



Faculteit Bedrijf en Organisatie

Een slimme spraakassistent bij het verlenen van eerste hulp bij ongevallen

Jorgé Reyniers

Scriptie voorgedragen tot het bekomen van de graad van
professionele bachelor in de toegepaste informatica

Promotor:
Jens buysse
Co-promotor:
Ronny Pringels

Instelling: —

Academiejaar: 2018-2019

Tweede examenperiode

Faculteit Bedrijf en Organisatie

Een slimme spraakassistent bij het verlenen van eerste hulp bij ongevallen

Jorgé Reyniers

Scriptie voorgedragen tot het bekomen van de graad van
professionele bachelor in de toegepaste informatica

Promotor:
Jens buysse
Co-promotor:
Ronny Pringels

Instelling: —

Academiejaar: 2018-2019

Tweede examenperiode

Woord vooraf

0.1 Verwijzing naar repository

Alle extra documenten kunnen gevonden worden in de repository van de bachelorproef. De link naar de repository is **<https://github.com/JorgeReyniers/bachelorproef-spraakassistent-in-gesprekscontext>**.

Samenvatting

Inhoudsopgave

0.1	Verwijzing naar repository	3
1	Inleiding	15
1.1	Probleemstelling	16
1.2	Onderzoeksvraag	16
1.3	Onderzoeksdoelstelling	16
1.4	Opzet van deze bachelorproef	16
2	Stand van zaken	17
2.1	Spraakgestuurde technologie	17
2.1.1	Spraakherkenning	17
2.1.2	Spraaksynthese	18
2.2	Spraakassistenten	19
2.2.1	De geschiedenis van spraakassistenten	20

2.2.2	De spraakassistenten van nu	22
2.2.3	Hoe werkt een spraakassistent	24
2.3	Bestaande eerste hulpapplicaties	24
2.3.1	De Vlaamse EHBO-app van het Rode Kruis	24
2.3.2	Andere EHBO-applicaties	25
3	Methodologie	27
3.1	De keuze van de spraakassistenten	27
3.2	Wat wordt er vergeleken	27
3.3	Gebruikte materialen	28
3.4	Het verloop van het onderzoek	28
3.4.1	Deel één: met de deelnemer	28
3.4.2	Deel twee: zonder de deelnemer	29
3.4.3	Ontwikkelen van een applicatie	30
3.5	Orthopedagogisch onderzoek	30
3.5.1	De zoektocht naar een gepaste doelgroep	30
4	Resultaten vergelijkend onderzoek	35
4.1	Vergelijking van de assistenten in spraakkwaliteit	35
4.1.1	Vergelijking van de assistenten per eigenschap	35
4.1.2	Vergelijking van de eigenschappen per assistent	36
4.2	Vergelijking van de assistenten in spraakherkenning	36
4.2.1	Een moeilijk te voeren onderzoek	36

5	De eerste hulp assistent	41
5.1	De functionaliteiten	41
5.2	Persona	41
6	Conclusie	43
A	Onderzoeksvoorstel	45
A.1	Introductie	45
A.2	State-of-the-art	46
A.3	Methodologie	47
A.4	Verwachte resultaten	47
A.5	Verwachte conclusies	48
B	Stappenplan proefafname	49
B.1	Deel 1	49
B.2	Deel 2	49
B.3	Deel 3	50
B.4	Deel 4	50
B.5	Deel 5	50
B.6	Deel 6	51
C	Stappenplan analyse	53
C.1	De spraaksynthese van de assistenten	53
C.1.1	Inlezen van csv-file	53
C.1.2	Installeren van de nodige packages	53

C.1.3	Vorbereiden van de data	53
C.1.4	Leeftijd van de deelnemers	54
C.1.5	Vergelijking van de assistenten per eigenschap	54
C.1.6	Vergelijking van de eigenschappen per assistent	56
C.1.7	De gemiddelde score en standaardafwijking van alle eigenschappen per assistent gesorteerd van hoog naar laag	56
C.2	De spraakherkenning van de assistenten	57
C.2.1	Inlezen van csv-file	57
C.2.2	Vorbereiden van de data	57
C.2.3	Een overzicht van de gevormde tekst	57
	Bibliografie	59

Lijst van figuren

2.1	een voorbeeld van een spectrogram (24)	18
2.2	William C. Derschs Shoebox deed eenvoudige berekeningen met spraak-commando's (9)	20
2.3	Watson versloeg twee kampioenen in Jeopardy live op televisie (13)	22
2.4	Hoe hoog ligt het intelligentieniveau bij slimme spraakassistenten (2)	24
3.1	De flowchart waar de applicaties voor de test op zijn gebaseerd	32
3.2	Een deelnemer ontvangt de vragen die hij zal stellen aan de assistenten.	33
3.3	De audiofragmenten worden beluisterd door de assistenten in een geluidsdichte studio.	34
4.1	De score die de deelnemers hebben gegeven op de emotionaliteit van de assistenten	35
4.2	De score die de deelnemers hebben gegeven op de levendigheid van de assistenten	36
4.3	De score die de deelnemers hebben gegeven op de menselijkheid van de assistenten	37

4.4 De score die de deelnemers hebben gegeven op het tempo van de assistenten	38
4.5 De score die de deelnemers hebben gegeven op de verstaanbaarheid van de assistenten	38
4.6 De gemiddelde score en standaardafwijking van alle assistenten en hun eigenschappen gesorteerd van hoog naar laag.	39

Lijst van tabellen

1. Inleiding

Sinds het bestaan van de computer zijn we het gewoon om externe apparatuur zoals een toetsenbord of muis te hanteren om ermee te communiceren. Dit zorgde doorheen de jaren voor heel wat frustratie en was het bijna een verplichting voor velen om hiermee te leren werken. Toch gebruiken we van mens tot mens een ander soort communicatiemiddel, een manier die ons van nature wordt aangeleerd, de gesproken conversatie. Een instinctieve kunst die mensen al duizenden jaren onder de knie hebben. Het is één van de eerste en belangrijkste vaardigheden die een kind leert. Een kunst, eigen aan de mens, die nooit hoofdzakelijk is gebruikt als communicatiemiddel met een computer. Toch werd hier al tientallen jaren geleden met geëxperimenteerd. Vandaag de dag komt de verwachting om op dezelfde manier te communiceren met de computer dan als we doen met mensen dichtbij.

Op 2 april 2019 kwam het Rode Kruis met het nieuws dat ze een mobiele applicatie hebben ontwikkeld voor het geven van eerste hulp bij ongevallen. Dit nieuws was dan ook de aanleiding om mijn onderzoek naar verschillende spraakgestuurde technologieën te koppelen aan het geven van EHBO.

- 1.1 Probleemstelling**
- 1.2 Onderzoeksvraag**
- 1.3 Onderzoeksdoelstelling**
- 1.4 Opzet van deze bachelorproef**

2. Stand van zaken

2.1 Spraakgestuurde technologie

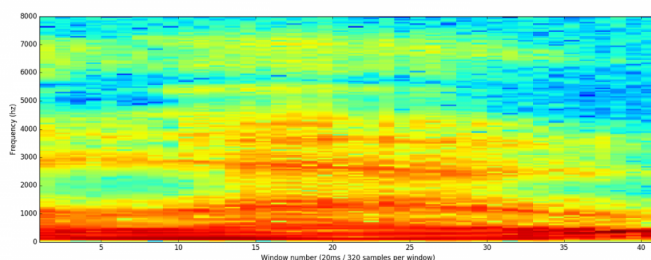
Er zijn enkele begrippen die met het thema te maken hebben, maar die niet hetzelfde omvatten. Taal- en spraaktechnologie is een verzamelnaam voor allerlei technieken waarmee de computer communiceert met zijn gebruiker door menselijke taal.[22] Het is de poging van de computer om de menselijke taal na te bootsen.

2.1.1 Spraakherkenning

Spraaktechnologie wordt vaak geassocieerd met spraakherkenning. Volgens [18] is spraakherkenning de kunst van de computer om gesproken taal te identificeren en om te zetten naar voor de computer leesbare machinetaal. Voor een computer succesvol spraak heeft omgevormd naar tekst, wordt er een lang en moeilijk proces doorlopen. Er wordt kort besproken wat de belangrijkste stappen zijn in het converteren van uitgesproken tekst naar tekst op een computerscherm. [24], [6] en [Woodford 2019] beschrijven hoe spraakherkenning in zijn werk gaat. Wanneer een persoon iets uitspreekt, wordt zijn stem als geluidsgolven opgenomen door een microfoon. De analoge signalen worden omgezet naar digitale door een techniek genaamd sampling. Op vaste intervallen wordt de amplitude van de geluidsgolf gemeten en gedigitaliseerd. Het geluid is omgezet naar bits.

Uit het gedigitaliseerd geluid wordt geprobeerd om zo veel mogelijk ruis te filteren. Daarnaast wordt het ook naar een vast volume en een gelijke snelheid gebracht, omdat niet iedereen even snel en even luid spreekt. De uiteindelijke bedoeling is om een neurale netwerk in te zetten om klanken en woorden te vormen uit het digitale geluid. Met een neurale netwerk wordt de techniek bedoeld uit de IT-wereld die de werking van de her-

senen gaat nabootsen om een computer zichzelf taken te leren. Uit deze verkregen verzameling van getallen is het voor een neurale netwerk nog steeds moeilijk om letters en woorden te herkennen. De bits worden eerst nog gegroepeerd in delen van ongeveer 20 milliseconden lang. Elke groep wordt dan opgesplitst in frequentiebanden waar telkens wordt nagegaan hoeveel energie er in vervat zit. Zo wordt een spectrogram gemaakt die een soort van vingerafdruk voorstelt van het geluid. In deze soort data kan een neurale netwerk gemakkelijker patronen herkennen. Dit is het moment waar het neurale netwerk



Figuur 2.1: een voorbeeld van een spectrogram [24]

zijn werk begint te doen. Hij gaat aan elk stukje data een spraakklank toekennen. Spraakklanken zijn de klanken die samen een taal vormt. Voor de Nederlandse taal bestaan er 40 verschillende spraakklanken. Klanken worden in woordenboeken fonetisch beschreven om duidelijk te maken hoe woorden worden uitgesproken. Het neurale netwerk gaat uit een fonetische lijst de spraakklank halen die de grootste waarschijnlijkheid heeft om correct bij een audiosignaal te horen. Je kunt nooit helemaal zeker zijn dat er een bepaalde klank is uitgesproken. Dit klinkt eenvoudiger dan het is. Elke klank wordt namelijk door elke mens op een andere manier uitgesproken. Door het netwerk te trainen met een grote hoeveelheid data, zal het beter leren omgaan met deze verscheidenheid.

