

MASTER'S THESIS

De robot als mens 2.0

Een onderzoek naar de verhouding tussen mens en robot

Moggré, A

Award date:

2023

Awarding institution:

Department of Cultural Studies

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 06. Dec. 2023

Open Universiteit
www.ou.nl



Open Universiteit

Faculteit Cultuurwetenschappen

Masterscriptie

De robot als mens 2.0

Een onderzoek naar de verhouding
tussen mens en robot

The robot as human being 2.0

An inquiry into the relationship between
human beings and robots

Studentnummer: x
Adres: x
x
Cursusjaar: 2022/2023
Cursuscode: CM0202162114
Begeleider: Prof. Dr. Eddo Evink
Examinator: Dr. Tom Giesbers
Versienummer: 4
Datum: 17-07-2023

Inhoudsopgave

1. Inleiding en probleemstelling	3
1.1. Probleemstelling	3
1.2. Onderzoeksvragen	5
1.3. Maurice Merleau-Ponty en de fenomenologie	7
1.4. De fenomenologische filosofie	8
2. Ontwikkelingen in de wereld van AI	10
2.1. Traditionele AI	10
2.2. Nieuwe ontwikkelingen	10
2.3. Brein en bewustzijn	11
2.4. Intelligentie	13
2.5. Wat is denken?	14
3. Verschillen tussen mens en robot	17
3.1. Waarnemen	18
3.2. Psychosomatisch karakter van waarnemen	21
3.3. Diepte, ruimte en beweging	23
3.4. Verschil tussen mens en robot in het waarnemen	28
3.5. Empirisme en rationalisme	30
3.6. De fenomenologische opvatting	31
4. Intersubjectiviteit	34
4.1. Perceptie als basis	35
4.2. Wisselwerking tussen de zichtbare wereld en subjecten	36
4.3. Wisselwerking en het onzichtbare	41
5. Emoties, gevoelens en empathie	52
5.1. Emoties en gevoelens	52
5.2. Het herkennen van emoties	57
5.3. Empathie	59
6. Het menselijk gevoel en robots; de filosofie van Levinas	64
6.1. De filosofie van Levinas	64
6.2. Intentionaliteit bij Levinas	68
6.3. Levinas versus Merleau-Ponty	70
6.4. De robot en het gevoel	73
7. Conclusies	75

De mens is niets dan een knooppunt van betrekkingen; het zijn alleen de betrekkingen die gelden voor de mens

A de Saint-Exupéry

1. Inleiding, probleemstelling en onderzoeksvragen

1.1. Probleemstelling

De ontwikkelingen in de wereld van de geautomatiseerde systemen gaan snel. De internetrevolutie die ongeveer 25 jaar geleden begon heeft ons leven en gedrag ingrijpend veranderd. Nieuwe ontwikkelingen dienen zich aan en één daarvan is kunstmatige intelligentie (verder afgekort tot AI van Artificial Intelligence). Op allerlei terreinen worden AI gestuurde robots ingezet. In de ouderenzorg wordt onderzoek gedaan naar de inzet van robots bij het bestrijden van eenzaamheid. De robot fungeert daarbij als gesprekspartner. In de medische wereld wordt onderzoek gedaan naar de inzet van robots bij het herkennen van ziektebeelden. Dit werk werd vrijwel altijd door artsen uitgevoerd maar robots lijken dit sneller en nauwkeuriger te kunnen. In de consumentenmaatschappij wordt AI ingezet als middel om koopgedrag te beïnvloeden door grootschalige analyse van het gedrag van mensen. In de vervoerssector wordt onderzoek gedaan naar zelfrijdende auto's. Door de overheid worden robots ingezet bij het assisteren van ambtenaren tijdens het contact met de burger. In meer autocratisch bestuurde samenlevingen worden robots ingezet om burgers te controleren en hun gedrag te beïnvloeden. De invloed van robots gaat in alle haarvaten van de samenleving een rol spelen en daarbij gaan robots steeds meer activiteiten van mensen overnemen. De vraag doemt op in welke mate robots het gedrag en de mogelijkheden van mensen kunnen overnemen. Wordt de robot een nieuwe en betere mens zoals sommigen ons willen doen laten geloven of zijn robots alleen maar slechte imitatoren van delen van het menselijk gedrag en het menselijk handelen? In dit werkstuk wordt de relatie tussen mens en robot verder onderzocht aan de hand van de volgende probleemstelling:

Op welke wijze verschillen de onderlinge relaties van mensen met de onderlinge relaties van mens en robot vanuit fenomenologisch perspectief?

Omdat de ontwikkelingen in de AI snel gaan moet een voorbehoud worden gemaakt met betrekking tot de mogelijkheden. Dit werkstuk gaat uit van de huidige en te voorziene ontwikkelingen. Er kunnen altijd nieuwe ontwikkelingen zijn die herziening van standpunten noodzakelijk maken.

1.2. Onderzoeksvragen

De opkomst en nieuwe mogelijkheden van AI dwingen ons na te denken over de wijze waarop wij ons als mens dienen te verhouden tot deze nieuwe revolutionaire technologie. Zeker als hiervan wordt beweerd dat deze het menselijk brein kan simuleren en in de toekomst zelfs onafhankelijk van mensen kan voortbestaan. Dit wordt gedaan aan de hand van de hiernavolgende vragen.

Het is allereerst van belang de technologische mogelijkheden van AI ten aanzien van het simuleren van het menselijk brein te onderzoeken.

Catherine Malabou heeft met een filosofische blik de wereld van de AI geduid.¹ Twee bekende boeken van haar zijn *What should we do with our brain* en *Morphing Intelligence*. In het eerste boek uit 2004 betoogt zij dat de hersenen van de mens uniek zijn, met name door de plasticiteit van de synapsen in de hersenen. Deze plasticiteit zorgt ervoor dat de hersenen zich ontwikkelen en zich voortdurend aanpassen aan veranderende omstandigheden. Als gevolg hiervan kunnen hersenen niet nagebootst worden door computers en algoritmen.² In het tweede boek *Morphing Intelligence* uit 2019 komt zij hierop terug omdat er inmiddels computers kunnen worden gemaakt die eenzelfde mogelijkheid tot plasticiteit hebben waarbij geheugen en processoren in een neurale netwerk hetzelfde gedrag kunnen simuleren en als zodanig deze plasticiteit nabootsen.³

Ten tweede doet zich de vraag voor wat de verschillen en overeenkomsten zijn tussen een door AI bepaalde robot en een mens. In deze scriptie worden aan hand van de filosofie van Merleau-Ponty deze overeenkomsten en verschillen geanalyseerd. Maurice Merleau-Ponty is een Franse filosoof die in de traditie van de fenomenologie de menselijke waarneming heeft onderzocht en hij

¹ Catherine Malabou is een Franse filosofe. Zij heeft o.a. bij de filosoof Derrida gestudeerd en is professor filosofie aan de Kingston Universiteit in Londen en professor vergelijkende literatuurwetenschappen aan de Universiteit van Californië in Irvine.

² Catherine Malabou, *Wat te doen met ons brein?*. Jeanne Hollerhoek vert. (Amsterdam 2011) 29-31.

³ Catherine Malabou, *Morphing Intelligence, from IQ measurement to artificial brains*. Carolyn Shred vert. (herz. editie; New York 2021) 82.

betoogt dat de relatie tussen mens en wereld en geest en lichaam nauw met elkaar samenhangen en niet van elkaar gescheiden kunnen worden bestudeerd.

Ten derde gaat het erom in hoeverre de relatie tussen mensen anders is dan die tussen mens en robot. De relatie tot de andere mens behoort tot het meest wezenlijke van het menselijk bestaan en het mens zijn. Het onderzoek naar de verschillen in de relaties tussen mensen en de relaties met robots laat zien op welke wijze robots verschillen van mensen en wat dit betekent voor onze verhouding tot robots. Merleau-Ponty heeft aan onze relatie tot de medemens een belangrijk deel van zijn onderzoek besteed en zijn ideeën hierover vormen het uitgangspunt voor de analyse. In haar boek *Merleau-Ponty and the ethics of intersubjectivity* analyseert Anya Daly de wijze waarop Merleau-Ponty deze relatie definieert en wat dit betekent voor ons als mensen. Zij gaat met name in op het begrip wisselwerking bij Merleau-Ponty. De wisselwerkingsthese houdt in dat 'zelf', 'de ander' en 'de wereld' relationeel moeten worden opgevat en dat deze relaties ethisch van karakter zijn. Dit houdt in dat de ander (subjecten) en de wereld essentieel zijn voor het zelfbewustzijn en omgekeerd dat het zelfbewustzijn medebepalend is voor de ander en de wereld. Deze analyse wordt als uitgangspunt gebruikt om de verschillen in deze relaties te definiëren.⁴

Als laatste wordt de filosofie van Merleau-Ponty vergeleken met de filosofie van Levinas.⁵ De vraag dient zich aan of de filosofie van Levinas niet een radicaler onderscheid laat zien tussen mens en robot dan de filosofie van Merleau-Ponty. Bij Levinas gaat het niet direct om waarnemen maar staat de relatie van de mens met zijn medemens centraal. Levinas fundeert de ethiek als bepalend onderdeel voor de menselijke relatie. Levinas ziet de mens in eerste instantie als een organisme dat leeft in de genieting. 'Leven van...' zegt Levinas is genieting. In zijn boek *Totaliteit en Oneindigheid* werkt hij vanuit dit basisprincipe toe naar een filosofie over de intersubjectiviteit en uiteindelijk naar een wat hij noemt fundering van de ethiek. Het gaat in dit werkstuk met name om zijn ideeën over het genieten als basis voor zijn filosofie. De ideeën van beide filosofen worden bekeken in het licht van de ontwikkelingen in de AI.

⁴ Anya Daly doet onderzoek op het gebied van de fenomenologie en geeft les aan de universiteit van Tasmanië.

⁵ Emmanuel Levinas was een tijdgenoot van Merleau-Ponty en filosofeerde eveneens in de traditie van de fenomenologie.

1.3. Maurice Merleau-Ponty en de Fenomenologie

Merleau-Ponty verloor op vijfjarige leeftijd zijn vader maar had desondanks een gelukkige jeugd. Hij had een blijvende nauwe band met zijn moeder die overleed in 1953. Tijdens de tweede wereldoorlog diende hij als luitenant in het Franse leger tot hij in 1940 gewond raakte. Hij trouwde met Suzanne Jolibois en kreeg één dochter met haar. Na de oorlog richtte hij samen met Jean Paul Sartre en Simone de Beauvoir het tijdschrift *Les Temps Moderne* op. Hij verzorgde voor dit tijdschrift politieke artikelen en commentaren. In 1952 werd hij gekozen tot hoofd filosofie bij het Collège de France, de meest vooraanstaande positie in Frankrijk voor een filosoof. Hij raakte gebrouilleerd met Sartre en De Beauvoir omdat hij kritiek had op hun revolutionaire Marxisme waar hij zich steeds meer van distantieerde. Hij weigerde een benoeming tot lid van het Legioen van Eer als protest tegen de inhumane acties van Frankrijk tijdens de koloniale oorlog tegen Algerije. Hij overleed in 1961 op drieënvijftigjarige leeftijd aan een hartstilstand.

Zijn hoofdwerk is *Fenomenologie van de waarneming*. Hierin laat hij zien dat de fenomenologie zich bezig dient te houden met onze lichamelijke ontmoeting met de wereld.⁶ Hij heeft veel te danken aan de grondlegger van de fenomenologie Edmund Husserl. In 1929 volgt hij de lezingen van Husserl aan de Sorbonne en in 1939 bezoekt hij het Husserl archief in Leuven. Hij heeft ook kritiek op Husserl en staat een andere fenomenologie voor.

⁶ Corijn van Mazijk, *De wereld als verschijning. Fenomenologie en de twintigste eeuw* (Amsterdam 2021) 131.

1.4. De fenomenologische filosofie

De fenomenologie is een filosofische stroming die aan het begin van de vorige eeuw is ontstaan. De fenomenologie bestudeert de fenomenen zoals die zich aan de menselijke geest voordoen. De fenomenologie onderzoekt de menselijke geest en alles waar het bewust van is. De fenomenologie is ontstaan als reactie op het naturalisme. De grondlegger Husserl vond dat de natuurwetenschappen een betere fundering nodig hadden. De aard van de natuurwetenschappelijke kennis werd niet goed begrepen en het natuurwetenschappelijke wereldbeeld voldeed volgens hem niet om de mens en zijn culturele wereld beter te begrijpen. Een nieuwe filosofie was nodig om dit begrip te krijgen. Hij legde daarbij veel nadruk op een filosofische methode en zag de filosofie als een wetenschap die op een systematische manier nieuwe resultaten moest boeken.⁷ Het bewustzijn moest volgens hem bestudeerd worden door te reflecteren op het bewustzijn zelf en niet op extern observeerbaar gedrag.⁸ Fenomenen zijn alle ervaringen van ons bewustzijn. Dit zijn de ervaringen van de dingen in de wereld om ons heen maar ook alle acten van denken, voelen, waarnemen en inleven.⁹ Het bewustzijn is het meest wezenlijke van de mens en bevat de wereld zoals die op een bepaald moment als fenomeen aan ons verschijnt. Het begrip bewustzijn is ook te lezen als ervaring of subjectiviteit. Het is dus mijn ervaring.¹⁰ Een belangrijk idee van Husserl is dat bewustzijn altijd een bewustzijn van iets is. Het is dus een act. Dit 'iets' kunnen fysieke objecten zijn maar ook ideële objecten zoals bijvoorbeeld historische gebeurtenissen of wiskundige stellingen. Het gaat daarbij dus niet alleen om de natuurlijke wereld maar ook om de culturele wereld. De wereld van de gedachten en ideeën.¹¹ Voor de bestudering hiervan is een fenomenologische houding nodig die de aandacht richt op de activiteit van het bewustzijn en niet op het object van het bewustzijn. Hij legde er de nadruk op dat de verschijningen van de fenomenen aan ons bewustzijn zo objectief en systematisch mogelijk onderzocht dienen te worden. Hiervoor voerde hij de methode van de fenomenologische reductie in.

