

Sociale Robots

**Een interdisciplinair onderzoek naar de sociale relatie tussen mens en
robot**



Universiteit Utrecht

2018-2019

Liberal Arts & Sciences

LAS3V11003

Coordinator: Herman Hendriks

09-11-2018

Maria Deetman (5743613)

Cognitieve en Neurobiologische Psychologie

Vakreferent: Dr. David Terburg

Perpe Luijben (5496020)

Kunstmatige Intelligentie

Vakreferent: Dr. Stefan van der Stigchel

Sophie van Kekem (5644577)

Taalkunde

Vakreferent: Prof. dr. Norbert Corver

Abstract.

De rol van sociale robots in de huidige maatschappij wordt onderzocht door middel van interdisciplinair onderzoek. Vanuit Kunstmatige Intelligentie, Taalkunde en Cognitieve en Neurobiologische psychologie is onderzocht welke aspecten nodig zijn om de sociale relatie tussen mens en robot te kunnen verbeteren. De disciplines zijn tot verschillende inzichten gekomen. Wanneer deze naast elkaar worden gelegd kan worden geconcludeerd dat voor een prettige omgang het robots de robot het beste zich zo menselijk mogelijk gedraagt en een zo menselijk mogelijk uiterlijk heeft. Wanneer dieper wordt ingegaan op een mogelijke toepassing hiervan, blijkt dat de interdisciplinaire inzichten leiden tot conflicten. Het *uncanny valley effect* heeft negatieve invloed op de manier waarop een robot wordt waargenomen en wordt veroorzaakt door een onvolledig perfecte uitvoering van de menselijke verschijning. Er wordt daarom geconcludeerd dat er moet worden afgedaan op het menselijke uiterlijk van de robot. Daarentegen moet de robot zich qua gedrag wel zo menselijk mogelijk uiten tegenover de gesprekspartner.

Inhoudsopgave

Introductie	p. 3
Hoofdstuk 1: Emotie in non-verbale communicatie	p. 6
Hoofdstuk 2: Emotie en Taal	p. 15
Hoofdstuk 3: vormgeving van de robot	p. 26
Hoofdstuk 4: Integratie	p. 37
Conclusie	p. 44
Literatuurlijst	p. 46

Introductie

Sociale relaties zijn heel belangrijk voor gezondheid en bij een gebrek aan sociale relaties kan eenzaamheid ontstaan. Eenzaamheid heeft grote invloed op de kwaliteit van leven van een mens en kan de gezondheid verslechteren (Ekwall, Sivberg & Hallberg, 2005). Technologie speelt daarentegen een steeds grotere rol in ons leven en zal dit in de toekomst waarschijnlijk nog meer gaan doen. In dit onderzoek wordt gekeken naar de mogelijkheden van technologie om een sociale functie te verrichten en hiermee eenzaamheid te verminderen en de kwaliteit van leven te verbeteren.

Een eerste keuze die daarbij gemaakt moet worden is of hierbij gebruik gemaakt kan worden van ‘simpele’ technologie als een tablet, of ‘complexere’ technologie zoals een robot. Uit een onderzoek naar het opvolgen van ontspanningsinstructies van Broadbent et al. (2015), is naar voren gekomen dat mensen het prettiger vinden om te interacteren met een robot dan met een tablet. Ook wordt een robot meer vertrouwd dan een tablet (Broadbent et al., 2015). Er is dan ook gekozen om in dit onderzoek te focussen op sociale relaties en de rol die robots daarin kunnen spelen. Dit brengt ons tot de volgende hoofdvraag: Hoe kan de sociale relatie tussen mens en robot worden verbeterd? Er zal worden onderzocht of robots een gelijkwaardige sociale relatie kunnen opbouwen met mensen die vergelijkbaar is met de sociale relatie tussen mensen. Om dit te kunnen onderzoeken is een duidelijke definitie van het begrip sociale relatie van belang.

Een sociale relatie heeft als basis de interacties die mensen met elkaar kunnen hebben. Ensie (2017) illustreert dat sociale relaties op veel verschillende manieren in elkaar kunnen zitten. De relatie die iemand met een winkelmedewerker heeft is niet dezelfde als die met een goede vriend. In dit onderzoek zal een *beter* sociale relatie gedefinieerd worden door een sociale relatie waarin mensen een beter begrip van elkaar hebben en elkaar zo beter kunnen leren kennen. Dit wederzijdse begrip kan alleen tot stand komen als aan bepaalde voorwaarden voldaan wordt. Zo kan er ook een sociale relatie bestaan tussen mens en dier, maar dieren hebben minder of andersoortige cognitieve vaardigheden. Ze missen bijvoorbeeld het complexe taalvermogen dat mensen hebben en empathische en sociale vaardigheden zijn minder ontwikkeld. Hierdoor is de sociale relatie tussen mens en dier minder hoogwaardig.

De opgestelde definitie van sociale relatie dwingt meteen tot het maken van een keuze wat betreft de vormgeving van een robot; van een machine, dier of mens. Omdat wij trachten een sociale relatie te onderzoeken op een zo hoog mogelijk niveau, is ervoor gekozen om de robot op een mens te baseren. Daarbij is het zo dat bij het ontwerpen van een robot geldt dat hoe menselijker een sociale robot is, hoe groter de kans dat de gebruiker verwacht dat deze robot zich menselijk zal gedragen en dus een sociale interactie met de robot zal aangaan (Reuten, Van Dam & Naber, 2018).

Om een *goede* sociale relatie te hebben, zoals deze tussen vrienden, is het van belang dat er wederzijds begrip is. Wederzijds begrip kan niet bestaan zonder empathie. Empathie kan worden gedefinieerd als de mogelijkheid om je te kunnen inleven in een ander en hun emotionele ervaringen te begrijpen (Decety, 2015). Dit hangt nauw samen met het herkennen van emotie in de ander. Empathie is de basis van het begrijpen van je gesprekspartner, en is daarom essentieel in een sociale relatie tussen mensen en daarom ook in de relatie tussen mens en robot. Om deze reden zal worden onderzocht hoe mensen emoties bij elkaar herkennen en hoe dit geïmplementeerd kan worden in robots. De opzet van deze scriptie is interdisciplinair. Dit komt voort uit de complexiteit van de vaardigheden die mensen bezitten om sociale relaties met elkaar aan te gaan. Deze hebben meerdere aspecten die niet vanuit een enkele discipline onderzocht kunnen worden. Zo dient onderzocht te worden hoe sociale interactie tussen mensen werkt, welke rol taal en lichaamstaal hierbij spelen, maar ook welke uiterlijke kenmerken mensen prettig vinden met betrekking tot een sociale gesprekspartner. Op basis van de antwoorden op deze vragen kan worden bepaald aan welke eisen een robot zou moeten voldoen om een sociale rol op zich te kunnen nemen.

De disciplines in dit onderzoek, namelijk Cognitieve en Neurobiologische Psychologie, Taalkunde en Kunstmatige Intelligentie hebben veel raakvlakken met elkaar. Om deze reden zijn de deelvragen zo gekozen dat de disciplines ieder hun eigen bijdrage konden leveren die het meest waardevol zou zijn met betrekking tot de hoofdvraag. Vanuit het perspectief van Cognitieve en Neurobiologische Psychologie wordt niet-talige communicatie tussen mensen geanalyseerd en wordt bestudeerd hoe emotie hierbij wordt overgebracht. Vanuit de discipline Taalkunde wordt communicatie door middel van taal bestudeert en hoe mensen door middel van interactie hun

gesprekspartner beter leren kennen. Hierbij wordt gefocust op hoe emotie en persoonlijkheid uit de taal worden herkend. Beide disciplines leveren inzichten over belangrijke aspecten van sociale communicatie tussen mensen. Deze kennis kan vervolgens worden gebruikt om robots menselijke emoties te laten herkennen. Vanuit Kunstmatige Intelligentie wordt gekeken naar de vormgeving van robots. Onderzocht wordt welke aspecten van deze vormgeving positief kan bijdragen aan de sociale relatie tussen mens en robot. Dit resulteert in de volgende disciplinaire hoofdstukken. In hoofdstuk 1 zal Cognitieve en Neurobiologische Psychologie onderzoeken hoe mensen emoties interpreteren bij elkaar door middel van non-verbale communicatie. Daarna zal in hoofdstuk 2 Taalkunde bestuderen hoe emoties en persoonlijkheid tot uiting komen in taalgebruik. Vervolgens zal in hoofdstuk 3 Kunstmatige Intelligentie onderzoeken hoe een robot moet worden vormgegeven om er een relatie mee op te kunnen bouwen. Tot slot zal de informatie die voortkomt uit de verschillende disciplinaire literatuurstudies bij elkaar gevoegd worden om een interdisciplinaire integratie te creëren die nieuwe inzichten geven met betrekking op de onderzoeksvraag. Wanneer de inzichten uit de disciplines zouden worden geïmplementeerd in een robot, zou het echter kunnen dat bepaalde inzichten niet goed gecombineerd kunnen worden. Om deze conflicten tegen te gaan zal door middel van de methoden uit Repko (2016) een *common ground* worden gecreëerd om daaruit vervolgens een *more comprehensive understanding* te creëren. Dit zal ervoor zorgen dat de kennis die voortkomt uit de literatuurstudies zo volledig mogelijk kan worden toegepast. Samen zullen ze voor het beste begrip zorgen van dit veelzijdige onderwerp.

Hoofdstuk 1: Emotie in non-verbale communicatie

Disciplinair hoofdstuk Cognitieve en Neurobiologische Psychologie door Maria Deetman

Robots zijn al lange tijd een toekomstdroom en zijn te zien in bijvoorbeeld populaire media. Alleen niet al deze robots worden op dezelfde manier weergegeven. Aan de ene kant zijn er logisch redenerende robots zoals C3PO, die te zien is in Star Wars. Aan de andere kant zijn er robots zoals Wall-E (2014) van Disney die de harten heeft veroverd van kinderen en volwassenen over de hele wereld. Ondanks dat C3PO meer op een mens lijkt wanneer gekeken wordt naar het uiterlijk, is Wall-E een robot waar mensen zichzelf beter in kunnen herkennen. Waarom? Omdat Wall-E een menselijke kant laat zien in de vorm van emoties en empathie. Hij is blij wanneer hij een mooie bloem ziet en wordt verliefd nadat hij zich lang eenzaam heeft gevoeld. Het zijn deze sociale vaardigheden die een robot menselijker maken. De capaciteiten van robots komen echter voort uit geprogrammeerde intelligentie die zich aan vaste regels houdt, zoals de robots die op dit moment werkzaam zijn in fabrieken en daar een vaste routine doorlopen. Mensen zijn vaak sceptisch tegenover het idee dat robots een sociale rol kunnen spelen in de maatschappij (Lim, 2015). Het idee van een superslimme, compleet logische robot in onze samenleving vinden mensen onwennig. Het kan een science-fiction idee oproepen waar robots de wereld gaan overnemen, zoals in films als de Terminator (1984). Omdat ze menselijke emoties niet snappen, zouden ze het verschil tussen goed en slecht niet kunnen inzien. De oplossing lijkt simpel; zorg ervoor dat robots onze emoties begrijpen.

Onze emoties maken ons spontaan en onvoorspelbaar, iets wat een robot allesbehalve is. Als we robots een plek in onze samenleving willen geven, zullen ze onze onvoorspelbaarheid en emoties moeten begrijpen, respecteren en erop in kunnen spelen (Thomaz, 2015). Mensen hebben het vermogen om elkaar te begrijpen zonder expliciet hun bedoelingen uit te hoeven spreken. Hierbij is non-verbale communicatie erg belangrijk. Om robots emoties te laten herkennen en begrijpen, kan ter inspiratie gekeken worden hoe deze processen bij mensen te werk gaan. Het vakgebied *developmental robotics* houdt zich hiermee bezig en onderzoekt de ontwikkelingsprocessen en mechanismes van kinderen om deze vervolgens toe te passen in de programmering van

robots (Lim & Okuno, 2014). Dit zou de meest authentieke manier zijn om deze sociale capaciteiten bij een robot te implementeren (Lim, 2015).

In dit hoofdstuk zal daarom de volgende deelvraag worden beantwoord: hoe herkennen mensen emotie uit non-verbale communicatie? Ten eerste zal worden omschreven hoe emoties worden gedefinieerd binnen de cognitieve en neurobiologische psychologie. Ten tweede zal worden beschreven hoe gezichten worden verwerkt en welke hersengebieden hier een rol bij spelen. Ten derde zal worden uitgelegd hoe emoties worden geïnterpreteerd door ze neurologisch te spiegelen. Ten vierde zal de rol van oogcontact in sociale communicatie worden besproken. En tot slot zal een conclusie worden getrokken met betrekking tot de deelvraag. Door het hoofdstuk heen zullen ook mogelijke toepassingen van de opgedane kennis op sociale robots worden aangedragen.

Wat is emotie?

Om te begrijpen hoe emoties herkend worden, zal eerst gekeken moeten worden naar wat emoties precies zijn. Zelfs binnen de cognitieve neurowetenschappen worden emoties op verschillende manieren gedefinieerd. Om emoties niet alleen als iets subjectiefs te zien, maar ook meetbaar te maken, kunnen deze het beste worden vertaald in een combinatie van drie dingen; gevoel, gedrag en fysiologie (Purves et al., 2013). Deze drie onderdelen zullen ieder toegelicht worden.

Ten eerste manifesteren emoties zich in de gevoelens die ze mensen geven. Emoties geven positieve of negatieve waarde aan stimuli uit onze omgeving en kunnen zo zorgen voor motivatie. Het onderscheid tussen emotie en motivatie is soms moeilijk te maken, omdat ze nauw samenhangen. Het verschil is dat emotie de basis vormt om iets als positief of negatief te ervaren, wat onder andere op basis van gevoel, terwijl motivatie het verschil betreft tussen iets willen of niet willen (Rolls, 2005). Een voorbeeld van motivatie is honger, wat inhoudt dat iemand wil eten. Aan de andere kant kan het zien van voedsel blijdschap in iemand naar boven brengen, wat een emotie is. Beide kunnen iemand vervolgens tot actie doen overgaan. Dit brengt ons op het volgende aspect van emotie, namelijk gedrag.