De laatste stap is het omzetten van de klanken naar woorden. De klanken worden aan elkaar gelinkt tot woorden. Ook hierbij helpt trainingsdata om accurater woorden en zinnen te vormen. Als het netwerk bijvoorbeeld twijfelt tussen de woorden hallo, aloo en haylow, dan zal het waarschijnlijkheid kiezen voor 'hallo' omdat dit waarschijnlijk vaker voorkomt in de trainingsset. Daarnaast houdt het ook rekening met de waarschijnlijkheid dat een bepaald woord volgt op een woord. Ter illustratie, een netwerk zal door te trainen begrijpen dat de kans groter is dat het woord 'voorbeeld' gevolgd is op woorden zoals 'als, een of goed' dan op woorden zoals 'inktvis of tafel'. Voor neurale netwerken populair werden werd hiervoor een andere techniek gebruikt, namelijk het Hidden Markov Model. Hoe dit model precies werkt is buiten de scope van dit onderzoek.

2.1.2 Spraaksynthese

Een andere techniek, die net het omgekeerde is van spraakherkenning, is spraaksynthese. Volgens [18] is spraaksynthese menselijke spraak dat is gevormd door een computer. Spraaksynthese is de basis voor elk Text-To-Speech systeem. Het wordt gebruikt om geschreven tekst om te zetten naar gesproken taal, geproduceerd door de computer. Spraaksynthese is aanwezig in ons dagelijkse leven. In automatische telefoongesprekken, de

luchthaven, gps-systemen, op de bus en natuurlijk in digitale assistenten. [20] vertelt dat er twee soorten van methodes zijn voor Text-To-Speech. Concatenative TTS, waar korte audiofragmenten aaneengeschaakeld worden, is daar één van. Het is goed verstaanbaar omdat de woorden zijn opgenomen in hoge kwaliteit, maar het klinkt niet natuurlijk. De andere methode, parametric TTS bedenkt de spraak gebaseerd op enkele parameters. Het haalt taalkundige kenmerken uit de tekst. Daarnaast extraheert het vocoder kenmerken die het corresponderende spraaksignaal representeert. Hier wordt niet verder op ingegaan omdat er nu een derde methode is opgedoken. Deep Learning heeft ervoor gezorgd dat de Text-To-Speechsoftware geavanceerder werd. Eenvoudigweg is deep learning een benaming voor complexe neurale netwerken.

–Hier nog uitleg over deep learning TTS– –Hier nog uitleg over spraakhsynthese– Ook iets vermelden over Lyrebird, waar je zelf al met je eigen stem text to speech kan doen in het Engels. iets over dat dat de nieuwe Deep Learning methode is.

Spraaktechnologie omvat naast deze twee begrippen nog meer. Denk maar aan spelling- en grammaticacontrole, spraak- en tekstanalyse, automatische vertalingen, enzovoort.

Spraakherkenning en spraaksynthese zijn nodig in het ontwikkelen van een Voice User Interface (VUI), of stemgestuurde gebruikersomgeving, waar de gebruiker de computer als het ware bedient met zijn stem in plaats van bijvoorbeeld een toetsenbord of aanrakingen. De computer moet gesproken taal van de gebruiker begrijpen (spraakherkenning) en moet een gepast antwoord teruggeven (spraaksynthese). Spraakassistenten zoals Alexa, Siri of Google Assistant zijn voorbeelden van VUI's.

Omdat een assistent de spraak kan omzetten naar tekst betekent dit nog niet dat hij begrijpt wat iemand heeft verteld. Assistenten worden ontwikkeld met Natural Language Processing. Het is een methode om ongestructureerde data gebaseerd op de natuurlijke taal te verwerken tot een vorm die de computer kan begrijpen. Het zorgt ervoor dat de betekenis of het doel wordt achterhaald van wat iemand zegt. NLP valt onder de noemer van Artificiële Intelligentie en het maakt gebruik van deep learning modellen. Modellen die getraind zijn om patronen te gaan herkennen in de menselijke taal door grote hoeveelheden aan data van bijvoorbeeld conversaties en berichten door te nemen. In principe is het vergelijkbaar met hoe een kind de taal leert, namelijk door naar voorbeelden te luisteren. [17] Volgens [5] kan NLP voornamelijk onderverdeeld worden in twee niveaus, syntaxis en semantiek. Syntaxis is de grammatica van de tekst leren begrijpen. Het splitsen van zinnen of woorden en elk deel identificeren is één van de vele functies. Semantiek is de betekenis van de tekst leren begrijpen. Algoritmen worden gebruikt om bijvoorbeeld woorden te interpreteren en te classificeren als persoonsnaam of plaatsnaam.

2.2 Spraakassistenten

Een spraakassistent, ook wel een virtuele, persoonlijke of slimme assistent genoemd, voert taken uit via verbale instructies van een gebruiker. Het is vooral aanwezig in smartphones, maar het wordt ook geïntegreerd in smart speakers, auto's of wearables.

Dit onderzoek vergelijkt twee van de meest prestigieuze assistenten, Googles Assistant en Amazons Alexa. Daarnaast zijn ook Apple's Siri, Microsoft's Cortana en Samsung's Bixby bekende voorbeelden.

2.2.1 De geschiedenis van spraakassistenten

De slimme spraakassistenten zijn vandaag gekend bij het grote publiek. Ze zijn ingebouwd in onze smartphones en slimme luidsprekers. Steeds paraat om ons de vertragen te melden op de weg, het weer te voorspellen voor morgen of onze favoriete muziek te spelen. Het is iets van deze tijd, maar toch hebben ze al een lang pad van tientallen jaren bewandeld. Dit is hoe het allemaal begon en hoe we zijn geëvolueerd naar de bekende assistenten van vandaag.

Jaren 50 - 60

De eerste systemen die ietwat leken op een spraakassistent waren gefocust op het louter herkennen van de menselijke spraak. In [26] wordt geschreven hoe in de Bell Laboratories in 1952 het "Audrey" systeem werd ontwikkeld. Audrey begreep de getallen 0 tot 9 op voorwaarde dat de sprekers tussen elk getal een pauze lieten. In theorie kon het gebruikt worden om met de stem een telefoonnummer in te geven. Onder andere de kost en omvang van de machine was groot. Het intoetsen van de telefoonknoppen bleef efficiënter, dus het effectieve gebruik van Audrey bleef uit.

[9] onthulde in 1962 de "Shoebbox", een machine die met spraakcommando's eenvoudige berekeningen kon uitvoeren. De uitvinder William C. Dersch demonstreerde voor televisie hoe het apparaat, zo groot als een schoendoos, naast de getallen 0 tot 9 ook zes woorden zoals plus en totaal kon herkennen.



Figuur 2.2: William C. Dersch's Shoebbox deed eenvoudige berekeningen met spraakcommando's [9]

Jaren 70 - 80

Spraakherkenning in de jaren 70 werd vooral gekenmerkt door het departement voor defensie in de Verenigde Staten. Uit interesse voor spraakherkenning financierden ze een vijfjarig project over het thema. Volgens [15] en [11] heeft dit geleid tot de ontwikkeling van Harpy in 1976. Harpy begreep 1011 woorden en kreeg vooral betekenis door haar efficiëntere zoekmethode, de “Beam-search”, om logische zinnen te gaan herkennen.

In [15] staat dat in de jaren 80 er een grote doorbraak kwam door de ontwikkeling van het hidden Markov model. Dit model gebruikt statistieken om een woord te herkennen in een onbekend geluid. Dit werd gedaan door het berekenen van de waarschijnlijkheid dat het onbekend geluid staat voor een bepaald woord. De woordenschat van de spraakherkenningssoftware bleef groeien tot een paar duizend woorden en had dankzij onder andere het hidden Markov model het potentieel om ongelimiteerd woorden te gaan herkennen. Onder andere dankzij deze ontwikkelingen bleven ook de commerciële toepassingen niet uit. In 1987 kwam de Worlds of Wonder’s doll Julie uit. Kinderen konden de pop trainen om te reageren op hun uitspraken. Dit staat zo beschreven in [15], waar je ook een reclamespot voor de pop kan bekijken. De technologie groeide snel, maar had wel een grote zwakte. De zin moest gedictieerd worden. Na elk woord werd dus een korte pauze verwacht.

Jaren 90

Volgens [11] kwam in de 90’s automatische spraakherkenning in een eerste vorm zoals we het vandaag kennen. De doorbraak in die tijd heette Dragon. De eerste versie werd gelanceerd in 1990 onder de naam Dragon Dictate en had een woordenschat van 80 000 woorden. Daarnaast kon het iets nieuws, iets wat in de huidige spraakassistenten nog steeds gebruikt wordt, natural language processing. Zinnen moesten niet meer gedictieerd worden, maar Dragon kon oorspronkelijk 30 tot 40 woorden per minuut herkennen.

Volgens een artikel uit 1998 [16] is Dragon verantwoordelijk voor een doorbraak in spraakherkenningssoftware. De opvolger van de Dragon Dictate, Dragon NaturallySpeaking laat gebruikers spreken in een microfoon, aangesloten op de computer, en laat de woorden direct verschijnen op het computerscherm. Indien het een fout maakte, kon je het zelf corrigeren en kon de software leren uit zijn fouten. Het was ook de eerste spraakherkenningssoftware die toeliet om op een normale manier te praten.

Van 2010 tot nu

In [9] is te lezen hoe een mijlpaal werd bereikt door de Watson machine die won in Jeopardy. Watson was zo goed in taalverwerking dat hij 2 kampioenen in Jeopardy heeft verslaan live op televisie. Jeopardy is een Amerikaans spelprogramma waar de kandidaten het antwoord kregen en ze zelf de bijpassende vraag moesten geven. Watson was niet alleen goed in het samenstellen van correcte vragen, maar kon die ook telkens hardop uitspreken.



Figuur 2.3: Watson versloeg twee kampioenen in Jeopardy live op televisie [13]

2.2.2 De spraakassistenten van nu

–Kostprijs alexa en GA + ondersteuning talen meegeven– Kort na deze gebeurtenis in 2011 werd Siri gebouwd in de Iphone 4S en werd zo de eerste spraakassistent voor het grote publiek uitgebracht. Siri is Apple's variant op de slimme spraakassistent en is tegenwoordig beschikbaar op meerdere apparaten met een IOS-besturingssysteem. Het grootste aandeel van de gebruikers kent Siri van op zijn Iphone, maar daarnaast is de assistent ook geïntegreerd in de Mac computer, de Apple Watch of de Apple TV. Ondertussen heeft het ook zijn eigen Smart Speaker, de HomePod.

Google gaf hierop een antwoord in 2012 door Google Now uit te brengen, de voorloper van de Google Assistant van vandaag. Volgens Google is de Google Assistant jouw eigen persoonlijke Google, die altijd bereid is om je te helpen wanneer je maar wilt. De Google Assistant bestond eerst onder de naam Google Now en was aanwezig in smartphones met een Android besturingssysteem. De Google Assistant van vandaag is te vinden in veel meer omgevingen. Smartphones, auto's, laptops, tablets, tv's, smartwatches en in hun eigen smart speaker, de Google Home. Deze speaker heeft ook een variant gekregen met een scherm, de Smart Display.

Tijdens de Microsoft BUILD conferentie in 2013 werd Cortana geïntroduceerd als de spraakassistent van Microsoft. Cortana is ontwikkeld voor onder andere Windows 10, Windows Phone, Xbox One en in de slimme speaker Invoke.