⁷ Ibidem, 33.

⁸ Ibidem, 39.

⁹ Ibidem, 9.

¹⁰ Ibidem.

¹¹ Ibidem, 45.

De fenomenologische reductie houdt in dat de filosoof zich moet concentreren op de fenomenen die aan ons bewustzijn verschijnen. Deze moeten zo onafhankelijk mogelijk worden bestudeerd. Dit kan door de werkelijkheid als het ware tussen haakjes te zetten. De werkelijkheid buiten ons bewustzijn doet er voor het onderzoek niet toe. Het wordt dan mogelijk te onderzoeken hoe de fenomenen aan ons bewustzijn verschijnen. En hier dient het volgens Husserl om te gaan.¹²

Merleau-Ponty is op vele manieren verbonden met het werk van Husserl. Toch had hij ook kritiek op zijn filosofie en op zijn methode. In tegenstelling tot Husserl denkt hij dat de reductie niet perfect is uit te voeren. Merleau-Ponty betoogt dat de fenomenoloog zelf opgenomen is in de wereld en dat deze dus geen standpunt onafhankelijk van de wereld kan innemen.¹³ Een gevolg hiervan is dat er geen filosofisch systeem kan worden gedacht dat alle filosofische problemen oplost omdat zowel de mens als de fenomenen steeds veranderen. Nadat de aandacht voor de fenomenologie aan het eind van de vorige eeuw afnam en overging in het poststructuralisme en postmodernisme is er tegenwoordig weer een hernieuwde belangstelling voor deze stroming. Ook in de moderne wetenschappen zoals medische wetenschap, de neurowetenschap en AI wordt gebruik gemaakt van de inzichten van deze stroming.¹⁴

De fenomenologie van Merleau-Ponty vormt een goed uitgangspunt om te laten zien wat de verschillen tussen mens en robot vooralsnog zijn en tevens dat de fenomenologie een meer vruchtbare benadering oplevert dan meer traditionele theorieën zoals empirisme en rationalisme.

¹² Ibidem, 49.

¹³ Ibidem, 138.

¹⁴ Ibidem, 178.

2. Ontwikkelingen in de wereld van de AI

Wat zijn op dit moment de te verwachten mogelijkheden van AI? Menselijke robots, als de meest aansprekende mogelijkheid van AI, zullen steeds meer menselijke activiteiten overnemen en menselijk gedrag kunnen nabootsen. De vraag is dus hoe wij ons kunnen verhouden tot deze robots en wat de voorziene mogelijkheden hiervan zijn?

2.1. Traditionele AI

Zowel de kennis op het gebied van de werking van ons brein als de kennis op het gebied van AI heeft de laatste decennia een enorme ontwikkeling doorgemaakt. Tot in de jaren vijftig van de vorige eeuw werd gedacht dat het brein als een soort von Neumann computer werkt. Het brein wordt hierbij beschouwd als een centrale processor die de door de zintuigen aangeleverde informatie verwerkt en combineert met de in het geheugen opgeslagen informatie. Karakteristiek aan deze opvatting is dat het brein statisch is en niet verandert. Het bewustzijn is iets apart en wordt gevoed door het brein. De opvatting van Kant over het bewustzijn is dat het transcendentaal is. Het bestaat als aparte entiteit en wordt vormgegeven door de categorieën als kwantiteit, kwantiteit, relatie en modaliteit. Als wij de wereld waarnemen worden wij geleid door de categorieën van ruimte en tijd. Wij kunnen daarom de dingen niet werkelijk zien zoals ze zijn maar wij zien ze gestructureerd en gevormd in onze geest als binnen de ruimte en de tijd. Onze geest heeft beperkingen en hieruit volgt dat wij de dingen ervaren binnen de beperkingen van onze geest. Kant probeerde op deze wijze een brug te slaan tussen het rationalisme (alle kennis wordt verkregen door het denken) en het empirisme (alle kennis is ervaringskennis). Wat bewustzijn precies is en hoe bewustzijn tot stand komt is echter nog onderwerp van discussie. Bewustzijn kan voorlopig alleen indirect worden afgeleid.

2.2. Nieuwe ontwikkelingen

De ontwikkelingen op het gebied van de neurowetenschappen geven aanleiding tot nieuwe ideeën over het brein en het bewustzijn. Malabou schrijft in haar boek *Wat te doen met ons brein*:

De huidige staat van het onderzoek en van de observaties maakt inderdaad voor cognitieve wetenschappers de conclusie mogelijk dat gedachten, kennis, verlangens en affecten een neurale, dat wil zeggen een biologische basis hebben en dat de mentale beelden van het geestesleven tot stand komen in de hersenen.¹⁵

Verder schrijft zij dat het hierbij gaat om een filosofische en epistemologische standpuntbepaling die ook politieke, maatschappelijke en economische consequenties heeft en die belangrijke nieuwe vragen oproept. Voor ideeën over een ziel die als onafhankelijke entiteit los van het brein functioneert is in deze opvatting bijvoorbeeld geen plaats. Een belangrijke nieuw ontdekte eigenschap van het brein is plasticiteit. Plastisch materiaal kan als het eenmaal is veranderd niet meer zijn oorspronkelijke vorm terugkrijgen maar behoudt wel de oorspronkelijke sporen en kenmerken. Het is dus zowel stevig als vervormbaar. Het brein wordt in eerste instantie gevormd in de baarmoeder maar is bij de geboorte nog niet af en wordt in de kindertijd verder gevormd. Hierbij spelen genetische factoren een rol maar in een later stadium wordt de rol van de buitenwereld op deze vorming steeds belangrijker. In het volwassen stadium bestaat het brein uit een netwerk van ongeveer 20 miljard neuronen. Deze neuronen worden door middel van tientallen miljarden synapsen met elkaar verbonden en deze verbindingen wijzigen zich voortdurend door o.a. inwerking van buiten en de eigen innerlijke processen¹⁶. Mentale beelden zijn het resultaat van deze verbindingen. Het is dus zo dat ons brein voortdurend in beweging is en zich aanpast aan de omstandigheden. Een opvallende eigenschap van deze werking van de neuronen en synapsen zorgt er bijvoorbeeld voor dat de hersenen zich na een trauma weer kunnen herstellen.

2.3. Brein en bewustzijn

Volgens de huidige inzichten zetelt de basis voor het bewustzijn in het brein. Daarmee is echter nog niet duidelijk hoe bijvoorbeeld denkprocessen tot stand komen en hoe deze worden beïnvloed. Volgens Daniël Dennett moeten we hierbij denken aan de werking van hardware en software. Deze functioneren op verschillende abstractieniveaus. Denkprocessen zijn als software op de neurale

¹⁵ Catherine Malabou, *Wat te doen met ons brein?*. Jeanne Hollerhoek vert. (Amsterdam 2011) 90.

¹⁶ *Ibidem*, 35-37.

processen.¹⁷ Denk hierbij aan een computerspelletje dat uiteindelijk bestaat uit fysieke elektrische signalen in een computer (de hardware). Het programma (de software) geeft op een ander abstractieniveau instructies aan de computer. Om het programma te kunnen schrijven is het voldoende de programmeertaal te kennen en hoef je niet te begrijpen hoe de hardware dit programma uitvoert. Om het spel te kunnen spelen moet je de regels van het spel uiteindelijk beheersen en hoef je niet te weten hoe het programma werkt (het derde niveau). Deze niveaus zijn dus wel met elkaar verbonden maar kunnen onafhankelijk van elkaar worden begrepen en het is zelfs niet mogelijk om vanuit de elektrische signalen in de computer de regels van het spel te destilleren. In die zin functioneren de niveaus onafhankelijk van elkaar maar ondanks dat zijn ze toch op een andere manier afhankelijk van elkaar. Als we deze parallel doortrekken zal het dus ook een illusie blijken te zijn om uit de basale kennis van de werking van de neuronen en synapsen te destilleren hoe denkprocessen uiteindelijk werken en wat bewustzijn is. Er bestaat veel optimisme dat we op termijn deze processen in detail zullen begrijpen maar of dit gegrond is, blijft, voorlopig althans, nog ongewis. Met de decodering van de genetische code dacht men ook de heilige graal tot alle kennis over de mens gevonden te hebben maar ook hier blijkt dat deze kennis weer nieuwe vragen oproept.

In haar boek *Wat te doen met ons brein* gaat Malabou ervan uit dat computers een architectuur hebben die gebaseerd is op de von Neumann structuur.¹⁸ Een von Neumann computer werkt op basis van één centrale verwerkingseenheid en een geheugen. De geprogrammeerde opdrachten worden door deze processor één voor één verwerkt. Menselijke intelligentie gedefinieerd als het vermogen om problemen op te lossen, werkt op basis van de neurale processen. De neurale architectuur is veel flexibeler en kan meerdere opdrachten gelijktijdig verwerken. Een computer gebaseerd op deze structuur kan veel meer aan dan een von Neumann computer ooit zal kunnen presteren. Kunstmatige intelligentie die het moet doen met de von Neumann architectuur zal deze intelligentie dus nooit kunnen benaderen.¹⁹ In haar boek *Wat te doen met ons brein* verdedigt zij de stelling dat het op grond van deze architectuur

¹⁷ Ibidem, 22.

¹⁸ Ibidem, 123-129.

¹⁹ Ibidem.

dus nooit mogelijk zal zijn het brein te imiteren en dat het dus onmogelijk zal zijn een robot met menselijke eigenschappen te ontwikkelen.

2.4. Intelligentie

Echter de ontwikkelingen op het gebied van de kunstmatige intelligentie staan niet stil en inmiddels zijn er computers gemaakt met een vergelijkbare architectuur als de neurale netwerken van het brein. Sterker nog, de nieuwe computerarchitectuur is ontstaan als gevolg van de nieuwe kennis over het brein. Dat dit een nieuwe computerarchitectuur tot gevolg heeft met nieuwe mogelijkheden is duidelijk. Malabou concludeert in haar boek *Morphing Intelligence* dat haar eerste conclusie niet deugt en ze gaat er inmiddels van uit dat er dus wel robots kunnen worden ontwikkeld die gebaseerd zijn op de architectuur van een menselijk brein en ook een vorm van intelligentie kunnen hebben die mede gebaseerd is op menselijke intelligentie.²⁰ In haar nieuwe boek bespreekt ze het ontstaan van het begrip intelligentie en de moeilijkheden die dit begrip oproept. Intelligentie wordt begrepen als het vermogen problemen te herkennen en op te lossen, het zich kunnen oriënteren in de ruimte, plannen te maken en activiteiten te organiseren. Er wordt vanuit gegaan dat deze vermogens bij de mens kunnen worden gemeten middels gestandaardiseerde IQ testen. Intelligentie wordt uitgedrukt in een getal dat als basis dient voor het vaststellen van de mate van intelligentie. IQ voldoet al lang niet meer en intelligentie wordt door o.a. Dewey als een proces gezien en als een mogelijkheid tot ontwikkeling van de categorieën van het denken.²¹ Howard Gardner geeft in zijn boek *Frames of Mind* acht vormen van intelligentie weer waaronder muzikale intelligentie, lichamelijk kinesthetische intelligentie, interpersoonlijke en intra-persoonlijke intelligentie. Of dat intelligentie is of menselijke vermogens waarvan intelligentie deel uitmaakt blijft een kwestie van definitie. De ontwikkelingen op het gebied van AI geven volgens Malabou ook weer aanleiding tot nieuwe vormen van intelligentie.²² Computers zijn in staat grote hoeveelheden ongestructureerde data razendsnel tot voor de mens

²⁰ Catherine Malabou, *Morphing Intelligence, from IQ measurement to artificial brains*. Carolyn Shred vert. (herz. editie; New York 2021) 82.

²¹ Ibidem, 101 en 107-109.

²² Ibidem, 151.

bruikbare informatie om te werken. Deze processen zijn niet gebaseerd op inventiviteit maar op automatische creatie wat dus een nieuwe vorm van intelligentie is die bij de mens niet voorkomt. Malabou ziet hierin mogelijkheden voor AI om de menselijke intelligentie voorbij te streven alhoewel deze andere vormen zal hebben. De discussie in haar boek geeft echter aan dat het begrip intelligentie niet eenduidig is gedefinieerd. Er is te weinig inzicht in de werking van mentale processen om uitspraken te doen over de werking van menselijke intelligentie. Wat computers kunnen is imposant maar uiteindelijk is de werking van de computer bekend. Algoritmes mogen voor veel mensen een duister gebied zijn, voor specialisten zijn ze dat niet. Zelflerende computerprogramma's kunnen voor de programmeur ondoorzichtig worden omdat zij zelf programmacode kunnen produceren. Het is echter ook mogelijk AI in te zetten om deze ondoorzichtigheid te voorkomen of de code weer om te zetten in voor mensen begrijpelijke taal. Alle computer intelligentie is gebaseerd op de causale logica van het beslissingsmodel 'als x dan y' waarbij x en y logische en/of rekenkundige functies zijn. Uiteindelijk leert een computer alleen door te vergelijken en door het toepassen van statistische, wiskundige en logische functies op data middels algoritmen. Ook al gaat dit steeds slimmer en sneller en zal het menselijk brein steeds beter worden gesimuleerd.

