

Thesis bachelor Kunstmatige Intelligentie

Is het opvolgen van aanwijzingen uit T-Scan een goede manier om teksten te verbeteren in de context van woningcorporaties en is dit een bruikbare tool is om te implementeren bij woningcorporaties zelf?

Auteur

Michiel Schouten
m.schouten5@student.uu.nl

Eerste lezer

M.P. Schraagen
m.p.schraagen@uu.nl

Tweede lezer

F.W. Adriaans
f.w.adriaans@uu.nl

Kunstmatige Intelligentie
7,5 ECTS



**Utrecht
University**

Department Geesteswetenschappen
Universiteit Utrecht
Nederland
23 april 2025

Samenvatting

Hier komt de abstract van de scriptie.

Inhoudsopgave

| | | |
|----------|---|-----------|
| 1 | Introductie | 4 |
| 1.1 | Onderzoeksvraag | 5 |
| 1.2 | Relevantie aan AI | 5 |
| 1.3 | Maatschappelijke relevantie | 5 |
| 1.4 | Theoretisch kader | 6 |
| 1.4.1 | Inbedding huidig wetenschappelijk onderzoek | 6 |
| 1.4.2 | Wat houdt leesbaarheid in? | 8 |
| 1.4.3 | Hoe kan je een tekst aanpassen ter vereenvoudiging? | 8 |
| 1.4.4 | Wat houdt T-Scan/LiNT in? | 9 |
| 1.4.5 | Schatting van taalniveau in woningcorporaties | 10 |
| 2 | Methode | 11 |
| 2.1 | Brieven | 11 |
| 2.2 | Leesbaarheidsinstrument | 12 |
| 2.3 | Schrijfprotocol | 12 |
| 2.4 | Herschrijven van brieven | 13 |
| 2.5 | Enquête | 14 |
| 2.5.1 | Gebruikte teksten | 14 |
| 2.5.2 | Opzet | 14 |
| 2.5.3 | Proefpersonen | 16 |
| 2.6 | Verwerking enquête | 16 |
| 3 | Resultaten | 20 |
| 3.1 | Begrijpelijkheid en betuttelend | 20 |
| 3.2 | Suggesties op T-Scan | 23 |
| 3.3 | Bewerkte teksten behouden of weglaten? | 25 |
| 4 | Conclusie | 27 |
| 5 | Discussie | 29 |
| 5.1 | Beperkingen | 29 |
| 5.2 | Vervolgonderzoek | 30 |
| | Appendices | 32 |

| | | |
|----------|---------------------------------------|-----------|
| A | Brieven | 33 |
| A.1 | Gebruikte brieven overzicht | 33 |
| A.2 | Overzicht variabelen | 33 |
| A.3 | Synoniemen en vaktermen | 33 |
| B | Gebruikte alinea's overzicht | 36 |
| B.1 | Tekst 1 | 36 |
| B.2 | Tekst 2 | 36 |
| B.3 | Tekst 3 | 37 |
| B.4 | Tekst 4 | 37 |
| B.5 | Tekst 5 | 37 |
| B.6 | Tekst 6 | 37 |
| B.7 | Tekst 7 | 38 |
| B.8 | Tekst 8 | 38 |
| B.9 | Tekst 9 | 38 |

Hoofdstuk 1

Introductie

De meeste teksten die overheden communiceren naar haar burgers zijn notoir lastig te lezen. Zo bleek uit onderzoek dat dat 10% van de brieven aangaande schuldhelpverlening lastig te lezen was voor hoogopgeleiden. 50% van de teksten was moeilijk te begrijpen voor laagopgeleide lezers volgens onderzoek van Pander Maat en van der Geest (2021).

Hierdoor begrijpen burgers vaak niet wat de overheid van hen vraagt of wat ze tegen hen wil zeggen. Burgers die (financieel) sterk afhankelijk zijn van de overheid zijn ook vaker laaggeletterd. Dit kan snel problemen opleveren. [bron]. Het aantal laaggeletterden in Nederland kan conservatief geschat worden op 15% (Maat, Kleijn & Frissen, 2023, p. 18).

De laatste jaren wordt er steeds meer bewustzijn gecreëerd over de leesbaarheid van teksten die de overheid communiceert naar haar burgers. Onder andere door de campagne Direct Duidelijk (tegenwoordig Gebruiker Centraal)¹. Semipublieke instellingen als het Openbaar Vervoersbedrijven en woningcorporaties zijn ook bezig om al hun teksten begrijpelijker te maken voor hun klanten.

Deze scriptie zal de focus leggen op de begrijpelijkheid van de teksten die woningcorporaties versturen naar hun klanten. Laaggeletterden hebben over het algemeen een lage socio-economische positie. Op hun beurt zijn mensen in zo'n positie oververtegenwoordigd in het klantenbestand van woningcorporaties. Ook zogenaamde nieuwkomers zijn de Nederlandse taal vaak niet machtig maar zijn wel aangewezen op sociale huurwoningen. Hierdoor is het extra belangrijk dat die teksten leesbaar en begrijpelijk zijn. Op dit moment zijn er teksten die nog niet op het gewenste niveau zijn, wat burgers in de problemen kan brengen.

¹<https://www.gebruikercentraal.nl/meedoen/netwerk-direct-duidelijk/duidelijke-overheidscommunicatie/>

1.1 Onderzoeksvraag

Er bestaan al jaren automatische tekstanalyse instrumenten die de leesbaarheid van een tekst kunnen analyseren en die kunnen aangeven welke teksteigenschappen ervoor zorgen dat een tekst moeilijk leesbaar is. Een zo'n tool is de T-Scan, ontwikkeld door Pander Maat et al. (2014). Uit de teksteigenschappen die de T-Scan teruggeeft kunnen tekstverbeteringen worden gedestilleerd. De belangrijkste teksteigenschappen uit de T-Scan die bepalend zijn voor de tekstcomplexiteit volgt uit onderzoek van Kleijn (2018) en Maat et al. (2023). Een tool hieromheen gebouwd is het Leesbaarheidsinstrument voor Nederlandse Teksten (LiNT). De vraag is of het opvolgen van aanwijzingen uit T-Scan een goede manier is om teksten te verbeteren in de context van woningcorporaties en of dit een bruikbare tool is om te implementeren bij woningcorporaties zelf. Deze vraag zal worden beantwoord door antwoord te verkrijgen op drie deelvragen.

1. Krijgt de bewerkte tekst een betere leesbaarheidsscore?
2. Signaleren mensen dezelfde, of andere problemen als de T-Scan?
3. In hoeverre dragen de gesignaleerde tekstproblemen van T-Scan bij aan het leesbaar maken van de tekst?

1.2 Relevantie aan AI

De T-Scan en de toegepaste versie LiNT zijn leesbaarheidstools. Ze rekenen verschillende

1. Doel is dus om de tool T-Scan te testen om daadwerkelijk leesbaardere teksten te maken in dit domein.

1.3 Maatschappelijke relevantie

In zekere zin is deze scriptie een voortgang op het onderzoek Monitor Begrijpelijke Overheidsteksten van Pander Maat en van der Geest (2021). Hier werd onderzocht hoe leesbaar teksten van regionale overheden zijn met betrekking tot schuldhulpverlening. Uit dit onderzoek bleek dat “De meeste geanalyseerde overeenkomsten zijn onbegrijpelijk voor lezers van alle opleidingsniveaus, terwijl ondertekening van die overeenkomsten grote consequenties heeft voor het doen en laten van de burger met schulden.” (Pander Maat & van der Geest, 2021, p. 20) Het gaat hier dus over overeenkomsten in de zin van het contractrecht. In dit onderzoek zal worden gekeken hoe dat is bij woningcorporaties gezien deze ook een zeer belangrijke rol hebben in de maatschappij.

De scriptie wordt uitgevoerd bij Zig Websoftware. Dit softwarebedrijf levert software voor woningcorporaties die 40% van de sociale huurwoningen beheren. Een groot deel van de Nederlanders die in sociale huurwoningen zitten maken gebruik van de software van Zig. Dit maakt dat deze scriptie een grote maatschappelijke impact kan maken. Indien T-Scan bruikbaar is voor het verhogen van de leesbaarheid, dan kan dit ervoor zorgen dat 2 miljoen Nederlanders begrijpelijkere brieven ontvangen.

1.4 Theoretisch kader

1.4.1 Inbedding huidig wetenschappelijk onderzoek

Het onderzoek en onderzochte vragen staan niet op zichzelf. Veel onderzoek in dit veld in de context van de Nederlandse taal is gedaan door Dr. H.L.W. Pander Maat geaffilieerd onderzoeker bij de Universiteit Utrecht. Vele studies van Pander Maat liggen in lijn met het onderzoek in deze scriptie, hierom zal hij veelvuldig worden aangehaald. Hij was zelf echter niet betrokken bij dit onderzoek.

Er wordt al vele jaren onderzoek gedaan naar de leesbaarheid van teksten en welke factoren daar aan bijdragen zie bijvoorbeeld Freebody en Anderson (1983). Freebody en Anderson deden onderzoek naar de samenhang van het vocabulaire van een tekst en het tekstbegrip van de lezer. Hier bleek dat de (on)bekendheid van woorden in een tekst een grote voorspellende factor is voor het begrijpen van een tekst. Hoewel het niet het hoofddoel is van dit onderzoek zal dit effect nogmaals bestudeerd worden.

Recenter onderzoek naar automatische tekstanalyse is onder andere gedaan voor de Portugese taal Aluisio, Specia, Gasperin en Scarton (2010). Hier werd naar 59 teksteigenschappen gekeken om een tekst te kwalificeren op de non-lineaire schaal: *Rudimentair*, *Basis* of *Geavanceerd*. *Rudimentair* houdt in dat iemand expliciete informatie kan vinden in een korte tekst. *Basis* houdt in dat iemand teksten met een gemiddelde lengte kan lezen en hier informatie uit kan destilleren. *Geavanceerd* betekent dat iemand een tekst volledig begrijpt, meningen van feiten kan onderscheiden, de tekst kan samenvatten en er iets over de inhoud kunnen zeggen. Als onderdeel van dit onderzoek hebben ze op basis van de 59 teksteigenschappen een Machine Learning model getraind die met redelijke zekerheid kon zeggen of een tekst was bedoeld voor een *Rudimentaire*, *Basis* of *Geavanceerde* lezer. In combinatie met dit model is er ook een andere tool ontwikkeld: SIMPLIFICA Scarton, Oliveira, Candido Jr, Gasperin en Aluísio (2010). Dit is een programma die automatisch tekst kan versimpelen op basis van vooraf bepaalde regels. Een regel kan zijn om een eenvoudiger synoniem te gebruiken voor een moeilijk woord. De tool doet op zinsniveau lexicale en syntactische suggesties. Het is aan de gebruiker om de wijzigingen over te nemen. Over

deze tool is helaas weinig engelstalige literatuur te vinden, in hoeverre dit programma bijdraagt aan de leesbaarheid van teksten is voor mij onbekend.