In 2015 kwam de eerste slimme luidspreker op de markt. De Echo van Amazon, voorzien met hun slimme spraakassistent, Alexa. Amazon is één van 's werelds grootste bedrijven in het online verkopen van goederen. Het grote verschil met Google is dat de assistent voor het eerst werd gebruikt in de Echo, Amazon's smart speaker.

[12] besprak verschillende functionaliteiten van spraakassistenten Google Assistant, Alexa en Siri, op correctheid en natuurlijkheid bij 8 ondervraagden. In de administratieve cate-

gorie, zoals agendabeheer, to-dolijsten en alarmen kwam de Google Assistant als minst correct en minst natuurlijke assistent uit de bus, maar prijkt in de veelzijdige categorie (nieuws, weer, verkeer, woordbetekenissen, rekenen,, enz.) dan weer ver bovenaan op beide vlakken. Algemeen werd de Google Assistant als de meest natuurlijke ervaren, onder andere door de toon van de stem die verwondering, onzekerheid en vreugde uitte. Voor [23] stelden 100 personen allerlei vragen aan voice assistants Google Assistant, Alexa, Siri en Cortana. Ze gaven telkens een score op spraakherkenning en contextueel inzicht. Google Assistant kwam als grote winnaar uit het onderzoek door 59,80 % van de vragen te beantwoorden. Een verschil van 15,82 % met Siri, die de op één na nauwkeurigste bleek in dit onderzoek. Google Assistant was vooral leider in categorieën als reizen, mailing, navigatie, vertalingen en begreep volgens het onderzoek goed de verschillende variaties in de stemmen van de onderzochte personen.

Volgens [12] is de Echo Dot de favoriete smart speaker als het aankomt op het aankopen van artikelen. Dit is geen verrassing omdat hij oorspronkelijk ontworpen is om te winkelen en zelfs de enige spraakassistent is waarmee je online kan shoppen. In [23] bleek Alexa de minst nauwkeurige assistent te zijn met 7,91 % nauwkeurigheid.

Google spendeert veel middelen aan zijn assistent. Het heeft indruk gemaakt tijdens de recente Google I/O conferentie van dinsdag 7 mei 2019, waar ze onverwachts nieuws brachten. De volgende versie van de Google Assistant zal namelijk opvallend veel sneller gaan omdat ze de AI-modellen die verantwoordelijk zijn voor NLP offline beschikbaar hebben gemaakt. Dat wilt zeggen dat een commando van een gebruiker niet meer helemaal naar een server in Amerika moet transporteren om ervoor te zorgen dat de assistent het kan begrijpen. Vanaf de volgende versie zal deze logica afgehandeld worden door uw toestel zelf, omdat Google erin geslaagd is het geheugen van die modellen zo te reduceren dat het kan opgeslagen worden op uw apparaat. Het belooft dus dat binnenkort de gebruiker na het stellen van een vraag amper nog zal moeten wachten op een antwoord.

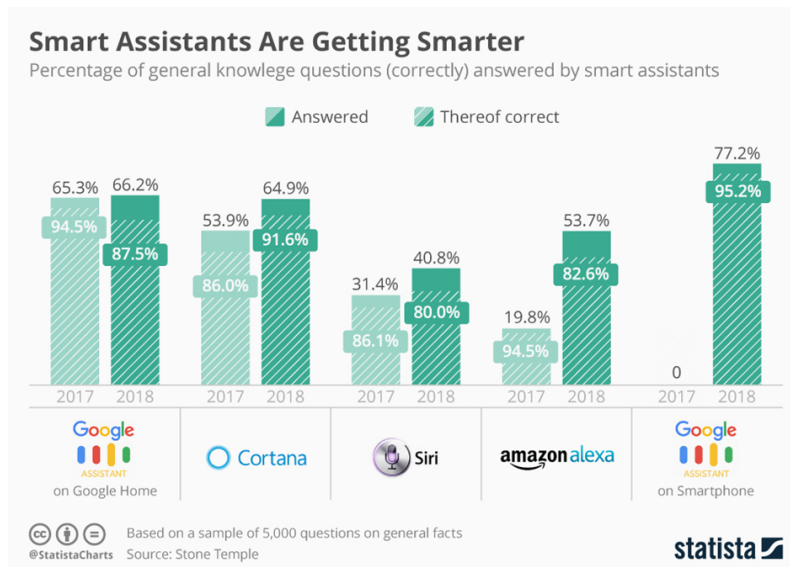
VRT NWS meldt op 28 mei dat de Belgische variant van de Google Assistant wordt gelanceerd. [1] Voorlopig nog alleen met een Nederlands accent. Wat er wel bijkomt is de samenwerking met verschillende Belgische bedrijven. Zo kan iedereen binnenkort via de Google Assistant naar het radionieuws van VRT NWS luisteren, aan de NMBS vragen wanneer de volgende trein komt of artikelen van de Colruyt toevoegen aan zijn lijstje.

Slimme spraakassistenten worden alleen maar slimmer. [2] heeft onderzocht hoe hoog het intelligentieniveau is van 4 slimme assistenten, namelijk Google Assistant, Microsoft Cortana, Amazon Alexa en Apple Siri in 2017 en 2018. De geanalyseerde gegevens zijn de antwoorden van de assistenten op 5000 algemene vragen. De beste prestatie werd verricht door Google Assistant die op 77,2 procent van de vragen een antwoord kon bieden, waarvan 95 procent correct. Bij alle assistenten zie je een verhoging van de intelligentie in vergelijking met het vorige jaar.

–Hier nog verder doen met overige bronnen over assistenten vandaag–

Als het gaat over de spraakqualiteit, ook meer uitleggen wat er precies is beoordeeld. ->

De Text-to-speech of speech synthesis, bij Google Assistant met WaveNet: <https://towardsdatascience.com/>



Figuur 2.4: Hoe hoog ligt het intelligentieniveau bij slimme spraakassistenten [2]

google-assistants-voice-synthesizer-a168e9af13b1

2.2.3 Hoe werkt een spraakassistent

–De flow uitleggen van de gebruiker die iets vraagt tot de assistent die een antwoord heeft gegeven–

2.3 Bestaande eerste hulpapplicaties

2.3.1 De Vlaamse EHBO-app van het Rode Kruis

Op 2 april 19 kwam het Rode Kruis met het nieuws dat ze een app hebben ontwikkeld die kan helpen bij het geven van eerste hulp bij ongevallen. 80 procent van de Vlamingen weet niet wat hij moet doen als een nabije persoon begint te stikken, een hartstilstand krijgt of hevig begint te bloeden. Uit angst om iets fouts te doen, gebeurt er dan ook vaak niks. Met de app willen ze zoveel mogelijk mensen in staat stellen om hulp te verlenen. [4]

Het Rode Kruis benadrukt dat de applicatie de opleiding niet kan vervangen, maar dat het hulp kan bieden bij het geven van eerste hulp.

In de applicatie zijn er drie grote onderdelen, eerste hulp verlenen, eerste hulp leren en een AED-toestel vinden in de buurt. Er zijn ook nog enkele opties die je naar de website van het Rode Kruis brengen om informatie te verkrijgen over het geven van bloed of plasma, het doen van een gift, het volgen van een opleiding of het aanmelden als vrijwilliger.

Als je eerste hulp wilt verlenen kun je uit het overzicht een onderwerp over eerste hulp kiezen, waarna je informatie krijgt over wat je moet vaststellen en wat je nodig hebt. Daarnaast geeft de app ook een stappenplan van instructies wat je moet doen. De levensbedreigende situaties staan helemaal bovenaan en zijn voorzien van extra ingesproken instructies.

Als je eerste hulp wilt leren kun je eerst een onderwerp kiezen. Voorbeelden zijn een beroerte of alcoholvergiftiging. Daarna krijg je over het onderwerp vragen & antwoorden, informatieteksten en video's. Per leerdeel krijg je een quiz die je moet oplossen om bepaalde badges te verdienen.

Wanneer iemand in uw omgeving een hartstilstand krijgt dan kun je met de applicatie een kaart openen waar AED-toestellen staan op gesitueerd. Je kunt er ook een nieuwe AED melden of meer informatie lezen.

2.3.2 Andere EHBO-applicaties

Nederland heeft al langer een mobiele EHBO-applicatie. Deze verschilt niet zo veel met de Belgische versie. Ze heeft wel een zoekfunctie om sneller de instructies voor uw ongeval te vinden. Je kan er ook EHBO-kits en cursussen bestellen in de webshop.

–Nog iets over Engelstalige applicaties–

3. Methodologie

3.1 De keuze van de spraakassistenten

Omdat het gaat over een vergelijkend onderzoek, is het belangrijk in weinig veranderlijke omstandigheden te werken. Om te vermijden dat de hardware een invloed heeft op de resultaten worden de assistenten gebruikt op hetzelfde toestel, de Xiami Redmi Note 4. Hierdoor kregen bepaalde assistenten geen kans meer om opgenomen te worden in het onderzoek. Bixby is enkel beschikbaar op apparaten van Samsung zoals Siri enkel beschikbaar is voor Apple-gebruikers. Daarnaast heeft het ontwikkelen van eigen applicaties beperkingen voor beide assistenten.

–vertel hier nog over welke beperkingen precies–

Na eerder opgesomde redenen wordt beslist om enkel Amazon Alexa en Google Assistent verder te onderzoeken. Meer specifiek zal de Amerikaans-Engelse Alexa vergeleken worden met de Amerikaans-Engelse Google Assistent en de Nederlandstalige Google Assistent uit Nederland.

3.2 Wat wordt er vergeleken

–Hier nog over de twee dingen die worden onderzocht: spraakherkenning en spraaksynthese + verwijzing naar stand van zaken– –dan verdere onderverdeling: spraaksynthese aan de hand van de kwaliteit. 5 eigenschappen: verstaanbaarheid, menselijkheid, levendigheid, tempo en emotionaliteit en extra uitleg– –spraakherkenning wordt er vergeleken hoe goed de assistent spraak heeft omgevormd naar tekst door het aantal fouten te me-

ten. Hoe correct een assistent een bepaalde vraag van de gebruiker kan begrijpen door te meten hoe sterk zijn gevormde zin verschilt van de gestelde vraag–

3.3 Gebruikte materialen

Voor het onderzoek zijn volgende zaken gebruikt:

- Acer Aspire F 15 laptop
 - Intel Core I5
 - geïntegreerde microfoon
 - 2,5 jaar in gebruik
- Xiami Redmi note 4 smartphone
 - Android 7.0
 - Octa core processor
 - 4 GB RAM
 - 2 jaar in gebruik
- Ultimate Ears BOOM 2 Speaker
 - 12,5 watt vermogen
 - 90dB gevoeligheid
 - 3,5mm mini-jack (AUX) ingang
 - 1,5 jaar in gebruik
- 3.5mm Jack kabel
- Google Assistant app voor Android
- Alexa Beta app voor Android

3.4 Het verloop van het onderzoek

3.4.1 Deel één: met de deelnemer

De steekproef bestaat uit 30 Vlamingen tussen 19 en 60 jaar die Nederlands spreken als moedertaal en enige kennis hebben van de Engelse taal. Het onderzoek is afgenomen in twee grote delen.