2.5. Wat is denken?

Er is op dit moment geen zicht op fundamenteel nieuwe manieren van het verwerken van data. In dit verband is het voorbeeld van de Chinese kamer illustratief. Het experiment van de Chinese kamer is een gedachte-experiment dat werd bedacht door de filosoof John Searle in 1980.²³ In het experiment wordt een persoon voorgesteld die zich bevindt in een afgesloten kamer en geen Chinees spreekt. Deze persoon heeft echter wel toegang tot een set regels in het Engels die hem vertellen hoe hij Chinese karakters moet manipuleren. Buiten de kamer is er een persoon die een vraag in het Chinees stelt, en de persoon in de kamer moet dan de juiste reactie in het Chinees

²³ Howard Gardner, *The mind's new science, A history of the cognitive revolution* (New York 1987) 171.

produceren door de regels te volgen. Volgens Searle is het mogelijk dat de persoon in de kamer perfecte antwoorden kan produceren zonder daadwerkelijk Chinees te begrijpen. Het enige wat hij doet, is het volgen van de regels die hem zijn gegeven. Searle stelt echter dat dit niet betekent dat de persoon in de kamer daadwerkelijk begrip heeft van de taal die hij manipuleert. Dit voorbeeld wordt door Searle gebruikt om aan te geven dat van een systeem dat alleen maar regels volgt en syntactische operaties uitvoert niet gezegd kan worden dat het denkt.²⁴

Van het onlangs geïntroduceerde programma CHATGPT, wat min of meer op deze manier werkt, kan dus niet gezegd worden dat het denkt ook al geeft het antwoord op alle vragen die worden gesteld. Dus is er nog een groot verschil tussen het denken van de mens en het 'denken' van een chatbot. Echter als je alleen de antwoorden leest en niet weet dat er een computer achter die antwoorden zit dan lijkt het toch alsof de antwoorden van een welbespraakt mens komen. Het programma is goed in staat gedachten te verwoorden en zelfs veel beter en uitgebreider dan mensen dat kunnen. Hieruit blijkt dat de begrippen denken, intelligentie en bewustzijn nog de nodige onduidelijkheden bevatten. Het zal zo zijn dat we eerst meer inzicht moeten krijgen in de werking van de menselijke geest en de interactie tussen geest, lichaam en omgeving voordat we deze begrippen verder in kunnen vullen. Bewustzijn als het vermogen om de wereld waar te nemen, te begrijpen en het zelfbewustzijn als de mogelijkheid tot zelfreflectie en de uitdrukking hiervan in taal zijn grotendeels nog onbegrepen fenomenen. Zolang hier nog geen inzicht in is zal een uitspraak over de toekomstige robots met menselijke eigenschappen een lage waarschijnlijkheidswaarde hebben. Voor die kennis is het nodig dat er ook nieuwe methoden worden gevonden om het brein op andere abstractieniveaus te analyseren. Het maken van steeds nauwkeuriger hersenscans zal daar waarschijnlijk niet genoeg voor zijn omdat het denken op een ander abstractieniveau gebeurt. Het is echter wel zo dat de ontwikkelingen zo snel gaan dat er steeds meer mogelijk is op het gebied van AI en er zeker robots

²⁴ Ibidem, 172-174.

komen die veel van wat mensen presteren, en waarbij de intelligentie een grote rol speelt, zullen overnemen. Ook zal het zoeken naar nieuwe modellen om de menselijke geest te modelleren, inzicht kunnen geven in de bewustzijnsprocessen. Door intelligentie vanuit het onderzoek naar de hersenen en vanuit de modelmatige simulering van menselijke vermogens te benaderen, zal er ongetwijfeld veel vooruitgang worden geboekt. Het denken van een mens staat ook niet op zich. Gedachten komen tot stand in interactie met de omgeving en vooral door relaties met anderen in een intersubjectief proces. Nieuwe kennis wordt gegenereerd door voort te borduren op bestaande kennis en door discussie en kritiek. Het denkproces vindt niet plaats in één brein maar is een gezamenlijk proces geworden. Het vindt als het ware plaats op een abstractieniveau dat nog hoger ligt dan het niveau van het individuele brein. Robots zullen uiteindelijk in staat zijn met elkaar en met mensen deel te nemen aan een interactief denkproces. Hierdoor dragen robots op een intelligente manier bij aan het denken op dit hogere abstractieniveau. Is er op dit niveau dan sprake van een echte menselijke robot? Het is vooralsnog niet duidelijk tegen welke grenzen AI gaat oplopen bij het simuleren van het menselijk brein. Echter bij robots zal de organische lichamelijke ontbreken waardoor hun 'denken' fundamenteel anders zal zijn dan dat van mensen.

3. Verschillen tussen mens en robot

Merleau-Ponty heeft in zijn boek *Fenomenologie van de waarneming* een filosofisch onderzoek gedaan naar de wijze waarop de mens interacteert met zijn omgeving. Zijn filosofie kan een licht werpen op de huidige bestaande verschillen tussen mens en robot. Het AI-programma ChatGPT vat zijn opvattingen over intelligentie als volgt samen:

"Phenomenology of Perception" (Nederlands: *"Fenomenologie van de Waarneming"*) is een boek geschreven door de Franse filosoof Maurice Merleau-Ponty, gepubliceerd in 1945. Het boek is een belangrijk werk in de fenomenologische filosofie en onderzoekt de relatie tussen het lichaam en de geest, en de manier waarop we de wereld om ons heen waarnemen en begrijpen.

Merleau-Ponty stelt dat onze waarneming niet los kan worden gezien van ons lichaam en onze ervaringen. We ervaren de wereld niet alleen met onze zintuigen, maar ook met ons lichaam en onze geest. Onze waarneming is altijd gebaseerd op onze ervaringen en context, en is daarom subjectief.

Het boek bespreekt verschillende onderwerpen, waaronder de fenomenologie van de waarneming, de waarneming van kleur en vorm, de waarneming van tijd en beweging, en de waarneming van andere mensen. Merleau-Ponty benadrukt dat we onze ervaringen niet kunnen scheiden van onze waarneming, en dat we de wereld om ons heen begrijpen door middel van onze lichamelijke en emotionele betrokkenheid.

Een belangrijk concept in het boek is het idee van de "lichamelijke intentionaliteit". Dit verwijst naar de manier waarop ons lichaam gericht is op de wereld om ons heen, en hoe onze waarneming en ervaringen voortkomen uit deze gerichtheid. Merleau-Ponty stelt dat we de wereld om ons heen niet alleen waarnemen, maar ook actief betrokken zijn bij het creëren van onze ervaringen.

In het algemeen kunnen we stellen dat "Phenomenology of Perception" een diepgaande analyse is van de manier waarop wij als mens de wereld om ons heen waarnemen en begrijpen. Merleau-Ponty 's werk heeft een grote invloed gehad op de filosofie en andere disciplines zoals de psychologie, de antropologie en de kunst.²⁵

Merleau-Ponty zegt over de menselijke geest dat het een niet begrepen menselijke mogelijkheid is en daarmee bedoelt hij dat we niet alleen niet weten

²⁵ Dit voorbeeld laat zien hoe krachtig huidige AI-programma's zijn op het gebied van linguïstische taken. Het programma is nog niet zo lang beschikbaar en is in staat zelf te leren. Het zal interessant zijn om te zien in hoeverre deze zoekopdracht in de loop van de tijd meer verfijnde informatie zal opleveren. Om meer gedetailleerde informatie te verkrijgen blijft het overigens nog wel nodig zelf de goede vragen te formuleren.

wat de menselijke geest precies inhoudt maar dat het bewustzijn mogelijkheden kan hebben die door het subject nog niet gekend zijn.²⁶

Het onderzoek naar de verschillen tussen mens en robot met als uitgangspunt de fenomenologische benadering van Merleau-Ponty zal worden uitgevoerd aan de hand van de volgende vragen:

Wat betekent waarnemen in de filosofie van Merleau-Ponty?

Hoe wordt de menselijke motoriek geïnterpreteerd door Merleau-Ponty?

Wat zijn de verschillen tussen mens en robot bij het waarnemen en bewegen?

Wat betekent de kritiek van Merleau-Ponty op het rationalisme en empirisme in het licht van de mogelijkheid van een menselijke robot?

3.1. Waarnemen

Merleau-Ponty heeft veel kritiek op de menswetenschappen. Hij vindt dat deze een te beperkt beeld geven van de mens omdat zij onder andere voor het onderzoek te veel gebruik maken van de objectiverende methoden van de natuurwetenschappen. Deze methoden zijn ontstaan vanuit het empirisme en rationalisme. Hij heeft daarom ook veel kritiek op deze stromingen als het om hun mensbeeld gaat. In zijn boek *Fenomenologie van de Waarneming* onderzoekt hij op een andere manier de mens en zijn wereld en gaat daarbij uit van de waarneming van de mens. In de waarneming ontstaat betekenis die voorafgaat aan wetenschappelijke interpretaties. De verschillende mensbeelden hebben consequenties voor de mogelijkheden van menselijke robots. In dit hoofdstuk wordt de filosofie van de waarneming besproken en de consequenties voor de mogelijkheden van de menselijke robot. Dit mensbeeld

²⁶ Maurice Merleau-Ponty, *Fenomenologie van de waarneming*. Douwe Timersma en Rens Vlasblom vert. (3^e druk; Amsterdam 2022) 17.

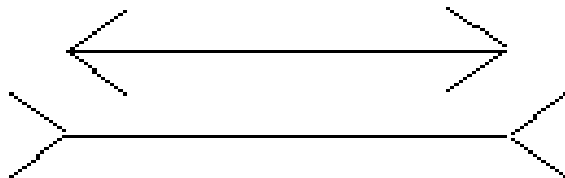
wordt vergeleken met het beeld dat rationalisme en empirisme geven en wat hiervan de gevolgen zijn voor een menselijke robot.

In de inleiding van het boek gaat Merleau-Ponty in op het begrip waarneming zelf. De eerste opmerking hierover is dat het waarnemen van mensen altijd meer is dan een bepaald object of een aantal objecten waarnemen. Merleau-Ponty heeft het over het veld van waarnemen.²⁷ Wij nemen niet de afzonderlijke objecten één voor één waar maar de waarneming bevat veel meer dan het object. Er is een voorgrond en een achtergrond. Mijn aandacht gaat misschien uit naar een bepaald object maar daarmee is de achtergrond niet verdwenen. Voor mij staat een laptop en mijn aandacht gaat naar de laptop. Tegelijkertijd neem ik het bureau waar waarop de laptop staat en waar ik aan zit te werken en de lamp die op het bureau mijn werkzaamheden bijlicht. Zou ik eerst alleen de laptop waarnemen en daarna het bureau en dan de lamp dan zou ik ook apart de verbinding tussen deze drie objecten moeten leggen terwijl deze verbinding direct in de waarneming is gegeven en logisch is geordend. Verplaats ik mijn aandacht naar de lamp dan blijft de laptop in de waarneming aanwezig alleen verdwijnt de laptop naar de achtergrond. Door middel van de waarneming staan wij dus in verbinding met de wereld. Door middel van het zien gaan wij een betrekking aan met de wereld. Horen is bijvoorbeeld een andere manier om in verbinding met de wereld te staan en laat een andere betrekking tot de wereld toe. De verbinding bestaat niet alleen tussen mij en het object maar bestaat uit een heel veld waar alles een plaats heeft en die mij een bepaalde logische geordendheid laat ervaren. De laptop staat op het bureau en wordt dus door het bureau gedragen. Dat alles is in de waarneming mede gegeven. Omdat wij op onze eigen wijze op de wereld zijn betrokken, hebben wij ook het bewustzijn van een ruimte die ook buiten het directe waarnemen aanwezig is. Ik ben mij ook bewust van de grond waarop het bureau en de bureaustoel staan alhoewel ik deze grond niet zie. Ik ben mij er ook bewust van dat de ruimte groter is dan ik nu kan waarnemen. Ik zit in de ruimte die zich laat benoemen als studeerkamer. “Er is een onbepaald zien, een zien van ik weet niet precies wat”²⁸ zegt Merleau-Ponty. In een wereld op zich, d.w.z. een objectieve wereld buiten mij waar ik geen betrekking mee heb, is alles bepaald. Maar zo werkt het

²⁷ Ibidem, 10.

²⁸ Ibidem, 12.

waarnemen niet. Er is dan ook een onbepaaldheid als een positief verschijnsel in het waarnemen.²⁹ Volgens Merleau-Ponty is het een misvatting dat fysiologische gebeurtenissen op zich staan. Zij gehoorzamen reeds aan biologische wetten en psychologische wetten³⁰. De bureaustoel staat recht op achter het bureau zodat ik erop kan plaatsnemen. Recht op en achter zijn betekenisvolle dimensies van de ruimte, dus van het waarnemingsveld. Stond de stoel ergens anders dan moet ik hem eerst verplaatsen voordat ik aan het werk kan. Daarom zijn gewaarwordingen geen dingen maar betekenisvolle gehelen die zich laten vormen door een context. Context is belangrijk voor het waarnemen en is de achtergrond waarbinnen de objecten zich betekenisvol tonen. Een duidelijk voorbeeld van het belang van de context is de Mueller Lyer tekening. Twee gelijke lijnen lijken verschillend van lengte door een andere context.



Er zijn meer voorbeelden van optische illusies maar het laat zien dat hersenen en ogen nauw met elkaar samenwerken. In dit geval worden wij gefopt door verschillende contexten en lijken de lijnen niet van gelijke lengte.

Het belang van de context is duidelijk te zien in de wereld van AI met betrekking tot zelfrijdende voertuigen. Was Elon Musk in 2014 er nog van overtuigd dat in 2015 de eerste zelfrijdende auto's zich op de weg zouden begeven, anno 2023 zijn de vooruitzichten op het kunnen realiseren van zelfrijdende voertuigen een stuk voorzichtiger. Dit heeft alles te maken met het niet begrijpen van het fenomeen waarnemen zoals Merleau-Ponty dat definieert. Een typisch voorbeeld kan dit verduidelijken. Een probleem bij zelfrijdende auto's is het interpreteren van rood licht. Als ergens een reclamezuil rood neonlicht verspreidt zal de robot dit kunnen interpreteren als een stoplicht en gaan staan wachten tot het licht op groen springt. Een ander voorbeeld is een overstekende voetganger die op het trottoir staat en gezien wordt door een zelfrijdende auto.

²⁹ Ibidem, 13.

³⁰ Ibidem, 16.