Het automatisch herschrijven van teksten met behulp van Large Language Models (LLMs) zoals ChatGPT² is sinds een aantal jaar ook mogelijk. Zo is onderzocht of tekst daadwerkelijk leesbaarder en begrijpelijker wordt als een LLM de opdracht krijgt om dit te doen (van Raaij, Kolkman en Podoyntsina (2024)). van Raaij et al. kwamen tot de conclusie dat de beste manier om een tekst zo te versimpelen is door ChatGPT de volgende prompts in te laten voeren. Eerst moet een tekst samengevat worden in bulletpoints. Het volgende prompt is om die bulletpoints eenvoudig uit te schrijven tot een lopende tekst. Met deze methode werd de meeste informatie behouden in de versimpelde versie en werd het tekstbegrip verhoogt. Het tekstbegrip werd gemeten door participanten open en gesloten vragen te stellen over de tekst. Dit waren vragen als "Wat houdt een persoonsgebonden budget in?". Daarnaast werd ook gevraagd wat participanten vonden van de toon van een tekst. In deze scriptie wordt hier ook naar gevraagd. Deze scriptie maakt echter geen gebruik van LLM's om teksten te versimpelen.

Zoals gezegd wordt in deze scriptie ook onderzoek gedaan naar de toon van een versimpelde brief. Er wordt onderzocht of mensen een eenvoudigere tekst betuttelender of neerbuigender vinden dan eenzelfde tekst met een hoger leesniveau. Hier is in het verleden al uitgebreid onderzoek gedaan. Uit de literatuur blijkt dat tekstvereenvoudiging onder, ook hoogopgeleide, lezers beter bevalt Maat en Gravekamp (2022) dan complexere teksten. De aanname is dat als iemand een complex onderwerp eenvoudig kan uitleggen deze persoon meer kennis over dit onderwerp heeft dan iemand die alleen met complexe tekst een complex onderwerp kan uitleggen. Of een versimpelde tekst betuttelend is wordt gevraagd in de enquête.

Er is ook al onderzoek gedaan naar welke tekstkenmerken verschillen tussen eenvoudige en complexe teksten die hetzelfde verhaal vertellen (Pander Maat en Ditewig (2017)). Uit dit onderzoek werden de verschillen wel onderzocht maar bleef het onduidelijk welk tekstkenmerk precies bijdroeg aan verbetert tekstbegrip onder lezers.

Deze scriptie gaat meer in op de voorspellende/voorschrijvende aard van T-Scan dan de beschrijvende aard zoals bij bovengenoemd onderzoek.

Verder leesbaarheidsonderzoek is gedaan om te onderzoeken welke teksten ten opzichte van elkaar eenvoudiger dan wel complexer zijn D Clerq et al. (2014). Met behulp van zowel twee taalexperts als gemiddelde mensen uit de samenleving werd de leesbaarheid getoetst door teksten onderling te vergelijken. Zowel de rangschikking van de experts als de resultante van de inbreng van alle deelnemers lag zeer op één lijn. Er werd in dit onderzoek niet gevraagd naar specifieke teksteigenschappen die de tekst eenvoudiger of moeilijker maakten ten opzichte van een andere tekst. De teksten die

²<https://openai.com/chatgpt/overview/>

vergeleken werden verschilden qua onderwerp van elkaar, het ging dus niet om een moeilijke en vereenvoudigde versie van eenzelfde tekst.

1.4.2 Wat houdt leesbaarheid in?

Leesbaarheid is op verschillende manieren te meten. Het is zeer lastig om leesbaarheid te meten van een tekst omdat uiteindelijk mensen de tekst moeten begrijpen en interpreteren. Er blijft dus altijd ruimte voor onnauwkeurigheden, gezien leesbaarheid sterk afhangt van de mens zelf.

Leesbaarheid kan zowel worden ingevuld als het gemak waarmee een tekst wordt gelezen als de begrijpelijkheid ervan. Dit laatste wordt ook wel tekstbegrip genoemd. Tekstbegrip kan zowel gemeten worden door inhoudelijke vragen te beantwoorden over de gelezen tekst. De vragen kunnen gaan over de hoofdlijnen zoals gedaan in Pander Maat en Ditewig (2017); van Raaij et al. (2024). Er wordt dan onderzocht of lezers meer of minder goede antwoorden geven op deze begripsvragen bij de ene versie ten opzichte van de andere versie. Crossley, Skalicky en Dascalu (2019) bepaalde het tekstbegrip door simpele ja-nee vragen te stellen over twee teksten, welke sneller gelezen was en welke begrijpelijker was.

Een andere manier om tekstbegrip te meten is door een zogenaamde Cloze-test uit te voeren. In een Cloze-test worden bepaalde woorden weggelaten uit een doorlopende tekst. Het is aan een participant om het juiste woord in te voegen. Een juist woord wordt meestal gekwalificeerd als een woord die dezelfde semantiek bezit. Deze manier wordt gezien als een goede indicator om het tekstbegrip te onderzoeken Kleijn (2018).

Anders dan de eerder genoemde manier om de leesbaarheid van een tekst te bepalen en om te kijken of een versimpelde tekst daadwerkelijk leesbaarder is, wordt in deze scriptie de leesbaarheid bepaald doordat gebruikers zelf aangeven in welke mate ze een tekst leesbaar vinden of niet.

1.4.3 Hoe kan je een tekst aanpassen ter vereenvoudiging?

>Lexicaal en syntactisch en vergroten cohesie Een brief is op verschillende manieren aan te passen om de boodschap beter over te brengen. Ten eerste kan het helpen als er infographics in een brief staan die de boodschap van de brief ondersteunt. Daarnaast is het mogelijk om de tekstopbouw aan te passen. Zo kan het doel en de gevolgen voor de lezer eerst helder uitgelegd worden om vervolgens met een onderbouwing daarvoor te komen. Dit raakt al aan het inhoudelijk aanpassen van een tekst. Er kan worden gekozen om bepaalde details weg te laten opdat de lezer het grote geheel beter begrijpt. Daarnaast zijn er op een abstract niveau drie tekstkenmerken die gewijzigd kunnen worden die in hoge mate invloed hebben op het tekstniveau zoals beschreven door Pander Maat en Ditewig in 2017. Dit zijn (1) lexicale aanpassingen, denk aan woordgebruik, (2) syntactische aanpassingen dit gaat

om zinsopbouw en (3) cohesie-ingrepen, dit gaat om tekstverbanden en signaalwoorden. Deze aanpassingen blijken vaak samen te gaan (Pander Maat & Ditewig, 2017, p. 254). Om een lange zin op te knippen in meerdere kleinere zinnen zijn er verwijswaarden nodig naar die andere zinnen. Ook worden werkwoordsvormen dan aangepast. Uit onderzoek van Freebody en Anderson (1983) blijkt ook al dat woordgebruik een hoge voorspeller is van tekstbegrip.

1.4.4 Wat houdt T-Scan/LiNT in?

In veel van de onderzoeken hierboven wordt T-Scan en of LiNT gebruikt. T-Scan is een automatische tekstanalyse ontwikkeld in 2014 door Pander Maat et al. en Maat et al. (2017). Deze tool meet of berekent 472 teksteigenschappen. Dit gebeurt op basis van vooraf gestelde regels, anders dan de tool van Aluisio et al. (2010) is het geen Machine Learning (ML) model. Het Leesbaarheidsinstrument voor Nederlandse Teksten (LiNT) wordt dubbelzinnig gebruikt. Enerzijds bestaat er de LiNT-formule, deze formule berekent op basis van de woordfrequentie, concrete naamwoorden, de afhankelijkheidslengte en de deelzinslengte van een tekst de leesbaarheid van een tekst Maat et al. (2023). Deze formule is duidelijk, het is geen zogeheten blackbox-model waar een tekst ingaat en een getal uitrolt. Dit maakt het voor tekstschrijvers eenvoudig om te achterhalen wat er nog verbeterd moet worden aan hun tekst. De LiNT-formule geeft een score van 0 tot 100 terug. 0 betekent dat het heel leesbaar is, 100 betekent dat een tekst heel moeilijk is om te begrijpen. Deze schaal is onderverdeeld in 4 niveaus. De niveau's zijn:

1. Het streefniveau, LiNT-score 0-34. Dit niveau is alleen voor de 15% laaggeletterden in Nederland is dit niveau te moeilijk om te begrijpen.
2. LiNT-score 34-46. 30% van de volwassen Nederlanders kan dit niveau niet goed aan.
3. LiNT-score 46-60. 55% van de volwassen Nederlanders kan dit niveau niet goed aan.
4. LiNT-score 60-100. 82% van de volwassen Nederlanders kan dit niveau niet goed aan.

(Maat et al., 2023, p. 19)

Daarnaast is er het leesbaarheidsinstrument. Dit instrument is in feite een grafische invulling van de LiNT-score en T-Scan. De maker geeft zelf aan dat dit instrument onder andere gebruikt kan worden voor het bepalen van de complexiteit van een tekst, probleemopsporing binnen een tekst en originele en herschreven teksten vergelijken (Maat et al., 2023, p. 27). In dit onderzoek wordt naar al deze drie aspecten gekeken.

Een verdere uitleg van T-Scan is al gedaan, evenals onderzoek naar de effectiviteit van de LiNT is al gedaan Maat en Gravekamp (2022). Dit laatste onderzoek heeft ook beargumenteerd waarom T-Scan het beste instrument is voor tekstanalyse.

1.4.5 Schatting van taalniveau in woningcorporaties

Om te onderzoeken hoe groot het probleem van moeilijke brieven in de sociale huursector is, moet ook duidelijk zijn in welke mate de huurders hiervan geletterd zijn. 15% van de Nederlanders is dan wel laaggeletterd, maar of deze groep onder- of oververtegenwoordigd is onder de klanten van woningcorporaties maakt uit voor de grootte van het probleem. Hier is geen duidelijk onderzoek naar gedaan. Wel zijn er een aantal onderzoeken gedaan in de afgelopen 2 decennia.

Een factsheet van de gemeente Amsterdam uit 2008 stelt dat slechts 20% van sociale huurders hoger onderwijs heeft gevolgd ³. Uit welke bron dit getal komt is niet duidelijk. Wat hoogopgeleid precies inhoudt en of dit betekent dat 80% laagopgeleid is, is niet duidelijk.

Deze scriptie draagt bij aan het wetenschappelijk discours door te onderzoeken of T-Scan gebruikt kan worden om teksten handmatig te verbeteren en door te onderzoeken of de aanpassingen daadwerkelijk verbeteringen zijn.

³<https://www.fairhuurvoorverhuurders.nl/files/literatuur/de-particuliere-woonsector-en-zijn-bewoners.pdf>

Hoofdstuk 2

Methode

2.1 Brieven

Om te onderzoeken of brieven van woningcorporaties daadwerkelijk leesbaarder kunnen worden gemaakt door gebruik te maken van automatische tekstanalyse zijn er daadwerkelijk brieven nodig. In totaal zijn er voor dit onderzoek in 173 brieven onderzocht van acht verschillende woningcorporaties uit Nederland. Dit zijn brieven die zij op dagelijkse basis versturen vanuit hun systemen. De brieven worden gepersonaliseerd verstuurd naar klanten. Hierom bevatten de brieven variabelen, die door de computer worden ingevuld bij het versturen van de brief. Deze variabelen zijn voor dit onderzoek handmatig ingevuld zie Appendix A.2.