De onderzoeker volgt voor dit deel een vast stappenplan die te vinden is in B. De vragenlijst die tijdens dit deel wordt ingevuld is te vinden onder de map bijlagen in de repository beschreven in 0.1. Er is een kleine applicatie ontwikkeld voor elke assistent waar de gebruiker EHBO-gerelateerde vragen aan kan stellen en telkens een vast antwoord van terugkrijgt.

Voor het eerste deel stellen de deelnemers drie Engelse en drie Nederlandse EHBO-gerelateerde vragen aan een fictieve spraakassistent. De gestelde vragen worden opgenomen door een laptop. Om te beginnen leest de deelnemer een korte introductietekst over spraakassistenten en vult hij enkele algemene vragen in. Daarna krijgt hij zes blade-

ren met op elk blad een vraag. Hij wordt aangespoord door de onderzoeker om elke vraag eerst in stilte te lezen om hem daarna te stellen aan de assistent. De aanwezigheid van een spraakassistent wordt geveinsd omdat de manier waarop de deelnemer de vragen stelt vergelijkbaar moet zijn aan de manier waarop hij ze zou stellen aan een echte assistent. De participant krijgt nooit een antwoord te horen, waardoor de mogelijkheid ontstaat dat hij steeds meer zijn best doet de volgende vraag duidelijker uit te spreken. Het verkregen resultaat zijn opgenomen audiofragmenten die later worden gebruikt in het tweede deel.

Na het stellen van alle vragen wordt de deelnemer op de hoogte gebracht van de assistent zijn afwezigheid en krijgt hij een volgende taak. Hij moet nu aan drie bestaande assistenten telkens dezelfde drie vragen stellen en te luisteren naar de drie antwoorden. Hij krijgt de antwoorden van de assistent te horen door een speaker die met een kabel is aangesloten aan een smartphone. Hij kan de vragenlijst over de assistenten hun spraakkwaliteit al eens doornemen. De participant wordt door de onderzoeker aangespoord om goed te luisteren. Hij mag vragen opnieuw stellen tot hij antwoord krijgt van de assistent. Telkens nadat een assistent de drie vragen heeft beantwoord wordt de vragenlijst over die assistent ingevuld. De deelnemer krijgt na elke beoordeling de kans om vorige beoordelingen van assistenten aan te passen. Op het einde vult de participant nog enkele algemene vragen in.

3.4.2 Deel twee: zonder de deelnemer

De vragenlijst die tijdens dit deel wordt ingevuld is te vinden onder de map bijlagen in de repository beschreven in 0.1.

Tijdens het tweede deel van het onderzoek luisteren de assistenten naar de vragen van de deelnemer. Er wordt gemeten hoe de assistenten deze vragen van spraak omzetten naar tekst. Er werden maatregelen genomen om de beïnvloeding van veranderlijke factoren te beperken. Een persoon zal nooit twee keer een vraag op exact dezelfde manier uitspreken. Als de deelnemer elke vraag al kan stellen zonder versprekingen dan nog zal er altijd een verschil zijn in onder meer volume, snelheid en intonatie. Daarnaast is er ook nog de extra beïnvloeding van externe ruis.

Een oplossing kan zijn om drie toestellen te gebruiken die elk voorzien zijn van een andere assistent. De gebruiker kan dan de vraag eenmalig stellen terwijl de drie assistenten tegelijkertijd luisteren. Ondanks dat dit op het eerste zicht een goede oplossing lijkt, zijn er toch enkele bedenkingen. Wegens financiële redenen is het voor de onderzoeker niet mogelijk om drie nieuwe smartphones aan te kopen. Moest dit toch mogelijk zijn dan kan er nog steeds een verschil van kwaliteit in de microfoon aanwezig zijn. Elk apparaat is uniek. Als er met drie gebruikte smartphones wordt gewerkt zal het verschil zeker aanwezig zijn door slijtage of ongelijk model.

Dit is de reden waarom de vragen zijn opgenomen in deel één. Een geregistreerd audiofragment maakt het mogelijk om de identieke vraag drie keren af te spelen terwijl telkens één assistent meeluistert op hetzelfde apparaat. Dit gebeurt in een geluidsdichte opnamestudio om externe ruis tijdens dit proces zoveel mogelijk te beperken. Er kan wel ruis mee opgenomen zijn tijdens deel één, maar die is dan aanwezig bij elke assistent die het

audiofragment hoort. De fragmenten worden afgespeeld door een speaker die aangesloten is op een laptop met een kabel.

Zowel Alexa als Google Assistant maken het mogelijk om een overzicht van uw interacties te bekijken. Na de audiofragmenten van de 30 gebruikers af te spelen voor elke assistent werd dit overzicht gebruikt om de vragenlijst in te vullen. In de vragenlijst wordt de omgezette tekst van de assistent genoteerd samen met het aantal fouten die het bevat.

3.4.3 Ontwikkelen van een applicatie

Welke middelen kan je gebruiken om het te ontwikkelen?

Hoe uitgebreid is de documentatie?

Welke programmeertalen worden aangeboden?

Kan de applicatie automatisch een nummer bellen?

3.5 Orthopedagogisch onderzoek

3.5.1 De zoektocht naar een gepaste doelgroep

Ondersteuning van de begeleider in de jeugdzorg

De aanleiding van dit onderzoek was een onderzoek van -citeer onderzoek mr. Buysse- over faciliterende IT bij individuele begeleidingsgesprekken in de jeugdzorg. Uit het onderzoek ontstond er twijfel over het gebruik van een spraakassistent als ondersteuning bij het individuele begeleidingsgesprek tussen de persoonlijke begeleider en het kind. Door de twijfel werd er dan ook beslist om niet verder te gaan met spraaktechnologie, maar met andere digitale tools.

Om er toch zeker van te zijn dat er geen mogelijkheden waren, ontstond deze bachelorproef. Echter, na een eerste gesprek met Iris Storme, docent orthopedagogie binnen Hogeschool Gent en tevens mede-researcher van -citeer onderzoek mr. Buysse- werden voor mij de beweegredenen voor het afkeuren van spraaktechnologie binnen hun onderzoek snel duidelijk.

Als je denkt aan mensen met een visuele of fysieke beperking komen er snel mogelijkheden naar boven. Denk maar aan het controleren van apparaten met een eenvoudig stemcommando. Deze personen kunnen technologie als een mogelijke oplossing zien, waardoor zij, en de begeleiders, dit gemakkelijker kunnen omarmen. Daartegenover staat de bijzondere jeugdzorg, waar de spraakassistent eerder ondersteuning zou bieden in de emotionele problematiek en de jongeren net hun façade nodig hebben om overeind te blijven. Deze doelgroep stelt zich niet zo graag kwetsbaar op en ervaart het praten over

gevoelens eerder als een drempel. De bijzondere jeugdzorg lijkt op het eerste zicht een minder relevante doelgroep.

Dit werd allemaal vastgesteld door het intuïtieve gevoel van de ondervraagde. Dit was voor mij persoonlijk voldoende om na te gaan denken over een nieuwe doelgroep.

Ondersteuning van personen met het syndroom van Down

Ik wijzigde mijn doelgroep naar personen die geboren zijn met trisomie 21, ook wel het syndroom van Down genoemd. Uit een eerste opzoeking stelde ik de volgende mogelijkheden.

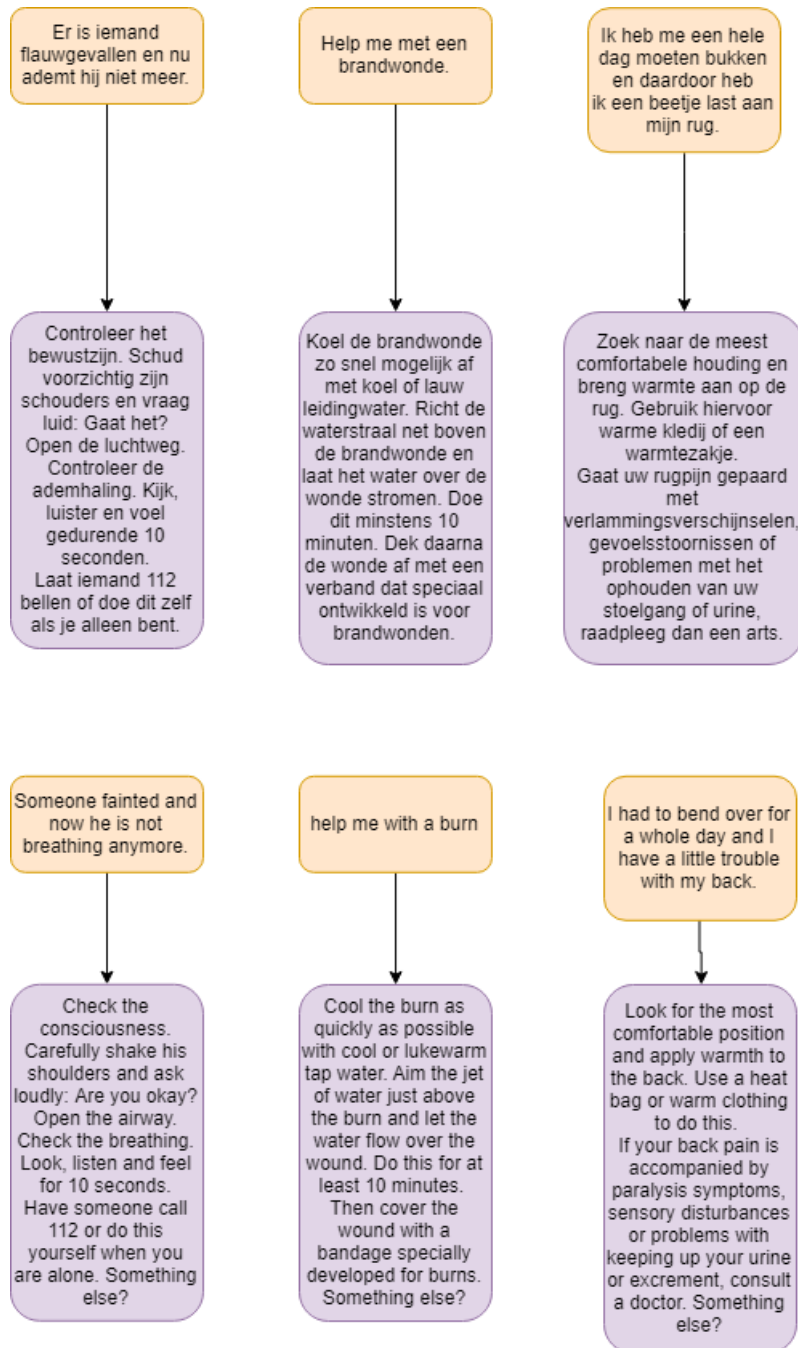
Personen met het Downsyndroom worden geboren met een verstandelijke beperking. Er kan gekeken worden naar welke noden uit die verstandelijke beperking vloeien, bijvoorbeeld moeite met rekenen, en hoe een spraakassistent hier ondersteuning kan bieden. Dit kan ook veel verder gaan als in vb. het helpen met zelfstandig wonen.