Emoties kunnen qua gedrag worden beschreven als reacties op onze omgeving die flexibeler zijn dan een reflex, maar meer instinctief zijn dan een doordachte handeling

(Adolphs, 2017). Door hun sterke link met gevoel lijken emoties ons soms te weerhouden van logisch nadenken. Emoties gaan echter niet per definitie tegen logica in. Emoties zijn ontstaan vanuit de evolutie en een drang om te overleven. Om dit zo succesvol mogelijk te doen, hebben mensen emoties nodig om snel te kunnen reageren op dingen die mogelijk schadelijk kunnen zijn of om hen juist een goed gevoel te geven tegenover belonende stimuli, om de kans op overleven te kunnen verhogen (Öhman, Flykt, & Esteves, 2001).

Tot slot is er de fysiologie. Emoties hebben zowel innerlijke als uiterlijke fysiologische kenmerken. Innerlijke kenmerken zijn bijvoorbeeld verandering in hartslag, hormoonniveaus, of ademhaling (Purves, et al., 2013). Deze kenmerken zijn vaak onbewust en hebben weinig invloed in de communicatie met andere mensen. Uiterlijke kenmerken zijn echter voor dit onderzoek meer van belang. Hierbij kan worden gedacht aan gezichtsexpressies, zoals een lach, of lichamelijke bewegingen, zoals een fysieke actie of een lichaamshouding. Dit zijn de kenmerken die mensen gebruiken om bij elkaar intentie en emotie af te lezen (Crivelli & Fridlund, 2018). Dat doen ze door gezichtsexpressies in interpreteren (Ekman, 1992), iemands blik te volgen (Frishen, Bayliss, & Tipper, 2007) en zelfs op grond van de opmaak van iemands gezicht persoonlijke eigenschappen toe te kennen (Swaddle & Reiersen, 2002). Op deze manier proberen mensen elkaar beter te begrijpen. Dit is dus ook de manier waarop robots menselijke emoties beter zouden kunnen begrijpen. Robots zullen dit moeten doen op basis van uiterlijke fysiologische kenmerken. In dit onderzoek zal vooral worden gekeken naar emotionele communicatie door middel van het gezicht.

Gezichten

Om emoties uit het gezicht te kunnen aflezen, zullen allereerst gezichten moeten worden gefilterd uit het totale visuele materiaal dat binnenkomt. Mensen laten een voorkeur zien voor gezichten boven anderen objecten en hebben de neiging om erop te focussen wanneer ze in hun gezichtsveld komen. Al kort na de geboorte laten baby's deze voorkeur zien door langer naar gezichten te kijken dan naar andere objecten (Johnson, Dziurawiec, Ellis, & Morton, 1991). Deze voorkeur voor gezichten groeit naarmate een kind ouder wordt, wat toont dat kinderen leren dat het gezicht een belangrijk focuspunt is

(Frank, Vul & Johnson, 2009). Dus ook voor een robot zal het automatisch focussen op gezichten de eerste stap zijn in de herkenning van emotie.

Het interpreteren van een gezicht gebeurt in een aantal stappen. De hersenen hebben aparte gebieden die ieder een belangrijk aspect van een gezicht verwerkt (Ward, 2016; Ishai, 2008). Ten eerste is er de *occipital face area*, die wordt geactiveerd wanneer er een gezicht in het visuele veld beland en wanneer er een verandering plaatsvindt in het gezicht (Pitcher, Walsh, & Duchaine, 2011). Dit gebied ligt binnen de visuele cortex en heeft de meest basale rol in het proces van gezichtsherkenning.

Wanneer het gezicht als gezicht herkend is, wordt het verder verwerkt. Uit een gezicht zijn verschillende dingen af te lezen. Één daarvan is de identiteit van de persoon in kwestie. Hiervoor hebben mensen naast de *occipital face area*, de *fusiform face area* (Kanwisher, McDermott, & Chun, 1997). Hier wordt de structuur van het gezicht gekoppeld aan een representatie van een bekend gezicht in het geheugen. Zo wordt het verband gelegd tussen het visuele beeld dat binnenkomt en de identiteit. Naast gezichten laat dit gebied ook activiteit zien op het gebied van andere expertises. Zo kan het zijn dat iemand met veel kennis over bijvoorbeeld verschillende automerken of vogelsoorten ook activatie laat zien in dit gebied bij het identificeren van auto's of vogels. De expertise die mensen hebben in het linken van gezichten aan identiteit, lijkt het resultaat te zijn van de vele gezichten die men in hun leven te zien krijgen. Door de aangeboren voorkeur die mensen voor gezichten hebben, krijgen ze veel ervaring in het onderscheiden ervan. Zo worden we van jongs af aan experts in het herkennen van gezichten en de *fusiform face area* activeert wanneer we deze kennis moeten gebruiken (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). De ontwikkeling van gezichtsherkenning in robots is al goed op weg, bijvoorbeeld als gekeken wordt naar de mogelijkheid om je telefoon te ontgrendelen door middel van je gezicht.

Hoewel de meeste mensen kunnen worden gezien als experts op het gebied van gezichtsherkenning, zijn er mensen die niet in staat zijn de identiteit van een gezicht te achterhalen. Mensen met de aandoening *prosopagnosia* - ook wel gezichtsblindheid genoemd - herkennen een gezicht wel als een gezicht, maar kunnen de identiteit ervan niet koppelen (Behrmann & Avidan, 2005; Bodamer, 1947). Uit het onderzoek van Tranel, Damasio, & Damasio (1988) bleek dat mensen die niet in staat waren om de

identiteit van bekenden te achterhalen door middel van het gezicht, wel in staat waren om andere informatie af te lezen van gezichten, waaronder geslacht, leeftijd en emotie. Deze aandoening laat zien dat het herkennen van individuele gezichten en de het aflezen van emotionele informatie twee gescheiden processen zijn.

Spiegelen

Het lijkt er dus op dat er een aparte weg is om emotie te herkennen. De cognitieve verwerking van emotieherkenning is echter minder duidelijk aan te wijzen dan die van identiteit. Er zijn verschillende theorieën over hoe de perceptie van emotie precies te werk gaat in de hersenen. Op dit moment zal gefocust worden op de *facial feedback hypothesis*. Deze hypothese is gegrond in het *facial mimicry* proces. *Facial Mimicry* is het proces waarbij mensen de emoties die ze waarnemen bij een ander automatisch spiegelen (Dimberg, 1982). Deze spiegeling gebeurt snel (binnen 400 ms) en op kleine schaal (Dimberg & Thunberg, 1998). Deze reactie gebeurt automatisch en onbewust en mensen kunnen het daarom ook niet tegenhouden (Dimberg et al., 2002). Volgens de *facial feedback hypothesis* is de activatie van de gezichtsspieren als reactie op deze spiegeling een belangrijke basis voor empathisch vermogen (Dimberg & Thunberg, 2012). Volgens deze theorie zien mensen de gezichtsuitdrukking bij een ander en kopiëren de uitdrukking. Mensen interpreteren dan wat hun gezichtsspieren doen en koppelen dat terug aan de bijpassende emotie. Vervolgens wordt vanuit daar ook geanticipeerd hoe de andere persoon zal reageren. Dit een bottom-up theorie van empathie. Er zijn echter ook top-down theorieën ten opzichte van empathie. Deze stellen dat de mimiek die optreedt bij het zien van een emotionele uitdrukking niet voortkomt uit het simpelweg overnemen van de uitdrukking, maar dat de uitdrukking zorgt voor een gelijkmatig gevoel bij de waarnemer, wat vervolgens wordt geuit door middel van dezelfde gezichtsuitdrukking. Het is lastig te oordelen welke van de twee opties het beste weergeeft wat er echt bij mensen gebeurt bij het interpreteren van emotie.

De neurologische processen van *facial mimicry* lijken gevestigd te zijn in het *mirror neuron system* (Likowsky et al., 2012; Iacoboni & Dapretto, 2006). Dit systeem is de collectieve naam van wanneer het zien van een actie dezelfde neurale paden activeert als het zelf doen van die actie. Dit systeem is niet alleen te zien bij emoties, maar is

bijvoorbeeld ook aanwezig bij lichamelijke acties, zoals het oppakken van een object (Iacoboni & Dapretto, 2006). Hieruit volgt dat er niet een enkel hersengebied gespecialiseerd is in het interpreteren van iemand anders zijn gemoedstoestand, maar dat de geactiveerde neurale netwerken geheel afhangen van de emoties die worden geïnterpreteerd. Ondanks dit lijken een aantal gebieden een consistente rol te spelen in het systeem. Drie van deze gebieden zullen kort worden toegelicht om hun rol in het *mirror neuron system* duidelijk te maken. Deze gebieden zijn de *superior temporal sulcus*, de *somatosensory* gebieden en de *orbital* en *ventromedial prefrontal cortex*. De *superior temporal sulcus* activeert wanneer kleine veranderingen in een gezicht zichtbaar zijn, zoals de richting van de blik of de verandering van uitdrukking. Het opmerken van deze veranderingen in expressie is een belangrijke basis voor wanneer de expressie moet worden gespiegeld (Iacoboni & Dapretto, 2006). De emotie die waargenomen wordt, kan worden gekoppeld aan onze eigen ervaringen, zoals te zien is in de *somatosensory* gebieden van de hersenen. Deze gebieden activeren op basis van fysiek gevoel. Gallo et al. (2018) liet zien dat wanneer mensen zich inleven in iemand die pijn heeft, dezelfde *somatosensory* gebieden activeren als wanneer wij zelf pijn ervaren. Op deze manier ervaart de ontvanger in zekere zin hoe de ander zich voelt door de ervaring te spiegelen in zichzelf. Ten slotte hebben de *orbital* en *ventromedial frontal cortex* als rol om de consequenties van acties te voorspellen (Ward, 2016). Deze voorspellende capaciteiten kunnen erop wijzen dat deze gebieden zich bezighouden met het anticiperen van wat de andere persoon gaat doen op basis van hun uitdrukking. Deze verschillende hersengebieden werken hoogstwaarschijnlijk samen met de neurale netwerken van onze eigen emoties en hun gezichtsuitdrukkingen om de emoties die worden herkend in het gezicht na te boosten om zo de innerlijke staat van de ander beter te kunnen begrijpen.

Om robots op dezelfde manier emoties te laten verwerken als mensen, zullen zij deze processen ook moeten doorlopen. Het zal daarom nodig zijn dat robots emotionele processen geprogrammeerd hebben zodat deze kunnen worden geactiveerd wanneer ze een emotie zien bij hun gesprekspartner. Dit proces zal automatisch op moeten treden wanneer robots een emotioneel gezicht zien, aangezien mensen dit ook al vanaf jongs af aan doen. Bij baby's van 4 maanden is het spiegelen van gezichtsuitdrukkingen al aanwezig. Dit spiegelen treedt echter vooral op wanneer de kinderen direct aangekeken

worden (de Klerk, Hamilton, & Southgate, 2018). Dit brengt ons op het volgende belangrijke aspect in sociale communicatie, namelijk oogcontact.

Oogcontact

De meeste mensen kijken in een driehoek naar een gezicht, met de focus op de ogen en de mond. Het lijkt erop dat mensen op deze manier een gezicht scannen, omdat daar de meeste sociale informatie vandaan gehaald kan worden. Uit de “*reading the mind in the eyes*” test van Baron-Cohen et al (2001) blijkt dat mensen veel informatie uit de ogen kunnen halen. In deze test worden plaatjes weergegeven van enkel een paar ogen, waaruit mensen gevraagd worden de emoties af te lezen. Deze test is ontworpen om inlevingsvermogen te testen en kan worden gebruikt om bijvoorbeeld Autisme te identificeren. Autisme is een aandoening die wordt gekarakteriseerd door een gebrek aan sociale cognitie (American Psychiatric Association, 2013). Mensen met deze aandoening hebben daarom vaak moeite om emotie en intentie af te lezen bij anderen. Het automatisch naar de ogen kijken zou hier een grote rol in kunnen spelen. Gemiddeld gezien kijken mensen met autisme namelijk minder snel naar de ogen, maar meer naar bijvoorbeeld het voorhoofd, de wangen of opvallende objecten vlakbij het gezicht (zoals oorbellen) (Purves, et al., 2013).

De amygdala is een van de hersengebied dat een grote rol speelt in het automatisch focussen op de ogen van een ander. Dit bleek uit de studie van Adolphs, Gosseling, Buchanan, Tranel, Schyns, & Damasio (2005). Zij lieten de rol van oogcontact bij emotieherkenning blijken met een patiënt met schade aan de amygdala. De patiënt had door haar aandoening veel moeite met het herkennen van angst in de gezichten van anderen. Het bleek dat zij niet naar de ogen keek van de gezichten die werden weergegeven. Wanneer haar geïnstrueerd werd zich te focussen op de ogen, normaliseerde haar vermogen om angst te herkennen in de stimuli. Deze vindingen laten het belang zien van het maken van oogcontact in de herkenning van emoties en de rol van de amygdala daarin. Om deze reden zal het nodig zijn om bij robots een algoritme te implementeren die ze automatisch naar de ogen van hun gesprekspartner laat kijken, zoals de amygdala dat in zekere mate bij mensen doet.

De ogen zijn niet alleen een belangrijk onderdeel in het herkennen van emotie, maar ook van perspectief en intenties. Dit zijn ook belangrijke aspecten in een sociale relatie. Kinderen leren vlak voor hun eerste verjaardag om iemands blik te volgen (Corcum & Moore, 1998). Wanneer we dit kunnen, volgen we iemands blik onbewust en automatisch (Frishen, Bayliss, & Tipper, 2007), wat ons veel informatie geeft over iemands perspectief en focus. Dit proces heet ook wel *joint attention*. Dit proces gebruiken mensen bijvoorbeeld wanneer zij samenwerken. Zoals aan het begin van dit hoofdstuk vermeld is, zullen robots moeten kunnen inspelen op de spontane natuur van mensen. Het zal dus voor robots belangrijk zijn om ook automatisch naar de ogen van hun gesprekspartner te kijken om te kunnen voorspellen wat deze persoon zal doen en erop in te kunnen spelen. Daarbij zal dit ook andersom een waardevolle bijdrage zijn voor mens-robot relatie. Dat wil zeggen, wanneer robot iets gaan doen, is het prettig als ze kijken naar wat ze zullen gaan doen, zodat de mensen eromheen de acties van de robot kunnen voorspellen (Thomaz, 2015).