Daarnaast zijn de brieven verder bewerkt volgens de regels in de T-Scan-handleiding (Maat et al., 2017, p. 7-12). Inhoudelijk zijn er een aantal zaken aan de brieven aangepast. Zo moet T-Scan de aanhef, koptitels en afsluitende groet overslaan. T-Scan ziet de titels van vaak maar enkele woorden als een zin. Hoewel dit feitelijk kan kloppen met een titel als "Heeft u nog vragen?", brengt het de gemiddelde zinslengte sterk naar beneden. Zeker omdat in de gebruikte teksten relatief weinig zinnen staan, hebben de aanhef en bijvoorbeeld de afsluitende groet hier veel invloed op. Inhoudelijk geven ze weinig informatie mee en dragen ze niet bij aan het meer of minder leesbaar maken van een tekst. Zo vond Pander Maat en Ditewig (2017) geen aanwijzing dat kopjes tekstbegrip bevorderen, wel de reproduceerbaarheid van een tekst.

Uiteindelijk blijven er 14 teksten over die herschreven worden, deze hadden de slechtste LiNT-scores. Deze teksten scoorden onder andere slecht op de onbekendheid van woorden. De inhoudswoorden van deze teksten, dat wil zeggen de zelfstandige naamwoorden, bijvoeglijke naamwoorden, werkwoorden en bijwoorden, komen slechts 5000 tot 50000 keer op 1 miljard woorden voor. Ter vergelijking, in teksten in gemiddeld roddelnieuwsblad komen de inhoudswoorden bijna 80000 keer voor op 1 miljard woorden. Ook op de

abstractie van de zelfstandig-naamwoorden wordt slecht gescoord ($M=68,3$, $SD=19,6$) . Deze laatste tekst bestaat wel uit slechts 1 zin. Daarnaast werden er redelijk veel nieuwe woorden gebruikt. Dit wordt uitgedrukt in het percentage inhoudswoorden dat niet voorkwam in de 50 woorden daarvoor ($M=91,5$, $SD=4,2$). De teksten scoorden daarentegen wel goed op het voorkomen van mensen in een tekst. Het gaat hier om verwijzingen naar personen, het aantal persoonlijk voornaamwoorden, het aantal aanspreking en in mindere mate het percentage werkwoorden in de actieve vorm. Een uitleg van alle tekstenkenmerken die LiNT teruggeeft aan de gebruiker is te vinden in tabel 2.1. In deze tabel zijn ook referentiewaarden te vinden voor een makkelijkere en een moeilijker tekst.¹

De teksten scoren al redelijk goed op het aantal bijvoeglijke bepalingen per deelzin en het aantal opsommingen per zin. De teksten zijn al heel persoonlijk geschreven, zo scoren de brieven uitstekend op het aantal persoonlijkvoornaamwoorden en het aantal aansprekingen. Op de rest van de kenmerken scoort het gemiddeld of slecht, dat wil zeggen, het maakt de tekst moeilijker.

2.2 Leesbaarheidsinstrument

De tekstkwaliteit van de brieven zal worden onderzocht aan de hand van T-Scan, zoals ook gebruikt in het onderzoek van Pander Maat en van der Geest (2021). De T-Scan registreert welke eigenschappen van de tekst moeilijk te begrijpen zijn. Hieruit kunnen suggesties volgen. Een meer inzichtelijke toepassing van de T-Scan is LiNT zoals eerder beschreven in hoofdstuk 1.4.4.

2.3 Schrijfprotocol

Eerst zal er een protocol worden geschreven over hoe om te gaan met tekstsuggesties. Het protocol wordt geschreven aan de hand van de eerste ervaring met het omschrijven van een brief. Uit deze eerste ervaring zullen de meeste dilemma's opkomen die in bij het herschrijven van volgende brieven relevant zijn. Het protocol geeft een richtlijn aan het herschrijven. Dit protocol zorgt ervoor dat de studie reproduceerbaar is.

Anders dan eerder onderzoek Pander Maat en Ditewig (2017), worden teksten hier herschreven met de handen op de rug. Alleen specifieke tekst wordt aangepast met zo min mogelijk inhoudelijke wijzigingen. Ook wordt de volgorde van zinnen niet aangepast. Daarnaast wordt er geen extra uitleg van processen toegevoegd, ook al kan dat intuïtief wel nodig voelen. Wel

¹Uitleg van grammaticale termen is online raadpleegbaar: <https://lint.hum.uu.nl/uitleg/5-minuten-grammatica>

kan een vakterm worden uitgelegd indien er geen adequaat synoniem voor bestaat.

Het protocol is zogezegd opgesteld tijdens het schrijven van een eerste brief. Zo werd duidelijk welke tekstkenmerken worden aangepast en hoe die reproduceerbaar gedaan kan worden. Na een extra verfijning zijn aan de hand van het volgende protocol de 14 brieven met de laagste LiNT-score herschreven.

1. Analyseren van de resultaten van LiNT
 - (a) Moeilijke woorden
 - (b) Moeilijke zinsconstructies
2. Via `ishetb1.nl`² een mogelijke synoniem zoeken die eenvoudiger is
 - (a) De zin zo omschrijven dat het synoniem er in past.
 - (b) Lijst bijhouden met alle vervangingen.
3. Bij speciale (vak)termen een duidelijke korte uitleg toevoegen.
 - (a) Lijst bijhouden welke dit zijn.
4. Lange zinnen opknippen in kleinere zinnen.
5. Als er bijzinnen zijn, dan moeten deze zoveel mogelijk worden herschreven tot aparte zinnen.
6. Zinnen omschrijven naar een actieve werkwoordsvorm
7. Zinnen zoveel mogelijk omschrijven met verwijzingen naar mensen.
8. Reflectie opschrijven over gemaakt keuzes en dilemma's

De lijsten met gebruikte synoniemen en de uitgelegde vaktermen zijn te vinden in Appendix A.3. Zeker omdat het

2.4 Herschrijven van brieven

Het resultaat van de herschrijvingen zijn te vinden in Tabel A.1. De brieven die 10 punten of meer verbeteren op de LiNT-score zullen worden geselecteerd om vragen over te beantwoorden. Uit elke tekst wordt uiteindelijk een alinea geselecteerd voor in de enquête.

Net als blijkt in het onderzoek (Pander Maat & Ditewig, 2017, p. 254), hebben de herschreven hebben de teksten meer verwijzingswoorden nodig.

²<https://ishetb1.nl>

Opgeknippte zinnen moeten nog wel naar elkaar verwijzen. Voornamelijk bijzinnen moeten verwijzen naar de oorspronkelijk hoofdzin. Na het herschrijven zijn de LiNT-scores flink verbeterd, zie hiervoor de laatste twee kolommen van Tabel A.1. Een uiteenzetting van alle tekstkenmerken die LiNT heeft berekend zijn te vinden in tabel 2.3.

Delen van de herschreven brieven zijn te vinden in Appendix B.

Bij het herschrijven zijn voor de uiteindelijk de volgende wijzigingen gedaan.

>>>Tabel invoegen welk type wijziging hoe vaak voorkomt<<<

2.5 Enquête

2.5.1 Gebruikte teksten

Vervolgens is er een enquête uitgestuurd met daarin 4 vragen over elk van de brieven. Om te voorkomen dat de enquête te lang zou duren is er gekozen om vragen te stellen over slechts 8 van de 14 brieven die oorspronkelijk de slechtste LiNT-score hadden. De 8 brieven hadden de grootste LiNT-score verbetering met minstens 10 punten ($M=16.13$, $SD=8.89$) verschil ten opzichte van de originele teksten. Het grote contrast zou het eenvoudiger moeten maken voor respondenten om verschillen te zien tussen tekstversies en om deze te kunnen benoemen.

Om enquête vermoeidheid te voorkomen was het doel om de enquête niet langer dan 30 minuten te laten duren [bron?]. Hierdoor was het niet mogelijk om alle 8 brieven volledig te tonen. Hierom is er een selectie gemaakt uit deze brieven. Na een korte empirische studie bleek dat de volledige enquête te maken zou zijn binnen 30 minuten als er over maximaal 9 alinea's vragen zouden worden gesteld. Om de diversiteit tussen de teksten te behouden is ervoor gekozen om uit alle 8 brieven 1 alinea te selecteren. De brief genaamd ThuisInLimburg heeft 2 alinea's geleverd aangezien dit een lange brief is die verschillende onderwerpen behandelt. Net als in eerder onderzoek D Clerq et al. (2014) zijn de tekst groottes zo'n 100 tot 200 woorden. De teksten zijn op zichzelf te begrijpen. De specifieke alinea's zijn te vinden in B.

2.5.2 Opzet

De enquête is gemaakt in de Qualtrics Experience Management Platform³. De enquête start met een scherm waarop de respondent welkom wordt geheten, waarop er uitleg komt over de inhoud van de enquête alsook informatie over het gebruik van de data. Om door te gaan met de enquête moet de respondent aangeven of deze vrijwillig meedoet, 18 jaar of ouder is en diens data wil delen voor het onderzoek.

³<https://survey.uu.nl>

Hierna volgt een vraag over het opleidingsniveau van de respondent. Dit kan voor verdere analyse interessant zijn.

Daarna volgen 9 vragenblokken. Deze 9 vragenblokken worden per respondent in een willekeurige volgorde laten zien. Elk blok gaat over 1 alinea tekst. Eerst wordt per blok bepaald of de vragen over de originele tekst of over de bewerkte tekst, de kans hierop is 50%. De enquête-software is zo ingesteld dat het streeft naar een gelijk verdeling vragen over beide versies. De respondent weet niet welke versie van de brief die te lezen krijgt.

De respondent krijgt vervolgens de bepaalde alinea te lezen. Hierna volgen de volgende 3 vragen:

Vraag 1 In hoeverre begrijpt u de brief?

Vraag 2 Vindt u de brief betuttelend?

Vraag 3 Wat zou u aanpassen aan de tekst om hem leesbaarder te maken?
Herschrijf dit gedeelte van de tekst.

De eerste vraag wordt beantwoord door middel van een Likert-schaal Likert (1932). Deze gaat van 'erg goed leesbaar', met waarde 1 naar 'zeer moeilijk leesbaar' met waarde 5. De tweede vraag kan worden beantwoord met 'Ja', 'Nee' en 'Weet ik niet'

Na het invullen van deze eerste 3 vragen volgt er een vierde vraag, de antwoorden op de eerste drie vragen kunnen dan niet meer gewijzigd worden. Deze vraag is als volgt:

Vraag 4 Hier ziet u een aangepaste zin uit de bovenstaande tekst, welke wijzigingen zou u nog toevoegen of er juist uitlaten?

Bovenaan ziet de respondent eerst de versie van de alinea waar die de voorgaande drie vragen over heeft beantwoord. De titel bij deze alinea is "Originele versie". Vervolgens staat onder deze alinea de titel "Bewerkte versie" met daarbij de tegenovergestelde versie van de tekst waar de voorgaande vragen over zijn gegaan. Indien de eerste drie vragen over de daadwerkelijke originele tekst gingen, dan gaat deze vraag over de herschreven versie en vice versa.

Om de respondent te helpen bij het gerichter antwoord geven op deze vraag zijn de verschillende wijzigingen voorzien van verschillende achtergrondkleuren. De kleuren van de wijzigingen in de ene versie corresponderen met de wijzigingen in de bewerkte versie. De alinea's met de kleuren zijn terug te vinden in bijlage B.

De antwoorden op vragen één, twee, drie en vier worden vergeleken met de resultaten vanuit de T-Scan.

