, NU.nl en De Telegraaf, is. Ook is onderzocht of er tussen deze nieuwssites een verschil in leesbaarheid is (1b) en of er een verschil in leesbaarheid per nieuwsrubriek is (1c), namelijk show, binnenland, sport, buitenland en economie. Daarnaast is onderzocht wat het gebruik van afkortingen op de vijf populairste nieuwssites van Nederland is (2a) en of het gebruik van afkortingen per nieuwssite (2b) en per nieuwsrubriek (2c) verschilt.

Uit het onderzochte corpus kwam naar voren dat het taalniveau van de vijf populairste nieuwssites dicht bij taalniveau B1 ligt met een gemiddelde leesbaarheidsscore van 58.58, waarbij een score tussen de 60.00 en 70.00 gelijk is aan taalniveau B1. Desondanks komt de leesbaarheidsscore nog overeen met taalniveau B2. De leesbaarheid tussen artikelen verschilt onderling sterk, met een minimum van 29.00 en een maximum van 87.00. Daarbij is de standaarddeviatie 10.72. Dit betekent dat het taalniveau tussen de nieuwsartikelen verschilt, maar kan er wel voorzichtig geconcludeerd worden dat het taalniveau in nieuwsartikelen meer de moeilijke kant op gaat, dan de makkelijke.

De leesbaarheid tussen de vijf populairste nieuwssites blijkt van elkaar te verschillen. RTL Nieuws en AD zijn van de vijf nieuwssites het makkelijkst te lezen op basis van de Flesch-Douma formule. Tevens blijkt dat de andere nieuwssites niet van elkaar verschillen in leesbaarheid en dus ongeveer het zelfde taalniveau aanhouden. RTL Nieuws is op basis van descriptieve statistieken normaal leesbaar en NOS, AD, Telegraaf.nl en NU.nl zijn allemaal vrij moeilijk leesbaar. Ook blijkt de leesbaarheid per nieuwsrubriek te verschillen. De nieuwsrubriek shownieuws blijkt het makkelijkst te lezen. Uit dit onderzoek blijkt dat in Nederland het taalniveau tussen economische en buitenlandse nieuwsartikelen niet van elkaar verschillen, maar dat in deze nieuwsrubrieken wel het moeilijkste taalniveau gebruikt wordt in vergelijking met de overige nieuwsrubrieken. Sport en binnenlandse nieuwsartikelen zijn makkelijker te lezen dan economische en buitenlandse nieuwsartikelen.

In het gebruik van afkortingen op de vijf populairste nieuwssites van Nederland, de onderlinge verschillen per nieuwssite en per nieuwsrubriek blijken geen significante verschillen te bestaan. Op basis van descriptieve statistieken blijkt dat in 78 van de 100 nieuwsartikelen afkortingen worden gebruikt. Gemiddeld wordt over al deze artikelen 3.67 afkorting gebruikt.

Hieruit kan geconcludeerd worden dat in het overgrote deel van de nieuwsartikelen afkortingen worden gebruikt.

3.2 Discussie

Dit onderzoek kan de bevindingen uit de literatuur deels bevestigen. Uit dit onderzoek blijkt dat alleen RTL Nieuws een taalniveau aanhoudt dat aansluit bij taalniveau B1. De overige nieuwssites schrijven nieuwsartikelen op een hoger taalniveau, namelijk op mbo en eind havo/vwo niveau. Dit komt overeen met taalniveau B2. Dit sluit dus aan bij de bevindingen van Losttuter (1949) die vond dat journalisten veelal te moeilijk schrijven en dat deze complexiteit toeneemt. Daarentegen werd onderzocht door Heij en Visser (2006) dat Nederlandse publiekscommunicatie op taalniveau C1 wordt geschreven, maar uit dit onderzoek blijkt dat journalisten taalniveau B2 aanhouden. Dit heeft als gevolg dat nieuwsartikelen van NU.nl, AD, NOS en De Telegraaf voor 60% van de Nederlanders niet begrijpelijk zijn. Deze mensen hebben in dit geval dus onvoldoende toegang tot de nieuwsmedia en zullen zichzelf niet volledig kunnen laten informeren over wereldwijde gebeurtenissen, hun opinie laten ontwikkelen en voorzien worden in extra informatie die nodig is bij politieke en maatschappelijke ontwikkelingen door nieuwsmedia (Van Keulen, Korthagen, Diederik & Van Boheemen, 2018).