Daarnaast zijn er ook mogelijke bijkomende aandoeningen zoals een minder goed geheugen, coeliakie, slaapapneu, oogafwijkingen of een gedragsstoornis. Hier kan spraakassistentie mogelijks ook ondersteuning in bieden. Ik denk aan bijvoorbeeld interactieve activiteiten voor het stimuleren van de motoriek, het geheugen of het spraakvermogen, helpen herinneren aan belangrijke taken, helpen herinneren aan wat ze wel of niet mogen eten, stimuleren van een vast slaappatroon, enzovoort.

–Bronnen vermelden–

Dit waren nog maar losse ideeën die ontstonden uit een eerste verkenning over personen met het gendefect. Het werd duidelijk dat hier zeker mogelijkheden waren, dus was de volgende stap om hierin gaan te verdiepen door interviews af te nemen van mensen die een persoonlijke ervaring hebben met deze doelgroep.

Echter, mijn eerste aanvraag voor een interview aan iemand wiens dochter geboren is met trisomie 21 stootte direct op weerstand. Het antwoord dat ik kreeg, ging over het gevaar van veralgemenen. Het is een complexe materie omdat het gaat over een combinatie van samenkomende symptomen en kenmerken. De effecten van het gendefect kunnen in verschillende gradaties voorkomen, waardoor elke persoon die het gendefect bezit uniek moet bekeken worden.



Figuur 3.1: De flowchart waar de applicaties voor de test op zijn gebaseerd



Figuur 3.2: Een deelnemer ontvangt de vragen die hij zal stellen aan de assistenten.



Figuur 3.3: De audiofragmenten worden beluisterd door de assistenten in een geluids-dichte studio.

4. Resultaten vergelijkend onderzoek

Wie in detail wilt bekijken hoe de resultaten zijn gekomen, kan meer informatie vinden bij C.

4.1 Vergelijking van de assistenten in spraakkwaliteit

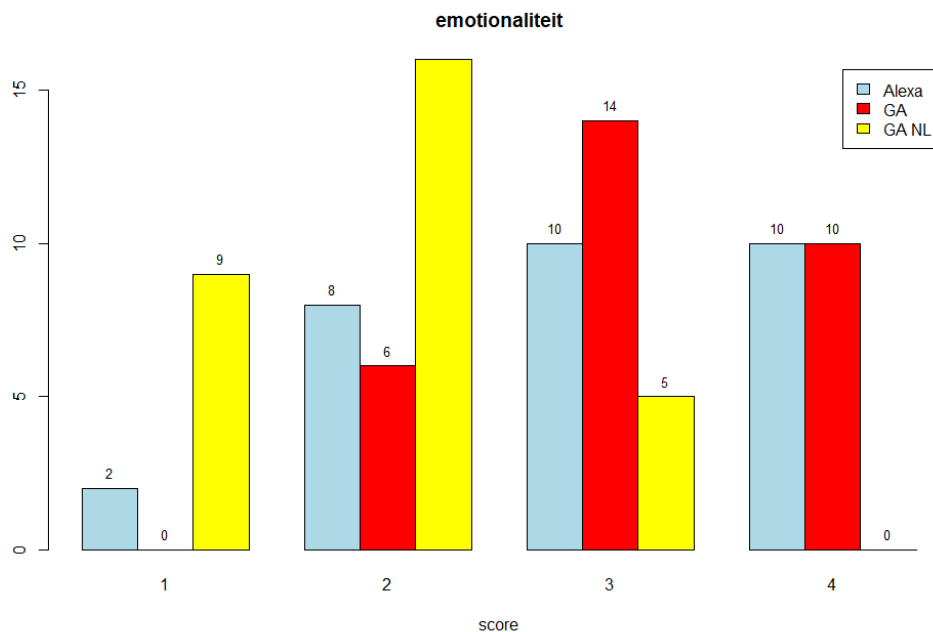
4.1.1 Vergelijking van de assistenten per eigenschap

4.1.2 Vergelijking van de eigenschappen per assistent

De resultaten van de uitgevoerde t-testen zijn te vinden onder de map onderzoek/onderzoeksresultaten in de repository beschreven in 0.1.

De p-waarde ligt bij elke test duidelijk onder het significantieniveau van 0.05 dus kunnen we de nulhypothese verwerpen. De volgende stellingen zijn statistisch verantwoord.

- Alexa scoort significant hoger op emotionaliteit dan GA NL.
- GA scoort significant hoger op emotionaliteit dan GA NL.
- Alexa scoort hoger op levendigheid dan GA NL.
- GA scoort hoger op levendigheid dan GA NL.
- Alexa scoort hoger op menselijkheid dan GA NL.
- GA scoort hoger op menselijkheid dan GA NL.
- GA scoort hoger op tempo dan Alexa.
- Alexa scoort hoger op tempo dan GA NL.
- GA scoort hoger op verstaanbaarheid dan GA NL.



Figuur 4.1: De score die de deelnemers hebben gegeven op de emotionaliteit van de assistenten

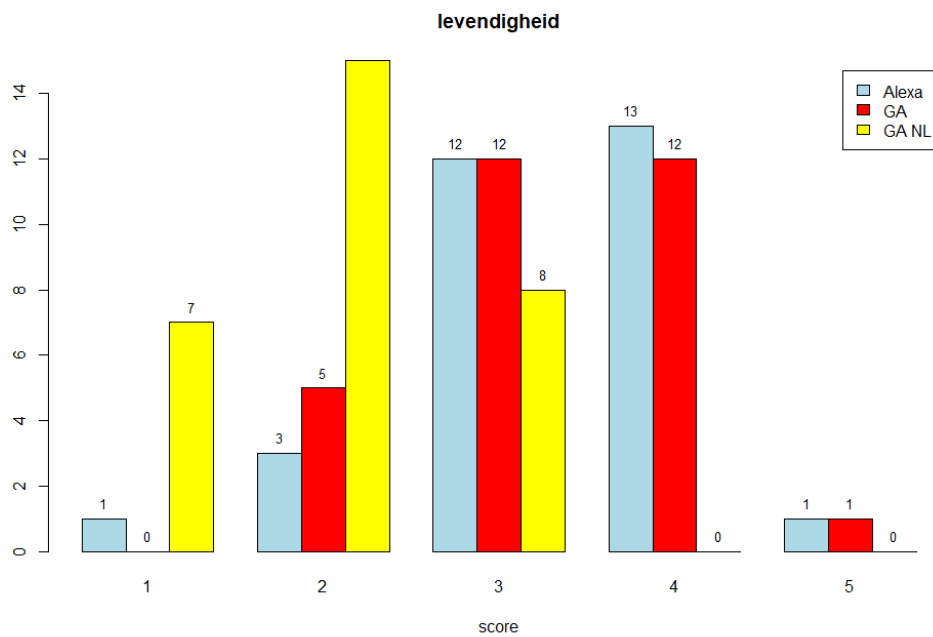
4.2 Vergelijking van de assistenten in spraakherkenning

4.2.1 Een moeilijk te voeren onderzoek

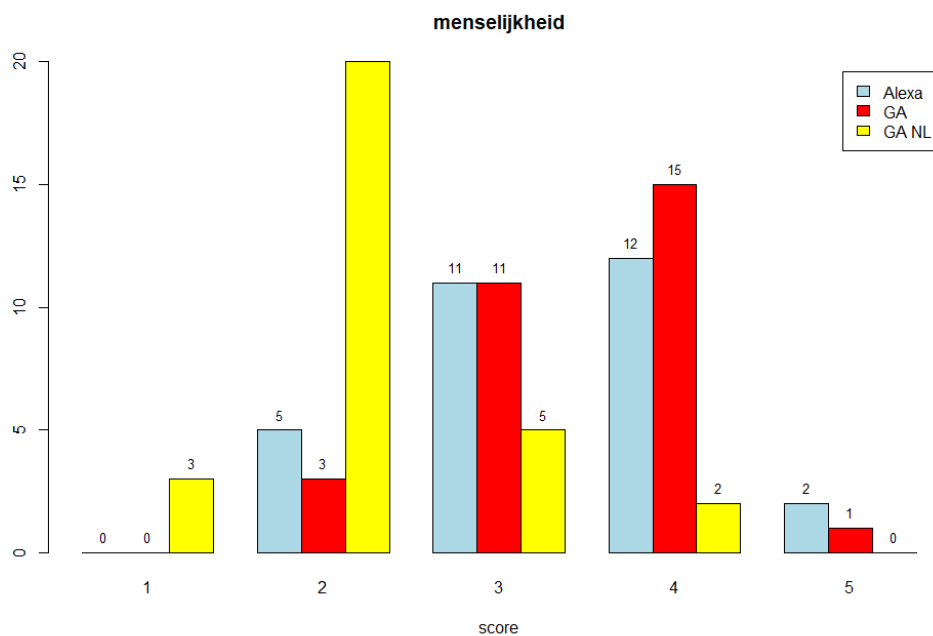
Uitleg over hoe moeilijk het was om dit onderzoek correct te voeren ondanks de genomen maatregelen beschreven in de methodologie.

Een eerste bemerking bij het onderzoek is dat elke fout gelijk meetelt. Stel dat twee assistenten de uitspraak 'help me with a burn' omvormen naar tekst. De ene assistent begrijpt de vraag als 'help us for a burn' en de andere als 'tell me with a good'. Alhoewel het gevoel zegt dat de eerste assistent de vraag beter heeft begrepen dan de tweede hebben ze toch allebei evenveel fouten gemaakt. De eerste assistent mist de woorden 'me' en 'with', terwijl de tweede assistent de woorden 'help' en 'burn' mist. Om de correctheid van een gevormde zin beter te interpreteren zou aan elk woord een soort van hoofdzakelijkheid voor het begrijpen van de zin moeten toegekend worden. Er bestaan echter geen vaste regels om dit aan woorden toe te wijzen.