2.5.3 Proefpersonen

De enquête is volledig ingevuld door 20 mensen over een periode van twee weken. De enquête is voornamelijk beantwoord door hoger opgeleiden, waarvan kan worden aangenomen dat zij in hoge mate geletterd zijn. Een overzicht van het opleidingsniveau van de participanten is uiteengezet in tabel 2.4. 90% van de participanten is dus hoog opgeleid.

Het hoge opleidingsniveau heeft zowel voordelen als nadelen. Een groot voordeel is dat deze groep zowel een zin kan begrijpen als kan aangeven wat er wel of niet moeilijk of makkelijk is aan deze zin. Daarbij komt het voor dat laag geletterden niet durven aan te geven dat ze een tekst niet begrijpen om zo een moment van schaamte te voorkomen. Het nadeel hiervan is dat de mensen die daadwerkelijk een brief moeten begrijpen niet gevraagd worden naar wat zij vinden. Het zou wel meer representatieve resultaten opleveren. Echter is de tijd voor dit onderzoek maar zeer beperkt. Dit maakt het praktisch onmogelijk om zowel langs de ethische commissie van Universiteit Utrecht te gaan voor goedkeuring van het onderzoek als het aanschrijven van woningcorporaties of hun huurders benaderd mogen worden. Vervolgonderzoek zou zich vooral hierop kunnen toespitsen. Verder is het nog interessant om te onderzoeken of de gebruikte woordfrequenties door T-Scan overeenkomen met de woordenschat voor mensen van wie het Nederlands de tweede taal is.

2.6 Verwerking enquête

1. Hoeveel antwoorden op elke vraagtype

Tabel 2.1: Overzicht variabelen in LiNT met referentiewaarden en -teksttypes

| LiNT type | Uitleg | Makkelijker ^{Type} | Moeilijker ^{Type} |
|-----------------------|--|-----------------------------|----------------------------|
| w__onbekend | Log onbekende woorden frequentie (lager is beter) | 1,1 ¹ | 1,92 ⁵ |
| w__abstract | Percentage abstracte zelfstandig-naamwoorden (lager is beter) | 32,9 ² | 77 ⁵ |
| w__nieuw | Percentage inhoudswoorden die niet voorkwam in de vorige 50 woorden (lager is beter) | 86,9 ³ | 92,9 ⁶ |
| z__dzlen | Lengte van deelzin in woorden (lager is beter) | 7,9 ³ | 11,8 ⁵ |
| z__afhankelijk | Afhankelijkheidslengtes in woorden (lager is beter) | 4,5 ³ | 10,8 ⁵ |
| z__bijvoeglijk | Aantal bijvoeglijke bepaling per deelzin (lager is beter) | 0,88 ¹ | 2,18 ⁵ |
| z__opsomming | Aantal opsomming in doorlopende zinnen (lager is beter) | 0,14 ¹ | 0,49 ⁵ |
| z__bijzin | Aantal bijzinnen per zin (lager is beter) | 0,37 ³ | 1,4 ⁵ |
| z__len | Aantal woorden per zin (lager is beter) | 10,7 ³ | 24,9 ⁵ |
| p__men | Aantal woorden verwijzend naar persoon (per 1000 woorden) (hoger is beter) | 160 ¹ | 41 ⁶ |
| p__pvn | Aantal persoonlijk voornaamwoorden per 1000 (hoger is beter) | 75 ¹ | 21 ⁵ |
| p__aspr | Aantal aansprekingen in 2e persoon per 1000 woorden (hoger is beter) | 27,1 ⁴ | 1,1 ⁵ |
| p__actief | Percentage actieve vormen onder vervoegde werkwoorden (hoger is beter) | 94,94 ¹ | 82,51 ⁵ |

Uitleg types tekst 1) Roddelnieuws/-berichten, 2) Reisblogs, 3) Vmbo-boeken, 4)

Medisch advies, 5) Wetenschap, 6) Partijprogramma's.

Tabel 2.2: Overzicht LiNT resultaten originele teksten met referentiewaarden.

| LiNT type | Gemiddeld | Standard deviatie | Makkelijker | Moeilijker |
|----------------|-----------|-------------------|-------------|------------|
| w__onbekend | 1,84 | 0,24 | 1,1 | 1,92 |
| w__abstract | 68,27 | 19,63 | 32,9 | 77 |
| w__nieuw | 91,54 | 4,17 | 86,9 | 92,9 |
| z__dzlen | 9,72 | 1,33 | 7,9 | 11,8 |
| z__afhankelijk | 6,08 | 2,63 | 4,5 | 10,8 |
| z__bijvoeglijk | 1,09 | 0,43 | 0,88 | 2,18 |
| z__opsomming | 0,19 | 0,19 | 0,14 | 0,49 |
| z__bijzin | 0,73 | 0,6 | 0,37 | 1,4 |
| z__len | 15,29 | 3,73 | 10,7 | 24,9 |
| p__men | 109,06 | 24,54 | 160 | 41 |
| p__pvn | 92,48 | 28,47 | 75 | 21 |
| p__aspr | 68,06 | 22,86 | 27,1 | 1,1 |
| p__actief | 79,95 | 27,19 | 94,94 | 82,51 |

Tabel 2.3: Overzicht LiNT resultaten bewerkte teksten met referentiewaarden.

| LiNT type | Gemiddeld | Standard deviatie | Makkelijker | Moeilijker |
|----------------|-----------|-------------------|-------------|------------|
| w__onbekend | 1,47 | 0,21 | 1,1 | 1,92 |
| w__abstract | 61,85 | 17,01 | 32,9 | 77 |
| w__nieuw | 89,03 | 5,12 | 86,9 | 92,9 |
| z__dzlen | 8,13 | 0,97 | 7,9 | 11,8 |
| z__afhankelijk | 4,45 | 1,17 | 4,5 | 10,8 |
| z__bijvoeglijk | 0,78 | 0,25 | 0,88 | 2,18 |
| z__opsomming | 0,11 | 0,12 | 0,14 | 0,49 |
| z__bijzin | 0,37 | 0,15 | 0,37 | 1,4 |
| z__len | 10,28 | 1,06 | 10,7 | 24,9 |
| p__men | 128,69 | 32,5 | 160 | 41 |
| p__pvn | 114,29 | 35,29 | 75 | 21 |
| p__aspr | 86,78 | 28,62 | 27,1 | 1,1 |
| p__actief | 92,37 | 12,88 | 94,94 | 82,51 |

Tabel 2.4: Overzicht hoogste genoten opleidingsniveau participanten

| Opleidingsniveau | Aantal |
|-------------------------------|--------|
| Middelbaar onderwijs afgerond | 1 |
| Hoger onderwijs, geen diploma | 1 |
| Bachelordiploma | 8 |
| Universitair diploma | 9 |
| Geen antwoord | 1 |

Hoofdstuk 3

Resultaten

"De tekst is qua vorm denk ik eigenlijk best prima. Lekker direct en beknopt. De woordkeuze is alleen fucked up [...]"
-- Participant

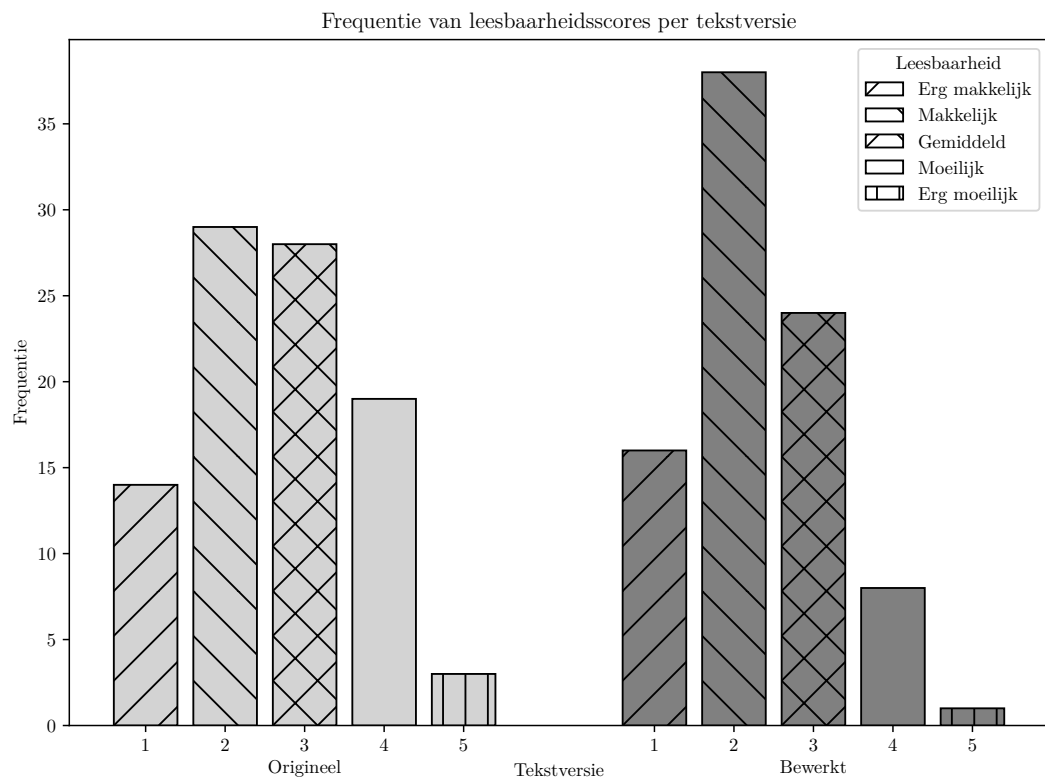
In dit hoofdstuk staan de resultaten van de enquête uiteengezet. De resultaten zullen gebruikt om antwoord te geven op de drie deelvragen. De antwoorden op de vraag in hoeverre een tekst leesbaar is en of een tekst betuttelend zijn in principe ordinaal van aard. Het is immers niet te zeggen of de stap tussen *Erg makkelijk leesbaar* en *Makkelijk leesbaar* even groot is als de stap tussen *Moeilijk leesbaar* en *Erg moeilijk leesbaar*.

In deze resultaten sectie wordt deze data echter afwisselend als ordinale en interval data beschouwd. Dit laatste maakt het mogelijk om een indicatie te geven of er een verbetering is en in hoeverre die bestaat. Dit is voornamelijk om het onderzoek toepasbaar te maken voor de maatschappij. Blijkens eerder onderzoek van de Winter en Dodou (2010) is in de meeste gevallen het verschil tussen beide benaderingen niet heel groot.