Eerder werd onderzocht door Wasike (2018) dat de leesbaarheid per nieuwsrubriek verschilt. Uit zijn onderzoek bleek dat de nieuwsrubriek sport het makkelijkst leesbaar is en dat economische artikelen het moeilijkste leesbaar zijn. In dit onderzoek bleek ook dat de leesbaarheid per nieuwsrubriek verschilt, maar dat shownieuws het makkelijkst leesbaar is en dat economische en buitenlandse artikelen even moeilijk leesbaar zijn. Dit verschil is te verklaren door het feit dat het onderzoek van Wasike (2018) een Amerikaans onderzoek is. Om deze reden zijn de resultaten uit dit onderzoek wel representatief voor Nederlandse nieuwsartikelen.

Uit onderzoek van Vandeghinste en Bulté (2019) bleek dat afkortingen een tekst moeilijk leesbaar maken. Afkortingen in een tekst kunnen ervoor zorgen dat lezers in de war brengen of ontmoedigen om verder te lezen (1991). Er wordt geadviseerd geen afkortingen te gebruiken die lezers niet snel zouden herkennen. Deze theoretische bevindingen worden niet ondersteunt in dit onderzoek. Uit dit onderzoek blijkt namelijk dat in 78 van de 100 nieuwsartikelen afkortingen worden gebruikt. Dit sluit daarentegen wel aan bij de literatuur dat journalisten nieuwsartikelen steeds complexer schrijven (Losttuter, 1947). Daarnaast sluit dit

ook aan bij de bevindingen van Nolan (1991) dat steeds meer gebruikt wordt gemaakt van afkortingen om ruimte te besparen.

Een discussiepunt in dit onderzoek is de grootte van de standaarddeviatie ten opzichte van de gemiddelden van de hoeveelheid afkortingen in een nieuwsartikel. Deze grote standaarddeviatie zorgt voor een grote onzekerheid over de hoeveelheid afkortingen in een nieuwsartikel. Voor iedere nieuwssite en nieuwsrubriek zijn twintig artikelen per cel getoetst. Dit kan er toe hebben geleid dat er te weinig statistische power aanwezig is geweest.

De optimale grootte van het corpus van nieuwsrubrieken en nieuwssites om de variabele afkortingen te meten kan bepaald worden met de betrouwbaarheidsinterval methode (McClave & Sinchich, 2017). Voor dit onderzoek is het significantieniveau 5%. De populatiegrootte n , welke twintig is voor het bepalen van het aantal afkortingen in een nieuwsartikel, is voldoende als $n \geq n_i$. Dit houdt in dat de gebruikte populatiegrootte groter moet zijn dan de minimale populatiegrootte voor statistische significantie. Uit berekeningen blijkt dat er in het extreemste geval een steekproefgrootte van 2920 nodig is om binnen buitenland nieuws met 95% zekerheid te stellen dat de hoeveelheid afkortingen binnen 10% van het betrouwbaarheidsinterval ten opzichte van het gemiddelde ligt. Hieruit blijkt dat het bestuderen van twintig artikelen per nieuwssite en nieuwsrubriek onvoldoende is om een statistisch significante uitspraak te doen over de hoeveelheid afkortingen in een nieuwsartikel. Om te verifiëren of er wel een significant effect zou kunnen zijn, is vervolgonderzoek met een grotere steekproef nodig.

Zoals eerder genoemd is er sprake van een grote standaarddeviatie, zowel bij de leesbaarheidsscore als afkortingen. Een mogelijke verklaring hiervoor is dat er in de totale corpus een aantal uitschieters geweest kan zijn. Vervolgonderzoek kan zich daarom richten op deze uitschieters, door bijvoorbeeld te kijken wat er gebeurt als de uitschieters uit het onderzoek gehaald worden. Tevens bleek uit de literatuurstudie dat men minimaal 90% van de woorden in een tekst moet kennen, om een tekst begrijpelijk te kunnen lezen (Bossers, Van der Werf & Van Veen, 2012). Het advies is daarom om bij vervolgonderzoek een experiment uit te voeren waarbij proefpersonen worden getest op woordkennis. Hiermee kan dieper onderzoek worden gedaan naar de leesbaarheid van nieuwsartikelen.