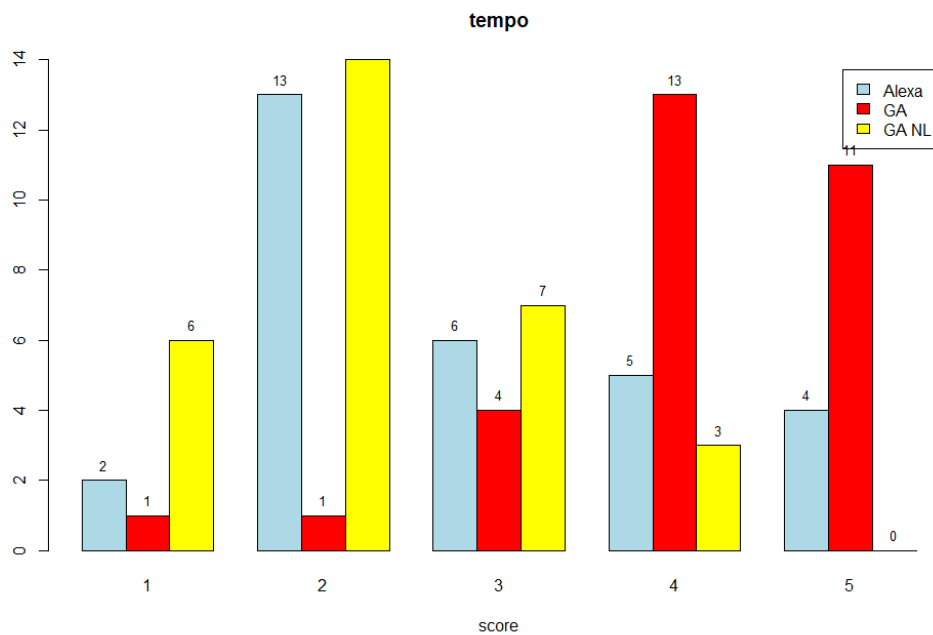
Ondanks dat er verschillende maatregelen zijn genomen om ervoor te zorgen dat elke assistent identiek dezelfde vraag krijgt, is deze opzet toch niet helemaal geslaagd. Alexa neemt elke conversatie op en bewaart ze in uw geschiedenis. Door enkele conversaties van tijdens het onderzoek te beluisteren is er opgemerkt dat hier en daar gekraak aanwezig is. Het gekraak was tijdens het onderzoek niet hoorbaar, waardoor de oorzaak waarschijnlijk bij de microfoon van de smartphone ligt. Het onderzoek is niet correct gevoerd omdat de sterkte van het gekraak kan variëren en mogelijks invloed heeft op de prestaties van de Speech-To-Text-functie van een assistent.



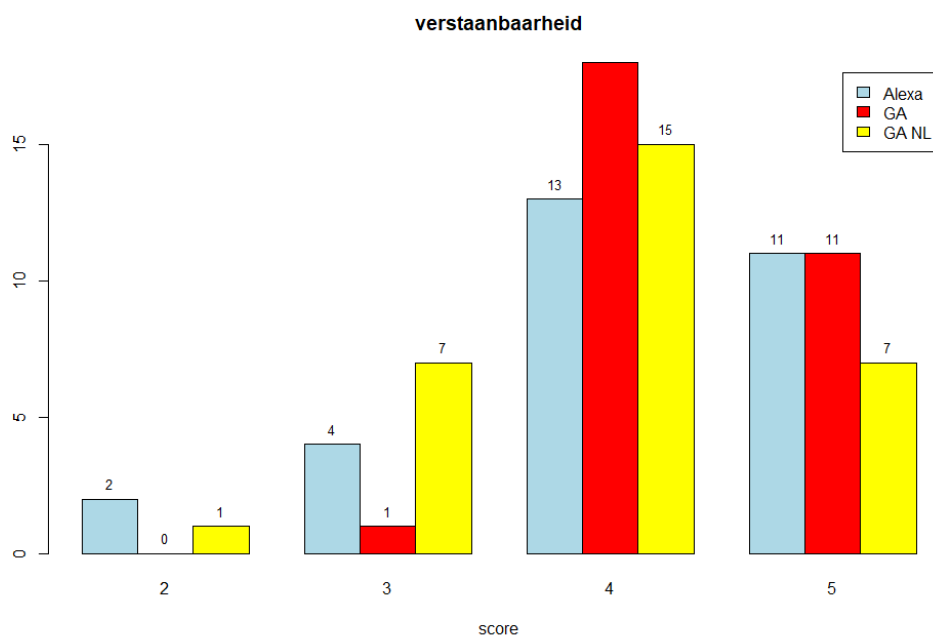
Figuur 4.2: De score die de deelnemers hebben gegeven op de levendigheid van de assistenten



Figuur 4.3: De score die de deelnemers hebben gegeven op de menselijkheid van de assistenten



Figuur 4.4: De score die de deelnemers hebben gegeven op het tempo van de assistenten



Figuur 4.5: De score die de deelnemers hebben gegeven op de verstaanbaarheid van de assistenten

eigenschap	assistent	mean	sd
verstaanbaarheid	GA	4.333333	0.5466723
verstaanbaarheid	Alexa	4.100000	0.8847365
tempo	GA	4.066667	0.9802650
verstaanbaarheid	GA NL	3.933333	0.7849153
menselijkheid	GA	3.466667	0.7302967
menselijkheid	Alexa	3.366667	0.8502873
levendigheid	Alexa	3.333333	0.8441823
levendigheid	GA	3.300000	0.7943768
emotionaliteit	GA	3.133333	0.7302967
emotionaliteit	Alexa	2.933333	0.9444332
tempo	Alexa	2.866667	1.1957780
tempo	GA NL	2.233333	0.8976342
menselijkheid	GA NL	2.200000	0.7143842
levendigheid	GA NL	2.033333	0.7183954
emotionaliteit	GA NL	1.866667	0.6814454

Figuur 4.6: De gemiddelde score en standaardafwijking van alle assistenten en hun eigenschappen gesorteerd van hoog naar laag.

5. De eerste hulp assistent

5.1 De functionaliteiten

5.2 Persona

6. Conclusie

A. Onderzoeksvoorstel

Het onderwerp van deze bachelorproef is gebaseerd op een onderzoeksvoorstel dat vooraf werd beoordeeld door de promotor. Dat voorstel is opgenomen in deze bijlage.

A.1 Introductie

Kinderen in jeugdcentra hebben een individuele begeleider nodig. De begeleider is verantwoordelijk voor het kind de aandacht, begeleiding en zorg op maat te geven die hij of zij nodig heeft. De begeleider is de vertrouwensfiguur voor het kind. Een probleem dat zich voordoet is dat de begeleider alleen aanwezig is tijdens zijn werkuren. De vertrouwenspersoon kan voor een periode wegvallen en tijdens deze periode is het kind zijn steunfiguur kwijt.

Spraakgestuurde technologie is aan een opmars bezig. In de Verenigde Staten hebben één op de vijf volwassenen toegang tot een smart speaker met stemassistent. [14] In België is het gebruik van een slimme luidspreker nog niet van de grond. Dit kan sinds de komst van de Nederlandse versie van Google Assistent wel eens gaan veranderen. Ook Google's slimme luidspreker, Google Home, kwam eind oktober voor het eerst op de Nederlandse markt. [8] Het apparaat bevat de nieuwe Google Assistent en is daardoor de eerste smart speaker die Nederlands begrijpt. Consumenten hebben een heel andere band met een voice assistant dan met een ander apparaat. Ze spreken over hun apparaat alsof het een mens is.[19] Een gesprek voeren met een apparaat kan zorgen voor een persoonlijke band, iets wat bij de bekende smartphone ontbreekt.

Het commerciële gebruik van deze technologie in ons land is dus gloednieuw. Het onder-

zoek over wat we met deze technologie allemaal kunnen verwezenlijken, zorgt voor veel mogelijkheden.

Doel van dit toegepast onderzoek is om na te gaan of stemgestuurde technologie gebruikt kan worden om kinderen in jeugdcentra te ondersteunen. Het is niet de bedoeling om een apparaat te ontwikkelen dat de begeleider in de toekomst zal vervangen. Het kan het kind helpen ondersteunen wanneer zijn individuele begeleider niet aanwezig is in het jeugdcentrum. Het onderzoek bevat volgende onderzoeksvraag: Hoe kan stemgestuurde technologie helpen met de begeleiding van kinderen in jeugdcentra bij afwezigheid van de individuele begeleider? De onderzoeksvraag bestaat uit twee deelvragen, elk binnen zijn aparte domein. Op orthopedagogisch vlak luidt de vraag: Welke noden zijn er om een stemgestuurde applicatie te ontwikkelen? Vertrekkende vanuit deelvraag 1 gaat het binnen het technologische domein over: Welke spraaktechnologie biedt de meeste mogelijkheid voor een goede oplossing? Een goede oplossing is een oplossing die zo veel mogelijk voldoet aan de vereisten die bepaald zijn in de eerste deelvraag.

A.2 State-of-the-art

De laatste jaren wordt er meer en meer onderzoek gevoerd naar hoe IT en zorg kunnen samenwerken. Zo kwam de overheid met e-health, een elektronisch platform waar alle betrokkenen in de volksgezondheid gegevens kunnen uitwisselen.[3]

Een belangrijk begrip waar onderzoek naar gedaan wordt is Blended Care. Geestelijke gezondheidszorg, ondersteund door IT. Het beste proberen gebruiken van beide werelden. De patiënt krijgt online een behandeling, maar wordt daarnaast ook nog steeds ondersteund door een begeleider. De mix van online en face-to-face therapie heeft al bewezen veel voordelen te hebben. Uit het SROI-verslag van [21] concludeert men dat gemiddeld over vijf jaar, de investeringen voor Blended Care 2,2 keer dit bedrag aan maatschappelijke baten opleveren. De cliënt krijgt de mogelijkheid om zelf aan zijn geestelijke gezondheid te werken tussen sessies met de begeleider door. Wat er voor zorgt dat de cliënt vertrouwen krijgt in het zelfstandig omgaan met zijn gezondheid. [27] Uit het net vermelde onderzoek blijkt ook dat er wel nog meer onderzoek nodig is om te bepalen welke precieze mix de voorkeur krijgt van cliënt en begeleider in bepaalde situaties.

Eén van de grootste behaalde projecten die technologie in de zorgsector toepast is Zora, de zorgrobot. Een slimme robot die nu wordt ingezet in verschillende zorginstellingen. Het wordt beklemtoond dat het niet de bedoeling is om mensen in de zorg te vervangen, maar te begeleiden. [7] Er verschijnen veel positieve berichten over Zora en haar functionaliteiten, maar er is wel zeker één groot hekelpunt, de kostprijs. Die ligt namelijk rond de vijftienduizend euro. [10]

Hoewel het dus bewezen is dat zorg en technologie samengaan, is er nog geen specifiek onderzoek gedaan naar het gebruik van een smart speaker in de jeugdzorg. Dit nicheproduct kan mogelijks toegevoegde waarde bieden aan het grotere geheel. Zo is ook al bewezen dat IT ervoor zorgt dat de drempel om de stap naar hulp te zetten lager wordt door de

anonimiteit die ermee gepaard gaat. [21] Het gebruik van een smart speaker biedt veel mogelijkheden om de zorgsector te verbeteren.

Een andere vraag is of mensen klaar zijn om gebruik te maken van deze technologie. De jeugd staat meer open voor het gebruik van nieuwe technologische middelen, maar ook de begeleiders moeten hier mee akkoord gaan. De heer Buysse, mijn promotor voor deze bachelorproef, is aan een project bezig over faciliterende IT bij individuele begeleidingsgesprekken in de jeugdzorg. Uit navraag bij ex-cliënten en begeleiders bleek dat IT geen oplossing is om het gesprek te vervangen. De mogelijkheden om het gesprek te ondersteunen zijn er wel. Vandaar ook dit specifieke onderzoek naar stemgestuurde technologie als ondersteuning bij afwezigheid van de individuele begeleider.