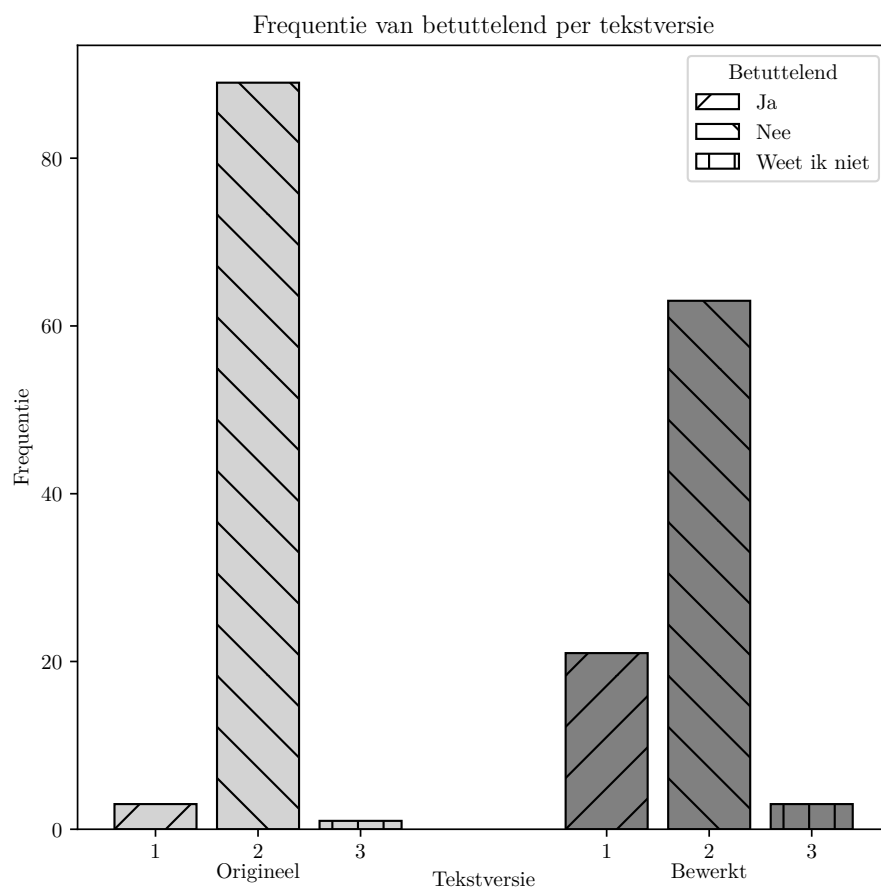
3.1 Begrijpelijkheid en betuttelend

Om te bepalen of er een significant verschil zit tussen de opgegeven leesbaarheidsscores van de originele en bewerkte versie is er een Mann-Whitney U test (Mann en Whitney (1947)) uitgevoerd. Er is een significant verschil tussen de leesbaarheidsscores van beide versies ($p = 0.026$). Deze teksten worden dus als makkelijker leesbaar ervaren. Deze test is ook uitgevoerd op de vraag of de tekst betuttelend is. Ook dit resultaat is significant ($p = 0.001$). Er kan dus worden aangenomen dat de teksten wel als betuttelender worden ervaren in de bewerkte versie. De precieze uiteenzetting van deze data is te vinden in Figuur 3.1 en Figuur 3.2.

Als de data wordt benadert als interval data is ook uit te rekenen in hoeverre de gemiddelden verschillen. Over alle teksten is dat voor de



Figuur 3.1: Frequentie van leesbaarheidsscores per tekstversie.



Figuur 3.2: Frequentie van betuttelend per tekstversie

originele versie 2,66 tegenover 2,31 in de bewerkte versie op een schaal van 1 tot 5. Hoe dichterbij de 1 hoe leesbaarder een tekst wordt ervaren. Een t-test tussen beide populaties geeft een significantie van $p = 0.022$. De variance van de leesbaarheid van originele teksten bedraagt 1.141 tegenover 0.844 in de bewerkte versie. In tabel 3.1 is de leesbaarheidsscore per tekst en tekstversie uiteengezet. In deze tabel is ook te zien hoeveel respondenten een bepaalde tekstversie hebben gezien, wat de gemiddelde leesbaarheidsscore is en de standaard deviatie.

Bij beide benaderingen is te zien dat de bewerkte versie een betere leesbaarheidsscore heeft, oftewel deze teksten worden als eenvoudiger beschouwd. Tegelijkertijd geven de respondenten wel vaker aan dat deze teksten als betuttelend zijn. Sommige respondenten hadden graag gezien dat de toon van de brief zakelijker was.

Tabel 3.1: Overzicht leesbaarheidsscores per vraag en tekstversie

| Tekst | Origineel | | | Bewerkt | | |
|---------|-----------|--------------|------|---------|--------------|------|
| | Aantal | Leesbaarheid | SD | Aantal | Leesbaarheid | SD |
| Tekst 1 | 11 | 2,55 | 1,04 | 9 | 1,78 | 0,97 |
| Tekst 2 | 12 | 3,58 | 1,08 | 8 | 2,12 | 0,64 |
| Tekst 3 | 10 | 2,5 | 1,08 | 10 | 2,8 | 1,14 |
| Tekst 4 | 10 | 2,1 | 0,88 | 10 | 2,4 | 0,84 |
| Tekst 5 | 9 | 2,22 | 1,09 | 11 | 2,09 | 0,83 |
| Tekst 6 | 10 | 2,5 | 1,08 | 10 | 2,8 | 0,79 |
| Tekst 7 | 11 | 2,73 | 1,1 | 9 | 2,11 | 0,93 |
| Tekst 8 | 9 | 3,11 | 0,78 | 11 | 2,55 | 1,04 |
| Tekst 9 | 11 | 2,45 | 0,93 | 9 | 2 | 0,71 |
| Totaal | 93 | 2,66 | 1,07 | 87 | 2,31 | 0,92 |

Leesbaarheid Lagere score is beter.

3.2 Suggesties op T-Scan

De derde vraag van de enquête was *Wat zou u aanpassen aan de tekst om hem leesbaarder te maken? Herschrijf dit gedeelte van de tekst.* Deze vraag was vrij open en hier is niet altijd antwoord op gegeven. Sommigen herschreven de tekst compleet, anderen slechts een zin en weer anderen noemden slechts een woord. Het kon ook voorkomen dat er slechts werd gezegd dat er iets mis was aan de tekst, maar niet precies wat. Om de reacties eenduidig te turven zijn er 4 categorieën suggesties opgesteld. Dit zijn.

Cat. 1 Er is een concrete wijziging genoemd.

- Cat. 2** Er is wel een wijzigingstype genoemd maar niet concreet hoe die wijziging moet zijn.
- Cat. 3** Er wordt aangegeven dat er iets gewijzigd moet worden maar niet wat. (Bijvoorbeeld: “ik snap deze zin niet”)
- Cat. 4** Overig, een participant wil extra informatie toevoegen of delen van de tekst compleet weglaten.

Deze categorisering van de resultaten werkt goed als participanten suggesties gaven aan de originele tekst. Elke suggestie is namelijk direct een suggestie op de tekst en in de meeste gevallen een suggestie op een specifiek wijzigingstype vanuit T-Scan. In de bewerkte versie zijn alle teksteigenschappen al in lijn met T-Scan. Als een participant een suggestie doet op de bewerkte tekst dan betekent dit niet gelijk dat de suggestie vanuit T-Scan geen verbetering brengt. Een participant kan bijvoorbeeld een wijziging voorstellen die ofwel tegen de aanwijzingen van T-Scan zijn, zoals een moeilijker woord of een ander woord die even eenvoudig is.

Om deze reden zijn de suggesties per tekstversies verschillend uiteengezet.

In tabel 3.2 staan de suggesties uiteengezet die participanten gaven die overeenkomen met de suggesties vanuit T-Scan. De tabel laat zien dat participanten maar weinig suggesties hebben gedaan ten opzichte van alle mogelijke suggesties vanuit T-Scan. Voornamelijk het omschrijven van zinnen naar een actieve vorm (28,26%), de zinslengte verkorten (22,41%) en woordabstractie (15,71%) de meest genoteerde suggesties die overeenkomen met T-Scan. Gemiddeld werd slechts 12,92% van alle suggesties gesignaleerd. Suggesties van participanten die niet overeenkomen met de suggesties vanuit T-Scan zijn er ook. Participanten gaven 8 keer aan dat er extra informatie in de brief moest. Bijvoorbeeld de locatie van een gesprek of een verwijzing naar een website. Twee maal is er een grammaticale fout gevonden. Een aantal keer is de tekst volledig omgeschreven, zo werden zinsvolgordes aangepast of totaal andere terminologie gebruikt. Deze suggesties zijn niet opgenomen in tabel 3.2. Dit zijn wijzigingen die T-Scan zelf niet zou kunnen doen, aangezien die functionaliteit daar niet in is ingebouwd. Drie maal is er een eenvoudiger woord gesuggereerd die niet door T-Scan als *onbekend* is gesignaleerd.

De wijzigingen die participanten hebben gemaakt op de bewerkte tekst zijn niet gecategoriseerd. Telkens als er een specifieke suggestie uit T-Scan werd aangepast of naar werd verwezen door een participant werd dit geregistreerd, zonder een categorie te specificeren. Dit om voornoemde redenen. Het overzicht van het aantal gesignaleerde wijzigingstypes is te vinden in tabel 3.3. Buiten deze opmerkingen zijn er nog 40 opmerkingen gemaakt niet niet in de categorieën van T-Scan passen. 17 daarvan bestonden uit het toevoegen van extra informatie, zoals het toevoegen van een plaats en

Tabel 3.2: Overzicht gesignaleerde problemen in de originele teksten per categorie uiteengezet.

| Type (Aantal) | Aantal voorkomens | Categorie | | | | Totaal | |
|-------------------|----------------------|-----------|---|---|---|--------|------------|
| | | 1 | 2 | 3 | 4 | Som | Gezien (%) |
| overig (6) | 61 | 1 | 0 | 0 | 0 | 1 | 1,639 |
| p_actief (4) | 46 | 12 | 1 | 0 | 0 | 13 | 28,261 |
| p_aspr (8) | 82 | 9 | 0 | 1 | 0 | 10 | 12,195 |
| w_abstract (7) | 70 | 9 | 2 | 0 | 0 | 11 | 15,714 |
| w_onbekend (29) | 295 | 31 | 5 | 0 | 0 | 36 | 12,203 |
| z_afhankelijk (1) | 12 | 0 | 0 | 0 | 0 | 0 | 0 |
| z_bijzin (15) | 149 | 17 | 0 | 0 | 0 | 17 | 11,409 |
| z_len (5) | 58 | 8 | 4 | 1 | 0 | 13 | 22,414 |
| z_opsomming (1) | 9 | 0 | 0 | 0 | 0 | 0 | 0 |

datum, of het omgooien van de zinsvolgorde. 4 keer is de toon van de brief aangepast, naar een zakelijker vorm, 3 maal werd de context onvoldoende bevonden om iets over de tekst te zeggen.

Tabel 3.3: Overzicht gesignaleerde problemen in de bewerkte tekst uiteengezet.

| Type (Aantal) | Aantal voorkomens | Gesignaleerd | Gezien (%) |
|-------------------|----------------------|--------------|------------|
| overig (6) | 59 | 1 | 1,695 |
| p_actief (4) | 34 | 10 | 29,412 |
| p_aspr (8) | 78 | 14 | 17,949 |
| w_abstract (7) | 70 | 7 | 10 |
| w_onbekend (29) | 285 | 22 | 7,719 |
| z_afhankelijk (1) | 8 | 0 | 0 |
| z_bijzin (15) | 151 | 17 | 11,258 |
| z_len (5) | 42 | 4 | 9,524 |
| z_opsomming (1) | 11 | 2 | 18,182 |

3.3 Bewerkte teksten behouden of weglaten?

De laatste vraag van de enquête bestond uit het beantwoorden van de vraag *Hier ziet u een aangepaste zin uit de bovenstaande tekst, welke wijzigingen zou u nog toevoegen of er juist uitlaten?*. De kleuren corresponderen met een specifieke wijziging die is gemaakt op aangeven van LiNT. Welke kleuren dit zijn is te vinden in Appendix B. In totaal zijn er 76 wijzigin-

gen gemaakt in de gebruikte teksten. In tabel 3.4 is te zien hoe vaak een type wijziging is voorgekomen, hoe vaak deze is tegengekomen door de respondenten en of dit type wijziging als een positieve of als een negatieve wijziging werd gezien. Als een respondent een kleur (en daarmee een wijziging) niet noemde of als werd aangegeven dat de kleur noch een verbetering noch een verslechtering was werd deze neutraal gecodeerd. In deze tabel geldt $R = pos - neg$. Neutrale antwoorden zijn hier niet in meegenomen. Het resultaat, oftewel suggesties voor de originele en bewerkte is berekend voor de bewerkte versie. Dat wil zeggen: in hoeverre worden de type wijzigingen in de bewerkte versie als positief ervaren. Het aantal negatieve resultaten heeft de formule $Neg_{resultaat} = Pos_{origineel} + Neg_{bewerkt}$. Een positieve reactie op de onbewerkte tekst betekent dus een negatieve reactie op de bewerkte tekst. Het aantal positieve resultaten heeft de gelijkaardige formule: $Pos_{resultaat} = Neg_{origineel} + Pos_{bewerkt}$.