Maar als deze voetganger eerst achter een object verdwijnt (bijvoorbeeld een geparkeerde vrachtauto) alvorens over te steken zal het zelfrijdende voertuig hier niet op bedacht zijn omdat het de voetganger eerst zag verdwijnen. Om een AI-systeem alle situaties die kunnen voorkomen te laten leren is zonder contextinterpretatie waarschijnlijk een lastige opgave. Het zou ongelukken met zelfrijdende voertuigen hebben geschied als de wereld van de AI te rade was gegaan bij Merleau-Ponty want hij laat zien dat waarnemen iets anders is dan een camera ergens op richten, de informatie op te slaan en te gebruiken voor vergelijking. In het eerste voorbeeld wordt, in tegenstelling tot een AI-systeem, bij een mens de context in de waarneming direct duidelijk. Bij het tweede voorbeeld kan een mens ook de waarneming interpreteren en zal zij begrijpen dat de voetganger eerst verdwijnt om daarna weer te verschijnen. Door de moeilijkheden van AI met het verwerken van contextuele informatie, wordt duidelijk dat waarnemingen betekenisvolle gehelen binnen een context zijn en door zowel biologische als psychologische structuren worden gevormd. Contextuele waarneming zal uiteindelijk essentieel blijken bij het ontwerp van zelfrijdende auto's. Dit wordt overigens wel al erkend en toegepast in de wereld van AI.³¹

3.2. Psychosomatisch karakter van waarnemen

Volgens Merleau-Ponty is zien altijd zien vanuit een bepaald perspectief zonder in dit perspectief te worden opgesloten.³² Ik zie de lamp voor mij vanuit mijn huidige positie achter het bureau. Tegelijkertijd ben ik mij ervan bewust dat de lamp ook een achterkant heeft die, als ik om mijn bureau heenloop, zichtbaar wordt. Ik weet dus dat er andere perspectieven mogelijk zijn. Dat bedoelt Merleau-Ponty met het niet opgesloten zijn in een bepaald perspectief. Als de blik verandert, veranderen daarmee gelijktijdig achtergrond en voorgrond in een continu proces. De scheiding tussen object en omgeving is niet altijd goed aan te geven.³³ Bewustzijn van de wereld ontstaat via het lichaam maar er is meer nodig dan alleen het lichaam. Verbinden met de wereld gebeurt bij de mens door vaste organen en zenuwcircuits. Hierdoor ontstaat een beeld van de

³¹ Chris Fields, 'How humans solve the frame problem', *Journal of Experimental & Theoretical Artificial Intelligence*, 25, (2013) 441-456, aldaar 443.

³² Merleau-Ponty, *Fenomenologie*, 81.

³³ *Ibidem*, 92.

ruimte. Maar dit beeld op zich is betekenisloos. De interactie tussen de geest en het lichaam zorgt ervoor dat dit beeld betekenis krijgt. De ruimte krijgt een geordende inhoud die betekenis heeft en waardoor ik de ruimte kan betreden en er bijvoorbeeld doorheen kan lopen zonder overal tegen aan te stoten. Maar het lichaam als zodanig neemt niet waar. Er is een geest nodig om deze waarneming betekenis te geven. Tussen het fysieke en mentale doen zich uitwisselingsrelaties voor die zich onmogelijk als psychisch of somatisch laten omschrijven. Als wij "naar de wereld zijn" betekent dit dat wij actief gericht zijn op de wereld.³⁴ Dit is een mentale activiteit. Deze activiteit zorgt ervoor dat wij de wereld buiten ons kunnen ervaren. Dit gebeurt door het zien met de ogen wat op zich een lichamelijke gewaarwording is. De tijd die verstrijkt tijdens deze activiteit ervaren wij op deze manier zowel fysiek als mentaal. Als wij zeggen "het is drie uur" dan staat de zon op een bepaalde plek aan de hemel, maar het heeft voor ons ook een bepaalde betekenis (bijvoorbeeld dat het theepauze is). Op deze wijze wordt door Merleau-Ponty de ervaring van tijd en ruimte bij de mens verklaard als een psychosomatische ervaring. Deze ervaring is subjectief en voltrekt zich bij ieder mens op zijn eigen wijze. Dit gebeurt op zich bij dieren ook. Merleau-Ponty betoogt dat de ogen bij de mens zodanig gepositioneerd zijn dat wij in staat zijn dit bewust te doen.³⁵ Toch is het verschil tussen mens en dier hiermee niet verklaard. Bij mensen is bewust waarnemen een specifieke betekenis gevende activiteit. Hoe die betekenis precies tot stand komt wordt hiermee niet verduidelijkt. Dit betekenis geven is een mentale ervaring en lijkt daarmee prioriteit te hebben boven de fysieke ervaring van het zien. Anderzijds is betekenis geven zonder gebruik te maken van zintuigen niet voor te stellen. Ervaring van ruimte en tijd bij robots is moeilijk in te denken. Een robot zal wel in de tijd en de ruimte aanwezig zijn maar ontwerpt die niet zoals bij een mens gebeurt. Een robot kan wel zeggen hoe laat het is en waar het zich bevindt maar kan deze ruimte en tijd niet ontwerpen. Een robot heeft wel een interne klok. Deze klok is zelfs wezenlijk voor een robot om te kunnen functioneren. Maar een robot ervaart geen interne klok. Het ervaren van de werkelijkheid is dus essentieel voor organisch leven. Merleau-Ponty gaat zelfs zover om te beweren dat als wij geen lichaam zouden hebben er geen ruimte meer is.³⁶ Daarmee

³⁴ Ibidem, 104, Het "naar de wereld zijn" is een uitdrukking van Merleau-Ponty. Het is de vertaling van être au monde.

³⁵ Ibidem, 107.

³⁶ Ibidem, 117.

bedoelt hij dat er in dat geval geen ruimtelijke ervaring is al mens. Als een robot uitstaat of geen software heeft geladen dan verandert er niets aan zijn bestaan in ruimte en tijd. Het is een object met capaciteiten die weliswaar nieuw zijn voor machines en op veel gebieden de mens kunnen overtreffen (zoals veel apparaten dat op fysiek gebied kunnen) maar de betekenisvolle ervaring van ruimte en tijd is vooralsnog voorbehouden aan de mens.

3.3. Diepte, ruimte en beweging

Diepte

Ook met betrekking tot het zien van diepte is er vooralsnog geen overeenstemming tussen mens en robot. De mens ziet volgens de huidige inzichten diepte door middel van stereoscopische waarneming, het inschatten van grootte, overlapping en perspectief. Hierdoor wordt een beeld in de hersenen gevormd. Hulpmiddelen zoals het kunnen bewegen van het lichaam en de ogen dragen hieraan bij. Dit wordt vooralsnog aan de werking van de hersenen toegeschreven in samenwerking met de zintuigen. Merleau-Ponty analyseert het zien van diepte als een betrekking tussen mij en de dingen en niet als een (afstand) berekening. Diepte legt het directe verband tussen subject en ruimte bloot. Volgens hem spelen ervaring, bewustzijn en subjectiviteit hier ook een rol.³⁷ Het valt dus te bezien in hoeverre robots dit zullen kunnen evenaren. Zodra perceptie een rol gaat spelen en als beschrijvend middel gebruikt wordt, spelen bewuste en onbewuste psychosomatische processen een rol. Perceptie, het psychosomatisch waarnemen, is nog niet verklaard en nog onderwerp van controverse.

Ruimte

Als het om de ruimte gaat waarin de mens verkeert dan ziet Merleau-Ponty hierin een geleefde ruimte. De geleefde ruimte is een structuurverschijnsel van het subject en deze ruimte kan alleen worden begrepen binnen een waarnemingsveld.³⁸ Dat het een structuurverschijnsel is betekent dat ons brein als het ware al vooraf is ingesteld op de ervaring van de geleefde ruimte. Dit

³⁷ Ibidem, 306.

³⁸ Ibidem, 326.

betekent dat de ruimte zoals wij die zien door ons wordt ontworpen. De zintuigen en het verstand passen zich aan de gegeven ruimte aan. Een blinde heeft een ruimtelijke beleving maar die is anders dan bij ziende mensen. Alhoewel moet worden opgemerkt dat andere zintuigen voor een blinde plaatsvervangend voor het zien kunnen worden. De ruimte die wij waarnemen is automatisch het middelpunt van onze wereld. Als wij hier zijn dan beleven wij deze ruimte en niet de ruimte ergens anders. Wij kunnen een ruimte waar wij niet zijn voorstellen maar niet beleven. Dus concludeert Merleau-Ponty dat de beleefde ruimte een correlaat is van de subjectiviteit.³⁹ De ruimte die wij ervaren is onze ruimte en dus een existentiële ruimte en ons aanwezig zijn in de wereld is een ruimtelijke aanwezigheid. Een ruimte beleven kan dus alleen maar door een subject dat de ruimte percipieert. En perceptie is psychosomatisch waarnemen. Daar waar dit waarnemen in het geding is moet een robot het af laten weten. De ruimte die wij beleven en ervaren is niet de ruimte 'op zich', dat wil zeggen, een ruimte onafhankelijk van het subject, maar een individueel beleefde ruimte. Het percipiërend subject ontwerpt werelden in één enkele beweging die alle objectiviteit verborgen houdt.⁴⁰ Het is de vraag of een robot, ook al herkent het een object, dit nauwkeurig kan beschrijven omdat het aannames moet doen omtrent de totale constitutie. Als een mens een object ziet weet hij in één beweging dat er ook andere perspectieven bestaan die het object vormen. Een robot zal eerst de verscheidene perspectieven moeten ontwaren wil het in staat zijn het object in zijn totaliteit weer te geven en te beschrijven. Een robot kan wel op meerdere manieren waarnemen. Het zal niet moeilijk zijn robots uit te rusten met verrekijkers, microscopen of telescopen die de wereld anders ontwerpen en vormgeven. Dit laat zien dat de mens op zijn eigen wijze de wereld waarneemt maar daarin ook beperkt wordt door zijn eigen fysieke mogelijkheden. Andere manieren van waarnemen zouden tot andere ideeën over de wereld kunnen leiden. Onze manier van waarnemen en interpreteren is niet de enige juiste manier om de wereld waar te nemen. Het lijkt er bij Merleau-Ponty ook op dat er geen objectieve ruimte zou bestaan. Maar hij bedoelt dat wij de objectieve ruimte niet als objectieve ruimte kunnen

³⁹ Ibidem, 338-339.

⁴⁰ Ibidem, 340.

kennen omdat onze kennis altijd bewuste kennis is en via het bewustzijn tot ons komt. Objectieve kennis wordt voor mij mijn ervaringskennis.

Beweging

In de beweging komt het verschil tussen robot en mens in het ervaren van ruimte nog sterker tot uitdrukking.⁴¹ De beweging vindt niet plaats als er geen doel is. Dat houdt in dat het doel al bekend is voordat er bewogen wordt. Dit doel is vaak geen gereflecteerd doel. Als ik wandel hoef ik niet over elke stap na te denken. Zowel de tijd als de ruimte ervaar ik direct. Merleau-Ponty zegt dat de ruimte en de tijd actief worden opgenomen. De grijpbeweging is op magische wijze bij haar doel betrokken.⁴² Als het stoplicht op groen springt steek ik over zonder te reflecteren. De waarneming van het stoplicht vormt één geheel met de betekenis van het op groen springen. Merleau-Ponty zegt dat wij de ruimte niet apart waarnemen maar dat wij de ruimte bewonen. Onze ruimte is niet een betekenisloze ruimte. De werkelijke fysieke ruimte is zonder betekenis maar de ruimte die wij zien heeft betekenis. Wij beleven de ruimte. Het is als het ware een virtuele ruimte. Om de fysieke ruimte wordt een virtuele menselijke ruimte geplaatst en daardoor bewonen wij de ruimte.⁴³ Ik pak mijn laptop en klap hem dicht. In het pakken van de laptop is het doel al bekend. Dat doel is mij bekend en daarmee is het mijn wereld die ik ontwerp met het aan mij bekende doel om de laptop dicht te klappen. De ruimte wordt dan mijn ruimte. Ieder mens bewoont op zijn eigen wijze de ruimte. Het is onmogelijk de ruimte ervaring van een ander te hebben. In het doel waarop de handeling vooruitloopt komt alles samen. De fysieke aansturing van de lichaamsdelen, de zintuigen waardoor wij de buitenwereld kunnen ervaren, en het verstand dat als een besturingscentrum het geheel op elkaar afstemt. Het is dus niet alleen het bewustzijn dat in de beweging een rol speelt. Bewegen is geen intellectuele daad. Het is niet zo dat wij een voorstelling maken in ons bewustzijn van de beweging en dan gaan bewegen. Bewegen toont aan dat het fysieke en het mentale onlosmakelijk in elkaar verstrengeld zijn en dat de manier van bewegen ook persoonsgebonden is. Om dit te verklaren gebruikt Merleau-Ponty het concept van een lichaamsschema. Het lichaamsschema vervult een belangrijke

⁴¹ Ibidem, 121.

⁴² Ibidem.