A.3 Methodologie

Eerst en vooral zal er parallel onderzoek gedaan worden naar een antwoord op de twee deelvragen van het onderzoek. Op orthopedagogisch domein wordt er aan veldonderzoek gedaan. Er wordt zoveel mogelijk gekeken naar wat de vraag is bij begeleider en kind. Wat zijn de vereisten voor de applicatie? Dit kan gedaan worden door kwalitatief onderzoek in de vorm van interviews en/of enquêtes. Die kunnen afgenomen worden bij alle betrokkenen. Dit kan ver gaan, maar er zal vooral gefocust worden op de directe doelgroep, de begeleiders en kinderen in jeugdcentra. Er kunnen ook gegevens verzameld worden door te observeren.

Op technologisch domein wordt er kennis verworven over de verschillende mogelijkheden van spraakgestuurde technologie. Een vergelijkende studie die de voor- en nadelen van de bestaande technologieën afweegt, zal beslissen welke technologie er gekozen wordt. Met de gekozen optie zal er uiteindelijk verder gewerkt worden.

Nadat beide evaluaties worden gematcht, zal er een proof of concept opgesteld worden die zal beslissen of de applicatieontwikkeling wel of niet wordt gestart. Als de applicatie mag ontwikkeld worden zal er dikwijls naar feedback van de directe betrokkenen gepolst worden. Er kan geobserveerd worden hoe kinderen voorlopige versies van de applicatie beleven. De bedoeling is om dan tegen het einde van het onderzoek een eerste versie van de applicatie te ontwikkelen waarop later kan worden voortgebouwd.

A.4 Verwachte resultaten

Concrete resultaten zijn moeilijk te voorspellen aangezien er geen metingen en simulaties worden gedaan in het onderzoek. Uit interviews en observaties kunnen er onvoorspelbare en uiteenlopende resultaten ontstaan waar de onderzoeker zelf nooit zou kunnen opkomen.

Er kan wel al nagedacht worden over welke noden de betrokkenen kunnen hebben. Zo kan er nood zijn aan een smart speaker die een gesprek kan aangaan met het kind wan-

neer hij/zij er naar vraagt. Indien gewenst kan het gesprek opgenomen worden zodat dit later kan beluisterd worden door de begeleider. Het kind geeft beter eerst de toestemming om het gesprek op te nemen zodat het apparaat het vertrouwen van het kind niet schaadt. Het kan een optie zijn om het apparaat te doen reageren op crisismomenten van een kind. Inspelen op het moment dat het kind een crisismoment beleeft, kan een belangrijke verantwoordelijkheid van het apparaat worden. Het kan goed zijn dat het apparaat gegevens uit het verleden kan ophalen om het kind zo goed mogelijk te ondersteunen. De technologie die zal gebruikt worden lijkt vooral te neigen naar de Google Home omdat dit voorlopig de eerste Nederlandstalige Smart Speaker is op de markt.

A.5 Verwachte conclusies

Er wordt verwacht dat uit de interviews, enquêtes en observaties ideeën voortvloeien die de functionaliteiten van de applicatie zullen bepalen. Het zal vast en zeker een uitdaging worden om het apparaat als een vertrouwenspersoon te doen fungeren voor het kind. Kri-tiek kan een hindernis worden tijdens het onderzoek. Orthopedagogen kunnen ervan over-tuigd zijn dat een vertrouwensband alleen kan ontstaan tussen mensen. Dergelijke perso-nen kunnen weigerachtig staan tegenover het gebruik van technologie in hun domein. Dat er tegenstanders zijn bewijst ook het bestaan van de website www.zorgictzorgen.nl. Toch wordt er verwacht dat er een goede samenwerking tussen de IT'er en de orthopedagoog zal ontstaan en er een eerste productversie zal ontwikkeld worden tegen het einde van het onderzoek. Een verwachting in de toekomst, na de bachelorproef, is dat een instantie de ontwikkeling van de applicatie verder in handen neemt. In de eerste plaats denk ik dan aan Hogeschool Gent, maar er zijn meerdere mogelijkheden. Er kan bijvoorbeeld contact opgenomen worden met het zorglab van Vives. Het zorglab verdiept en verbreedt de ex-pertise over zorgtechnologie en interprofessioneel samenwerken en vertaalt deze kennis via kennisvalorisatie naar eindgebruikers, zorgverleners, bedrijven en het onderwijs. [25]

B. Stappenplan proefafname

De verschillende stappen die de onderzoeker heeft gevolgd tijdens elke afname van de proef zijn genoteerd en afgeprint. Hierdoor kon de onderzoeker deze steeds bij zich houden om zeker niet af te wijken van het vaste plan. Het stappenplan is opgenomen in deze bijlage.

B.1 Deel 1

- Vraag aan de deelnemer om de introductietekst te lezen over stemgestuurde assistenten.
- Vraag aan de deelnemer om de algemene vragen op de eerste pagina van het formulier in te vullen.
- Vijns dat de deelnemer drie Engelse commandos zal geven aan een stemgestuurde assistent. Zeg dat de assistent klaar staat op de laptop en dat je terwijl het commando ook opneemt. Plaats de laptop rechtover de deelnemer en vertel hem dat hij mag voorover buigen terwijl hij het commando geeft.

B.2 Deel 2

- Vraag aan de deelnemer om het eerste Engelse commando te lezen, maar benadruk dat hij het nog NIET mag uitspreken.
- Neem het eerste commando van de deelnemer op, terwijl hij zagezegd tegen de assistent praat.
- Vraag aan de deelnemer om het tweede Engelse commando te lezen, maar benadruk

dat hij het nog NIET mag uitspreken.

- Neem het tweede commando van de deelnemer op, terwijl hij zagezegd tegen de assistent praat.
- Vraag aan de deelnemer om het derde Engelse commando te lezen, maar benadruk dat hij het nog NIET mag uitspreken.
- Neem het derde commando van de deelnemer op, terwijl hij zagezegd tegen de assistent praat.

B.3 Deel 3

- Vertel dat het vreemd is dat hij geen enkel commando heeft gehoord, dus dat we het daarom eens in het Nederlands gaan proberen.
- Vraag aan de deelnemer om het eerste Nederlandse commando te lezen, maar benadruk dat hij het nog NIET mag uitspreken.
- Neem het eerste commando van de deelnemer op, terwijl hij zagezegd tegen de assistent praat.
- Vraag aan de deelnemer om het tweede Nederlandse commando te lezen, maar benadruk dat hij het nog NIET mag uitspreken.
- Neem het tweede commando van de deelnemer op, terwijl hij zagezegd tegen de assistent praat.
- Vraag aan de deelnemer om het derde Nederlandse commando te lezen, maar benadruk dat hij het nog NIET mag uitspreken.
- Neem het derde commando van de deelnemer op, terwijl hij zagezegd tegen de assistent praat.

B.4 Deel 4

- Verduidelijk aan de deelnemer dat hij niet tegen een assistent heeft gepraat, maar dat dit nu wel zal gebeuren.
- Vraag aan de deelnemer de vragen met beschrijving te lezen die hij zal moeten beantwoorden. Verduidelijk dat het gaat over de kwaliteit van de uitspraak van de assistent en de manier waarop hij praat en niet over de inhoud van het antwoord. Het antwoord zal ook op het scherm komen, maar raad aan om vooral heel goed te luisteren.

B.5 Deel 5

- Vraag aan de deelnemer de drie commandos te geven aan Alexa en goed te luisteren naar de antwoorden.
- Vraag aan de deelnemer de vragen over Alexa in te vullen.
- Vraag aan de deelnemer de drie commandos te geven aan Google Assistant en goed te luisteren naar de antwoorden.

- Vraag aan de deelnemer de vragen over Google Assistant in te vullen.
- Vraag aan de deelnemer de drie commandos te geven aan Google Assistant NL en goed te luisteren naar de antwoorden.
- Vraag aan de deelnemer de vragen over Google Assistant NL in te vullen.

B.6 Deel 6

- Vraag aan de deelnemer de algemene vragen op de laatste pagina in te vullen.

C. Stappenplan analyse

Om de huidige resultaten te verkrijgen zijn er instructies uitgevoerd in R, gebruikmakend van R Studio. Wanneer men de analyse opnieuw wilt uitvoeren, kunnen de stappen, opgenomen in deze bijlage, gevolgd worden. Om het overzichtelijk te houden wordt er af en toe verwezen naar externe scripts. Deze zijn te vinden onder de map onderzoek/scripts in de repository beschreven in 0.1.

C.1 De spraaksynthese van de assistenten

C.1.1 Inlezen van csv-file

```
results <- read.csv("[pad naar map]/ spraakkwaliteit stemgestuurde assistenten .  
csv", sep=",")
```

C.1.2 Installeren van de nodige packages

```
Install .packages(eeptools)  
Install .packages(reshape2)  
Install .packages(stringr)  
install .packages(formattable)
```

C.1.3 Voorbereiden van de data

```
Run script dataformat_leeftijden .R
Run script dataformat_vergelijking _ assistenten _boxplots .R
Run script dataformat_vergelijking _ assistenten _barplots .R
Run script dataformat_per_eigenschap .R
Run script dataformat_per_assistent .R
Run script dataformat_ttesten .R
```

C.1.4 Leeftijd van de deelnemers

```
boxplot(leeftijd, main="Leeftijd van de deelnemers", yaxt='n')
axis(side=2, at=seq(0, 100, by = 5))
```

C.1.5 Vergelijking van de assistenten per eigenschap

Boxplot van alle scores t.o.v. de assistenten

```
boxplot(coreResultsLong$score~coreResultsLong$assistant, main='Alle gegeven
scores op de spraakkwaliteit van de assistenten ', xlab="assistenten ", ylab
= 'score op vijf ')
```

Boxplots van de scores per eigenschap van de assistenten

Verstaanbaarheid

```
boxplot(verstaanbaarheid$score~verstaanbaarheid$assistant, main="Gegeven
scores op de verstaanbaarheid van de assistenten ", xlab="assistent ", ylab =
"score")
```

Menselijkheid

```
boxplot(menselijkheid$score~menselijkheid$assistant, main="Gegeven scores op
de menselijkheid van de assistenten ", xlab="assistent ", ylab = "score")
```

Levendigheid

```
boxplot(levendigheid$score~levendigheid$assistant, main="Gegeven scores op de
levendigheid van de assistenten ", xlab="assistent ", ylab = "score")
```

Tempo

```
boxplot(tempo$score~tempo$assistant, main="Gegeven scores op het tempo van de
assistenten ", xlab="assistent ", ylab = "score")
```

Emotionaliteit

```
boxplot(gevoel$score~gevoel$ assistant , main="Gegeven scores op de aanwezigheid  
van gevoel bij de assistenten ", xlab=" assistent ", ylab = "score")
```

Barplots van de scores per eigenschap van de assistenten**Verstaanbaarheid**

```
barplot <- barplot(bpVerstaanbaarheid, main="verstaanbaarheid", xlab="score",  
  col=c(" lightblue ", "red", "yellow"), legend = rownames(bpVerstaanbaarheid),  
  beside=TRUE)  
text(x = barplot, y=bpVerstaanbaarheid, label = bpVerstaanbaarheid, pos = 3,  
  cex = 0.8)
```

Menselijkheid

```
barplot <- barplot(bpMenselijkheid, main="menselijkheid", xlab="score", col=c("  
  lightblue ", "red", "yellow"), legend = rownames(bpMenselijkheid), beside=  
  TRUE)  
text(x = barplot, y=bpMenselijkheid, label = bpMenselijkheid, pos = 3, cex =  
  0.8)
```