Het slechtst scoren wijzigingen die te maken hebben met p_aspr , oftewel hoe lezers worden aangesproken. Als in de originele tekst bijvoorbeeld het volgende zinsdeel gebruikt wordt *U kan de woning bekijken*, betekent een p_aspr dat daar *U kan uw woning bekijken* van wordt gemaakt. De rest van de wijzigingstypes werden als positief ervaren, gemiddeld worden deze wijzigingen in een kwart van de gevallen uiteindelijk als een verbetering gezien.

Tabel 3.4: Beoordeling wijzigingstypes voor de originele tekst, de bewerkte test en het resultaat daarvan.

| Type | Origineel | | Bewerkt | | Resultaat | | |
|-------------------|-----------|--------------|---------|--------------|-----------|---------------|-------|
| | Aantal | R (neg,pos) | Aantal | R (neg,pos) | Aantal | R (neg,pos) | R (%) |
| overig (6) | 59 | -2 (21,19) | 61 | 36 (11,47) | 120 | 38 (30,68) | 31,7 |
| p_actief (4) | 34 | 6 (14,20) | 46 | 21 (10,31) | 80 | 15 (30,45) | 18,8 |
| p_aspr (8) | 78 | 22 (24,46) | 82 | 19 (28,47) | 160 | -3 (74,71) | -1,9 |
| w_abstract (7) | 70 | -2 (34,32) | 70 | 28 (18,46) | 140 | 30 (50,80) | 21,4 |
| w_onbekend (29) | 285 | -23 (122,99) | 295 | 133 (66,199) | 580 | 156 (165,321) | 26,9 |
| z_afhankelijk (1) | 8 | -4 (6,2) | 12 | 9 (1,10) | 20 | 13 (3,16) | 65,0 |
| z_bijzin (15) | 151 | -10 (67,57) | 149 | 82 (26,108) | 300 | 92 (83,175) | 30,7 |
| z_len (5) | 42 | -4 (14,10) | 58 | 29 (6,35) | 100 | 33 (16,49) | 33,0 |
| z_opsomming (1) | 11 | -1 (4,3) | 9 | 0 (3,3) | 20 | 1 (6,7) | 5,0 |

Omdat er weinig datapunten zijn is het moeilijk om de scores te vergelijken. Zo is $z_afhankelijk$ als positief ervaren in 65% van de gevallen, maar ging hier ook maar 1 tekst over.

Hoofdstuk 4

Conclusie

In deze scriptie is onderzoek gedaan naar of T-Scan een bruikbaar instrument is om brieven van woningcorporaties te vereenvoudigen. Door middel van literatuuronderzoek zoals beschreven in hoofdstuk 1.4 en, het toepassen van de suggesties vanuit T-Scan en de daaruit volgende enquête kan antwoord worden gegeven op de deelvragen.

Een probleem met de volgende conclusie is dat de populatie van de proefpersonen zeer klein is (20) en niet representatief voor mensen in woningcorporaties. 90% van de participanten van dit onderzoek is immers hoog opgeleid tegenover de 20% van de sociale huurders zoals genoemd in hoofdstuk 1.4.5.

Krijgt de bewerkte tekst een betere leesbaarheidsscore?

Allereerst is met zekerheid te zeggen dat de bewerkte teksten een betere leesbaarheidsscore krijgen. Gemiddeld scoren de teksten 0,35 punten beter op een schaal van 1-5, met een 2,31. De teksten zitten tussen *makkelijk leesbaar* (2) en *gemiddeld leesbaar* (3) in. Door zowel de *p-waarde* te bepalen met een t-test als een Mann-Whitney-U test, blijkt het verschil significant ($p < 0.05$). Wel werden de teksten als betuttelender ervaren. Participanten vonden dat een tekst vanuit een woningcorporatie zakelijk van aard moest zijn. Dit strookt dan weer niet met de conclusies van Maat en Gravekamp (2022).

Signaleren mensen dezelfde, of andere problemen als de T-Scan?

Of mensen dezelfde problemen signaleren als T-Scan is moeilijk te zeggen. Voornamelijk omdat teksten niet helemaal zijn overgeschreven. Gemiddeld werd zo'n zes procent van de problemen die T-Scan aangaf ook gerapporteerd door de participanten. Het is niet te zeggen of participanten wel de problemen hebben gezien maar niet de moeite hebben genomen om het aan te geven, of dat ze het niet als probleem zagen. Vervolgonderzoek zou deze

vraag in de enquête zich meer moeten toespitsen op het herschrijven van 1 zin. Zo blijft het voor participanten haalbaar om een concreet antwoord te geven en blijft het niet onduidelijk of een zin simpelweg is overgeslagen of dat er geen problemen zijn gesignaleerd in de betreffende zin.

Hebben mensen voor de makkelijkere versie nog suggesties die de T-Scan niet heeft?

Op basis van de resultaten uit hoofdstuk 3.3 kan worden geconcludeerd dat de herschrijvingen op basis van de gesignaleerde problemen vanuit T-Scan goed werken. Voornamelijk een zin met bijzinnen herschrijven naar opzichzelfstaande zinnen en de zinslengte verkorten werden goed ontvangen door de participanten. Dit ligt in lijn met eerder onderzoek (Maat et al. (2023)). Ook het vereenvoudigen van het vocabulaire door minder onbekende en minder abstracte woorden te gebruiken worden positief ontvangen. Ook deze conclusies worden ondersteund door eerder onderzoek (Freebody en Anderson (1983)). Het persoonlijker maken van de brieven, ofwel de aanspreekvormen persoonlijker maken wordt als enige wijziging als negatief ervaren door de participanten. Dit ligt in lijn met het feit dat de teksten als meer betuttelend worden ervaren.

Is het opvolgen van aanwijzingen uit T-Scan een goede manier om teksten te verbeteren in de context van woningcorporaties?

Al met al is het opvolgen van aanwijzingen uit T-Scan een redelijk goede manier om teksten te verbeteren in de context van woningcorporaties. Het opvolgen van wijzigingen maakt de tekst immers makkelijker leesbaar en de problemen die T-Scan aanstipt worden ook deels gesignaleerd door participanten zelf. De suggesties die participanten geven die T-Scan niet geeft gaan voornamelijk over de (onduidelijke) inhoud. Daar kan T-Scan echter niets aan doen.

Een implementatie van T-Scan in de vorm van LiNT zou dus kunnen worden gebruikt door woningcorporaties zelf, wel moet er dan ook goed worden gekeken naar de inhoud van de brieven.

Hoofdstuk 5

Discussie

In dit hoofdstuk staan een aantal beperkingen van dit onderzoek uiteengezet evenals een aantal aanbevelingen voor vervolgonderzoek.

5.1 Beperkingen

Zoals benoemd in de conclusie is de populatie proefpersonen zeer klein met maar 20 respondenten. Daarbij is deze niet representatief voor personen die in sociale huurwoningen zitten. Daar is slechts 20% hoogopgeleid terwijl dat in de populatie proefpersonen tenminste 80% is. Dit onderzoek kan dus niet zoveel zeggen over hoe de eigenlijke doelgroep van de gebruikte brieven de leesbaarheid van de brieven ervaart. Indien dit onderzoek wordt uitgezet binnen meerde woningcorporaties verspreid over verschillende regio's kan er meer gewicht aan de resultaten worden gehangen. Dan kan het ook duidelijker worden of het het waard is om het instrument in te zetten bij woningcorporaties zelf.

Een ander probleem die opkwam bij de respondenten was dat ze niet wisten of ze vragen moesten beantwoorden als zichzelf of als laaggeletterden. Soms begrepen ze de tekst of een woord wel, maar gingen ze er vanuit dat een laaggeletterde dit niet zou hebben. Dit kan de resultaten hebben beïnvloed. Het is niet te achterhalen welke data dit is en in hoeverre dat de algehele resultaten hebben beïnvloed. Indien de enquête nogmaals wordt uitgezet dan is het goed om te benadrukken dat iedereen de vragen moet beantwoorden voor diens eigen leesniveau.

Een veelgemaakte opmerking bij vraag 3 van de enquête was dat de context ontbrak om goed geïnformeerde opmerkingen over de tekst te maken. Men wist bijvoorbeeld niet in welke situaties een bepaalde brief gestuurd zou worden of of er in eerdere brieven al bepaalde zaken zouden zijn aangehaald. Denk aan een betalingsherinnering. De toon van zo'n brief kan bepaald worden door de omvang van een probleem. Indien de omvang niet duidelijk is, kan de toon als misplaatst worden gekenmerkt. Het toevoegen

een inleiding met context per alinea zou dit probleem verhelpen. Net als in het onderzoek van [...] invoegen. Niet meer zeker welke...].

Wat het herschrijven van de brieven betreft kan er ook het een en ander worden gewijzigd. Zo is de schrijver van deze scriptie geen taalexpert die alles van de Nederlandse taal af weet. De schrijver kan hierin onnodige fouten hebben gemaakt. Net als in het onderzoek van D Clerq et al. (2014) kan er voor het herschrijven gebruik gemaakt worden van taalexperts. Aan de andere kant zullen corporatiemedewerkers ook met dit instrument moeten kunnen werken, deze mensen zijn ook niet allen taalexperts. Daarbij kunnen LLM's worden ingezet om teksten te herschrijven, zoals ook is gedaan in het onderzoek van van Raaij et al. (2024). Wellicht kan als onderdeel van de prompt de waardes uit T-Scan worden meegegeven om tot een specialistischer oordeel te komen.

5.2 Vervolgonderzoek

Het was een goed idee om de teksten niet inhoudelijk te wijzigen. Hierdoor konden de resultaten duidelijker worden geïnterpreteerd. Bij het beantwoorden van de vragen van de enquête konden participanten alleen iets zeggen over de wijzigingen vanuit T-Scan en niet zozeer over andere aspecten van de tekst. Het is daarom goed om dit te behouden in vervolgonderzoek.