Conclusie

Het is op dit moment nog moeilijk om te zeggen hoe mensen precies de emoties van een ander kunnen aflezen. In dit hoofdstuk is uitgegaan van de *facial feedback hypothesis*. Deze stelt dat het spiegelen van iemands emotie de basis vormt van ons empathisch vermogen om iemand gemoedstoestand te begrijpen. De belangrijkste aspecten hierbij zijn de automatische processen die mensen laten zien wanneer zij in contact staan met een ander. Dit zijn de automatische focus op het gezicht en op de ogen. Vanuit daar nemen mensen het perspectief over van de gesprekspartner. Door gezichtsuitdrukkingen te spiegelen in zichzelf, kunnen zij de gemoedstoestand van een ander achterhalen. Daarnaast wordt door het volgen van iemands blik intentie ontdekt.

Voor robots om een sociale relatie aan te gaan met mensen zal het nodig zijn dat zij deze processen ook zullen gebruiken. Door de expressie te herkennen en het te spiegelen zal de bijbehorende emotie moeten worden geactiveerd in de robot. De robot zal op deze manier de link kunnen leggen dat een lach de uiting is van een blijde emotie, ook als deze lach heel klein is. Zo zou een robot miljoenen lachende mensen niet allen als

blij kunnen herkennen maar ook kunnen begrijpen, omdat de robot deze blijde emotie ook zelf bezit.

Hoofdstuk 2: Emotie en persoonlijkheid in talige interactie

Disciplinair hoofdstuk Taalkunde door Sophie van Kekem

Om een sociale relatie op te bouwen is interactie nodig. Zonder enige vorm van interactie zouden mensen immers geen relatie kunnen opbouwen. Interactie kan simpelweg gedefinieerd worden als het uitzenden en ontvangen van informatie. In de interactie tussen mensen vindt echter meer plaats dan alleen dit. Er zijn twee vormen van interactie: directe interactie vindt plaats wanneer personen fysiek bij elkaar in de buurt zijn en indirecte interactie doet zich voor wanneer personen materiële hulpmiddelen of tussenpersonen gebruiken (Ensie, 2017). Tegenwoordig zijn de begrippen ‘direct’ en ‘indirect’ moeilijker van elkaar te onderscheiden doordat de technologie zich steeds verder ontwikkelt en videogesprekken de natuurlijke situatie van een ‘face to face’ gesprek erg dicht kunnen benaderen. Aangezien de beoogde relatie tussen robots en mensen gebaseerd wordt op de relatie die mensen met elkaar hebben is het belangrijk om te onderzoeken hoe interactie ofwel communicatie tussen mensen in elkaar steekt. Om te kunnen communiceren is er een communicatiemiddel nodig. Taal is zo’n middel dat kan worden opgedeeld in geschreven en gesproken taal. Doven communiceren niet met gesproken taal, maar met gebarentaal. In deze scriptie zal dit buiten beschouwing worden gehouden, maar voor de toekomst kan het interessant zijn ook te onderzoeken op wat voor manieren robots kunnen communiceren met dove mensen. Daarnaast wordt er ook gecommuniceerd door middel van lichaamstaal, zoals gezichtsuitdrukkingen en gebaren. Dit communicatiemiddel is in het vorige hoofdstuk besproken.

Behalve dat met taal expliciete boodschappen kunnen worden overgebracht, kan er ook impliciete informatie in taal schuilgaan. Met taal kan niet alleen iets expliciets gezegd worden; ook geeft de manier waarop dingen gezegd worden informatie over de spreker. Het gebruiken van bepaalde woorden is gecorreleerd aan bepaalde persoonlijkheidskenmerken (Mohammed & Kiritchenko, 2013). Zo kunnen verschillende personen dezelfde boodschap anders verwoorden. Stel dat iemand bijvoorbeeld wil dat het raam dichtgedaan wordt. De ene persoon zou dit kunnen formuleren als: *Wil je het raam dicht doen?*, terwijl een ander misschien zou zeggen: *Het is wel koud hier*. Emotioneel gedrag is sterk gekoppeld aan de persoonlijkheid van een persoon. Sommige

mensen reageren immers op een andere manier op situaties dan andere mensen. Dezelfde woorden kunnen op verschillende manieren gezegd worden. De ene persoon gebruikt een vrolijke stem om te vragen of het raam dicht mag omdat hij denkt dat dit sneller resultaat oplevert, terwijl een ander misschien een boze stem zou gebruiken omdat hij chagrijnig wordt van de kou. De manier waarop iemand reageert hoeft niet direct aan zijn persoonlijkheid gerelateerd te zijn maar kan ook te maken hebben met de gesprekspartner. Zo zal iemand anders reageren op een kind dan op een volwassene. Ook de emotionele staat waarin iemand zich bevindt, kan invloed hebben op hoe hij reageert. Wanneer iemand een slechte dag heeft zal hij wellicht eerder negatief reageren dan iemand die een hele goede dag heeft. Emoties worden beschouwd als een veranderlijk fenomeen, terwijl persoonlijkheid als een meer constant verschijnsel wordt gezien (Mohammed & Kiritchenko, 2013). Er zijn ook emoties die van langere duur zijn of regelmatig terugkeren. Dit wordt een gemoedstoestand genoemd. Een gemoedstoestand kan soms zo lang duren dat het zelfs een onderdeel van de persoonlijkheid wordt.

Er is veel controverse over de exacte definitie van emotie. Ensie (2015, Emotie, par. 1) definieert het begrip als volgt:

Emotie is een innerlijke beleving of moeilijk te omschrijven gevoel. De vier bekendste emoties noemt men angst, vreugde, boosheid en verdriet. Elk mens toont emoties. Deze kunnen plotseling optreden of tijdens een gebeurtenis geuit worden... Emotie is altijd een combinatie van lichaam, gevoel en cognitie. Het uiten van emoties geeft veelal ook lichamelijke reacties als trillen, gespannen of juist ontspannen spieren, zweten en hartkloppingen.

Volgens een andere definitie omvatten emoties cognitieve evaluaties, subjectieve veranderingen, autonome en neurale opwinding, en impulsen tot actie (Plutchik, 1982). Dit lijkt te impliceren dat cognitie, subjectieve ervaringen en lichamelijke ervaringen een belangrijke rol spelen.

Persoonlijkheid wordt volgens Corr en Matthews (2009) gedefinieerd als karakteristieken die zich uiten in gedrag, cognitie en emotionele patronen. Een

veelgebruikte persoonlijkheidstest is *The Big Five* (Goldberg, 1992). Deze test maakt gebruik van vijf dimensies van persoonlijkheid: (1) extraversie en introversie, (2) meegaandheid en dominantie, (3) orde en wanorde, (4) emotionele stabiliteit en instabiliteit en (5) autonomie en niet-autonomie (Goldberg, 1992).

Uit onderzoek is gebleken dat niet alleen taalgebruik gecorreleerd is aan persoonlijkheidskenmerken maar dat ook emoties die naar voren komen in taalgebruik aanwijzingen kunnen vormen voor persoonlijkheid (Mohammed & Kiritchenko, 2013). Zo zal iemand met een meegaande persoonlijkheid meer gebruik maken van positieve emotiewoorden en zal iemand met de tendens van emotionele instabiliteit meer negatieve emotiewoorden gebruiken. Nu komt emotie niet heel expliciet naar voren in taal. Zoals Corver (2016) al stelde zijn er geen specifieke voor- of achtervoegsels die aan woorden kunnen worden toegevoegd om aan te geven hoe iemand zich voelt terwijl hij spreekt. Als die er wel waren zouden veel misverstanden die optreden in communicatie via technologie als e-mail of chatten voorkomen kunnen worden. Tegenwoordig kan er ook gebruikt gemaakt worden van emoticons wanneer via technologie gecommuniceerd wordt, maar lang niet iedereen gebruikt deze emoticons en ze zijn ook niet altijd dekkend voor de lading van een emotie die overgebracht dient te worden. De centrale vraag die gesteld wordt in dit taalkundige disciplinaire hoofdstuk is als volgt: hoe komen aan de ene kant emoties en aan de andere kant persoonlijkheidseigenschappen tot uiting in taalgebruik? Als duidelijk is hoe mensen deze kenmerken bij elkaar herkennen, kan deze informatie gebruikt worden om robots te leren deze kenmerken ook te herkennen zodat wellicht de sociale relatie tussen mens en robot verbeterd kan worden. De vraag die dit hoofdstuk richting geeft luidt, met andere woorden: hoe kunnen talige coderingen van emotie en persoonlijkheid gebruikt worden om de sociale relatie tussen mens en robot te verbeteren?

Om deze vraag te kunnen beantwoorden, zal gebruik worden gemaakt van literatuuronderzoek. Eerst zal in paragraaf 1 gedefinieerd worden wat emotie in taalgebruik precies inhoudt. Daarna zal in de volgende twee paragrafen gekeken worden naar de manier waarop emotie en persoonlijkheid naar voren komen in taal en hoe dit geanalyseerd zou kunnen worden. De gezochte informatie bevindt zich op verschillende niveaus van taalgebruik, maar met name op lexicaal en prosodisch niveau. In paragraaf 2

zal onderzocht worden hoe emoties naar voren komen in taalgebruik en hoe de verschijnselen op prosodisch niveau geanalyseerd zouden kunnen worden. Vervolgens zal in paragraaf 3 onderzocht worden hoe emoties en persoonlijkheid naar voren komen in taalgebruik, maar dan op lexicaal niveau. Tot slot volgt een conclusie waarin zal worden teruggekomen op de disciplinaire deelvraag: hoe komen emoties en persoonlijkheid tot uiting in taalgebruik? Er zal ook een advies worden gegeven over welke kenmerken in taalgebruik het meest bruikbaar zijn om de sociale relatie tussen robots en mensen te verbeteren.

Emotie in taalgebruik gedefinieerd

Emotionele expressie heeft als functie de evaluatieve reactie van een individu te kunnen communiceren naar de sociale omgeving (Scherer, 1995). Zoals al eerder gezegd, komt emotie niet heel expliciet naar voren in taalgebruik. In Foolen (1997) worden verschillende niveaus van taal beschreven waarin emotie terugkomt. Het eerste niveau is fonologisch en betreft met name intonatie (Batliner, in Foolen, 1997). Intonatie is het verloop van de toonhoogte die de onderliggende bedoeling van een uiting weergeeft. Het tweede niveau is morfologisch, het kleinste taalniveau dat een eigen isoleerbare betekenis heeft, en uit zich bijvoorbeeld in het gebruik van verkleiningsuitgangen in veel talen (Dressler & Merlini Barbaresi, in Foolen, 1997). Ten derde is er het lexicale niveau, waarop bijvoorbeeld vloekwoorden en koosnamen aan bod komen (Carpenter, in Foolen, 1997; Jay, in Foolen, 1997; Saviç, in Foolen, 1997). Het vierde niveau is dat van interrupties (Ehlich, in Foolen, 1997; Drescher, in Foolen, 1997) en partikels (Pos, in Foolen, 1997). Het vijfde niveau is syntactisch, en heeft met name betrekking op expressieve zinstypes (Corver, 2016; Elliot, in Foolen, 1997; Radford, in Foolen, 1997; Rosengren, in Foolen, 1997; Maynard, in Foolen, 1997). Het laatste niveau is pragmatisch; waarbij expressief taalgebruik wordt ingezet als waarschuwing, dreiging, felicitatie of condoleance.

Hoe zouden deze uitingen van emoties op verschillende niveaus in taalgebruik gecategoriseerd kunnen worden? Hoe kan worden achterhaald welk aspect op welke emotie wijst? De basisemoties van Ekman zouden hiervoor op het eerste gezicht een goed uitgangspunt zijn. Deze taxonomie van emoties lijkt echter niet gereflecteerd te zijn in de bouw van taal (Corver, 2016). Dat wil zeggen: er zit bijna nooit één pure emotie in een

taaluiting. De *appraisal theory* kijkt anders naar emoties (Ellsworth & Scherer, 2003). Deze kent geen basisemoties, maar stelt dat elke emotie is opgebouwd uit meerdere componenten. Zo kan een emotie bestaan uit negatieve en/of positieve waardeoordelen. Als die theorie klopt, wekt het geen verbazing dat de basisemoties niet terugkomen in taalgebruik (Corver, 2016).

Volgens Fries (1995) zijn er twee dimensies belangrijk in expressief taalgebruik: enerzijds intensiteit en anderzijds positief-negatief affect (Foolen, 1997). Het positief-negatief affect sluit aan op de appraisal theory. Intensiteit is een onderdeel van emotie in het onderzoek van Corver (2016). Hierin wordt intensiteit gelinkt aan het concept *unexpectedness* (onverwachtheid). Zo worden onverwachte positieve dingen positiever gewaardeerd dan verwachte positieve dingen en worden onverwachte negatieve dingen negatiever gewaardeerd dan verwachte negatieve dingen. Er bestaat een positieve correlatie tussen onverwachtheid en de intensiteit van emotie. Op deze twee dimensies van expressief taalgebruik zal hieronder worden teruggekomen bij het bespreken van enkele niveaus waarop taalgebruik emotie laat zien.

Prosodie

Op prosodisch niveau kan informatie uit taalgebruik worden gehaald. “Onder prosodie verstaan we het geheel van eigenschappen van een spraak uiting die niet herleid kunnen worden tot de opeenvolging van klinkers en medeklinkers” (Rietveld & van Heuven, 2009, p. 277). Het woord *regent* kan op twee manieren worden opgevat afhankelijk van waar de klemtoon in het woord wordt gelegd. Klemtoon is een aspect van taal dat zich op het prosodische niveau bevindt. Later in deze paragraaf zullen nog meer aspecten van prosodie aan bod komen.

Wellicht een van de belangrijkste voordelen van het kunnen herkennen van prosodie is het zogeheten *cocktail party effect*. Het *cocktail party effect* is het menselijk vermogen om de luistervaardigheden te focussen op één spreker terwijl er meerdere sprekers of andere bronnen van rumoer in een ruimte aanwezig zijn (Arons, 1992). Dat mensen in staat zijn om één stem uit meerdere stemmen te kunnen oppikken kan het gevolg zijn van karakteristieken van menselijke spraak, het auditief systeem of perceptie en taalverwerking (Arons, 1992). Ook een combinatie van deze aspecten zou een mogelijke verklaring kunnen zijn. Voor dit onderzoek is het belangrijk om op te merken

dat robots hopelijk zo geprogrammeerd kunnen worden dat ook zij één stem in een rumoerige ruimte kunnen herkennen. Dit zal de robot en zijn gebruiker in staat stellen om bijvoorbeeld ook met muziek op de achtergrond met elkaar te spreken en ervoor zorgen dat een gesprek niet stopt wanneer iemand er doorheen praat.