Daarnaast is in dit onderzoek enkel gekeken naar de aanwezigheid van afkortingen en de hoeveelheid afkortingen, omdat er sprake was van verkennend onderzoek. In het theoretisch kader zijn verschillende soorten afkortingen genoemd, welke niet zijn meegenomen in dit onderzoek. Om deze reden zou vervolgonderzoek zich kunnen richten op het gebruik van verschillende soorten afkortingen in nieuwsartikelen. In dit onderzoek is de aanname gedaan dat afkortingen die bekender zijn dan het volledige woord een tekst niet complexer maken.

Daarnaast is aangenomen dat afkortingen van instanties, zoals RIVM, juist wel van invloed zijn op tekstbegrip. Een ander soort vervolgonderzoek zou zich daarom kunnen richten op de invloed van het soort afkorting op tekstbegrip. Eerder in de literatuurstudie werd genoemd dat 69% van de journalisten het gevoel heeft dat rapportagetechnieken de complexe wereld moeten weerspiegelen (Willnat & Weaver, 2014). Aan de andere kant werd juist gesteld dat afkortingen vermeden moeten worden (Nolan, 1991). Vervolgonderzoek zou zich daarom kunnen richten op de vraag of er een toename is in het aantal afkortingen.

Tevens zijn in dit onderzoek de titels van de nieuwsartikelen niet meegenomen. Het is daarom interessant om bij vervolgonderzoek te onderzoeken of nieuwsmedia afkortingen gebruiken in titels, wat de leesbaarheid hiervan is en of dit van invloed is op het tekstbegrip. Daarnaast is een interessant vervolgonderzoek of de leesbaarheid van de titel een voorspeller zou kunnen zijn voor de leesbaarheid van het gehele nieuwsartikel.

Deze studie draagt bij aan maatschappelijke implicaties. Resultaten van deze studie kunnen van praktisch nut zijn voor Nederlandse journalisten. Uit dit onderzoek blijkt dat het gemiddelde nieuwsartikel een relatief moeilijker taalniveau heeft dan taalniveau B1, daarnaast verschilt het taalniveau per nieuwssite en per nieuwsrubriek. Willnat en Weaver (2014) stelden dat journalisten vinden dat rapportagetechnieken de complexe wereld moeten weerspiegelen. Indien journalisten er naar streven dat nieuwsartikelen begrijpelijk zijn voor het merendeel van de Nederlandse samenleving, is het advies deze complexiteit niet te laten weerspiegelen in het taalniveau van nieuwsartikelen en nieuwsartikelen te gaan schrijven op taalniveau B1.