⁴³ Ibidem, 164.

rol in de filosofie van Merleau-Ponty. Het is het *beseft* van het eigen lichaam. Wij weten hoe wij in de ruimte zijn gepositioneerd en wat de mogelijkheden zijn van ons lichaam en hoe wij dit lichaam kunnen gebruiken. Dit maakt het bijvoorbeeld mogelijk verfijnde motorische handelingen uit te voeren zoals schrijven en het zit als het ware ingebakken in ons DNA. Het is de bewuste gevoelservaring van ons lichaam die ervoor zorgt dat alle motorische taken direct worden uitgevoerd zonder tussenkomst van het voorstellende verstand. Het maakt de typisch menselijke bewuste ervaring mogelijk van de wereld om ons heen en zorgt ervoor dat wij ons kunnen verhouden tot de wereld. In hoeverre het lichaamsschema van mens tot mens uniek is, wordt niet helemaal duidelijk. Bij mensen is er een grote overlap als het om het ervaren van de ruimte gaat. Als wij door een woonwinkel lopen dan zullen wij allemaal de obstakels ontwijken die daar zijn opgesteld en wij interpreteren de tekens die aangeven hoe er gelopen moet worden ook op dezelfde manier. Wij mensen behoren tot de soort mens. Als mens hebben wij bepaalde eigenschappen gemeen. In hoeverre de soort bepalend is voor onze specifieke vermogens en wat individueel is bepaald of door ons specifieke menselijk DNA wordt veroorzaakt blijft onduidelijk. Het concept van het lichaamsschema is voor Merleau-Ponty wel een argument dat gebruikt wordt tegen het rationalisme en empirisme. Tegen het rationalisme omdat het ons in staat stelt voorbewust op de wereld te reageren en ons tot de wereld te verhouden. Tegen het empirisme omdat het betekent dat het meer is dan alleen zintuiglijk of instinctmatig te reageren op onze omgeving. Wij leven niet als een stimulus response model. Merleau-Ponty omschrijft dit als volgt “De motoriek is de primaire sfeer waarin de zin van alle betekenissen in het gebied van de voorgestelde ruimte ontstaat.⁴⁴ De identiteit van een object is gedurende de waarnemingsactiviteit een aspect van mijn eigen lichaam door alle verkennende bewegingen heen. Het lichaam projecteert zich in de beweging naar de dingen en de anderen toe.”⁴⁵

Het is in dit verband niet verbazingwekkend dat het frameprobleem een nog onoverwonnen probleem is in de wereld van AI. Het frameprobleem is een probleem dat ontstaat wanneer een robot moet beslissen welke veranderingen

⁴⁴ Ibidem, 164-172.

⁴⁵ Ibidem, 216.

in de omgeving van belang zijn om rekening mee te houden en welke niet. De robot moet de huidige situatie begrijpen en anticiperen op veranderingen in de toekomst. Het is echter niet altijd duidelijk welke veranderingen van belang zijn en welke veranderingen verwaarloosbaar zijn. Dit probleem wordt versterkt als de robot moet bewegen. Wat zijn de consequenties voor de omgeving als de robot beweegt? Hoe verandert de omgeving door deze beweging? Het vermogen om te kunnen omgaan met onzekerheid en tegenstrijdige informatie is essentieel. Onderzoekers hebben verschillende oplossingen geprobeerd voor het frameprobleem. Eén voorbeeld van een onderzochte oplossing is het gebruik van niet monotone logica. Deze logica incorporeert onzekerheden die kunnen leiden tot een herziening van eerder getrokken conclusies. Echter tot nu toe is deze aanpak nog niet succesvol gebleken. De vraag die gesteld kan worden is of robots onze manier van waarnemen moeten of kunnen overnemen. Dit zou inhouden dat een robot over de werkelijke wereld een menselijke betekenisvolle wereld heen legt. De vraag is of dit mogelijk is. Er moeten misschien andere manieren van waarnemen worden gevonden die wel betrouwbaar zijn (en die mensen misschien zouden moeten overnemen). Ook dieren hebben een lichaamsschema alhoewel daar het bewuste symbolische aspect aan lijkt te ontbreken. Bij dieren wordt het gedrag geleid door het instinct. Daarvan wordt verondersteld dat het geen bewust gedrag is maar dat het een stereotiepe reactie is op een zintuiglijke impuls. Het lichaamsschema bestaat ook voor dieren omdat zij ook de wereld op de voor hen relevante manier moeten kunnen waarnemen. Het is de basis is voor het naar de wereld zijn en is dus onderdeel van levende organische wezens. Robots moeten op zich in staat zijn het doel van hun handelen in de beweging in te voegen. Daarvoor hebben ze kennis nodig van hun eigen fysieke bouw en hun relatieve plaats in de ruimte. Sensorische eigenschappen kunnen goed worden gesimuleerd door robots. Zij zullen ook motorische eigenschappen moeten hebben. Robots zullen dus ook een soort lichaamsschema moeten hebben. De combinatie van doelhandelingen, sensorische eigenschappen en motorische eigenschappen zijn belangrijke onderdelen van het lichaamsschema. Er kan dus sprake zijn van een lichaamsschema bij robots. Het is goed voorstelbaar dat dit deel van het lichaamsschema kan worden overgedragen op robots. Van andere delen van het lichaamsschema zoals betekenisvol waarnemen, is het moeilijk voorstelbaar dat een robot dit zal kunnen.

De filosofie van Merleau-Ponty geeft precies aan waarom het frameprobleem moeilijk is op te lossen. De mens neemt geen objecten waar maar contexten en begrijpt die contexten niet alleen met de geest maar ook met het lichaam. Tijdens het waarnemen veranderen object en achtergrond voortdurend. Leven is voor een mens bestaan en het lichaam drukt de existentie uit zoals de taal het denken.⁴⁶ De existentie is een zingevende activiteit. Het is dus geen toestand en het bewustzijn en het lichaam spelen een belangrijke rol in onze existentie. Een professionele turner zal een turntoestel anders waarnemen dan een stijve hark. Voor de turner is het een mogelijkheid zijn vaardigheid te tonen en voor de ander een mogelijkheid geblesseerd te raken. De existentie, de persoonlijke wijze van in het leven staan, wordt bepaald door onze fysieke mogelijkheden en onze ervaringen in de natuurlijke en culturele wereld. Het is een zingevende activiteit dat wil zeggen dat wij door deze activiteit de wereld om ons heen duiden en betekenis geven. Merleau-Ponty zegt dat het een activiteit is waardoor wat eerst geen zin had, zin verkrijgt en dat de feitelijke situatie wordt hernomen en omgezet in een betekenisvolle situatie.⁴⁷ Elk mens herneemt dus de natuurlijke wereld op zijn eigen wijze. Schaamte, verlangen en liefde worden onbegrijpelijk als de mens een door natuurwetten geregeerde machine zou zijn.⁴⁸ De verstrengeling van lichaam en geest zijn wezenlijk voor de menselijke existentie. Alles wat we beleven of denken heeft meervoudige zin en wordt gestuurd door een lichaam-subject dat altijd al verbonden is met de wereld. Dit druist in tegen de filosofische opvatting dat de mens een combinatie is van geest en ervaring van de buitenwereld. In deze opvatting is een mens niet meer dan een intelligente data verwerkende processor die in een omhulsel van vlees en botten gevangen zit. Hij kan weliswaar de buitenwereld tot voorstelling maken maar het zouden uiteindelijk algoritmen van de geest zijn die bepalen hoe hij dat doet.

3.4. Verschil tussen mens en robot in het waarnemen

Robots zijn tegenwoordig al goed in staat de buitenwereld op te nemen en te interpreteren. Zij doen dit door middel van verschillende technieken. Eén daarvan is bijvoorbeeld de Lidar techniek waarbij met behulp van lasers de

⁴⁶ Ibidem, 188-193.

⁴⁷ Ibidem, 197-199.

⁴⁸ Ibidem, 194-197.

vorm van een object, de afstand en relatieve snelheid wordt berekend. Ook voor meer abstracte of relatieve gegevens zoals boven en onder zijn technieken voorhanden. Boven en onder wordt bijvoorbeeld met behulp van zwaartekrachtmetingen bepaald. Dat de menselijke geest anders werkt wordt door Merleau-Ponty beschreven aan de hand van experimenten met een bril die boven en onder omdraait. Als een proefpersoon deze een aantal dagen draagt zal deze persoon de wereld als rechtop gaan ervaren en denken zelf ondersteboven te leven.⁴⁹ Ook bij een proef met een schuine spiegel vindt er een heroriëntatie plaats. Als een proefpersoon een object laat vallen valt het in eerste instantie schuin maar al na twee minuten vindt er een heroriëntatie plaats waardoor de wereld door de spiegel evenwijdig aan de proefpersoon wordt. Er treedt dus een nieuw kader op ten opzichte waarvan het gezichtsveld opnieuw verschijnt maar dan rechtop. In eerste instantie wordt het beeld van ons in de spiegel vervormd. Maar voor onze hersenen klopt dat niet omdat wij rechtop staan en het object recht naar beneden valt. De spiegel geeft als het ware een vervormd beeld. Onze hersenen zijn in staat het beeld te corrigeren. Door het beeld aan te passen aan onze ervaring van de werkelijkheid kunnen wij de wereld op onze manier interpreteren. Daardoor zijn we in staat ons aan te passen aan nieuwe situaties en omgevingen. Dit toont aan dat wij als het ware een virtueel lichaam hebben. Dit virtuele lichaam is aangepast aan de situatie waarin wij ons bevinden en zorgt ervoor dat de waarneming overeenkomt met onze ervaren werkelijkheid.⁵⁰ Het virtuele lichaam verdringt het reële lichaam zozeer dat de proefpersoon zich niet meer in de wereld voelt waar zij zich daadwerkelijk bevindt. Zo is er een zekere toe-eigening van de wereld door het lichaam.⁵¹ Merleau-Ponty concludeert hieruit dat het lichaam fungeert als agens en een rol speelt bij de totstandkoming van het kader. Wij zijn in de ruimte gesitueerd. De ruimte is in ons contact met de wereld voorondersteld.⁵² Er wordt erkend dat subjectiviteit en begrip van een zelf daarin een rol spelen. Een robot heeft geen beeld van zichzelf ten opzichte van de ruimte en ook geen bewustzijn van zijn omgeving. Hiermee komt een fundamenteel verschil tussen mens en robot aan het licht. Subjectiviteit en bewustzijn zijn vreemd aan een

⁴⁹ Ibidem, 282.

⁵⁰ Ibidem, 288.

⁵¹ Ibidem, 287.

⁵² Ibidem, 291.

robot. Het kan wel een spiegel als zodanig herkennen en de objecten in de spiegel als gespiegelde objecten behandelen maar kan zijn eigen gecreëerde kader niet veranderen. Wat overigens niet wil zeggen dat dit een eigenschap is die niet overdraagbaar zal zijn op robots.

Als Merleau-Ponty het hier over waarnemen van ruimte en object heeft, formuleert hij tegelijkertijd een kritiek op het empirisme en rationalisme. Hij zegt dat zowel bij empirisme als rationalisme bijvoorbeeld kleuren als vaste kwaliteiten worden gezien die in een reflexieve instelling verschijnen. Echter als een mens de ruimte en de objecten daarin percipieert en (be)leeft dan is kleur in een levende waarneming een toegang tot het object. We zien hier dat rationalisme en empirisme nauwer aansluiten bij de werking van een robot dan bij de fenomenologische mensbeschouwing. Ook het classificeren van waargenomen objecten zal bij een robot meer problemen opleveren. Bijvoorbeeld het onderscheid tussen natuurlijke objecten en culturele objecten zijn voor een mens direct in de waarneming gegeven. Een stoeptegel is een cultuurobject maar het is ook een steen en stenen zijn natuurlijke objecten. Omdat bij mensen ook de context direct is gegeven is deze kennis direct in de waarneming gevat.

3.5. Empirisme en Rationalisme

Volgens het empirisme komt alle kennis voort uit de ervaring van de wereld. In eerste instantie komen de ervaringen binnen via onze zintuigen. Kennis wordt afgeleid uit de door de zintuigen verkregen informatie. Deze kennis is in eerste instantie passief. Maar door vergelijking met ideeën die al in het geheugen voorhanden zijn, worden nieuwe ideeën gevormd. Kleuren bijvoorbeeld worden gekend door vergelijking met wat in het geheugen voorhanden is. Nieuwe kennis komt tot stand door nieuwe ervaringen die samensmelten met wat al in het geheugen aanwezig is. Het aanleren van taal bijvoorbeeld gebeurt in de empiristische visie door imitatie en herhaling. Kinderen zien hun ouders een taal spreken en gaan dit gedrag imiteren en door dit steeds te herhalen worden zij taalvaardig.

Volgens Merleau-Ponty kan vanuit het empirisme het bewustzijn niet goed worden verklaard. Binnen het brein wordt, door specifieke processen, het beeld,

dat via de waarneming wordt verkregen, opgebouwd. Het bewustzijn is in beginsel passief en de ervaringen worden omgezet tot ideeën. Maar het empirisme beschrijft slechts blinde processen die nooit het equivalent van het kennen kunnen worden omdat er niemand is die ziet. Er ontbreekt een niet te reduceren zin die in het bewustzijn als een veld voorhanden is.⁵³ Hierdoor ontlenen bijvoorbeeld cultuurobjecten en gezichten van mensen hun fysionomie aan overdracht en aan projecties van herinneringen. Er ontbreekt een niet te reduceren zin die in het bewustzijn als een veld voorhanden is en waardoor iets bijvoorbeeld treurig of levendig is.⁵⁴ In de wereld van AI wordt de empiristische visie op de manier van waarnemen als uitgangspunt genomen die dus door Merleau-Ponty wordt afgewezen omdat deze geen juiste beschrijving geeft van de wijze waarop het bewustzijn werkt.

Het rationalisme is de tegenpool van het empirisme. Bij de rationalisten heeft de geest aangeboren kwaliteiten en alle ideeën ontstaan door de werking van de geest en niet door de ervaring die via de zintuigen tot ons komt. Logica en wiskunde zijn bekende voorbeelden die door de rationalisten worden gebruikt om hun argumenten kracht bij te zetten. Logica komt niet tot ons via de zintuigen maar het toepassen van logica is aangeboren en behoort tot de kernkwaliteiten van de menselijke geest. Het leren en gebruiken van taal is een aangeboren kwaliteit. De taalstructuren zijn al bij de geboorte in de geest van het kind aanwezig.

Ook deze tegenpool van het empirisme, het rationalisme voldoet volgens Merleau-Ponty niet. Ook deze stroming blijft op afstand van de waarneming staan in plaats van zich ermee te verbinden.⁵⁵ Zij gaat uit van de vruchtbare werking van de aandacht. Door de werking van de aandacht verwerf ik de waarheid van het object. Het object moet een intelligibele structuur bevatten die door de aandacht wordt vrijgemaakt. Het bewustzijn constitueert alles. Er wordt in deze opvatting voorbijgegaan aan de contingentie van de aanleiding tot het denken.