Ook kunnen mensen bij het horen van alleen een spraakopname van een persoon de emotie erin herkennen (Scherer, 1995). In een uitgebreide studie van Scherer (1981) is dit fenomeen onderzocht door acteurs, zowel amateurs als professionals, te vragen om verschillende emoties uit te drukken terwijl zij een standaardzin uitspraken (in Scherer, 1995). Wanneer deze uitingen aan participanten werden voorgelegd om te beoordelen om welke emotie het ging hadden zij dit gemiddeld in 60% van de gevallen goed. Als het percentage correct geïdentificeerde emoties in de uitingen een kwestie van toeval zou zijn geweest, had dit volgens Scherer (1995) ongeveer 12% moeten zijn. Interessant is dat zich onder de te identificeren emoties ook emoties als liefde, trots en jaloezie bevonden die niet tot de fundamentele basisemoties behoren, maar toch geïdentificeerd konden worden (Scherer, 1995). Ook op prosodisch niveau blijkt dat de basisemoties van Ekman niet volstaan om de complexe weergave van emoties in taalgebruik te beschrijven.

De vraag is op grond waarvan mensen emoties herkennen en welke emoties robots zouden moeten kunnen herkennen om een sociale relatie met mensen mogelijk te maken. Ook dit heeft Scherer (1995) onderzocht. Door luidheid, toonhoogte, variatie en reikwijdte in toonhoogte, duur en de structuur van accenten uit echte uitingen te manipuleren werd onderzocht wat het effect van deze aspecten is op de identificatie van emotie in deze uitingen. Van de verschillende aspecten die bestudeerd zijn had reikwijdte in toonhoogte het sterkste effect op de identificatie van emoties. Een kleine reikwijdte was een markeerder voor verdriet of voor afwezigheid van een specifieke houding van de spreker. Een grote reikwijdte was een markeerder voor opwinding, ofwel voor sterk negatieve emoties zoals irritatie of woede, ofwel voor positieve emoties als betrokkenheid of empathische klemtonen (Scherer, 1995). Daarnaast was een hoog spreektempo gecorreleerd met uitingen van blijdschap en was een laag spreektempo gecorreleerd met uitingen van verdriet (Scherer, 1995).

Aangezien aspecten van toonhoogte de belangrijkste aspecten waren die in de studie van Scherer (1995) naar voren kwamen, zal worden ingegaan op de toonhoogte

kenmerken van een aantal specifieke emoties. Bij woede is er een stijging in gemiddelde toonhoogte en gemiddelde luidheid. Volgens sommige studies is een grotere variatie en reikwijdte van toonhoogte ook een markeerder van woede. Daarnaast is het spreektempo vaak hoger wanneer iemand boos is (Scherer, 1995). Bij angst is er sprake van een stijging in de gemiddelde toonhoogte en in de reikwijdte van toonhoogte. Ook hier is het spreektempo vaak hoger dan gemiddeld. Interessant is dat een hogere gemiddelde toonhoogte in sommige studies ook optrad bij zwakkere versies van angst zoals zorgen of zenuwen (Scherer, 1995). Bij blijdschap is er eveneens sprake van een verhoging in de gemiddelde toonhoogte, een grote reikwijdte en variatie van toonhoogte en een verhoogd spreektempo in het taalgebruik (Scherer, 1995). Daarentegen is er bij verdriet juist een vermindering van de gemiddelde toonhoogte en reikwijdte. Ook het spreektempo is lager bij verdriet (Scherer, 1995).

Prosodie is niet alleen belangrijk om individuele verschillen tussen mensen zichtbaar te maken maar ook om sociale cohesie in beeld te brengen (Hymes, 1997). Mensen die veel tijd met elkaar doorbrengen en een goede relatie hebben spreken op ongeveer dezelfde manier. Zo kunnen ze een vergelijkbare prosodie gebruiken, maar ook vergelijkbare woorden en een vergelijkbare zinsopbouw. Om tot een betere sociale relatie tussen robot en gebruiker te komen kan het daarom van belang zijn dat de robot de spreekstijl aan de gebruiker aanpast.

Lexicon

Op het lexicale niveau is het meest voor de hand liggende expressieve taalgebruik het gebruik van emotiewoorden, ofwel vloekwoorden. Volgens Sterkenburg (2007) hebben emotiewoorden een belangrijk betekenisaspect, dat zij hevigheid noemen. Daarbij is het ene vloekwoord krachtiger dan het andere. Dit betekenisaspect komt overeen met het eerder genoemde intensiteit. De eigenschap van intensiteit die kenmerkend is voor expressief taalgebruik, komt niet terug in andere woorden, zoals bijvoorbeeld *rododendron* of *Rotterdam* wat de Bond tegen vloeken graag als vervanging van de *godverdomme* vloek wil aandragen (Sterkenburg, 2007). Vloeken voelt daardoor eigenlijk alleen aan als vloeken wanneer er taboewoorden gebruikt worden, maar niet wanneer er willekeurige woorden gebruikt worden. Wat voor iemand als een taboewoord aanvoelt kan per persoon en per situatie verschillen. Voor de ene persoon is *sodeju* een

taboewoord en voor de ander *godverdomme*. Of het woord aan de eettafel met familie of alleen zonder bijstanders wordt gezegd kan ook uitmaken.

Daarnaast heeft het gebruik van taboewoorden volgens Jay & Janschewitz (2007) nog een andere functie. Het is een manier om fysieke expressie van emoties zoals schreeuwen, slaan of bijten om te zetten in symbolische expressie. Vloeken zou ontstaan zijn vanwege het evolutionair voordeel dat er daardoor minder fysiek geweld voorkwam (Jay & Janschewitz, 2007). Een belangrijke bijdrage van Jay & Janschewitz (2007) aan de kennis over vloekwoorden is het inzicht dat vloeken kan voorkomen in elke emotionele staat, niet alleen bij woede en frustratie. De taaluitingen *Fuck ik verbrand m'n hand* en *Fuck wat ben ik blij om jou te zien* hebben hele andere boodschappen. Zo kan het gebruik van vloekwoorden bij elk onderwerp aangeven dat iemand sterke gevoelens heeft over dat onderwerp, ongeacht de gevoelens waar het om gaat. Intensiteit is een onderdeel van het gebruik van vloekwoorden en het betreffende waardeoordeel kan zowel negatief als positief zijn. Als robots deze informatie kunnen detecteren kunnen zij wellicht beter inspelen op wat de menselijke gebruiker bedoelt. Dit zou de communicatie en daarmee de sociale relatie tussen mens en robot kunnen verbeteren.

Vloekwoorden zijn relatief duidelijke uitingen van emotie. Ook andere woorden kunnen impliciet iets vertellen over zowel de emotionele staat als persoonlijkheidskenmerken van de persoon die ze gebruikt. In Pennebaker, Mehl en Niederhoffer (2003) wordt onderzocht hoe partikels, de lijm tussen zelfstandig naamwoorden en werkwoorden, als markeerders kunnen dienen voor emotionele staat, sociale identiteit en cognitieve denkwijze. Daarbij zijn voornaamwoorden, lidwoorden, voorzetsels, voegwoorden en hulpwerkwoorden van belang. Dankzij technologische ontwikkelingen kunnen deze talige uitdrukkingsmiddelen door middel van tekstanalyse programma's gedetecteerd worden (Pennebaker, Mehl & Niederhoffer, 2003). Het zou interessant zijn als robots deze methoden ook zouden kunnen toepassen op spontane spraak en daarbij zelf een conclusie uit de informatie kunnen trekken en niet hoeven te vertrouwen op menselijke interventie.

Ter illustratie, in het onderzoek van Pennebaker en King (1999) wordt een eerdere versie van het Linguistic Inquiry and Word Count (LIWC) uit Pennebaker, Francis en Booth (2001) gebruikt om linguïstische correlaties met de persoonlijkheid te detecteren.

Hierbij blijkt dat meegaandheid, een van de vijf persoonlijkheidskenmerken van The Big Five persoonlijkheidstest, gekarakteriseerd wordt door het gebruik van meer positieve emotiewoorden en minder lidwoorden. Emotionele instabiliteit daarentegen wordt gekarakteriseerd door meer negatieve emotiewoorden en meer eerste-persoon voornaamwoorden zoals *ik*, *mijn* of *de mijne* (Pennebaker & King, 1999).

Een ander onderzoek naar persoonlijkheidsdetectie in taalgebruik is Mohammed en Kiritchenko (2013). Hierin worden verschillende studies gebruikt om aspecten van woordgebruik te classificeren in termen van de vijf persoonlijkheidskenmerken uit *The Big Five*. Zij doen dit zo dat aan elk persoonlijkheidsaspect een ja- of nee-label toegekend kan worden. Er wordt gebruik gemaakt van drie verschillende bestaande lexicons waaronder het *NRC hashtag emotion lexicon*. Dit is een lexicon die is gecreëerd op basis van 775000 tweets waarin emotiewoorden als hashtags gebruikt worden. Daarnaast wordt ook de *Osgood dimension lexicon* gebruikt. Dit lexicon kijkt naar aspecten van *evaluation*, *potency* en *activity*. Het is ontwikkeld voor de analyse van sentiment, maar Mohammed en Kiritchenko (2013) gebruiken het om persoonlijkheid te detecteren. De *Osgood dimension lexicon* hanteert in vergelijking met de *NRC hashtag emotion lexicon* een grovere manier van classificeren (Mohammed & Kiritchenko, 2013). Als laatste gebruikten Mohammed en Kiritchenko (2013) het *specificity lexicon*. Hierin wordt de specificiteit van woorden die voorkomt uit de inhoud van dat woord vastgelegd. Entiteit is bijvoorbeeld een erg algemeen concept, terwijl balpen een erg specifiek concept is. Het lexicon geeft aan hoe algemeen of specifiek het concept is waarnaar verwezen wordt. In Mohammed en Kiritchenko (2013) wordt zo een lexicon gemaakt met gebruik van WordNet, een lexicale database van de Engelse taal waarin veel semantische informatie over woorden is opgeslagen.

Deze lexicons worden gebruikt om een uitgebreide lijst van taalaspecten bij te houden, waaronder: aantal woorden, woorden per zin, token ratio, woorden langer dan zes letters, ontkenningen, toestemmingen, lidwoorden, voorzetsels, telwoorden, voornaamwoorden (eerste persoon, tweede persoon, derde persoon), emotiewoorden, cognitiewoorden (inzicht, hypothetisch), gevoel- en perceptiewoorden (zien, horen), woorden voor sociale begrippen (praten, vriend), woorden van tijd, woorden van ruimte, woorden van beweging, interpunctie en vloekwoorden (Mohammed & Kiritchenko, 2013). Al deze

aspecten kunnen gebruikt worden om emotie en persoonlijkheid in teksten te ontdekken. Volgens Mohammed en Kiritchenko (2013) worden dit soort programma's al gebruikt om persoonlijkheid te detecteren in blogs en Facebook- of Twitterposts. Voor deze scriptie zou het vooral interessant zijn als deze programma's ook in robots geïmplementeerd zouden kunnen worden. Een belangrijk aspect hiervan zou zijn dat robots ook iets met deze informatie kunnen. Wanneer taal als marker dient voor stabiliteit moet een robot hier ook mee om kunnen gaan. Zo is er misschien iets mis wanneer iemand na steeds tekenen van stabiliteit te hebben vertoond opeens tekenen van instabiliteit gaat vertonen. De robot kan dan doorvragen en naar eventuele bestaande problemen informeren. Als dit soort programma's geperfectioneerd zouden worden zou het zelfs mogelijk zijn dat robots beter worden dan mensen in het detecteren van emoties en atypisch gedrag voor een bepaalde persoonlijkheid.

Conclusie

Als conclusie kunnen we stellen dat er een aantal manieren zijn waarop aan de ene kant emoties en aan de andere kant persoonlijkheid tot uiting komen in taal. Mensen hebben het vermogen om emotionele kenmerken van de gesprekspartner te detecteren. Het zou de communicatie en daarmee de sociale relatie tussen mens en robot ten goede komen als robots dit vermogen ook zouden hebben. Hierbij is het belangrijk dat robots net als mensen in staat zijn om een stem op basis van prosodie te volgen: het zogeheten *cocktail party effect*. Bovendien is het goed voor de communicatie als een robot emoties kan herkennen in prosodie en aan de hand hiervan kan afleiden wat in een bepaalde situatie de beste reactie is. Daarnaast kan op lexicaal niveau worden bijgehouden wanneer en in welke context emotiewoorden als vloekwoorden gebruikt worden. Zo kan worden afgeleid over welke onderwerpen een persoon sterke gevoelens heeft. Er zijn al programma's die emoties en persoonlijkheidskenmerken in tekst kunnen herkennen. Deze programma's zouden gemakkelijk ook geïmplementeerd kunnen worden bij robots. Ten slotte is het van belang dat robots na verloop van tijd dezelfde woorden en spreekstijl gaan gebruiken als hun gebruiker. Dit zal voor een gevoel van sociale nabijheid zorgen, wat de sociale relatie tussen mens en robot kan versterken. Door de geringe grootte van dit disciplinaire hoofdstuk was het helaas niet mogelijk alle niveaus van taalgebruik te onderzoeken op markeerders van emotie en persoonlijkheid. Daarom is gekozen om

alleen deze niveaus uitgebreid te onderzoeken die naar verwachting te implementeren waren in robots: prosodisch en lexicaal.