In vervolgonderzoek is het ook handig om vraag 4 van de enquête te veranderen opdat deze meer in lijn ligt met de aanwijzingen in Cox en Cox (2008). In dit onderzoek werd er naar 2 dingen gevraagd: *welke wijzigingen zou u nog toevoegen of er juist uitlaten*. Sommige participanten beantwoordden of alleen de eerste of alleen de tweede vraag. Zo werd er soms alleen kleuren genoemd, zonder uit te leggen of dit een moest worden weggelaten of juist moest blijven. Uit het antwoord op vraag 3 kon dit vaak wel bepaald worden. Deze vraagstelling zorgde daarnaast voor 2 bijkomende problemen. Ten eerste duurde het lang voor participanten om alle kleuren op te schrijven en toe te voegen of ze deze wilden behouden of niet. Dit kan worden verholpen door hier een gesloten vraag te maken. Er komt dan in plaats van een leeg tekstveld een matrix tabel met kolommen *Kleur*, *Weglaten* en *Behouden*. Per kleur kan dan eenvoudig de voorkeur worden aangegeven. Dit vereenvoudigt daarbij de verwerking van de resultaten. Een tweede probleem is dat het voor kleurenblinden een zeer uitdagende opgave is. Zij konden kleuren ofwel niet goed onderscheiden of gaven afwijkende kleuren op in het tekstveld. Via een nummeringsstelsel per wijziging zou dit verholpen kunnen worden.

Indien dit onderzoek wordt uitgezet onder een grotere groep zal ook automatisch de enquête korter kunnen worden gemaakt. Na een rondgang onder een aantal participanten bleek dat de tijd die nodig was om de enquête in te vullen verschilde van een half uur tot 2,5 uur. De lengte van de enquête

werkt vermoeidheid in de hand evenals de wil om serieus naar een tekst te kijken. In dit onderzoek was de enigszins lange lengte gerechtvaardigd om genoeg datapunten te verzamelen met het oog op een beperkt aantal respondenten. Met minder respondenten zijn er minder vragen nodig om op een hoop datapunten te komen.

Een andere manier om te onderzoeken of het tekstbegrip wordt verhoogd is niet door middel van een enquête maar door een A-B test te doen en te onderzoeken welke tekstversie meer of minder problemen veroorzaakt voor een woningcorporatie. Een groep zou een jaar lang de originele brieven kunnen blijven ontvangen terwijl een andere groep een verbeterde versie van dezelfde brief krijgt. Aan het eind van het jaar kunnen dan verschillende indicatoren worden vergeleken zoals: achterstallig huur, aantal personen op signaleringslijsten, vragen bij het klant contact centrum, klachten etc. Indien de groep die een jaar lang de verbeterde versie van een brief heeft ontvangen bij deze indicatoren beter scoort dan de groep die de originele brief ontvangt, dan kan daaruit worden geconcludeerd dat de bewerkte versie beter werkt en leesbaarder is.

In dit onderzoek is er weinig rekening gehouden met het feit of mensen die gebruik maken van sociale huurwoningen bekend zijn met het jargon dat door woningcorporaties gebruikt wordt. T-Scan gaf aan dat bepaalde woorden onbekend (*w_onbekend*) waren. Zoals gezegd wordt de bekendheid van woorden bepaald op basis van het SubLeX corpus: Nederlandstalige ondertiteling van engelse televisie programma's. Het is aannemelijk dat er in dergelijke programma's weinig wordt gesproken over zaken als sociale huurwoningen, hierdoor kan de woordfrequentie van zulk jargon onterecht laag zijn ingeschat. Het is waardevol om te onderzoeken of mensen in woningcorporaties zulk jargon daadwerkelijk als *moeilijk* beschouwen of dat dit slechts een beperking van T-Scan is.

Daarbij is het interessant om de leesbaarheid en het tekstbegrip te onderzoeken van de contracten die woningcorporaties opstellen. In hoeverre begrijpen de ondertekenaars van dergelijke contracten daadwerkelijk waar ze voor tekenen, wat hun rechten en wat hun plichten zijn? Wellicht kunnen de juridische teksten vereenvoudigd worden opgeschreven of worden toegelicht in een begrijpelijke taal. In artikel 33 van Boek 3 van het Burgerlijk Wetboek staat: "Een rechtshandeling vereist een op een rechtsgevolg gerichte wil die zich door een verklaring heeft geopenbaard." De *verklaring* van de *wil* is het ondertekenen van het contract, maar als een huurder niet weet wat deze precies verklaart met het ondertekenen is het maar de vraag of die overeenkomt met diens *wil*. In hoeverre dit juridisch problematisch is op dit moment is onduidelijk.

Appendices

Bijlage A

Brieven

A.1 Gebruikte brieven overzicht

Dit zijn de gebruikte brieven:

A.2 Overzicht variabelen

>

A.3 Synoniemen en vaktermen

>

| Id | Tekst | complexity_level | complexity_level_100 |
|-----------|----------------------|-------------------------|-----------------------------|
| 0 | O_Woonzorg_x00 | 3 | 47 |
| 1 | N_Woonzorg_x00 | 2 | 39 |
| 2 | O_Acantus_x12 | 3 | 52 |
| 3 | N_Acantus_x12 | 2 | 38 |
| 4 | O_ThuisInLimburg_x34 | 3 | 51 |
| 5 | N_ThuisInLimburg_x34 | 2 | 43 |
| 6 | O_FrieslandHuurt_x04 | 3 | 51 |
| 7 | N_FrieslandHuurt_x04 | 2 | 38 |
| 8 | O_Woonzorg_x01 | 4 | 70 |
| 9 | N_Woonzorg_x01 | 1 | 31 |
| 10 | O_ThuisInLimburg_x26 | 3 | 50 |
| 11 | N_ThuisInLimburg_x26 | 2 | 37 |
| 12 | O_DeWoningZoeker_x11 | 3 | 49 |
| 13 | N_DeWoningZoeker_x11 | 2 | 39 |
| 14 | O_ThuisInLimburg_x03 | 3 | 50 |
| 15 | N_ThuisInLimburg_x03 | 2 | 37 |
| 16 | O_ThuisInLimburg_x22 | 3 | 49 |
| 17 | N_ThuisInLimburg_x22 | 1 | 32 |
| 18 | O_DeWoningZoeker_x02 | 3 | 49 |
| 19 | N_DeWoningZoeker_x02 | 2 | 39 |
| 20 | O_ThuisInLimburg_x38 | 3 | 46 |
| 21 | N_ThuisInLimbrug_x38 | 2 | 37 |
| 22 | O_DeWoningZoeker_x10 | 3 | 48 |
| 23 | N_DeWoningZoeker_x10 | 2 | 43 |
| 24 | O_FrieslandHuurt_x08 | 3 | 46 |
| 25 | N_FrieslandHuurt_x08 | 2 | 40 |
| 26 | O_DeWoningZoeker_x23 | 3 | 46 |
| 27 | N_DeWoningZoeker_x23 | 2 | 42 |

Tabel A.1: Proces van omschrijven

| Name_id | Voor bewerking | | Originele tekst | | Bewerkte tekst | |
|---------------------|------------------|----------------------|------------------|----------------------|------------------|----------------------|
| | complexity level | complexity level 100 | complexity level | complexity level 100 | complexity level | complexity level 100 |
| Acantus_x00 | 3 | 46 | 2 | 41 | | |
| Acantus_x05 | 3 | 47 | 2 | 43 | | |
| Acantus_x07 | 3 | 46 | 1 | 29 | | |
| Acantus_x12 | 3 | 52 | 3 | 52 | 2 | 38 |
| Acantus_x18 | 3 | 48 | 2 | 45 | | |
| DeWoningZoeker_x02 | 3 | 49 | 3 | 49 | 2 | 39 |
| DeWoningZoeker_x10 | 3 | 47 | 3 | 48 | 2 | 43 |
| DeWoningZoeker_x11 | 3 | 50 | 3 | 49 | 2 | 39 |
| DeWoningZoeker_x23 | 3 | 46 | 3 | 46 | 2 | 42 |
| FrieslandHuur_t_x04 | 3 | 52 | 3 | 51 | 2 | 38 |
| FrieslandHuur_t_x08 | 3 | 46 | 3 | 46 | 2 | 40 |
| ThuisInLimburg_x03 | 3 | 53 | 3 | 49 | | |
| ThuisInLimburg_x04 | 3 | 49 | 2 | 50 | 2 | 37 |
| ThuisInLimburg_x22 | 3 | 49 | 3 | 49 | 1 | 32 |
| ThuisInLimburg_x26 | 3 | 49 | 3 | 50 | 2 | 37 |
| ThuisInLimburg_x34 | 3 | 51 | 3 | 51 | 2 | 43 |
| ThuisInLimburg_x38 | 3 | 48 | 3 | 47 | | |
| ThuisKompas_x05 | 3 | 46 | 2 | 46 | 2 | 37 |
| ThuisKompas_x08 | 3 | 47 | 2 | 42 | | |
| Valburg_x03 | 3 | 46 | 2 | 45 | | |
| Woonzorg_x00 | 3 | 47 | 3 | 47 | 2 | 39 |
| Woonzorg_x01 | 4 | 70 | 4 | 70 | 1 | 31 |

Bijlage B

Gebruikte alinea's overzicht

B.1 Tekst 1

Origineel

Beste,

Hierbij herinneren wij u aan de documenten die moeten worden geüpload om de uiteindelijke aanbieding te kunnen accepteren.

Met vriendelijke groet,
Dirk

Herschreven

Beste,

U heeft een aanbieding ontvangen. U kan de aanbieding pas accepteren als wij een aantal documenten van u hebben gekregen. Ik wil u eraan herinneren om deze te uploaden op de website.

Met vriendelijke groet,
Dirk

B.2 Tekst 2

Origineel

Waarom dit bericht?

CORPORATIE is voor de verhuur van woningen aangesloten bij Stichting WONINGCORPORATIES. Indien het ter voorkoming van ongewenst huurdersgedrag nodig is om bepaalde personen onder de aandacht te houden, kan deze vanuit de woningcorporatie worden aangemeld op een signaleringslijst. Nadat een persoon geplaatst is op deze lijst, wordt de woningcorporatie vanuit WONINGCORPORATIES op de hoogte gehouden wanneer deze persoon zich laat inschrijven als woningzoekende. Plaatsing op de signaleringslijst is voor maximaal 5 jaar.

Meer informatie over de signaleringslijst of mogelijkheden tot bezwaar kunt u vinden op de website. Ook kunt u telefonisch contact opnemen met Dirk.

Herschreven

Waarom dit bericht?

WONINGCORPORATIE is aangesloten bij de Stichting WONINGCORPORATIES. Om te zorgen dat iedereen fijn kan wonen, houden we ons allemaal aan de regels. Mensen die dit niet doen, komen op de signaleringslijst. De signaleringslijst delen we met andere woningcorporaties. Als u zich inschrijft bij een andere woningcorporatie van Stichting WONINGCORPORATIES weten zij dus van uw problemen als huurder. U staat maximaal 5 jaar op deze lijst.

Op de website kunt u meer informatie vinden over de signaleringslijst. Wilt u hier niet opstaan? Dan kan u ons vragen om hier opnieuw naar te kijken. U kunt hiervoor telefonisch contact opnemen met Dirk.

B.3 Tekst 3

Origineel

Uw oorspronkelijke inschrijving met inschrijfnummer 1234567 is opgeheven. Met deze inschrijving kunt u niet meer inloggen. Voor het inloggen op www.website.nl gebruikt u vanaf nu de gegevens van de samengevoegde inschrijving. Daarover ontvangt u apart een bericht van ons.