Referentielijst

- Babalola, E.T. (2002). Newspapers as instruments for building literate communities: the Nigerian experience. *Nordic Journal of African Studies*, 11(3), 403-410.
- Bohenn, E., F. Jansen, C. Kuijpers, R. Thijssen, I. Schot en W. Stockmann (2007). *Raamwerk Nederlands: Nederlands in (v)mbo-opleiding, beroep en maatschappij* (1^e ed., pp. 54-55). 's Hertogenbosch: CINOP
- Bonset, H. (2009). Een kritische blik op het Referentiekader Taal. *Levende Talen Magazine*, 96(7), 6-10.
- Bossers, B., Van der Werf, R., & Van Veen, L. (2012). Zelfstandig werken aan taalvaardigheid met de Nieuwslezer. *Teksten op maat, woorden op maat*, 76, 16-20.
- Buisman, M., Allen, J., & Fouarge, D. (2013). PIAAC: Kernvaardigheden voor werk en leven. *Journal of Health Economics*, 31(3), 490-501.
- Burke, V. & Greenberg, D. (2010). Determining readability: How to select and apply easy-to-use readability formulas to assess the difficulty of adult literacy materials. *Adult Basic Education and Literacy Journal*, 4(1), 34-42.
- Christoffels, I., Baay, P., Bijlsma, I., & Levels, M. (2016). *Over de relatie tussen laaggeletterdheid en armoede*. 's Hertogenbosch: Expertisecentrum Beroepsonderwijs.
- Cloudt, S. (2010). De maatschappelijke onderneming als metafoor? *BTMO*, 27-30.
- Commissariaat voor de Media (2019). Mediamonitor 2019. *Reuters Institute Digital News Report Nederland 2019*. Geraadpleegd op 12 maart 2020 van <https://www.cvdm.nl/actueel/mediamonitor-2019-nederlanders-blijven-zich-breed-oriënteren-voor-nieuws>
- Dalecki, L., Lasorsa, D. L., & Lewis, S. C. (2009). The news readability problem. *Journalism Practice*, 3(1), 1-12. doi: 10.1080/17512780802560708
- Dictionary, O. S. (2014). Definition Of News.
- Evers-Vermeul, J., & Holtermann, M. (2013). Doorlopende leerlijnen: implicaties voor leveling van leer-en examenteksten voor het middelbaar onderwijs. *Tijdschrift voor Taalbeheersing*, 35(1), 1-24. doi: 10.5117/TVT2013.1.EVER
- Fink, K., & Schudson, M. (2014). The rise of contextual journalism, 1950s–2000s. *Journalism*, 15(1), 3-20. doi: 10.1177/1464884913479015
- Flesch, R. (2016). How To Write Plain English. *University of Canterbury*. Geraadpleegd op 11 maart 2020 van

- https://web.archive.org/web/20160712094308/http://www.mang.canterbury.ac.nz/writing_gui_de/writing/flesch.shtml
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3), 221-233. doi: 10.1037/h0057532
- Fumani, M. R. F. Q. (2016). Acronyms and Abbreviations Application in News Articles: A Comparative Study of Iranian and Indian Newspapers. *International Journal of English Language & Translation Studies*, 4(3), 81-90.
- Genootschap OnzeTaal (z.d.). *Adviezen over afkortingen*. Geraadpleegd op 19 april 2020 van <https://onzetaal.nl/taaladvies/thematisch/afkortingen/>
- Heij, K., & Visser, W. (2006). *Schrijven in eenvoudig Nederlands*. Den Haag: Sdu uitgevers.
- Houtkoop, W. (1999). Basisvaardigheden in Nederland. *Amsterdam: Max Goote Kenniscentrum*.
- Jansen, C. (2013). Taalniveau B1: de nieuwste kleren van de keizer. *Onze Taal*, 82(2), 56-57.
- Jansen, C. J. M., & Lentz, L. (2008). Hoe begrijpelijk is mijn tekst? De opkomst, neergang en terugkeer van de leesbaarheidsformules. *Onze Taal*, 1, 4-7.
- Jansen, C. J. M., Van Nistelrooij, M., Olislagers, K., Van Sambeek, M., R. & De Stadler, L. G. (2009). Tegen zo'n leger kun je maar beter geen oorlog hoeven voeren. Effecten van het gebruik van metaforen in voorlichtingsteksten over HIV/AIDS in Zuid-Afrika. *Studies in Taalbeheersing*, 3, 175-185.
- Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. *Institute for Simulation and Training*, 56.
- Kintsch, W., E. Kozminsky, W.J. Streby, G. McKoon & J.M. Keenan (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 14, 196-214. doi: 10.1016/S0022-5371(75)80065-X
- Kraf, R., Lentz, L., & Pander Maat, H. (2011). Drie Nederlandse instrumenten voor het automatisch voorspellen van begrijpelijkheid - Een klein consumentenonderzoek. *Tijdschrift voor Taalbeheersing*, 33(3), 249-265. doi: 10.5117/TVT2011.3.DRIE411
- Kraf, R., & Pander Maat, H. (2009). Leesbaarheidsonderzoek: Oude problemen, nieuwe kansen. *Tijdschrift voor taalbeheersing*, 31(2), 97-123. doi: 10.5117/TVT2009.2.LEES356
- Kwakernaak, E. (2010). Het ERK hinkt. *Levende Talen Magazine*, 97(5), 4-8.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago.
- Litjens, P., Kuijpers, C., Wijers, M., & Jonker, V. (2010). *Referentieniveaus taal en rekenen in*