Hoofdstuk 3: Vormgeving van sociale robots

Disciplinair hoofdstuk Kunstmatige Intelligentie door Perpe Luijben

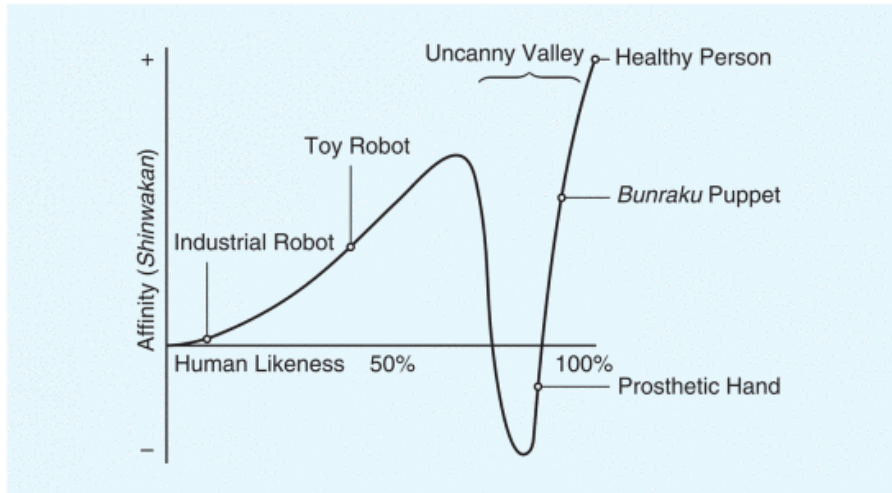
Voor het opbouwen van een sociale relatie met een robot is het van belang dat de gebruiker het prettig vindt om met de robot om te gaan. Daarvoor moet onder andere de vormgeving van de robot, het uiterlijk en de gedragingen, als prettig worden ervaren. In de inleiding van deze scriptie is al benoemd dat er is gekozen dit onderzoek te richten op een sociale robot met het uiterlijk van een mens. De reden hiervoor kan worden afgeleid in het fenomeen *antropomorfisme*: de neiging om niet-menselijke entiteiten menselijke eigenschappen toe te schrijven. Dit is een neiging die zou voortkomen uit de behoefte aan sociaal contact (Broadbent et al., 2013). De vormgeving van een robot bepaalt welke mentale staat, of eigenschappen, eraan wordt toegeschreven – nog voordat er interactie met de robot plaatsvindt. Onbewust worden er sociale heuristieken toegepast die berusten op stereotypering. Wanneer een robot bijvoorbeeld op een mens lijkt, worden er menselijke sociale heuristieken op toegepast; er ontstaan bepaalde verwachtingen over de robot. Hoe meer de robot op een mens lijkt, hoe waarschijnlijker het wordt geacht dat hij zich menselijk gedraagt (Powers & Kiesler, 2006). Wanneer er wordt gestreefd naar sociaal contact op een zo hoog mogelijk niveau, zoals gedefinieerd in de inleiding van deze scriptie, is een robot met een menselijke vormgeving een logische optie. Onder het mom hoe menselijker een robot er uit ziet, hoe menselijker en ‘socialer’ het contact met deze robot zal zijn. Robots die eruit zien als mensen behoren tot de categorie van antropomorfe robots (Li, Rau & Li, 2010). Binnen de categorie antropomorfische robots zijn er dan de subcategorieën van humanoïden, die zijn geïnspireerd op de menselijke vorm, en van androïden, waarbij is getracht deze zo menselijk mogelijk te laten lijken in zowel uiterlijk als gedrag (Ceh & Vanman, 2018). In dit hoofdstuk zal worden onderzocht of, en zo ja in hoeverre, deze menselijke vormgeving inderdaad de juiste keuze is voor de sociale robots. Er worden aspecten van vormgeving besproken die een duidelijke invloed hebben op de sociale relatie tussen mens en robot. Aan de hand van dit hoofdstuk wordt duidelijk hoe robotvormgeving kan bijdragen aan de verbetering van de sociale relaties tussen mensen en robots.

Vormgeving van het gezicht

Het eerste uiterlijke vormgevingsaspect dat wordt onderzocht is het gezicht. Bij sociaal contact tussen mensen is het gezicht erg belangrijk, het gezicht heeft namelijk grote invloed op de eerste indruk die wordt gevormd over een persoon. Het is erg waarschijnlijk dat de vormgeving van een gezicht bij robots ook belangrijk is voor de waarneming van de robot, zeker wanneer deze op een mens lijkt (Broadbent et al., 2013). Bovendien worden veel signalen die van belang zijn voor een goede sociale interactie afgelezen van het gezicht. Bij de vormgeving van het gezicht zijn de ogen erg belangrijk omdat hieruit de levendigheid van de gesprekspartner kan worden afgeleid (Looser & Wheatly, 2010). Levendigheid kan in dit geval letterlijk worden genomen; uit de ogen kan worden afgeleid of men te maken heeft met een levend persoon of niet. Een logisch gevolg hiervan is dat de ogen, en de rest van het gezicht belangrijk zijn voor het bepalen van de menselijkheid van het gezicht. Onderzoek naar gezichtsvormgeving bij robots toont bijvoorbeeld aan dat interactie met de robot met het meest menselijke gezicht als prettigst wordt ervaren (Broadbent et al., 2013). De robot zonder gezicht wordt daarna als prettigst ervaren en als minste het gezicht met vreemde of afwijkende kenmerken. Dit laatste gezicht heeft bijvoorbeeld zilverachtige huid en ogen. De meest menselijke gezichtsvormgeving wordt niet alleen als prettigst maar ook als meest levendig, sociaal en vriendelijk ervaren. Deze robot kreeg de meeste menselijke eigenschappen toegeschreven en wordt daardoor ook menselijker ervaren. Het vreemde gezicht werd zelfs als onprettig ervaren (Broadbent et al., 2013). Voor de vormgeving van een sociale robot zou daarom kunnen worden gekozen voor een zo menselijk mogelijk uiterlijk. Niet alleen is deze vormgeving prettiger, maar wordt deze ook logischer wijs als meer menselijk gezien door de gebruiker. Dit zou bevorderlijk kunnen zijn voor de relatie tussen mens en robot.

De reden dat de meest menselijke vormgeving als prettigst wordt ervaren en waarschijnlijk ook de reden dat de zilverachtige vormgeving juist als onprettig wordt ervaren kan worden verklaard met het *uncanny valley effect*. Dit effect beschrijft een toename van affiniteit voor een robot naarmate de menselijkheid van het uiterlijk ook toeneemt, tot op een bepaald punt, waarbij de gelijkenis met een mens ongeveer 80 tot 85

procent is, dan neemt de affiniteit fors af (Mori, Macdorman & Kageki, 2012). Het *uncanny valley effect* is hieronder weergegeven in figuur 1.



Figuur 1. Het uncanny valley effect (Mori, Macdorman & Kageki, 2012, p. 99)

Hoewel dit onprettige effect waarschijnlijk niet alleen wordt veroorzaakt door de menselijke gelijkheid van het uiterlijk (Rosenthal-Von der Pütten & Krämer, 2012), is het wel waarschijnlijk dat de vormgeving van het uiterlijk er in hoge mate aan bijdraagt. Het onprettige gevoel dat dit *uncanny valley effect* veroorzaakt, komt voort uit een intern conflict bij de gebruiker van een robot. Er is gelijktijdig het gevoel dat het om een mens gaat maar ook om een robot. De entiteit die wordt waargenomen, een mens, en de entiteit die gebruiker weet dat het is, een robot, vallen in verschillende categorieën. De grens tussen deze categorieën wordt echter overschreden en vager (MacDorman & Chattapadhyay, 2016). Een gevolg van deze overschrijding is dat een van de twee categorieën wordt onderdrukt. Het constante onderdrukken van het concept 'mens' kan als gevolg leiden tot afname van empathie voor de robot, iets wat bij het ontwerpen van een sociale robot moet worden voorkomen (Misselhorn, 2009). Het interne conflict dat het *uncanny valley effect* veroorzaakt is menseigen en kan daarom niet zomaar worden tegengegaan. Bij het ontwerpen van een sociale robot moet daarom gekeken worden naar hoe dit effect kan worden vermeden.

Vormgeving van het lichaam

Een goede vormgeving van robot omvat meer dan alleen het menselijk maken van het gezicht. De vormgeving van het gezicht, het gehele lichaam en ook de stem zijn bijvoorbeeld afhankelijk van het geslacht van de robot. Onderzoeken tonen aan dat het geslacht van de robot invloed kan hebben op hoe de robot wordt waargenomen (Tamagawa et al., 2011; Nass, Moon & Green 1997; Eyssell et al., 2012). Het geslacht is daarom een belangrijk aspect in de vormgeving van een menselijke sociale robot.

We kijken allereerst naar welke stem het beste zou werken voor een sociale robot. Mensen leiden cues af uit taal en stemgeluid die aanwijzingen geven over bijvoorbeeld de betrouwbaarheid, aangenaamheid en emotionele status van de spreker. Deze aspecten zijn belangrijk omdat ze ons iets vertellen over de spreker en daarmee bijdraagt aan wederzijds begrip (Tamagawa et al, 2011). Een monotone stem bevat weinig van dit soort cues en wordt over het algemeen als onprettig ervaren (Kuo et al, 2009). Synthetische stemmen worden als minder betrouwbaar en minder levendig ervaren dan menselijke stemmen (Stern, 2008). Mensen vinden menselijke stemmen bij robots prettiger, zelfs als ze genderneutraal zijn (Walters et al, 2008). Daarnaast kan zelfs het accent invloed hebben op de waardering van een robot. In veel gevallen zorgt een lokaal accent er bijvoorbeeld voor dat een robot positiever wordt waargenomen (Tamagawa et al, 2011). Bovendien is er ook het verschil tussen mannenstemmen en vrouwenstemmen. Onderzoek naar het gebruik van deze twee verschillende stemmen door een machine, wanneer alle andere geslachtscues afwezig zijn, levert een aantal interessante inzichten op. Over het algemeen wordt een mannenstem als bekwamer en vriendelijker ervaren dan een vrouwenstem (Nass, Moon & Green, 1997). Ook blijkt dat een vrouwelijk stem in een dominante rol als negatiever wordt ervaren. De mannenstem wordt als bekwamer ervaren wanneer er over ‘mannenzaken’ wordt gesproken, de vrouwenstem als er over ‘vrouwenzaken’ wordt gesproken, ook als zij identieke informatie verschaffen (Nass, Moon & Green, 1997). Bij deze typische ‘mannenzaken’ denkt men aan bijvoorbeeld aan gesprekken over computers, het repareren van apparaten of het bewaken van een huis. Bij typische ‘vrouwenzaken’ wordt bijvoorbeeld gedacht aan gesprekken over zorgtaken, liefde en relaties (Eyssell et al., 2012; Nass, Moon & Green, 1997). Daarnaast blijkt dat een robot met een stem van het eigen geslacht sterker worden geantropomorfiseerd; deze

robots krijgen in een grotere mate een denkvermogen toegeschreven. Ook identificeren mensen zich meer met een robot van hetzelfde geslacht (Eyssell et al., 2012).

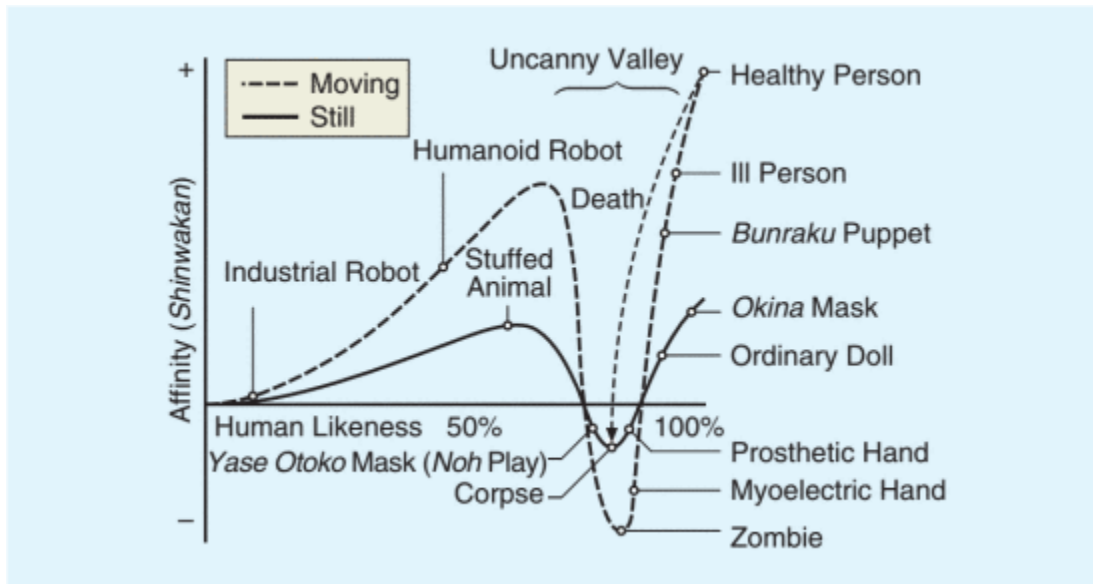
Het effect van geslacht op bekwaamheid komt ook terug wanneer naar de lichamelijke vormgeving in zijn geheel wordt gekeken. In een onderzoek naar bekwaamheid op basis van alleen het uiterlijk wordt een mannelijkere robot bekwaamer geacht wanneer het over ‘mannenzaken’ gaat. Een robot met een vrouwelijk uiterlijk wordt bekwaamer geacht wanneer het over ‘vrouwenzaken’ gaat (Eyssell et al, 2012). Ook de algemenere vorm van de robot kan invloed hebben op de waarneming van de robot. Robots met een groot, potig lichaam worden als bedreigender en mechanisch gezien. Deze lichamelijke vormgeving ondermijnt bovendien de menselijkheid van de robot. (Rosenthal-Von der Pütten & Krämer, 2012).

Deze onderzoeken naar de algemene vormgeving van het lichaam van een robot laten zien dat het van belang is om de robot niet te groot en dreigend te maken. Bovendien kan de juiste keuze van het geslacht van de robot een positieve invloed hebben op hoe deze wordt waargenomen. Deze juiste keuze van het geslacht is echter wel afhankelijk van de taak, bijvoorbeeld een sociale, die de robot moet uitvoeren en de kenmerken van zijn gebruiker.

Gedrag

De menselijkheid van een robot wordt naast de menselijkheid van zijn uiterlijk, ook bepaald door de bewegingen die de robot maakt. Het belangrijk om de bewegingen op de juiste manier te implementeren zodat het *uncanny valley effect* kan worden tegen gegaan. Zowel de aanwezigheid als de afwezigheid van de bewegingen in robots kunnen namelijk zo hun bijdrage hebben op het *uncanny valley effect*. De bewegingen die een robot maakt versterken namelijk de pieken en dalen die in de uncanny valley grafiek worden weergegeven. Wanneer een robot niet beweegt is de affiniteit voor de robot laag. Wanneer de snelheid, versnelling en remmingen van de robotbewegingen die van menselijke bewegingen benaderd neemt de affiniteit voor de robot echter toe. Als de bewegingen van een robot rond de 80 procent gelijkheid vertonen met de bewegingen van een mens, ontstaat er een onprettig gevoel. In dit geval is dit gevoel echter sterker dan

wanneer de robot stil zou staan (Mori, MacDorman & Kageki, 2012). Dit versterkte *uncanny valley effect* wordt weergegeven in figuur 2.