Herschreven

Uw persoonlijke inschrijving met het inschrijfnummer 1234567 is stopgezet. U kan niet meer inloggen met het account dat daarbij hoort. U kan nu inloggen op www.website.nl met de gegevens van uw gecombineerde inschrijving. U krijgt hierover van ons een apart bericht.

B.4 Tekst 4

Origineel

Beste meneer Jansen,

U heeft uw inschrijving met inschrijfnummer 12345678 gesplitst. Dit betekent dat u en uw medeaanvrager nu apart staan ingeschreven. U behoudt uw inschrijfnummer en inschrijfduur.

Herschreven

Beste meneer Jansen,

U heeft uw inschrijving met inschrijfnummer 12345678 gesplitst. U zoekt samen met iemand een woning. Jullie staan nu apart ingeschreven. U behoudt uw inschrijfnummer en hoelang u al ingeschreven staat.

B.5 Tekst 5

Origineel

Graag nodig ik u uit voor een oriënterend gesprek op 12 april 2025 om 12.00 uur met Dirk in ons kantoor aan het Wim Duisenbergplantsoen 41 in Maastricht.

Kunt u de volgende zaken meenemen naar het gesprek?

- verklaring van goed huurdersgedrag van uw huidige verhuurder,
- een inkomensverklaring van de Belastingdienst. Een inkomensverklaring vraagt u eenvoudig en gratis aan via de Belastingtelefoon, T 012345678. Houdt uw BSN-nummer bij de hand. De Belastingdienst stuurt de verklaring vervolgens binnen vijf werkdagen toe.
- jaaropgave van het afgelopen jaar,
- geldig legitimatiebewijs (paspoort/ identiteitsbewijs) en kopieën van de legitimatiebewijzen van alle meeverhuizende leden van uw huishouden,
- een kopie BRP uitdraai (Basisregistratie Personen) met hierin de adressenhistorie waar u heeft gewoond. Dit kunt u afhalen bij de gemeente.

Indien van toepassing dienen de gevraagde gegevens levens van de meeverhuizende partner te worden meegenomen.

Herschreven

Ik nodig u graag uit voor een gesprek op 12 april 2025 om 12.00 uur met Dirk. Het gesprek vindt plaats op ons kantoor aan het Wim Duisenbergplantsoen 41 in Maastricht.

Kunt u de volgende dingen meenemen naar het gesprek?

1. Een verklaring van goed huurdersgedrag. U vraagt dit aan via de verhuurder waar u nu een woning huurt.
2. Een inkomensverklaring van de Belastingdienst. Dit is een overzicht van uw inkomen. Een inkomensverklaring vraagt u makkelijk en gratis aan via de Belastingtelefoon (T 012345678). Houdt bij het bellen uw BSN-nummer bij de hand. De Belastingdienst stuurt de verklaring binnen vijf werkdagen toe.
3. Een jaaropgave van het afgelopen jaar.
4. Een geldig legitimatiebewijs (paspoort/ identiteitsbewijs).
5. Kopieën van de legitimatiebewijzen van alle meeverhuizende mensen van uw huishouden.
6. Een kopie van een BRP (Basisregistratie Personen) uitdraai. Op deze BRP-uitdraai staan alle adressen waar u heeft gewoond. Deze kunt u aanvragen bij de gemeente.

Als uw partner met u meeverhuist, dan moet uw partner ook alle documenten die hierboven staan meenemen.

B.6 Tekst 6

Origineel

Wij kunnen u de **bovengenoemde** woonruimte aanbieden **er van uitgaande dat uw inschrijving bij CORPORATIE correct is** en dat u de **Algemene Huurvoorwaarden** **die horen bij de af te sluiten huurovereenkomst**, **accepteert**.

Uiteraard is er gelegenheid om de woning **te bezichtigen** voordat u beslist of u naar deze woning wilt verhuizen. **Daarvoor kunt u contact** opnemen met de huidige bewoner. **Mocht dit onverhoopt problemen opleveren**, **laat ons dat dan even weten** zodat wij samen met u een oplossing kunnen bedenken.

Als u het op prijs stelt om samen met onze opzichter de woning te bezichtigen, **verzoeken wij u contact op te nemen met Dirk van onze woningstichting (06-12345678) voor het maken van een afspraak**.

Daar kunt u ook terecht voor verdere informatie over de **aangeboden** woning.

Herschreven

Wij kunnen u de **deze** woonruimte aanbieden. **Wij gaan er dan van uit dat uw inschrijving bij CORPORATIE goed is**. **Wij gaan er ook van uit dat u de Algemene Huurvoorwaarden van deze woning accepteert**.

U kan de woning bekijken voordat u beslist of u naar deze woning wilt verhuizen. **U moet dan contact** opnemen met de huidige bewoner. **Als dit niet lukt dan moet u contact met ons opnemen. Dan bedenken we samen een oplossing.**

Wilt u samen met onze opzichter de woning bezichtigen? Maak dan een afspraak met Dirk van onze woningstichting (06-12345678).

Daar kunt u ook terecht voor verder informatie over de woning.

B.7 Tekst 7

Origineel

Bijgevoegd treft u een antwoordformulier aan. Daarop kunt u aangeven of u de aangeboden woning wel of niet **wenst te accepteren**.

Wij verzoeken u het antwoordformulier **in te vullen** en te **retourneren** aan WONINGCORPORATIE voor 18 april 2025.

Als u de woning accepteert, ontvangt u binnenkort van ons een uitnodiging om het huurcontract te ondertekenen.

Over de **definitieve ingangsdatum** van de huurovereenkomst ontvangt u nog nader bericht.

Voor de volledigheid wijzen wij er nog op dat dit woningaanbod géén overeenkomst is.

Herschreven

In de bijlage vindt u een antwoordformulier. Op dit formulier geeft u aan of u de woning wil accepteren of niet.

Vul eerst het formulier in. Stuur het daarna **terug** naar WONINGCORPORATIE voor 18 april 2025.

Accepteert u de woning? Dan ontvangt u zo snel mogelijk een uitnodiging om het huurcontract te ondertekenen.

Over de **uiteindelijke startdatum** van de huurovereenkomst ontvangt u nog nader bericht.

Let op: dit woningaanbod is géén overeenkomst.

B.8 Tekst 8

Origineel

De genoemde **servicekosten** zijn inclusief Serviceabonnement. U betaalt voor het lidmaatschap van het Serviceabonnement **maandelijks**, **gelijk met uw huur**, **een bedrag van €5,24**. Op het acceptatieformulier kunt u aangeven of u wel of geen lid wilt worden. **Wordt u geen lid** dan **komen** deze maandelijks kosten **te vervallen**.

Herschreven

De genoemde **servicekosten** zijn inclusief Serviceabonnement. U betaalt voor het lidmaatschap van het Serviceabonnement. **Dit kost per maand €5,24. U betaalt dit tegelijk met uw huur**. Op het acceptatieformulier kunt u aangeven of u wel of geen lid wilt worden.

Wordt u geen lid? Dan hoeft u deze maandelijks kosten **niet te betalen**.

B.9 Tekst 9

Origineel

Bevestiging gemaakte afspraken

Het huurcontract tekent u bij WONINGCORPORATIE op 1 februari 2025 om 12.00 uur. **De eerste verhuurnota** moet u bij ons pinnen. U krijgt dan ook de sleutels **voor de woning**. U ontvangt **tijdig** van ons een mail met **de eerste verhuurnota**.

Herschreven

Bevestiging gemaakte afspraken

Het huurcontract tekent u bij WONINGCORPORATIE op 1 februari 2025 om 12.00 uur.

De huur en de servicekosten van de eerste maand moet u bij ons pinnen. U krijgt dan ook de sleutels **voor uw woning**. U ontvangt **op tijd** van ons een mail met daarin een **brief met de huur en de servicekosten**. **Deze brief heet de verhuurnota**.

Bibliografie

- Aluisio, S., Specia, L., Gasperin, C. & Scarton, C. (2010). Readability assessment for text simplification. In *Proceedings of the naacl hlt 2010 fifth workshop on innovative use of nlp for building educational applications* (pp. 1--9).
- Cox, J. & Cox, K. B. (2008). *Your opinion, please!: How to build the best questionnaires in the field of education*. Corwin Press.
- Crossley, S. A., Skalicky, S. & Dascalu, M. (2019). Moving beyond classic readability formulas: new methods and new models. *Journal of Research in Reading*, 42(3-4), 541-561. Verkregen van <https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9817.12283> doi: <https://doi.org/10.1111/1467-9817.12283>
- D Clerq, O., HOSTE, V., DESMET, B., VAN OOSTEN, P., DE COCK, M. & MACKEN, L. (2014). Using the crowd for readability prediction. *Natural Language Engineering*, 20(3), 293-325. doi: 10.1017/S1351324912000344
- de Winter, J. F. & Dodou, D. (2010, 1). Five-point likert items: t test versus mann-whitney-wilcoxon (addendum added october 2012). *Practical Assessment, Research, and Evaluation*, 15. Verkregen van <https://openpublishing.library.umass.edu/pare/article/id/1558/> doi: 10.7275/bj1p-ts64
- Freebody, P. & Anderson, R. C. (1983). Effects on text comprehension of differing proportions and locations of difficult vocabulary. *Journal of Reading Behavior*, 15(3), 19--39.
- Kleijn, S. (2018). Clozing in on readability. *How linguistic features affect*.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Maat, H. P. & Gravekamp, J. (2022). Kan een tekst te simpel zijn? hoe lager en hoger opgeleiden oordelen over eenvoudige taal [Journal Article]. *Tijdschrift voor Taalbeheersing*, 44(2), 62--90. Verkregen van <https://www.aup-online.com/content/journals/10.5117/TVT2022.2.002.MAAT> doi: <https://doi.org/10.5117/TVT2022.2.002.MAAT>
- Maat, H. P., Kleijn, S. & Frissen, S. (2023). Lint: een leesbaarheidsformule en een leesbaarheidsinstrument. *Tijdschrift voor Taalbeheersing*, 45(1),

2--39.

- Maat, H. P., Kraf, R., Dekker, N., van der Sloot, K., van der Klis, M., van den Bosch, A., ... Kleijn, S. (2017). *Handleiding t-scan*. Utrecht, Nijmegen, Tilburg.
- Mann, H. B. & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50--60.
- Pander Maat, H. & Ditewig, S. (2017). Hoe worden onderwijsteksten vereenvoudigd, en helpt dat? *Tijdschrift voor taalbeheersing*, 39(2), 245--263.
- Pander Maat, H., Kraf, R., Van den Bosch, A., Dekker, N., Gompel, M. v., Kleijn, S. d., ... Sloot, K. v. d. (2014). T-scan: a new tool for analyzing dutch text.
- Pander Maat, H. & van der Geest, T. (2021). Monitor begrijpelijke overheidsteksten.
- Scarton, C., Oliveira, M., Candido Jr, A., Gasperin, C. & Aluísio, S. (2010). Simplifica: a tool for authoring simplified texts in brazilian portuguese guided by readability assessments. In *Proceedings of the naacl hlt 2010 demonstration session* (pp. 41--44).
- van Raaij, N. B., Kolkman, D. & Podoyntsyna, K. (2024). Clearer governmental communication: Text simplification with chatgpt evaluated by quantitative and qualitative research. In *Proceedings of the workshop on determit! evaluating text difficulty in a multilingual context@ lrec--coling 2024* (pp. 152--178).