- het mbo: een vergelijking op hoofdlijnen.*'s-Hertogenbosch: CINOP.
- Lostutter, M. (1947). Some critical factors of newspaper readability. *Journalism Quarterly*, 24(4), 307. doi: 10.1177/107769904702400402
- McClave J. & Sinchich, T. (2017). *A First Course in Statistics*. New York City: Pearson Education Limited.
- Nolan, J. (1991). Acronyms in Newspapers: ASAP or SNAFU? *Newspaper Research Journal*, 12(1), 92–97. doi: 10.1177/073953299101200110
- Pelgrims, P. (2005). Journalistieke radio- en tv-teksten en schrijven. *Tielt: Uitgeverij Lannoo nv*.
- Redish, J. (2000). Readability formulas have even more limitations than Klare discusses. *ACM Journal of Computer Documentation*, 24(3), 132-137. doi: 10.1145/344599.344637
- Renkema, J. (1982). Leesbaarheidsformules, een overzichtartikel van de werkgroep Begrijpelijkheidsonderzoek. *Tijdschrift voor massacommunicatie*, 10, 115-123.
- Renkema, J. (1983). Onderzoek naar woordenrijkdom: Taalstatistische analyse via een microcomputer. *Tijdschrift voor Taalbeheersing*, 5(4), 275-289.
- Renkema, J. (1989a). Vermijd tangconstructies. *Communicatief*, 2, 34-40.
- Renkema, J. (1989b). Tangconstructies, experimenteel onderzoek naar leesbaarheid en attentiewaarde. *Tijdschrift voor neerlandistiek*, 12, 444-462.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The Percentage of Words Known in a Text and Reading Comprehension. *The Modern Language Journal*, 95, 26-43. doi: 10.1111/j.1540-4781.2011.01146.x
- Rijksoverheid (z.d.). *Taalniveau B1*. Geraadpleegd op 11 maart 2020 van <https://www.communicatierijk.nl/vakkennis/rijkswebsites/aanbevolenrichtlijnen/taalniveau-b1>
- Sikkema, T., Lentz, L., Pander Maat, H., & Jungmann, N. (2017). De schuld van incassodocumenten. *Tijdschrift voor Taalbeheersing*, 39(3), 273-295. doi: 10.5117/TVT2017.3.SIKK
- Team Taaladvies. (z.d.). *Afkortingen*. Geraadpleegd op 19 april 2020 van <https://taaltelefoon.be/spellingregels/afkortingen>
- Terada, A., Tokunaga, T., & Tanaka, H. (2004). Automatic expansion of abbreviations by using context and character information. *Information processing & management*, 40(1), 31-45. doi: 10.1016/S0306-4573(02)00080-8
- Thomas, G., Hartley, R. D., & Kincaid, J. P. (1975). Test-retest and inter-analyst reliability of the automated readability index, Flesch reading ease score, and the fog count. *Journal*

- of Reading Behavior*, 7(2), 149-154. doi: 10.1080/10862967509547131
- Vandeghinste, V., & Bulté, B. (2019). Linguistic Proxies of Readability: Comparing Easy-to-Read and regular newspaper Dutch. *Computational Linguistics in the Netherlands Journal*, 9, 81–100.
- Van Keulen, I., Korthagen, I., Diederens, P., & van Boheemen, P. (2018). *Digitalisering van het nieuws: online nieuwsgedrag en desinformatie en personalisatie in Nederland*. Den Haag: Rathenau Instituut
- Van Kleij-Van Kleffens, C., & Smit, R. (2016). Laaggeletterdheid steeds groter probleem. *Medisch contact*, 6(11), 38-40.
- Van Mulken, M., van Hooft, A., & Nederstigt, U. (2014). Finding the tipping point: Visual metaphor and conceptual complexity in advertising. *Journal of Advertising*, 43(4), 333-343. doi: 10.1080/00913367.2014.920283
- Wasike, B. (2018). Preaching to the choir? An analysis of newspaper readability vis-a-vis public literacy. *Journalism*, 19(11), 1570-1587. doi: 10.1177/1464884916673387
- Willnat, L., & Weaver, D. H. (2014). American Journalist in the digital age: Key findings. *School of Journalism, Indiana University*.
- Zamanian, M., & Heydari, P. (2012). Readability of Texts: State of the Art. *Theory & Practice in Language Studies*, 2(1), 53-53. doi: 10.4304/tpls.2.1.43-53