Figuur 2. Het uncanny valley effect en het uncanny valley effect bij beweging. (Mori, Macdorman & Kageki, 2012, p. 99)

Het *uncanny valley effect* wordt niet alleen veroorzaakt door grote bewegingen zoals bijvoorbeeld het optillen van een arm, maar ook door kleinere bewegingen. Als de gezichten van robots niet goed meebewegen met de emoties die zij door middel van spraak uiten, komt dit al snel als vreemd over. Als de gezichtsbevingen ook maar een klein beetje afwijken wordt dit opgemerkt en als onprettig ervaren. De reden het afwijken van deze bewegingen bijdraagt aan het *uncanny valley effect*, is de afwijkingen tussen gezichtsbevingen en spraak als alarmerend worden ervaren mogelijk zelfs als bedreigend (Tinwell et al, 2010). Bovendien is het goed kunnen uitvoeren van de juiste lichamelijke bewegingen nodig om emotionele uitingen te kunnen maken. Zo kunnen robots bijvoorbeeld houdingen aannemen die hun emoties weergeven. Onderzoek toont aan dat mensen uit de houdingen van robots de emoties van de robot kunnen afleiden (Beck et al, 2012). Houdingen die emoties uitdrukken zijn natuurlijke gedragingen voor mensen. Het goed kunnen uitdrukken van emoties door een robot kan daarom bijdragen aan hoe menselijk de robot overkomt. Bovendien zorgt het voor wederzijds begrip en

draagt daarmee bij aan een goede interactie met de robot, een belangrijk aspect bij sociale relaties.

Aanpassing aan de gebruiker

In een relatie tussen mensen is er sprake van wederzijdse afstemming. Bij een goede mens robot relatie zou hier ook sprake van moeten. Voor een goede omgang met robots is daarom aanpassing op de eigenschappen van de gebruiker erg belangrijk. Met welke eigenschappen van de gebruiker moet bij het ontwerpen van een sociale robot rekening worden gehouden?

Eén van die eigenschappen is de culturele achtergrond van de gebruiker. Mensen uit verschillende culturen hebben verschillende gedragsnormen. Deze gedragsnormen uiten zich wanneer er interactie plaatsvindt met een robot. Zo blijkt bijvoorbeeld uit onderzoek dat Koreanen en Chinezen een sociale robot als aardiger ervaren dan een Duitser. Een mogelijke verklaring hiervoor is dat in landen met een individualistische en zogenoemde mannelijke cultuur als bijvoorbeeld Duitsland, robots eerder worden gezien als een gereedschap en niet als een sociale entiteit. Robots met een sociale taak worden er minder gewaardeerd en daarom ook als minder aardig beschouwd (Li, Rau & Li, 2010).

Culturele verschillen kunnen ook invloed hebben op de manier waarop er met robots wordt gecommuniceerd. Zo blijken Amerikanen liever expliciet te communiceren met robots, maar Chinezen liever impliciet (Wang et al., 2010). Dit kan een grote invloed hebben op het succesvol aangaan van een vertrouwensband met een robot. Naast culturele verschillen in gedragsnormen zijn er andere manieren waarop cultuur invloed kan hebben op de relatie tussen mens en robot. Onderzoek toont aan dat mensen computeragents met een gezicht met dezelfde etniciteit als betrouwbaarder, intelligenter, overtuigender en aantrekkelijker beoordelen dan een gezicht met een andere etniciteit (Nass & Moon, 2000). Ook kan cultuur invloed hebben op hoe aardig de robot wordt ingeschat: de 'in-group' robot wordt aardiger gevonden dan de 'out-group' robot. Maar ook gezichtsuitdrukkingen en geslacht dragen mogelijk bij aan dit effect (Rosenthal-Von der Pütten & Krämer, 2012). Duidelijk is dat een robot ten behoeve van een goede relatie moet kunnen inspelen op culturele cues van de gebruiker van de robot.

Naast culturele verschillen zijn er nog andere factoren van invloed. Zo blijkt dat ouderen minder positieve gevoelens hebben over robots dan jongere mensen. Een mogelijke verklaring hiervoor is dat oudere mensen vaak minder kennis hebben van de technologie (Scollitini, Giuliani & Fornara, 2005), een probleem dat zich op den duur niet meer, of althans in mindere mate, zal voordoen, nu men al vanaf jonge leeftijd veel met technologie in aanmerking komt. Er zijn daarnaast nog veel meer factoren die van invloed kunnen zijn, denk bijvoorbeeld aan het verschil tussen introverte en extraverte mensen en het effect hiervan op sociale situaties. Introverte mensen hebben bijvoorbeeld liever korte gesprekken met robots waar extraverte mensen liever langere gesprekken voeren met robots (Ivaldi et al., 2016). Een sociale robot zal dit soort eigenschappen van mensen moeten kunnen achterhalen en hierop inspelen.

Vanwege de talloze culturele verschillen bij mensen maar ook de vele verschillen tussen mensen onderling is een eenduidig antwoord over de vormgeving in dit geval niet mogelijk. Veel culturele verschillen en menselijke eigenschappen vragen om verschillende vormen van omgang. Bij het ontwerpen van een sociale robot kan enkel worden geconcludeerd dat de robot zijn gebruiker moet begrijpen, leren kennen en dat de robot op een manier moeten worden vormgegeven die afgestemd is op bijvoorbeeld de culturele achtergrond van de gebruiker.

Vertrouwen

Hoewel zowel de hierboven besproken uiterlijke kenmerken en de besproken gedragingen en bewegingen van grote invloed zijn op de perceptie van robots is er een factor die nog veel belangrijker is bij een sociale relatie met robots: vertrouwen. Vertrouwen is belangrijk voor de sociale relatie tussen mensen (Kosfeld et al, 2005). Daarom kunnen we aannemen dat deze ook belangrijk zal zijn bij de sociale relatie tussen mensen en robots. Onderzoek van Hancock, Billings en Schaefer (2011) toont aan dat de performance en de betrouwbaarheid van een robot en het aantal fouten dat een robot maakt een grote invloed hebben op het vertrouwen in robots. Deze invloed is zelfs groter dan die van het antropomorfisme, het menselijke uiterlijk van de robot, en zijn persoonlijkheid (Hancock, Billings & Schaefer, 2011). Voor het opbouwen van een goede relatie en vertrouwen is naast twee entiteiten een betrouwbaar onderling

communicatiekanaal nodig. Via dit kanaal kan feedback worden uitgewisseld zodat de entiteiten zich kunnen aanpassen. Uit onderzoek blijkt dat de voorspelbaarheid, betrouwbaarheid, nabijheid en aanpassingsvermogen van de robot de grootste rol spelen bij het opbouwen en behouden van vertrouwen. Wanneer bijvoorbeeld de functies en beperkingen van de robot duidelijk zijn bij de gebruiker draagt dit bij aan het vertrouwen (Hancock, Billings & Schaefer, 2011).

Mogelijk draagt het voorkomen van de bovengenoemde *uncanny valley effect* bij aan het uitblijven of een mindere mate van vertrouwen. Wanneer bepaalde verwachtingen niet worden waargemaakt kan dit leiden tot een onprettig gevoel. Hoe menselijker een robot eruit ziet, hoe menselijker wij verwachten dat de robot is en zich gedraagt. Wanneer een mens een actie uitvoert voorspellen we als het ware de gevolgen. Wanneer die gevolgen niet overeenkomen met de verwachtingen is er sprake van een verwachtingserror (Saygin et al., 2012). Ook hier blijkt dat de *uncanny valley effect* een belangrijke bijdrage levert aan de waarneming van een robot.

Conclusie

In dit hoofdstuk zijn verschillende aspecten onderzocht van de vormgeving van een sociale robot. Het doel hierbij is het ontdekken welke aspecten van vormgeving bij kunnen dragen aan het verbeteren van de sociale band tussen mens en robot. Daarbij is vooral gekeken naar uiterlijke kenmerken en globale gedragingen van de robot. Onderzoek naar de vormgeving van het gezicht toont aan dat robotgezichten die het meest op mensen lijken als prettigst worden ervaren, maar dat bijna volledige menselijke gelijkenis juist leidt tot een negatief gevoel: het *uncanny valley effect*. Ook kan het *uncanny valley effect* worden waargenomen bij bepaalde vormgevingen van de stem van een robot. Het blijkt dat een natuurlijke stem als prettiger wordt ervaren dan een monotone en synthetische stem. Daarnaast blijkt dat ook voor de gedragingen van de robot geldt, menselijker is beter, tenzij het *uncanny valley effect* optreedt. Wanneer deze bewegingen niet overeenkomen met de wijze waarop een mens beweegt wordt dit opgemerkt en als vreemd en onprettig beschouwd.

Er kan worden geconcludeerd dat hoe natuurlijker of hoe menselijker de robot is of lijkt, zowel qua stem, gezichtsvormgeving, lichaamsbouw als qua bewegingen, hoe

beter dit is voor de relatie. Hierbij dient het *uncanny valley effect* zich steeds weer aan. Het is van voor een goede sociale relatie van groot belang dat dit effect wordt vermeden. Wanneer het *uncanny valley effect* zich voordoet schaadt dit namelijk het in de robot vertrouwen en daarmee wordt ook de kans op een goede sociale band kleiner. Het vermijden van dit effect is mogelijk als de vormgeving van de robot zo danig perfect zou zijn dat het onderscheid tussen en mens en de robot niet kan worden gemaakt.

Naast het omzeilen van het *uncanny valley effect* door de juiste vormgeving zijn er echter nog meer factoren mee moeten worden genomen bij het ontwerpen van een sociale robot. Voor een prettige omgang met een robot is zou een volledig op de gebruiker afgestemde robot het beste zijn. Daarvoor moet bijvoorbeeld een keuze gemaakt worden inzake het geslacht van de robot. Aan de ene kant kan worden gekozen voor een vrouwelijke robot, omdat deze bekwamer wordt geacht op het gebied van sociale zaken. Dit is een kwaliteit die mogelijk erg belangrijk is voor een sociale robot. Aan de andere kant is afstemming op het geslacht van de gebruiker ook een goede optie, omdat de gebruiker zich dan meer met de robot identificeert. Ook is er afstemming nodig op de gebruiker op cultureel en etnisch gebied. Robots met eenzelfde culturele achtergrond en zelfs taalaccent worden als leden van de 'in-group' beschouwd en meer gewaardeerd, wat een band bevordert.

Al met al luidt de conclusie dat de vormgeving van robots van erg veel factoren afhankelijk is. Afstemming op de gebruiker en een zo menselijke mogelijke vormgeving met vermindering van het *uncanny valley effect* zijn van groot belang. Wanneer dit tot stand wordt gebracht kan een menselijke gebruiker zich optimaal met de robot identificeren, waardoor een hoogwaardige sociale relatie kan worden gerealiseerd. Bij deze conclusie is het wel raadzaam om je als onderzoeker af te vragen of een hoogwaardige sociale connectie met een robot wel echt mogelijk is. Er zijn veel onderzoeken gedaan naar de vormgeving en gedragingen van robots en hoe mensen daarop reageren, wat de ontwikkelingen op dit gebied positief beïnvloed. Het laat echter ook zien dat de vormgeving van een robot zoals in de conclusie van dit hoofdstuk beschreven, dat wil zeggen een perfecte gelijkenis vertoont met een mens, nog niet mogelijk is. Bovendien is het niet zeker of mensen ooit emotioneel helemaal hetzelfde reageren op robots als op

andere mensen. Hoewel een robot er menselijk kan uitzien en zich menselijk kan gedragen, kan de wetenschap dat het niet om een echte mens gaat al genoeg zijn om een echt volwaardig menselijk sociaal contact in de weg te staan. Pas als mensen niet meer weten dat het om een robot gaat, heeft zo'n hoogwaardige sociale relatie bestaansmogelijkheid.

Integratie

Om de hoofdvraag van deze scriptie – Hoe kan de sociale relatie tussen mens en robot worden verbeterd? – te beantwoorden, is vanuit drie disciplines gekeken naar de volgende deelvragen: (1) Hoe herkennen mensen non-verbaal emoties bij elkaar? (2) Op welke manier komen emoties en persoonlijkheid tot uiting in taalgebruik? En (3) Hoe moet een robot worden vormgegeven om er een relatie mee op te kunnen bouwen?

In hoofdstuk 1 van Cognitieve en Neurobiologische Psychologie is gebleken dat de automatische processen die mensen laten zien wanneer zij in contact staan met een ander de belangrijkste basis is van emotieherkenning. In deze processen wordt automatisch gefocust op gezichten en ogen. De informatie die af te lezen is uit deze focuspunten spiegelen mensen vervolgens in zichzelf. Door gezichtsuitdrukkingen te spiegelen in zichzelf, kunnen mensen van een ander gemoedstoestand, intentie en perspectief achterhalen. Om robots een sociale rol te geven in onze maatschappij is het nodig dat zij deze automatische processen ook zullen gebruiken.

In hoofdstuk 2 van Taalkunde is naar voren gekomen dat er een aantal manieren zijn waarop aan de ene kant emoties en aan de andere kant persoonlijkheid tot uiting komen in taal. Mensen hebben het vermogen om emotionele kenmerken van de gesprekspartner te detecteren vanuit hun taalgebruik. Het zou de communicatie en daarmee de sociale relatie tussen mens en robot ten goede komen als robots dit vermogen ook zouden hebben. Hierbij is het belangrijk dat robots net als mensen in staat zijn om een stem op basis van prosodie te volgen: het zogeheten *cocktail party effect*. Bovendien is het goed voor de communicatie als een robot emoties kan herkennen in prosodie en aan de hand hiervan kan afleiden wat in een bepaalde situatie de beste reactie is. Daarnaast kan er op lexicaal niveau worden bijgehouden wanneer en in welke context emotiewoorden, zoals vloekwoorden, worden gebruikt. Zo kan worden afgeleid over welke situaties en dingen een persoon sterke gevoelens heeft. Er zijn al programma's die emoties en persoonlijkheidskenmerken in tekst kunnen herkennen. Deze programma's zouden ook geïmplementeerd kunnen worden bij robots. Ten slotte is het van belang dat robots na verloop van tijd dezelfde woorden en spreekstijl gaan gebruiken als hun

gebruiker. Dit zal voor een gevoel van sociale nabijheid zorgen, wat de sociale relatie tussen mens en robot kan versterken.