⁵³ Ibidem, 74.

⁵⁴ Ibidem, 77.

⁵⁵ Ibidem, 80.

3.6. De fenomenologische opvatting

Bewustzijn volgens de fenomenologische opvatting is niet passief en speelt zich niet alleen af op het vlak van de menselijke geest. Bewustzijn is altijd bewustzijn van iets. Dit iets kan zowel buiten het subject liggen als binnen in het subject. Het is het vermogen van het subject de wereld waar te nemen en daar betekenis aan te geven. Maar het subject kan ook zijn eigen innerlijke gedachten vormen en hierover reflecteren. Het is dus zowel naar buiten als naar binnen gericht. Dit actieve en creatieve proces wordt intentionaliteit genoemd waarmee men wil zeggen dat het bewustzijn altijd actieve gerichtheid is op de innerlijke of externe wereld.

Door het bewustzijn zijn wij in staat betekenis te geven aan de wereld om ons heen en onszelf in deze wereld te ervaren en te begrijpen.

De kritiek van Merleau-Ponty op deze fenomenologische opvatting is het ontbreken van de lichamelijkheid. Merleau-Ponty benadrukt dat bewustzijn een lichamenlijk bewustzijn is en dat hiermee ook de grenzen van het bewustzijn mede worden bepaald. Lichaam en geest functioneren als een eenheid. Elke waarneming is een verandering van de structuur van het bewustzijn. Aandacht is geen formele activiteit maar een schepping.

Bij alle opvattingen ontbreekt een precieze definitie van bewustzijn. De werking van bewustzijn wordt op verschillende manieren geïnterpreteerd maar hoe deze werking tot stand komt in de menselijke geest blijft onopgehelderd. Zolang hier geen duidelijkheid over bestaat is een menselijke robot een illusie. Als het idee van Merleau-Ponty juist is dat lichaam en geest als eenheid functioneren lijkt een menselijke robot niet te verwezenlijken. Als Merleau-Ponty zegt dat men nooit begrijpelijk zal kunnen maken hoe betekenis en intentionaliteit uit groepen cellen en molecuulstructuren kan worden begrepen geeft hij al aan wat later door Dennett wordt overgenomen als de hypothese van het bewustzijn. Dennett stelt dat bewustzijn zich op een ander abstractieniveau dan cellen en molecuulstructuren afspeelt en een eigen bestaanswijze heeft.⁵⁶ Echter dit idee van Dennett wijst in de richting van de rationalistische visie dat betekenis alleen

⁵⁶ Ibidem, 403.

in het brein ontstaat. De combinatie van beide ideeën, een brein dat op een hoger abstractieniveau ideeën produceert maar dat deze productie gestuurd en mede bepaald wordt door lichamelijke processen, is niet ondenkbaar.

4. Intersubjectiviteit

In de fenomenologie wordt een onderscheid gemaakt tussen objecten in de wereld en subjecten of de ander in de wereld. In dit hoofdstuk wordt de intersubjectiviteit behandeld zoals die door Merleau-Ponty verder wordt uitgewerkt. Zoals in de inleiding aangegeven behoort de relatie tot de andere mens tot het meest wezenlijke van het menselijk bestaan en het mens zijn. Om de verschillen tussen mens en robot te analyseren worden de volgende vragen uitgewerkt aan de hand van het boek *Merleau-Ponty and the ethics of intersubjectivity* van Anya Daly.

Welke uitgangspunten hanteert Merleau-Ponty bij zijn analyse van de intersubjectiviteit?

Wat houdt de wisselwerkingsthese in van Merleau-Ponty in relatie tot de zichtbare wereld van de fenomenen?⁵⁷

Wat betekent de wisselwerkingsthese in relatie tot de onzichtbare wereld van taal en denken?

Voor de uitwerking van de intersubjectiviteit introduceert Merleau-Ponty het begrip van de wisselwerking of chiasme. De wisselwerkingsthese is de these dat 'zelf', 'de wereld' en 'de ander' intern aan elkaar gerelateerd zijn. Er is een onderlinge afhankelijkheid. Intern wil zeggen dat de ander en de wereld essentieel zijn voor zelfbewustzijn en dat zelfbewustzijn essentieel is voor de ander en de wereld.⁵⁸ Anders gezegd: ik, de wereld, en de ander zijn in elkaar verstrengeld; met elkaar verweven. Merleau-Ponty werkt de

⁵⁷ In het boek van Daly staat de wisselwerkingsthese (reversibility thesis) centraal. In de voetnoot pagina 27 van haar boek verwijst ze naar het boek van Merleau-Ponty *The Visible and the Invisible*. Dit boek bevat teksten waaraan hij werkte maar die door zijn vroegtijdige dood niet konden worden afgemaakt. Zij verwijst daarbij naar de tekst waarin hij zegt "Chiasm that is reversibility" (VI 263, VI 317). Hoewel zij naast de term reversibility ook de term Chiasm gebruikt kan hieruit worden geconcludeerd dat Chiasm en Reversibility naar elkaar verwijzen. Op pagina 70 staat: "This reversibility, this chiasm, this identity within difference has crucial significance for Merleau-Ponty's project. Pagina 121: Chiasm, the intertwining, both the phenomenal world and language are two aspects of reversibility which is the ultimate truth. Omdat Chiasm beter vertaald kan worden met wisselwerking gebruik ik voor de term reversibility, wat in het Nederlands wordt vertaald met omkeerbaarheid, het woord wisselwerking. De term kan ook worden vertaald met wederkerigheid. Maar Daly gebruikt voor wederkerigheid een andere term namelijk reciprocity. Om dit onderscheid hier aan te houden, wordt wisselwerking gebruikt.

⁵⁸ Ibidem, 65.

wisselwerkingsthese uit voor de zichtbare wereld van de fenomenen en subjecten en voor de onzichtbare wereld van de taal en het denken.

4.1. Perceptie als basis

Merleau-Ponty gaat voor de uitwerking van de intersubjectiviteit uit van de perceptie als basis. De perceptie is fundamenteel voor alle bewustzijnsacten en dus ook voor de perceptie van de ander. De eerste vraag die aan de orde komt is: hoe weet ik dat er anderen zijn zoals ik? Er is hier door veel filosofische stromingen een antwoord op gezocht. Eén ervan is het argument van de analogie. Mijn toegang tot de ander is bemiddeld door het gedrag van de ander dat op mijn gedrag lijkt. Omdat dat gedrag op mijn gedrag lijkt concludeer ik dat de ander net als ik bewustzijn heeft. En uit zijn gedrag concludeer ik dat zijn interne leefwereld dezelfde is als die van mij.⁵⁹ Bij een sprekende robot zal dit argument niet meer geheel opgaan. De robot zal menselijk gedrag uiteindelijk in woord en daad kunnen imiteren maar daarmee is het nog geen mens. Wel laat dit zien dat robots het menselijk gedrag dicht kunnen benaderen. Als we dan het analogie argument gebruiken dan zullen robots voor mensen worden aangezien. Merleau-Ponty bekritiseert het argument van analogie. Het begrijpen van de binnenwereld van een ander is te complex. Bij hem komt het begrijpen van de ander voort uit de wederkerige intentionaliteit, dat is de wederkerige betrokkenheid op elkaar, die is gegeven door belichaamde perceptie, die zich voltrekt nog vóór de reflectie. Als voorbeeld noemt hij boosheid. Boosheid wordt direct door lichaamswaarneming begrepen. Het gebaar van boosheid laat mij niet op de ander haar boosheid reflecteren. Nee het gebaar is boosheid en ik ben in staat de boosheid direct, dus vóór de reflectie, te ervaren. De reflectie op de boosheid komt na de ervaring van de boosheid van de ander.⁶⁰ Bij hem zijn er dus twee onderscheiden momenten van de perceptie. De directe voorreflectieve ervaring van de boze ander en de reflectie erna. En in de voorreflectieve ervaring herken ik de ander als ander.

⁵⁹ Anya Daly, *Merleau-Ponty and the ethics of intersubjectivity*, (Melbourne 2016) 40.

⁶⁰ *Ibidem*, 41.

4.2. Wisselwerking tussen de zichtbare wereld en subjecten

4.2.1. Interne wisselwerking

Hij begint met de wisselwerking binnen het subject. Daarvoor dient het voorbeeld van het aanraken. Als ik mijzelf aanraak is er een verschil tussen het gevoel van aangeraakt worden en van het aanraken. Dit verschil is er echter niet in ruimte en tijd maar wel in de aandacht. En dat is op het niveau van de reflectie en dus achteraf. Gewaarworden in de modus van het aanraken gaat nooit samen met de gewaarwording van het aangeraakt worden. Dit betekent dat een soort van differentie binnen de eenheid van het lichaam van een subject plaatsvindt. Er is dus geen absoluut samenvallen maar ook geen absolute gescheidenheid van deze acten. Dit is de wisselwerking die binnenin het subject plaatsvindt.⁶¹ Deze wisselwerking doet zich voor bij alle zintuigen. Ziende voelen en voelende zien. De zintuigen zijn binnen het subject met elkaar verstrengeld en er is zowel onderlinge afhankelijkheid als onafhankelijkheid. Bij mensen die een zintuig missen, bijvoorbeeld blinden, kan het voelen toch door de hersenen worden verwerkt waardoor de blinde een idee krijgt van de ruimte en de objecten daarin. De schilder Cézanne bijvoorbeeld wilde voordat hij aan een schilderij begon alle zintuiglijke aspecten van een landschap in zich opnemen voordat hij aan een schilderij begon. Hij wilde dan zo schilderen dat je de landschappen als het ware ook kon ruiken.⁶²

4.2.2. Wisselwerking in de relatie tussen subject en object

In onze relatie met objecten ervaren wij de eenheid van ons lichaam pas echt. Al onze zintuigen zijn er in één beweging bij betrokken. De relatie van een subject met een object gaat altijd via het lichaam. Er is dus een eenheid tussen mens en wereld. Het lichaam van het subject en het object zijn beide in dezelfde wereld. Door de wereld en haar objecten waar te nemen ben ik in staat mijzelf als bewust(zijn) te begrijpen en ervaar ik bijvoorbeeld dat ik in de ruimte en de tijd existeer.⁶³ Ik ben in staat het specifieke object waar te nemen maar ook zijn meer algemene eigenschappen. Dus het algemene en het specifieke gaan samen in elke waarneming van een object. Er is dus een wisselwerking tussen

⁶¹ Ibidem, 67-70.

⁶² Ibidem, 73, Merleau-Ponty, Maurice, *Oog en geest*. Rens Vlasblom vert. (3^e druk; (Amsterdam 2017) 22.

⁶³ Ibidem, 71.

het algemene en het specifieke die in de waarneming is gegeven. Het samengaan van verschillende lichamelijke mogelijkheden zoals zien en bewegen wordt volgens Merleau-Ponty mogelijk gemaakt door de inbedding van het subject in de wereld.⁶⁴ Zien en bewegen worden bijvoorbeeld bij een jongleur, die ballen gelijktijdig in de lucht probeert te houden, simultaan uitgevoerd waarbij de zintuigen onderling afhankelijk zijn van elkaar. Als de jongleur te bewust wordt van zichzelf of van iets anders dan gaat het geheid fout. Er is dus een onderlinge afhankelijkheid van aandacht, denken, zien en bewegen. Dit zou in een wetenschappelijk wereldbeeld, waarbij wereld en subject geen eenheid zijn maar de wereld binnen het subject wordt gerepresenteerd, niet mogelijk zijn.⁶⁵ In de robotica wordt geprobeerd een robot als jongleur met behulp van AI deze activiteit te laten uitvoeren. Tot op heden is men hier niet in geslaagd. Het is echter wel denkbaar dat een robot dat zal kunnen. Zeker als er met behulp van neurale netwerken verschillende, met elkaar communicerende, algoritmen gelijktijdig kunnen worden uitgevoerd. Om de ballen, als ze gevallen zijn, dan weer te zoeken en op te rapen en opnieuw te beginnen zal waarschijnlijk nog wat moeilijker zijn. Het is niet ondenkbaar dat robots ramen gaan zemen of ballen in de lucht gaan houden. Dat is dan geen lichamelijke voorbewuste activiteit. Deze activiteit zal vooral door een combinatie van algoritmen en sensorische signalen vorm krijgen. Hiervoor is een voorbewuste lichamelijke perceptie zoals Merleau-Ponty die beschrijft niet meer noodzakelijk. En daarmee de these dat dit alleen door voorbewuste perceptie kan gebeuren ook niet meer als onontkoombaar te verdedigen. Evenals het gebruik van taal is het intelligent manoeuvreren van objecten in de wereld niet meer voorbehouden aan de mens alleen. Wel blijft overeind dat de zintuigen en de samenwerking van de zintuigen nog voor onoverkomelijke problemen zorgen in de wereld van de AI.

Vanuit filosofisch oogpunt is het van belang te doordenken wat de consequenties zijn voor ons als mensen. Als robots in staat zijn prestaties te leveren waarvan eerst gedacht werd dat die alleen aan de mens waren voorbehouden en dit ook nog eens beter en sneller kunnen, moeten we de consequenties hiervan voor onze omgang met elkaar en met de nieuwe

⁶⁴ Merleau-Ponty, *Oog en geest*, 39-41.