Bijlage A. Codeboek

Dit codeboek is het stappenplan voor de tweede codeur. De tweede codeur heeft dit stappenplan ontvangen voordat hij/zij begon met coderen.

1. Op de gezamenlijke drive is een bestand gezet met daarin 100 artikelen die gecodeerd moeten worden op aanwezigheid afkortingen en hoeveelheid afkortingen. Dit zijn de artikelen die eerder door de eerste codeur zijn gecodeerd.
2. Bij het coderen moet in gedachte worden gehouden dat er verschillende soorten afkortingen bestaan, namelijk ‘echte’ afkortingen, symbolen, initiaalwoorden, letterwoorden en verkortingen.
 - a. Een ‘echte’ afkorting is een verkorte vorm die wordt uitgesproken, zoals het woord of de woordgroep waarvoor de afkorting staat, zoals d.w.z. voor dat wil zeggen (Team Taaladvies, z.d.).
 - b. Een symbool is een notatie van een wetenschappelijk begrip, een eenheid, een grootte of een evaluatie, zoals cm voor centimeter (Team Taaladvies, z.d.; Genootschap OnzeTaal, z.d.).
 - c. Een initiaalwoord is een verkorte vorm die is samengesteld uit de beginletters van een woord en letter voor letter wordt uitgesproken, zoals EU voor Europese Unie (Team Taaladvies, z.d.; Genootschap OnzeTaal, z.d.).
 - d. Een letterwoord is een verkorte vorm die uit de beginletters van een woord is samengesteld, maar als een gewoon woord wordt uitgesproken, zoals aids voor acquired immune deficiency syndrome (Team Taaladvies, z.d.; Genootschap OnzeTaal, z.d.).
 - e. Een verkorting is een verkorte vorm die uit een of meer lettergrepen is samengesteld, zoals info voor informatie (Team Taaladvies, z.d.; Genootschap OnzeTaal, z.d.).
3. In dit onderzoek worden verkortingen en letterwoorden die zijn opgenomen in de Nederlandse taal niet meegenomen als afkorting, omdat er wordt aangenomen dat deze woorden bekender of even bekend zijn als de afkorting. Afkortingen van instanties zoals RIVM of FBI worden wel meegenomen, omdat er wordt aangenomen dat deze afkortingen zeer relevant zijn voor de context van een nieuwsartikel.

4. Klik nu op de link in het bestand op drive. Hierna kan de codeur beginnen met het lezen van de nieuwartikelen. Hierbij hoeven de titels van de artikelen niet gecontroleerd worden op afkortingen.
5. Kijk in de tekst naar klinkers en medeklinkers waar punten tussen staan of die eindigen met een punt. Denk hierbij aan afkortingen, zoals n.a.v. en bijv. Kijk in de tekst ook naar woorden die met hoofdletters worden geschreven, zoals FBI en RIVM.
6. Bij twijfel of er inderdaad sprake is van een afkorting moet een woordenboek worden geraadpleegd. Bij buitenlandse afkortingen moet een buitenlands woordenboek geraadpleegd worden en bij Nederlandse afkortingen een Nederlands woordenboek.
7. Indien er (een) afkorting(en) wordt geïdentificeerd in de tekst noteer een 1. Indien er geen afkorting(en) wordt geïdentificeerd in de tekst noteer een 0.
8. Wanneer er één of meerdere afkortingen in de tekst zijn geïdentificeerd moet er genoteerd worden hoeveel afkortingen er in het desbetreffende artikel voorkomen. Er hoeft niet genoteerd te worden om welke soort afkorting het gaat.