Uit hoofdstuk 3 van Kunstmatige Intelligente blijkt dat voor de vormgeving van de robot een natuurlijke, menselijke vormgeving het prettigst wordt ervaren. Wanneer de vormgeving echter te veel gelijkenis vertoont met mensen, rond de 80 procent, neemt de affiniteit voor robots drastisch af. Dit wordt ook wel het *uncanny valley effect* genoemd. Daarnaast blijkt dat een robot als prettiger wordt ervaren als deze wordt aangepast op de gebruiker op het gebied van onder andere geslacht, culturele achtergrond en etniciteit. Het vermijden van het *uncanny valley effect*, aanpassing op de gebruiker, het voldoen aan verwachtingen en betrouwbaarheid dragen allemaal bij aan het vertrouwen in de robot. Het vertrouwen in de robot is een belangrijke factor voor de sociale relatie tussen mens en robot.

Op basis van de inzichten die gevonden zijn in de disciplinaire hoofdstukken, zou de hoofdvraag beantwoord kunnen worden. De sociale relatie tussen mens en robot kan zou verbeterd kunnen worden door een robot te creëren die niet te onderscheiden is van een mens. Dit zou het *uncanny valley effect* tegengaan en zou de menselijke capaciteiten uit de hoofdstukken Cognitieve en Neurobiologische Psychologie en Taalkunde kunnen implementeren. De robot zou als het ware aan een ultieme Turing Test voldoen, omdat een mens de robot zou kunnen zien en ermee zou kunnen praten zonder dat deze doorheeft dat het om een robot gaat. Op deze manier geloven wij de meest hoogwaardige relatie, zoals gedefinieerd in de inleiding, te kunnen creëren tussen mens en robot. Er zijn echter een aantal conflicten die kunnen ontstaan wanneer de gevonden inzichten daadwerkelijk te geïmplementeerd zouden worden.

Conflicten

Elke deelvraag is ingegaan op een bepaald aspect van sociale communicatie en de mogelijke toepassing daarvan in een sociale robot. Zo heeft elk hoofdstuk een aparte bijdrage geleverd aan het beantwoorden van de hoofdvraag. Wanneer deze inzichten daadwerkelijk zouden worden geïmplementeerd, kunnen deze inzichten met elkaar conflicteren. Om een vollediger antwoord te kunnen geven op de hoofdvraag, zullen deze conflicten moeten worden opgelost. Op deze manier kan een *more comprehensive understanding* ontstaan. In deze integratie zullen deze conflicten worden gevonden en

opgelost met behulp van de integratiemethoden van Repko (2016). Uiteindelijk wordt de hiervoor getrokken conclusie mogelijk herzien en aangepast om volledig antwoord te kunnen geven op de hoofdvraag.

Het eerste conflict treedt op tussen de inzichten uit het hoofdstuk van Taalkunde en de assumpties in de hoofdstukken van Kunstmatige Intelligentie en Cognitieve en Neurobiologische Psychologie. Vanuit Taalkunde blijkt dat er bepaalde tekstanalyseprogramma's bestaan die wellicht geïmplementeerd zouden kunnen worden in robots. Hierdoor kunnen robots persoonlijkheidskenmerken afleiden uit het taalgebruik van de gebruiker. Aan de ene kant is dit positief, omdat dit zou betekenen dat een robot zijn gedrag kan afstemmen op de persoonlijkheid van de gebruiker. Aan de andere kant is het zo dat het onzeker is of dit systeem vergelijkbaar is met de manier waarop mensen bij elkaar persoonlijkheidskenmerken afleiden. Als dit systeem niet hetzelfde is, kan dat betekenen dat de robot zich op een niet-menselijke manier gedraagt. Dit zou ertoe kunnen leiden dat de robot tot andere conclusies komt dan wanneer een mens persoonlijkheidskenmerken afleidt. Dit is niet per definitie fout, maar het vormt wel een conflict met assumptie dat natuurlijk of menselijker beter is, wat een uitgangspunt is van zowel de disciplinaire hoofdstukken van zowel Kunstmatige Intelligentie als Cognitieve en Neurobiologische Psychologie.

Dit zal worden geïllustreerd aan de hand van een voorbeeld. Persoon 1 concludeert op basis van de woordkeuze van persoon 2, dat persoon 2 een optimistische persoonlijkheid heeft. Vervolgens kan het zijn dat persoon 1 zijn gedrag daarop aanpast. Als persoon 1 wordt vervangen door een robot met een taalanalyse-programma, kan het zijn de robot de tekst van persoon 2 op een andere manier analyseert dan persoon 1. Als het programma persoonlijkheid baseert op bijvoorbeeld syntax, in plaats van woordkeuze, kan het programma tot de conclusie komen dat persoon 2 niet optimistisch, maar juist depressief is, ondanks dat persoon 2 positieve woorden gebruikt om dit te maskeren. Het is hierbij niet per se van belang welke persoonlijkheid persoon 2 daadwerkelijk heeft, omdat dit iets is wat zowel persoon 1 als de robot niet zeker kunnen weten. Het komt erop neer dat de robot op een andere manier, en mogelijk op een andere conclusie komt dan persoon 1. Wanneer de robot zijn gedrag vervolgens aanpast op een conclusie die persoon 2 niet verwacht, kan de interactie een onnatuurlijke wending nemen. Dit kan een

negatieve invloed hebben op de sociale relatie en het vertrouwen. Vertrouwen is zeer belangrijk in de relatie tussen mens en robot, zoals blijkt uit het hoofdstuk van Kunstmatige Intelligentie.

De oplossing van dit conflict kan worden bewerkstelligd door het toepassen van een *multilevel causal integration* (Repko, 2016). Hierbij wordt het inzicht dat menselijk als beter wordt ervaren gesplitst naar twee niveaus. Het eerste niveau is het uiterlijk van de robot die wordt gezien door de menselijke gesprekspartner, waarvoor geldt: natuurlijk is beter, zoals blijkt uit het hoofdstuk van Kunstmatige Intelligentie. Het tweede niveau heeft betrekking tot de innerlijke processen van de robot. Deze vormen de basis van de uiterlijke gedragingen van de robot. De assumptie van Cognitieve en Neurobiologische Psychologie is dat niveau twee ook op basis van menselijke gedragingen tot stand moet komen. Om het conflict op te lossen, zal de assumptie van Cognitieve en Neurobiologische Psychologie moeten worden aangepast.

Om het inzicht uit het hoofdstuk van Taalkunde toe te kunnen passen zal het in niveau twee ook mogelijk moeten zijn om niet menselijke processen de basis te laten vormen van het gedrag van de robot. Niveau twee is daarmee uitgebreid. In plaats van de assumptie van Cognitieve en Neurobiologische Psychologie dat het op een menselijke manier moet, kan het dan ook door middel van niet menselijke processen. De assumptie dat natuurlijk beter is op niveau één moet wel opgaan, zodat deze voldoet aan de verwachtingen die mensen van elkaar hebben in sociale communicatie. Het maakt hierbij in zekere zin niet uit welke assumptie bij niveau twee gebruikt. Dit neemt niet weg dat niveau twee niet gebaseerd kan zijn op menselijke processen, maar dit hoeft geen criteria te zijn.

Om terug te komen op de theorie van Taalkunde, zou deze nu wel geïmplementeerd kunnen worden in de robot. Hierbij is het wel van belang de robot conclusies kan trekken die overeenkomen met de conclusies die mensen trekken, ongeacht hoe deze conclusies tot stand zijn gekomen. Het is echter nog niet duidelijk of de uitkomsten van tekstanalyse programma's overeenkomen met de analyses die mensen zelf doen. Om hierachter te komen is meer onderzoek nodig naar de implementatie van dit soort programma's en de reactie van mensen daarop.

Een tweede conflict dat naar voren komt, is tussen theorieën van Kunstmatige Intelligentie en Cognitieve en Neurobiologische Psychologie. De eerste theorie is *De facial feedback hypothesis* van Cognitieve en Neurobiologische Psychologie. *De facial feedback hypothesis* beschrijft dat het nodig is om gezichtsexpressies te spiegelen om iemand anders te kunnen begrijpen. De tweede theorie is het *uncanny valley effect* uit de Kunstmatige Intelligentie. Deze theorie stelt dat het uiterlijk van een robot als prettiger wordt ervaren naarmate deze meer op een mens gaat lijken tot op een bepaald punt; wanneer een robot bijna geheel op een mens lijkt, maar nog niet helemaal. Dit onwennige gevoel zou echter kunnen worden vermeden als de robot perfect op een mens lijkt.

De theorieën staan niet zozeer tegenover elkaar, maar in het ontwerpen van een daadwerkelijke robot ontstaat er een conflict. Om aan beide theorieën te voldoen zou de robot een gezicht moeten hebben dat perfect gezichtsuitdrukkingen kan spiegelen. Er is echter een grote kans dat deze perfecte uitvoering nog niet mogelijk is, omdat dit heel nauw komt en de technologie nog niet zo ver lijkt te zijn. Daaruit volgt dat het spiegelen op een manier zal moeten worden gerealiseerd die niet perfect is. Het *uncanny valley effect* komt op dit punt naar voren; omdat de expressies niet perfect worden uitgevoerd, wordt dit gemakkelijk opgemerkt en leidt dit tot een onprettig gevoel.

Om *common ground* te creëren is het nodig om theorie-extensie toe te passen op de *facial feedback hypothesis* (Repko, 2016). Volgens de *facial feedback hypothesis* is het nodig om de gezichtsuitdrukking van de gesprekspartner te spiegelen om hun innerlijke emotie te kunnen begrijpen. Deze hypothese is echter gebaseerd op menselijke interactie. Voor mensen is het voelen van een emotie en de bijpassende gezichtsuitdrukking onbewust en automatisch verbonden, zoals is uitgelegd in het hoofdstuk van Cognitieve en Neurobiologische Psychologie. Voor de toepassing in robots zou deze theorie aangepast kunnen worden, zodat de gezichtsuitdrukking van de robot, oftewel de uitingen van de emotie, los staat van de innerlijke ervaring van de emotie binnen de robot. Hieruit volgt de *more comprehensive understanding* dat de robot de emotie van de gesprekspartner wel intern moet spiegelen, maar dat de bijpassende gezichtsuitdrukkingen hierbij niet hoeven worden weergegeven in zijn fysieke vormgeving. Door de extensie van de *facial feedback hypothesis* hebben we het conflict opgelost door het negatieve gevoel zoals beschreven in het *uncanny valley effect* te

vermijden, door als het ware aan de andere kant van de *uncanny valley* te gaan zitten. Op deze manier kunnen beide theorieën worden toegepast in het ontwerpen van sociale robots.

Assumpties als basis van conflict

Er zijn naast de twee conflicten hierboven niet meer conflicten te vinden in de theorieën en concepten die naar voren zijn gekomen in ieder disciplinair hoofdstuk. Dit komt doordat de disciplines naar verschillende kanten van het probleem hebben gekeken en omdat hun conclusies zijn gebaseerd op dezelfde methoden. Bovendien zijn de gebruikte assumpties die ten grondslag liggen aan onze kennis vrijwel hetzelfde. Allereerst zijn er de epistemologische assumpties. Cognitieve en Neurobiologische Psychologie, Taalkunde en Kunstmatige Intelligentie hebben alle drie bijna geheel dezelfde assumpties over het vergaren van kennis. De drie disciplines maken over het algemeen gebruik van dezelfde methoden wanneer een theorie getest moet worden. De kennis die hieruit voortkomt zal daarom niet snel in conflict staan met elkaar, omdat de disciplines het eens zijn over de wanneer die conclusies mogen worden getrokken.

Over het algemeen geldt ook voor de *value-laden* assumpties dat er weinig conflict tussen de disciplines voorkomt. *Value-laden* assumpties zijn de assumpties die bepalen wanneer een onderwerp de moeite waard is om te onderzoeken. Het overgrote deel van de kennis die Cognitieve en Neurobiologische Psychologie en Taalkunde vergaren is op basis van fundamenteel wetenschappelijk onderzoek. Waar deze twee disciplines in verschillen is het gebied waarin ze deze fundamentele kennis proberen te vergaren. Het geeft inzicht in hoe de wereld in elkaar zit en specifiek bij deze vakgebieden hoe mensen in elkaar zitten. De kennis wordt vergaard omwille van de kennis. De theorieën en concepten die bij deze fundamentele onderzoeken naar voren komen, zijn daarbij ook voor een groot deel gebaseerd op dezelfde methoden, zoals eerder is beargumenteerd. Het is niet waarschijnlijk dat deze theorieën of concepten snel met elkaar in conflict komen te staan. De disciplines zullen het namelijk over het algemeen met elkaar eens wanneer deze conclusies kunnen worden getrokken. Deze kennis kan vervolgens worden toegepast wanneer dat nodig is, bijvoorbeeld in het

huidige onderzoek naar sociale relaties. Pas bij de toepassing van deze kennis, kan het voorkomen dat er assumpties worden gedaan die conflicteren, zoals is gebleken bij het eerste conflict.

Van de drie disciplines maakt Kunstmatige Intelligentie het meest gebruik van toegepast wetenschappelijk onderzoek, en gebruikt daarbij onder andere de kennis die is opgedaan in Cognitieve en Neurobiologische Psychologie en Taalkunde. Dit combineert het onder andere met logica, informatica en filosofie. De kennis die uit Kunstmatige Intelligentie naar voren komt, zal niet snel in conflict komen te staan met Cognitieve en Neurobiologische Psychologie of Taalkunde, omdat deze disciplines mede ten grondslag liggen aan de kennis die Kunstmatige Intelligentie produceert. Hierdoor worden conflicten over assumpties vermeden.

Conclusie

In deze scriptie is onderzocht hoe de sociale relatie tussen mens en robot kan worden verbeterd. Vanuit de disciplines Cognitieve en Neurobiologische Psychologie, Taalkunde en Kunstmatige Intelligentie zijn verschillende aspecten van deze sociale relatie onderzocht. Door de inzichten uit deze disciplines te combineren, is een eerste conclusie getrokken. In deze conclusie zou een verbeterde interactie gecreëerd kunnen worden tussen mens en robot wanneer de robot in zijn uiterlijk en gedrag niet te onderscheiden is van een mens. Op deze manier zouden robots moeten kunnen slagen voor een ultieme Turing Test.