⁶⁵ Daly, *Ethics*, 75.

technologie in kaart brengen. Dat is echter in veel gevallen niet mogelijk. We begrijpen de consequenties van nieuwe technologie vaak pas achteraf. De introductie van de mobiele telefoon is hiervan een voorbeeld. Dit is in feite een niet meer weg te denken robot die ons als partner behulpzaam is in een overvloed van menselijke activiteiten. De weg vinden doen we onder begeleiding van een vriendelijke stem uit de telefoon. We komen aan op de plaats van bestemming maar hebben geen idee meer waar we naar toe zijn gebracht omdat we zelf de weg ernaar toe niet meer hebben uitgestippeld. Muziek luisteren, met anderen communiceren, de wereld vastleggen en dit delen met anderen, partners zoeken, fit blijven. Voor alles wordt de mobiele telefoon ingeschakeld. De actieve gedragsbeïnvloeding van deze moderne robot gaat steeds verder zonder dat we tijdig adequate antwoorden formuleren voor de nieuw ontstane werkelijkheid. Er komen nu wel steeds meer regels en richtlijnen voor het gebruik van mobiele telefoons en de daarmee gegenereerde data. De Privacywet is daarvan een voorbeeld en ook de nieuwe DGA (Data Governance Act) waarin de omgang met niet persoonsgebonden data wordt geregeld. Maar tegen de inzet van de zogenaamde trollen, die opzettelijk foutieve informatie verspreiden en mensen daarmee een verkeerde voorstelling van zaken geven en op deze manier hun wereldbeeld proberen te beïnvloeden, is nog geen goed antwoord geformuleerd. Tal van conceptuele onduidelijkheden moeten daarvoor eerst worden opgelost. Vragen die beantwoord moeten worden zijn bijvoorbeeld: wat is foutieve informatie, wanneer is er sprake van opzet, welke verantwoordelijkheden zijn hier in het geding en wie is verantwoordelijk. Niemand had bij de introductie van de mobiele telefoon met alle nieuwe mogelijkheden rekening gehouden, laat staan de consequenties ervan begrepen. De onmogelijkheid om te voorzien wat de consequenties zijn van nieuwe technologie lijkt een structureel onoplosbaar probleem.

4.2.3. Wisselwerking en de ander

De wisselwerking is volgens Merleau-Ponty de sleutel tot de ander en garandeert tegelijkertijd de niet reduceerbaarheid van het zelf. Het duidelijkst komt dit tot uitdrukking in het voelen van de ander. Als ik de ander een hand geef voel ik hem en voelt hij mij, maar hij kan niet voelen hoe ik zijn handdruk ervaar en in die zin blijven wij vreemden voor elkaar. Er blijft een

onoverbrugbare afstand tot de ander. Dit bedoelt Merleau-Ponty met de niet reduceerbaarheid van het zelf.⁶⁶ Er is de gezamenlijkheid van de handdruk maar een verschil in het ervaren van de handdruk. Door mijn ervaring van de handdruk kan ik mij een voorstelling maken van het gevoel van het ervaren van de handdruk door de ander. Als ik een krachtige handdruk voel geeft mij dat een ervaring van bijvoorbeeld een sterk persoon. Door zelf een krachtige handdruk te geven wil ik deze ervaring bij de ander ook opwekken. Als ik de blik van de ander ontmoet worden wij beiden tegelijkertijd opgenomen in dit zien. Wij ervaren een gezamenlijkheid in deze blik. Het flirten is daarvan een versterkt voorbeeld. Omdat ik niet weet hoe de ander deze gezamenlijkheid ervaart blijft er ook een onoverbrugbare afstand. Maar mijn ervaring geeft de mogelijkheid de ervaring van de ander te interpreteren. Dit bedoelt Merleau-Ponty als hij zegt dat de eigen ondoorzichtigheid de ervaring van de ander mogelijk maakt.⁶⁷ Subjectiviteit bevat de anticipatie van de ander. De ander maakt mij bewust van mijn onreduceerbaarheid wat dus wil zeggen dat de ander mij niet kan ervaren zoals ik mijzelf ervaar en andersom. De uitgaande blik naar de ander en de terugkeer van de blik van de ander naar mij is een wisselwerking die mij met de ander verbindt maar de ander ook van mij onderscheidt. De ander wordt niet louter geconstitueerd door het zelf. Het niet louter geconstitueerd zijn door het zelf betekent dat mijn beeld van de ander niet de ander is zoals die door anderen kan worden ervaren. Hiermee wordt het beeld van de ander dus niet alleen door mijn relatie met de ander gedefinieerd. Daarmee opent zich ook een relatie naar andere subjecten in de wereld. Dit verzekert de unieke relatie van de ander tot de wereld. In tegenstelling tot een object daagt de ander mijn centrale plek in de wereld uit. Ik kan mijzelf niet zien zoals de ander mij ziet.⁶⁸ De relatie met de ander en met de wereld is een dynamische relatie. Het ik is geen constante maar constitueert zichzelf steeds opnieuw in relatie met de ander. De relatie openbaart een samengaan en een verschil. Er is een ervaring van gemeenschappelijkheid maar ieder heeft ook zijn eigen ervaring van de interactie. De eigen ervaring staat niet op zich maar wordt mede vormgegeven door al mijn ervaringen uit het verleden. Hij wordt als het ware ingebed in het geheel van wat mij een persoon maakt op een bepaald moment. Is er sprake van een interactie waarbij de habitus van mij en de ander overeenkomen dan

⁶⁶ Daly, *Ethics*, 78.

⁶⁷ Merleau-Ponty, *Fenomenologie*, 411.

⁶⁸ Daly, *Ethics*, 80.

zal de interactie anders verlopen dan bij een vreemde. Het is dus het geheel van al mijn ervaringen dat nieuwe ervaringen vormt.

4.2.4. Interactie met robots

De vraag die nu gesteld kan worden is of de interactie met een robot op een gelijke wijze verloopt of anders is. Stel ik loop in een stad en ben op zoek naar een restaurant. Ik ga naar de VVV en zie een VVV-medewerker met een uiterlijk als een Madame Tussauds look a like van een VVV-medewerker. Ik vraag aan de robot of het een goed restaurant weet. De robot legt mij in perfect Nederlands een aantal alternatieven voor en met een vriendelijke glimlach toont het alles wat ik moet weten om een goede keus te maken. Dan vraag ik of het eten in een bepaald restaurant lekker is. De robot antwoordt dat een bepaald menu erg lekker is. Er lijkt niets op te wijzen dat een gesprek van deze categorie anders verloopt dan elk ander gesprek met een medewerker van een VVV-kantoor. Er is een normale interactie tussen mij en de robot. Beiden lijken te worden opgenomen in een intersubjectieve betrekking waarvan een deel gemeenschappelijk is maar waar toch ieder zijn eigen identiteit behoudt. Echter de mededeling over het menu laat een vooralsnog onoverbrugbaar verschil zien. De robot komt tot het oordeel over het menu op basis van wat gegeven is over de menu's van alle restaurants via een algoritme. Het gaat niet af op de ervaring van het eten van het menu in het restaurant. Oordeelsvorming vindt louter plaats door redeneren. De 'als - dan' redenering van de robot is het enige waarover wordt beschikt. Men zou kunnen stellen volgens de rationalistische opvatting. Het contact met mij, vanuit de robot beschouwd, verloopt op een empiristische manier. Hij neemt mij waar en op basis van een algoritme ziet hij in mij een klant. Ik kan zien dat hij begrijpt wat ik wil omdat deze vraag binnen zijn referentiekader past. Op basis van een algoritme legt hij mij alles uit over het restaurant. Volgens de fenomenologische opvatting ontbreekt een belangrijk aspect en dat is oordeelsvorming gelijktijdig met en op basis van de ervaring. Er is ook geen gemeenschappelijke ervaring omdat de lichamelijke van de ervaring ontbreekt. Vooralsnog is er geen lichaam wat zich uitdrukt. Stel dat er ogen in de robot zitten dan zijn het geen levende ogen en het geven van een hand is niet een wederzijdse ervaring maar een enkelvoudige ervaring en in ieder geval van de kant van de robot geen lichamelijke ervaring.

Er kunnen uit dit gesprek een aantal conclusies worden getrokken. Een robot kan ver komen in het nabootsen van de mens maar zal vooralsnog alleen de mens nabootsen volgens de mensopvatting van het rationalisme en empirisme. Er is volgens de fenomenologische opvatting geen sprake van intersubjectiviteit omdat de ervaring, als zijnde een lichamenlijk en geestelijk samengaan van het in de wereld zijn, ontbreekt. De fenomenologische opvatting van Merleau-Ponty lijkt een menselijke robot vooralsnog uit te sluiten. De grote verschillen tussen mens en robot lijken vooral te vinden in oordeelsvorming en in het opdoen en verwerken van ervaringen. Lichamenlijke ervaringen zijn voor een robot niet mogelijk. Dat is alleen mogelijk bij organische lichamen en het lijkt vooralsnog ondenkbaar dat een robot in een levend lichaam kan worden geplaatst. Dit betekent dat robots fundamenteel anders zijn dan mensen en er ook geen werkelijke gemeenschappelijke wereld kan bestaan. Als we de eigen habitus van een subject definiëren als de sociale omgeving waar een subject zich thuis voelt (familie, vrienden, kennissen) zal een robot er weinig om geven waar het is. Alle lichamenlijke mogelijkheden zoals lekker eten en drinken, de zon op je huid voelen, pijn ervaren, zijn onmogelijk. Een besef van eigen menselijke identiteit lijkt voor een robot niet mogelijk juist vanwege het feit dat de lichamenlijkheid ontbreekt. De variëteit van robots zal, ook al kunnen ze zelfstandig leren, veel kleiner zijn want veel van de variëteit bij mensen heeft fysieke oorzaken die al in de baarmoeder voor onderscheid zorgen en tijdens de groei alleen maar verder toenemen. Een robot stelt wel vragen maar het zijn stereotype vragen. Er is geen besef dat het een vraag is die door een ander kan worden geïnterpreteerd.

4.3. Wisselwerking en het onzichtbare

Met het onzichtbare worden het denken en de taal bedoeld. Het gaat dan over het innerlijke leven van een mens. De manier waarop een mens denkt en hoe dit denken wordt gerealiseerd en uitdrukking vindt via de taal. Dit wordt aangeduid met interioriteit.⁶⁹ De vraag die wordt gesteld is: wat is innerlijk leven, wat is interioriteit? Het gaat niet alleen om het innerlijke leven, maar om het innerlijke leven in relatie tot de wereld. Exterioriteit is wat zich buiten het ik afspeelt, de buitenwereld. Het gaat hier om de eigen innerlijke gesteldheid in relatie tot de

⁶⁹ Ibidem, 88.

buitenwereld. Er is een onderlinge bepaaldheid van binnen en buiten. Dat wil zeggen dat het innerlijke leven en de buitenwereld van elkaar afhankelijk zijn. Merleau-Ponty verheldert hiermee ook de relatie tussen de oorspronkelijke, de natuurlijke wereld, en de culturele wereld. De natuurlijke wereld is de wereld die bestaat, ook als er geen mensen zouden zijn, en de culturele wereld is de wereld zoals wij die als mens ervaren en die door mensen wordt vormgegeven. Hiermee wordt de relatie van het zichtbare (de externe wereld) en het onzichtbare (denken in taal) verklaard.⁷⁰ Taal is volgens Merleau-Ponty het lichaam van het onzichtbare.⁷¹ Het gesproken woord is een werkelijk gebaar. Merleau-Ponty grijpt hier terug op de taaltheorie van de Saussure.⁷² Volgens de Saussure genereert de vocale expressie van taal betekenis door conventies. Taal is een conventie en de conventie staat centraal.. Dat wil zeggen dat het gebaseerd is op afspraken en dat de aard van het teken niet belangrijk is. De Saussure maakt onderscheid tussen betekenaar, betekende en het object waarnaar wordt verwezen. Bijvoorbeeld de klank van het woord schaap als betekenaar dat middels conventies is gekoppeld aan het begrip schaap. Het begrip schaap is het betekende. Betekenaar en betekende verwijzen samen naar het schaap in de wei. Hij zegt nu dat het woord 'schaap' niet louter wordt bepaald door de betekenaar maar dat het daar deels onafhankelijk van is. Schaap had ook 'paahcs' kunnen zijn en daarmee naar hetzelfde schaap verwezen als het woord schaap.⁷³ Om tot begrip te komen is het dus niet nodig dat woorden intrinsiek verbonden zijn met hun betekenaar in de wereld. Een duidelijk voorbeeld hiervan is een schaakspel. De vorm van de stukken is niet van belang. De schaakstukken hebben hun waarde te danken aan de regels van het spel en hun verhouding tot andere stukken in het spel. Hoe ze er uitzien is niet belangrijk, als ze maar van elkaar onderscheiden kunnen worden en volgens de regels van het spel gehanteerd.⁷⁴ Een belangrijke consequentie hiervan is dat het systeem niet uit de delen kan worden afgeleid. Merleau-Ponty gebruikt deze theorie en geeft er een fenomenologische interpretatie aan. Er is een radicaal onderscheid tussen gesproken woorden en sprekende woorden.⁷⁵ Gesproken woorden zijn taal. De taal als zodanig is een formeel systeem. Het is

⁷⁰ Ibidem, 89-90.

⁷¹ Ibidem, 90.

⁷² George Free, 'Language, Speech and Writing: Merleau-Ponty and Derrida on Saussure', *Human Studies*, 13 (1990) 293-307, aldaar 293.

⁷³ Ibidem, 294.

⁷⁴ Ibidem, 296.