Levendigheid

```
barplot <- barplot(bpLevendigheid, main="levendigheid", xlab="score", col=c("  
  lightblue ", "red", "yellow"), legend = rownames(bpLevendigheid), beside=  
  TRUE)  
text(x = barplot, y=bpLevendigheid, label = bpLevendigheid, pos = 3, cex = 0.8)
```

Tempo

```
barplot <- barplot(bpTempo, main="tempo", xlab="score", col=c(" lightblue ", "red"  
  , "yellow"), legend = rownames(bpTempo), beside=TRUE)  
text(x = barplot, y=bpTempo, label = bpTempo, pos = 3, cex = 0.8)
```

Emotionaliteit

```
barplot <- barplot( bpEmotionaliteit , main=" emotionaliteit ", xlab="score", col=  
  c(" lightblue ", "red", "yellow"), legend = rownames(bpEmotionaliteit), beside  
  =TRUE)  
text(x = barplot, y=bpEmotionaliteit , label = bpEmotionaliteit , pos = 3, cex =  
  0.8)
```

C.1.6 Vergelijking van de eigenschappen per assistent

Boxplots van de scores per eigenschap voor één assistent

Alexa

```
boxplot(alexa$score~alexa$eigenschap, main="Gegeven scores op de eigenschappen
van Alexa", xlab="eigenschap", ylab = "score")
```

Google Assistant

```
boxplot(ga$score~ga$eigenschap, main="Gegeven scores op de eigenschappen van
Google Assistant", xlab="eigenschap", ylab = "score")
```

Google Assistant NL

```
boxplot(ganl$score~ganl$eigenschap, main="Gegeven scores op de eigenschappen
van Google Assistant in het Nederlands", xlab="eigenschap", ylab = "score")
```

C.1.7 De gemiddelde score en standaardafwijking van alle eigenschappen per assistent gesorteerd van hoog naar laag

```
attach(coreResultsLong)
mean_sd_scores <- setNames(aggregate(x = score, by=list(eigenschap, assistant ),
FUN = function(x) c(mean = mean(x), sd = sd(x))), c("eigenschap", "
assistant ", ""))
mean_sd_scores
mean_sd_scores_ordered <- mean_sd_scores[order(-mean_sd_scores$mean),]
library( formattable )
formattable (mean_sd_scores_ordered)
```

t.testen

```
t. test ( emotionaliteit _scores_Alexa, emotionaliteit _scores_GANL,alternative = "
greater ", paired = TRUE)
t. test ( emotionaliteit _scores_GA,ttest_scores_GANL,alternative = " greater ",
paired = TRUE)
t. test ( levendigheid _scores_Alexa,levendigheid _scores_GANL,alternative = "
greater ", paired = TRUE)
t. test ( levendigheid _scores_GA,levendigheid _scores_GANL,alternative = " greater ",
paired = TRUE)
t. test ( menselijkheid _scores_Alexa,menselijkheid _scores_GANL,alternative = "
greater ", paired = TRUE)
```

```
t.test(menselijkheid_scores_GA, menselijkheid_scores_GANL, alternative = "greater", paired = TRUE)
t.test(tempo_scores_Alexa, tempo_scores_GANL, alternative = "greater", paired = TRUE)
t.test(tempo_scores_GA, tempo_scores_Alexa, alternative = "greater", paired = TRUE)
t.test(verstaanbaarheid_scores_GA, verstaanbaarheid_scores_GANL, alternative = "greater", paired = TRUE)
```

C.2 De spraakherkenning van de assistenten

C.2.1 Inlezen van csv-file

```
results <- read.csv("[pad naar map]/ spraakkwaliteit stemgestuurde assistenten .csv", sep=",")
```

C.2.2 Voorbereiden van de data

```
Run script dataformat_tabel_alle_teksten.R
```

C.2.3 Een overzicht van de gevormde tekst

```
library(formattable)
formattable(texttable)
```

Bibliografie

- [1] Loutfi Belghmidi. „Google Assistant vanaf nu ook in ons land: hoe werkt het en wat zijn de mogelijkheden?” In: *VRT NWS* (2019). URL: <https://www.vrt.be/vrtnws/nl/2019/05/26/virtueel-assistent-oke-google-vanaf-nu-ook-in-ons-land/>.
- [2] Mathias Brandt. „Smart Assistants Are Getting Smarter”. In: *Statista* (apr 2018). URL: <https://www.statista.com/chart/9580/how-smart-are-smart-assistants/>.
- [3] Datanews, red. *Krijtlijnen voor Vlaamse eHealth uitgezet*. 22 feb 2012. URL: <https://datanews.knack.be/ict/nieuws/krijtlijnen-voor-vlaamse-ehealth-uitgezet/article-normal-283363.html>.
- [4] Peter Decroubele. „EHBO voor iedereen: app van Rode Kruis helpt omstaanders in crisissituaties.” In: *VRT NWS* (apr 2019). URL: <https://www.vrt.be/vrtnws/nl/2019/04/02/ehbo-voor-iedereen-app-van-het-rode-kruis-helpt-omstaanders-in-c/>.
- [5] Michael J. Garbade. „A Simple Introduction to Natural Language Processing”. In: *Becoming Human / Medium* (2018). URL: <https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32>.
- [6] Adam Geitgey. „Machine Learning is Fun Part 6: How to do Speech Recognition with Deep Learning”. In: *Medium* (2016). URL: <https://medium.com/@ageitgey/machine-learning-is-fun-part-6-how-to-do-speech-recognition-with-deep-learning-28293c162f7a>.
- [7] Mieke Grypdonck. *Zora de zorgrobot, een aanwinst?* Red. door Mieke Grypdonck. 29 jan 2015. URL: <http://keepupthespirit.be/zora-de-zorgrobot-een-aanwinst/>.

- [8] Thomas Haenen. *Officieel: Google Home vanaf 24 oktober in Nederland te koop*. Red. door Thomas Haenen. 9 okt 2018. URL: <https://www.androidplanet.nl/nieuws/google-home-release-nederland-3/>.
- [9] IBM. „Pioneering Speech Recognition”. In: *IBM* (2011). URL: <https://www.ibm.com/ibm/history/ibm100/us/en/icons/speechreco/>.
- [10] W.J. Jongejan. *Zora: 15000 euro voor een kwetsbare buiksprekende "zorg"-robot*. Red. door W.J. Jongejan. 15 jun 2016. URL: <https://www.zorgictzorgen.nl/zora-15000-euro-voor-een-kwetsbare-buiksprekende-zorg-robot/>.
- [11] Jason Kincaid. „A Brief History of ASR: Automatic Speech Recognition”. In: *Medium* (jul 2018). URL: <https://medium.com/descript/a-brief-history-of-asr-automatic-speech-recognition-b8f338d4c0e5>.
- [12] Gustavo López, Luis Quesada en Luis A. Guerrero. „Alexa vs. Siri vs. Cortana vs. Google Assistant: A Comparison of Speech-based Natural User Interfaces”. In: *International Conference on Applied Human Factors and Ergonomics*. 2018. URL: https://www.researchgate.net/publication/318161646_Alexa_vs_Siri_vs_Cortana_vs_Google_Assistant_A_Comparison_of_Speech-Based_Natural_User_Interfaces.
- [13] John Markoff. „Computer Wins on Jeopardy!: Trivial, Its Not”. In: *New York Times* (2011). URL: <https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html>.
- [14] Linda Passies. *De voice assistant: dit is je nieuwe beste vriend*. Red. door Linda Passies. 27 jul 2018. URL: <https://www.frankwatching.com/archive/2018/07/27/de-voice-assistant-dit-is-je-nieuwe-beste-vriend/>.
- [15] Melanie Pinola. „Speech Recognition Through the Decades: How We Ended Up With Siri”. In: *pcworld* (nov 2011). URL: https://www.pcworld.com/article/243060/speech_recognition_through_the_decades_how_we_ended_up_with_siri.html.
- [16] Shaifali Puri. „Dragon Systems SPEECH-RECOGNITION SOFTWARE”. In: *Fortune magazine* (1998). URL: http://archive.fortune.com/magazines/fortune/fortune_archive/1998/07/06/244793/index.htm.
- [17] Margaret Rouse en Ed Burns. „natural language processing (NLP)”. In: *TechTarget* (2017). URL: <https://searchbusinessanalytics.techtarget.com/definition/natural-language-processing-NLP>.
- [18] Margaret Rouse en Karolina Kiwak. „Speech recognition”. In: *TechTarget* (dec 2016). URL: <https://searchcustomerexperience.techtarget.com/definition/speech-recognition>.
- [19] Robert Schueler en Pauline van der Wel. *The future of voice assistants in the Netherlands*. Red. door Robert Schueler en Pauline van der Wel. 1 jun 2018. URL: https://www.ipsos.com/sites/default/files/ct/news/documents/2018-06/the_future_of_voice_assistants_in_the_netherlands_v0.3_mg_0.pdf.
- [20] Jesús Seijas. „Into a better Speech Synthesis Technology”. In: *Becoming Human / Medium* (2018). URL: <https://becominghuman.ai/into-a-better-speech-synthesis-technology-29411b64f2a2>.

- [21] Boris Stil, Nino Belleng en Rob Snoeren. *De werkzame principes van Blended Care in het sociaal domein*. Onderzoeksrap. VitaValley, 1 mrt 2016. URL: <https://www.kwadraad.nl/wp-content/uploads/2016/04/25-03-2016-SR0IBlendedCareVerslag.pdf>.
- [22] Nederlandse Taalunie. „Nederlands in nieuwe technologieën en toepassingen”. In: *Nederlandse Taalunie* (jun 2017). URL: <http://taalunieversum.org/inhoud/taal-en-spraaktechnologie>.
- [23] Amrita Sunil Tulshan en Sudhir Namdeorao Dhage. *Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa*. Tech. rap. Sardar Patel Institute of Technology, 2019. URL: https://www.researchgate.net/publication/330165159_Survey_on_Virtual_Assistant_Google_Assistant_Siri_Cortana_Alexa_4th_International_Symposium_SIRS_2018_Bangalore_India_September_19-22_2018_Revised_Selected_Papers.
- [24] Dorien Vervoort. „Hoe werkt spraakherkenning?” In: *Techpulse* (2017). URL: <https://www.techpulse.be/achtergrond/218906/hoe-werkt-spraakherkenningstechnologie/>.
- [25] Vives. *Zorglab: Interprofessioneel Centrum Zorgtechnologie*. Red. door Vives. URL: <https://www.vives.be/nl/labs/zorglab>.
- [26] Vox-Creative. „A Brief History Of Voice Assistants”. In: *The Verge* (2019). URL: <https://www.theverge.com/ad/17855294/a-brief-history-of-voice-assistants>.
- [27] Jobke Wentzel e.a. „Mixing Online and Face-to-Face Therapy: How to Benefit From Blended Care in Mental Health Care”. In: *JMIR Mental Health* (9 feb 2016). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4764785/>.