In de integratie van de verschillende disciplines is echter duidelijk geworden dat deze eerste conclusie niet haalbaar is. De conclusie zal daarom moeten worden aangepast. De complexe uiterlijke gedragingen van mensen, zoals *facial feedback*, maken het moeilijk om de vormgeving van de robot perfect op een mens te laten lijken. Dit komt vooral doordat de technische methoden nog niet verfijnd genoeg zijn om een perfect menselijk uiterlijk na te maken. Het *uncanny valley effect* stelt dat het echter niet slim is om te streven naar een perfect menselijk uiterlijk wanneer dit niet geheel haalbaar is. Daarom wordt hier gekozen om de menselijke gelijkenis van het uiterlijk minder dan 80 procent te laten zijn. Hierdoor wordt het *uncanny valley effect* vermeden. Dit betekent dat er ingeleverd wordt op de hoogwaardigheid van de sociale relatie, waarin een hoogwaardige relatie wordt beschouwd als die tussen twee mensen. Dit lijkt op dit moment de beste optie. We geven de voorkeur aan een minder-menselijk-uitziende robot die als prettiger wordt ervaren boven het streven naar mogelijk onhaalbare perfectie.

Desondanks zal wel worden gestreefd naar een sociale robot die emoties op een natuurlijke manier interpreteert. Om alle kennis die is opgedaan over de interpretatie van emotie te gebruiken, hebben we de assumptie “natuurlijk is beter” aangepast. Op deze manier zal deze beter aansluiten bij alle disciplines. De uitingen van de robot kunnen het beste natuurlijk zijn, maar het proces waardoor deze uitingen tot stand komen hoeven niet per definitie hetzelfde te zijn als bij mensen. Zo kunnen de inzichten die naar voren zijn gekomen bij de deelvragen worden geïmplementeerd.

Hoe sociale robots precies kunnen worden geïmplementeerd in onze maatschappij is een belangrijk vraagstuk voor nu en in de toekomst. De verschillende disciplines in

deze scriptie hebben inzichten gegeven in hoe de sociale relatie tussen mens en robot kan worden verbeterd. Deze inzichten zijn gecombineerd en kunnen worden geïmplementeerd in het ontwerpen van sociale robots. De mogelijke conclusies met betrekking tot dit onderwerp zullen echter meer moeten zijn dan de som der delen. Na het integreren van de verschillende disciplines in deze scriptie hebben we de conclusie moeten aanpassen. Dit geeft het belang van het samenbrengen inzichten van disciplines en interdisciplinair onderzoek weer. Het samenwerken van de verschillende vakgebieden zal het uiteindelijke toepassing van hun kennis waardevoller maken. Desondanks zal meer onderzoek nodig zijn om een nog completer beeld te geven van hoe de sociale relatie tussen mens en robot er precies uit zal zien. Verschillende disciplines, waaronder Taalkunde, Kunstmatige Intelligentie en Cognitieve en Neurobiologische Psychologie, maar mogelijk ook andere, zullen moeten blijven samenwerken om steeds meer inzichten te creëren om een compleet antwoord te vinden op complexe vraagstukken zoals die in deze scriptie.

Literatuurlijst

Inleiding

- Decety, J. (2015). The neural pathways, development and functions of empathy. *Current Opinion in Behavioral Sciences*, 3, 1–6. <https://doi.org/10.1016/j.cobeha.2014.12.001>
- Ensie. (2017). Sociale relatie. Grondbeginselen der Sociologie. Geraadpleegd op <https://www.ensie.nl/grondbeginselen-der-sociologie/sociale-relatie-m-weber>
- Ekwall, A. K., Sivberg, B., & Hallberg, I. R. (2005). Loneliness as a predictor of quality of life among older caregivers. *Journal of advanced nursing*, 49, 23-32.
- Mann, J. A., MacDonald, B. A., Kuo, I. H., Li, X., & Broadbent, E. (2015). People respond better to robots than computer tablets delivering healthcare instructions. *Computers in Human Behavior*, 43, 112-117.
- Reuten, A., van Dam, M., & Naber, M. (2018). Pupillary responses to robotic and human emotions: the uncanny valley and media equation confirmed. *Frontiers in psychology*, 9, 774.
- Repko, A. F., Szostak, R., & Buchberger, M. P. (2016). *Introduction to interdisciplinary studies*. Sage Publications.

Hoofdstuk 1: Cognitieve en Neurobiologische Psychologie

- Adolphs, R. (2017). How should neuroscience study emotions? By distinguishing emotion states, concepts, and experiences. *Social cognitive and affective neuroscience*, 12(1), 24-31.
- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, 433(7021), 68.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42(2), 241-251.
- Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: face-blind from birth. *Trends in cognitive sciences*, 9(4), 180-187.
- Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005, August). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on* (pp. 708-713). IEEE.
- Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental psychology*, 34(1), 28.
- Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in cognitive sciences*.
- Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology*, 19(6), 643-647.

- Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian journal of psychology*, 39(1), 39-45.
- Dimberg, U., & Thunberg, M. (2012). Empathy, emotional contagion, and rapid facial reactions to angry and happy facial expressions. *PsyCh Journal*, 1(2), 118-127.
- Dimberg, U., Thunberg, M., & Grunedal, S. (2002). Facial reactions to emotional stimuli: Automatically controlled emotional responses. *Cognition & Emotion*, 16(4), 449-471.
- Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, 6(3-4), 169-200.
- Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, 110(2), 160-170.
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133(4), 694-724.
- Gallo, S., Paracampo, R., Müller-Pinzler, L., Severo, M. C., Blömer, L., Fernandes-Henriques, C., ... & Avenanti, A. (2018). The causal role of the somatosensory cortex in prosocial behaviour. *eLife*, 7, e32740.
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform face area increases with expertise in recognizing novel objects. *Nature neuroscience*, 2(6), 568.
- Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, 7(12), 942.
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1-2), 1-19.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of neuroscience*, 17(11), 4302-4311.
- de Klerk, C. C., Hamilton, A. F. D. C., & Southgate, V. (2018). Eye contact modulates facial mimicry in 4-month-old infants: an EMG and fNIRS study. *Cortex*, 106, 93-103.
- Kraaijevanger, E. J., Hofman, D., & Bos, P. A. (2017). A neuroendocrine account of facial mimicry and its dynamic modulation. *Neuroscience & Biobehavioral Reviews*, 77, 98-106.
- Likowski, K. U., Mühlberger, A., Gerdes, A., Wieser, M. J., Pauli, P., & Weyers, P. (2012). Facial mimicry and the mirror neuron system: simultaneous acquisition of facial electromyography and functional magnetic resonance imaging. *Frontiers in human neuroscience*, 6, 214.
- Lim, A., & Okuno, H. G. (2014). The Recipe for Empathy. *Springer International Journal of Social Robotics*, 7(1), 35-49. DOI: 10.1007/s12369-014-0262-y
- Lim, A. (2015). Robots, Emotions & Empathy | Angelica Lim | TEDxKL [video file]. Retrieved from <https://www.youtube.com/watch?v=RT090R59DbE>
- Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130(3), 466-478.
- Pitcher, D., Walsh, V. & Duchaine, B. *Exp Brain Res* (2011) 209: 481. <https://doi.org/10.1007/s00221-011-2579-1>
- Pollak, S. D., Cicchetti, D., Hornung, K., & Reed, A. (2000). Recognizing emotion in faces: developmental effects of child abuse and neglect. *Developmental psychology*, 36(5), 679.

- Purves, D., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L., Woldorff, M. G., & Brannon, E. M. (2008). Emotion. In, Purves, D., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L., Woldorff, M. G., & Brannon, E. M. (Ed.), *Cognitive Neuroscience (2e editie)* (pp. 319-357). Sunderland: Sinauer Associates, Inc.
- Rolls, E. T. (2005). *Emotion explained*. Oxford: Oxford University Press.
- Swaddle, J. P., & Reiersen, G. W. (2002). Testosterone increases perceived dominance but not attractiveness in human males. *Proc. Roy. Soc. London B*, 269, 2285-89.
- Thomaz, A. (2015). The next frontier in robotics: social, collaborative robots | Andrea Thomaz | TEDxPeachtree [video file]. Retrieved from <https://www.youtube.com/watch?v=O1ZhWv84eWE&t=81s>
- Ward, J. (2016). *The student's guide to social neuroscience*. Psychology Press.

Hoofdstuk 2: Taalkunde

- Arons, B. (1992). A review of the cocktail party effect. *Journal of the American Voice I/O Society*, 12(7), 35-50.
- Corr, P. J., & Matthews, G. (Eds.). (2009). *The Cambridge handbook of personality psychology* (pp. 748-763). Cambridge, UK:: Cambridge University Press.
- Corver, N. (2016). Emotion in the build of Dutch. Deviation, augmentation and duplication. *Tijdschrift voor Nederlandse Taal-en Letterkunde*, 132, 232-275.
- van Dale. (2018). Persoonlijkheid. *Van Dale*. Geraadpleegd van: <https://www.vandale.nl/gratis-woordenboek/nederlands/betekenis/persoonlijkheid>
- Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. *Handbook of affective sciences*, 572, V595.
- Ensie. (2015). Emotie. *Psychologie begrippen omschreven*. Geraadpleegd van: <https://www.ensie.nl/psychologie/emotie>
- Ensie. (2017). Sociale Interactie. *Grondbeginselen der Sociologie*. Geraadpleegd van: <https://www.ensie.nl/grondbeginselen-der-sociologie/sociale-interactie>
- Foolen, A. (1997). The expressive function of language: Towards a cognitive semantic approach. *The language of emotions: conceptualization, expression, and theoretical foundation*, 15-32.
- Hymes, D. (1997). The scope of sociolinguistics. In *Sociolinguistics* (pp. 12-22). Palgrave, London.
- Jay, T., & Janschewitz, K. (2007). Filling the emotion gap in linguistic theory: Commentary on Potts' expressive dimension. *Theoretical Linguistics*, 33(2), 215-221.
- Mohammad, S. M., & Kiritchenko, S. (2013). Using nuances of emotion to identify personality. *Proceedings of ICWSM*.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001), 2001.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of personality and social psychology*, 77(6), 1296.

- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 54(1), 547-577.
- Plutchik, R. (1982). A psychoevolutionary theory of emotions. Rietveld, A. en V. van Heuven. 2009. *Algemene fonetiek*.
- Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of voice*, 9(3), 235-248.
- van Sterkenburg, P. G. J. (2007). Woorden van en voor emotie.

Hoofdstuk 3: Kunstmatige Intelligentie

- Beck, A., Stevens, B., Bard, K. A., & Cañamero, L. (2012). Emotional body language displayed by artificial agents. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(1), 2.
- Broadbent, E., Kumar, V., Li, X., Sollers 3rd, J., Stafford, R. Q., MacDonald, B. A., & Wegner, D. M. (2013). Robots with display screens: a robot with a more humanlike face display is perceived to have more mind and a better personality. *PloS one*, 8(8), e72589.
- Ceh, S., & Vanman, E. J. (2018). The Robots Are Coming! the Robots Are Coming! Fear and Empathy for Human-like Entities. *PsyArXiv*. June, 4.
- Eyssel, F., De Ruiter, L., Kuchenbrandt, D., Bobinger, S., & Hegel, F. (2012, March). 'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on* (pp. 125-126). IEEE.
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, 53(5), 517-527.
- Ivaldi, S., Lefort, S., Peters, J., Chetouani, M., Provasi, J., & Zibetti, E. (2017). Towards engagement models that consider individual factors in HRI: On the relation of extroversion and negative attitude towards robots to gaze and speech during a human-robot assembly task. *International Journal of Social Robotics*, 9(1), 63-86.
- Krämer, N. C., von der Pütten, A., & Eimler, S. (2012). Human-agent and human-robot interaction theory: similarities to and differences from human-human interaction. In *Human-computer interaction: The agency perspective* (pp. 215-240). Springer, Berlin, Heidelberg.
- Kuo, I. H., Rabindran, J. M., Broadbent, E., Lee, Y. I., Kerse, N., Stafford, R. M. Q., & MacDonald, B. A. (2009, September). Age and gender factors in user acceptance of healthcare robots. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on* (pp. 214-219). IEEE.
- Li, D., Rau, P. P., & Li, Y. (2010). A cross-cultural study: Effect of robot appearance and task. *International Journal of Social Robotics*, 2(2), 175-186.
- Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy: How, when, and where we perceive life in a face. *Psychological science*, 21(12), 1854-1862.

- MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*, *146*, 190-205.
- Misselhorn, C. (2009). Empathy with inanimate objects and the uncanny valley. *Minds and Machines*, *19*(3), 345.
- Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, *19*(2), 98-100.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues*, *56*(1), 81-103.
- Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of applied social psychology*, *27*(10), 864-876.
- Powers, A., & Kiesler, S. (2006, March). The advisor robot: tracing people's mental model from a robot's physical attributes. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction* (pp. 218-225). ACM.
- Rosenthal-von der Pütten, A. M., & Krämer, N. C. (2014). How design characteristics of robots determine evaluation and uncanny valley related responses. *Computers in Human Behavior*, *36*, 422-439.
- Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & Frith, C. (2011). The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social cognitive and affective neuroscience*, *7*(4), 413-422.
- Scopelliti, M., Giuliani, M. V., & Fornara, F. (2005). Robots in a domestic setting: a psychological approach. *Universal access in the information society*, *4*(2), 146-155.
- Tamagawa, R., Watson, C. I., Kuo, I. H., MacDonald, B. A., & Broadbent, E. (2011). The effects of synthesized voice accents on user perceptions of robots. *International Journal of Social Robotics*, *3*(3), 253-262.
- Tinwell, A., Grimshaw, M., & Williams, A. (2010). Uncanny behaviour in survival horror games. *Journal of Gaming & Virtual Worlds*, *2*(1), 3-25.
- Wang, L., Rau, P. L. P., Evers, V., Robinson, B. K., & Hinds, P. (2010, March). When in Rome: the role of culture & context in adherence to robot recommendations. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*(pp. 359-366). IEEE Press.

Integratie

- Repko, A. F., Szostak, R., & Buchberger, M. P. (2016). Introduction to interdisciplinary studies. Sage Publications.