⁷⁵ Ibidem, 299.

een institutie. De formele taal is gevormd door het verleden en is als het ware gesedimenteerd in een cultuur. Dat wil zeggen dat de specifieke cultuuruitingen die zich in de loop van de geschiedenis hebben gevormd, zijn neergeslagen of opgeslagen in de taal. De gesproken taal is de weerslag van hoe in het verleden met taal is omgegaan. De uitgesproken woorden zijn de talige uitdrukking van een mens die spreekt. Dat is de constitutieve mogelijkheid van de sprekende taal. De sprekende mens leeft alleen door uitdrukking. Hij is deze uitdrukking.⁷⁶ Als ik een telefoongesprek heb met mijn vriend Jan dan spreek ik met Jan die zich uitdrukt. Ik spreek met die persoon en niet met iets dat zich achter de taal verbergt waarbij de taal als een soort sluier voor Jan is opgetrokken. Nee, het is de persoon Jan waarmee ik spreek. Daarmee is taal het lichaam van het denken. Een woord vertalen in een ander woord is niet voldoende. De linguïstische tekens waarmee een object wordt aangeduid zijn geen attributen van het object maar aspecten van een formeel systeem. Vertalen is niet de simpele omzetting van het ene woord in het andere woord. De betekenis van een woord is betrokken op het gebruik in een taalgemeenschap. Woorden representeren de typische eigenheid van een gemeenschap. Woorden verwijzen dus ook naar een door een taalgemeenschap ervaren werkelijkheid. Ze verwijzen naar de natuurlijke en culturele werkelijkheid zoals die door een taalgemeenschap worden ervaren. Taalgemeenschappen verschillen van elkaar. Hun culturele verschillen worden vormgegeven door de verschillende manieren waarop zij zich uitdrukken. De taal is daarvan het meest in het oog springende voorbeeld. Dit komt niet alleen tot uiting in de grammaticale taalconstructies maar bijvoorbeeld ook in de klank en het ritme waarin de taal wordt gesproken. Dialecten zijn hiervan een voorbeeld. In een dialect onderscheidt een taalgemeenschap zich van naburige taalgemeenschappen. Het teken is inherent gekoppeld aan zijn situatie.⁷⁷ Merleau-Ponty tracht de levende conditie van de taal te vatten. Het sprekende subject leeft alleen door expressie. De levende mens is expressie.⁷⁸ Taal is geen middel, taal is zelfconstitutie.⁷⁹ Merleau-Ponty benadrukt dat taal ook een lichamelijke component heeft.

⁷⁶ Ibidem, 300.

⁷⁷ Daly, *Ethics*, 91,

⁷⁸ Merleau-Ponty, *Oog en Geest* 46-51.

⁷⁹ Daly, *Ethics*, 96.

Ook bij de wisselwerking van het onzichtbare worden drie situaties onderscheiden. Ten eerste de relatie tussen taal en subject, ten tweede de relatie tussen taal, object en wereld en ten derde de rol van taal in de relatie tussen subjecten.

4.3.1. De wisselwerking van taal en denken

Denken en spreken zijn dezelfde act en het denken wordt uitgedrukt in de taal die wordt gesproken.⁸⁰ Zonder te spreken wordt er niet gedacht. Ik denk aan mijn tuin en dat betekent dat ik daarover spreek. Ook al doe ik dat tegen mijzelf. Het is niet mogelijk het denken van het spreken te scheiden. Betekenis geven aan de dingen in de wereld doe ik door erover te spreken. Ik denk dus in gesproken zinnen over datgene wat mij bezighoudt. Het denken verdwijnt als het niet wordt gearticuleerd in taal. Betekenis wordt geboren in de relatie tussen de woorden. Denken en taal zijn twee momenten van dezelfde realiteit aldus Merleau-Ponty.⁸¹ De vraag is nu welke aspecten van deze taaltheorie overeind blijven in het gesprek met een robot? De robot hoort een vraag, begrijpt de vraag en denkt erover na middels de 'als - dan' methode. Het zoekt in zijn geheugen naar vergelijkbare situaties en komt tot een conclusie die het verwoordt in taal. Ook bij de robot krijgen de woorden betekenis door het verband tussen de woorden. De robot drukt zich uit en er ontstaat begrip. Maar dat begrip ontstaat bij mij en niet bij de robot. Als de robot aan de andere kant van de telefoon zit, hoor ik dan het verschil tussen een mens en een robot? Chatten met een robot gebeurt al volop. Deze gesprekken zullen in de loop van de tijd steeds beter verlopen. Robots zullen een groot deel van de gesprekken die wij mensen met elkaar voeren kunnen overnemen. Juist omdat taal de genoemde formele aspecten heeft, zal de robot kunnen spreken. Maar is er sprake van sprekende taal in de zin dat de robot zich als individu uitdrukt? Het is moeilijk hier een antwoord op te geven. De robot scheidt voor ons een wereld door de taal die hij spreekt. We kunnen vragen stellen en het geeft antwoord. De robot kan zelf ook vragen stellen. Zo kan er een levendig gesprek ontstaan. Maar dit is een levendig gesprek voor ons. Voor de robot is het geen gesprek. Er worden bij de robot geen gevoelens opgewekt. Spreken houdt ook verband

⁸⁰ Merleau-Ponty, *Fenomenologie*, 216.

⁸¹ Daly, *Ethics*, 93-95.

met het gevoel. En dit is misschien wel even belangrijk voor de betekenis als de in woorden uitgedrukte betekenis. Het verband tussen spreken en voelen is nog niet opgehelderd. Creativiteit wordt geboren in het veld van het voorreflectieve en het reflectieve.⁸² Als dit het geval is dan kunnen robots niet op een menselijke manier creatief zijn omdat bij hen dit veld niet aanwezig is en ook nooit zal zijn als zij niet over eigenschappen beschikken die voortkomen uit een levend organisme. Dat wil niet zeggen dat robots niet creatief kunnen zijn. Er zijn andere manieren denkbaar waardoor robots wel creatief kunnen zijn. Dit gebeurt feitelijk al op grote schaal door allerlei verschillende soorten grote dataverzamelingen te doorzoeken op verbanden die nooit eerder zijn gelegd. Hierdoor wordt kennis gegenereerd die voorheen niet voorhanden was en dus kan er gesproken worden over een creatief proces. Er is hier dus niet alleen sprake van een nieuwe vorm van intelligentie maar ook een nieuwe vorm van creativiteit. Bedrijven proberen hiermee een voorsprong te krijgen op de concurrentie en ook overheden gebruiken deze nieuwe manier van het analyseren van data en niet altijd op een transparante door burgers te controleren manier. Mensen kunnen dit soort processen niet uitvoeren omdat de capaciteit aan menskracht die daarvoor nodig zou zijn het onmogelijk maakt.

Aandacht is volgens Merleau-Ponty een andere modaliteit van het denken. Aandacht behoort tot de primaire indruk, het is pre-reflectief, het is concentratie. Denken behoort tot de retentie en de protentie. Retentie en protentie zijn begrippen die door Husserl zijn geïntroduceerd bij zijn ontwerp van de fenomenologie. Retentie is een fase in ons bewustzijn. De tijd waarbinnen een bepaalde voorstelling van een directe ervaring in ons geheugen aanwezig blijft. Het object is in ons bewustzijn opgenomen en er is een voorstelling van gemaakt. Protentie betreft de anticipatie op de volgende perceptuele ervaring. Als een zwemmer in het water duikt, nemen wij de zwemmer in ons op en tegelijkertijd anticiperen wij op het volgende moment dat hij het water raakt en zo verder. Door retentie en protentie hebben wij een persoonlijke ervaring van tijd. Denkend zelfbewustzijn bevat twee ervaringen namelijk primaire indruk en denken als retentie en protentie. Dus denken is noodzakelijk splitsend en afstand nemend. Deze splitsing tussen denken en direct ervaren zorgt voor een

⁸² Ibidem, 95.

verschil tussen datgene wat wordt geleefd en wat wordt begrepen.⁸³ Reflectie voorziet in een basis van vervreemding van het direct geleefde. Maar er is een wisselwerking tussen het direct geleefde en het denken.⁸⁴ Robots denken alleen door middel van het 'als – dan' principe en zijn daarin dus beperkter dan mensen.

4.3.2. De wisselwerking van taal en wereld

Hier gebruikt Merleau-Ponty het voorbeeld van de muziek.⁸⁵ Er is een relatie tussen muzikale ideeën en notenschrift. Toch kan het muzikale idee alleen worden begrepen als deze wordt gehoord. Het muzikale idee kan alleen worden doorgegeven door de lichamelijke ervaring en deze heeft een eigen autoriteit die niet via het denken verloopt. Zo heeft de linguïstische zin alleen betekenis als deze refereert aan iets buiten zichzelf. Onze innerlijke wereld is niet de wereld die men ziet. Maar wij bewonen de wereld wel. Het onzichtbare innerlijke leven bewoont de wereld. Is zich bewust van de wereld. Denken aan een object is denken aan een object in taal, door significatie, door de categorieën van de taal zoals significatie en fonetische presentatie. Verschillende talen refereren aan verschillende organisaties van de wereld. Zoals de muziek wordt gehoord en wordt begrepen zo is betekenis iets dat niet wordt opgelegd aan de wereld maar de wereld is zijn betekenis.⁸⁶ Hiermee drukt Merleau-Ponty uit dat bijvoorbeeld het leren van een vreemde taal nog niet betekent dat men de taal spreekt zoals in de betreffende culturele wereld deze taal wordt gesproken. De taal die is gestructureerd volgens die wereld en wordt opgelegd moet eerst begrepen worden om de taal op een natuurlijke wijze te kunnen spreken. Culturele verschillen worden uitgedrukt in de wijze waarop de taal wordt gesproken en dat hangt samen met de gebruiken, normen en waarden die in een bepaald taalgebied vigeren. De wisselwerking bestaat eruit dat we kunnen zeggen dat taal is gestructureerd volgens onze relatie met de wereld en dat onze relatie met de wereld is gestructureerd door taal. Robots kennen de wereld niet. Betekenis bestaat voor hen niet. Dat is de reden dat natuurlijke taal nog onbereikbaar is voor robots. Ook hier is het verband tussen het aanvoelen

⁸³ Ibidem, 96.

⁸⁴ Ibidem.

⁸⁵ Ibidem, 102.

⁸⁶ Ibidem, 101-102.

van taal en betekenis geven nog onvoldoende duidelijk om dit op robots over te brengen.

4.3.3. De wisselwerking bij linguïstische subjecten

Gesproken expressie gaat een stap verder dan het zichtbare omdat gesproken expressie de mogelijkheid van echte wisselwerking heeft. De taal neemt altijd zijn verleden mee. Spraak constitueert in een sediment een mogelijkheid van intersubjectief gebruik.⁸⁷ Dit sediment bestaat uit de ontwikkeling die de taal heeft doorgemaakt gedurende de loop van de geschiedenis en die het tot een onvervreemdbaar onderdeel van de gesproken en geschreven cultuur maakt. De spreker maakt gebruik van deze taal en voegt daar zijn eigenheid aan toe die ook door het verleden is vormgegeven en zo de taal ook weer beïnvloedt. Merleau-Ponty zegt dat levende taal verborgen gevoelens onthult en boven zichzelf uitgroeit. Gesproken woorden verrassen mijzelf en onderwijzen mij mijn denken.⁸⁸ Door spreken word ik gebracht in de aanwezigheid van de ander. Bij de ander voltrekt zich hetzelfde proces van waarnemen en reflectie. De wisselwerking die tussen mij en de ander ontstaat gaat verder dan een uitwisseling van ideeën. Er ontstaat een gezamenlijke wereld. Wij zijn tijdens het gesprek in een gemeenschappelijke wereld. Er is een directe wisselwerking in het samengaan om spraak zijn doel van betekenisvolle uitdrukking te geven. De luisteraar hoort toe maar de rol van spreker en luisteraar zijn uitwisselbaar. De luisteraar wordt spreker en de spreker wordt luisteraar. De stem bepaalt een co-aanwezigheid die sterker is dan zien. De luisteraar is betrokken in alle aspecten van de taaluiting. De conversatie spreekt zich uit in mij en bewoont mij.⁸⁹ In verband hiermee hanteert Merleau-Ponty ook het begrip stijl. Stijl wordt volgens hem gekenmerkt door een bepaalde mate van onderscheiden kenmerken die dienen als vergelijking met dezelfde manifestaties maar zich daar toch van onderscheiden. Men kan zich ermee onderscheiden van anderen. Stijl bij een kunstenaar is een manier van formuleren die herkenbaar is voor anderen maar niet erg herkenbaar voor de kunstenaar zelf. Het geeft het subject een eigen identiteit.⁹⁰ Robots zijn goed in staat een menselijke taal te spreken. Ze kunnen

⁸⁷ Ibidem, 106.

⁸⁸ Ibidem, 107.

⁸⁹ Ibidem, 107-108.

⁹⁰ Ibidem, 113-114.

dat uiteindelijk beter dan mensen. In ieder geval als het om het aantal talen gaat dat ze kunnen spreken maar ook hun uitdrukkingsvaardigheid zal steeds beter worden. Veel antwoorden van ChatGPT zijn al duidelijker geformuleerd dan veel docenten het kunnen verwoorden. Maar een specifieke stijl van uitdrukken die door de ontwikkeling als organisch wezen ontstaat hebben zij niet. Om met elkaar te spreken hebben robots geen menselijke taal nodig. De huidige stand van de techniek vindt plaats door computers die als rudimentaire robots kunnen worden gezien. Zij communiceren in netwerken constant met elkaar met de snelheid van het licht en via de ruimte over het gehele aardoppervlak. Er is geen plek op de wereld die niet door robots kan worden bereikt. Echter voor de robots zelf heeft deze communicatie geen enkele betekenis. Betekenis wordt pas toegekend als informatie aan mensen wordt getoond via de geschreven of gesproken taal. Stel dat twee robots in een menselijke taal communiceren. Als dit gesprek niet van betekenis wordt voorzien door een mens blijft het een zinloos gesprek. Een spelletje schaak tussen twee robots is interessant en de robots zullen er misschien betere spelers van worden. Echter het heeft alleen betekenis als de mens het spel interpreteert. Een mens kan betekenis geven maar een robot kan alleen betekenis weergeven. Als het gaat om de communicatie tussen mens en robot gaat er echter iets anders spelen. De mens neemt dan de robot waar en geeft betekenis aan de robot. Als de robot een menselijke vorm heeft (denk bijvoorbeeld aan een uiterlijk zoals bij Madame Tussauds, waar het verschil tussen echt en onecht nauwelijks is waar te nemen) zal de mens er een culturele betekenis in leggen. Het zal voor mensen nog een hele uitdaging worden een onderscheid te maken. De robot wordt door zijn spreken aan ons present gesteld. Wij horen zijn stem. Dat is een menselijke stem. Hoe zouden wij een onderscheid moeten maken? Als wij een emoticon zien, dan zien wij een gezichtsuitdrukking en niet een rondje met twee stippen en een streepje. Het vereist een mentale inspanning hier iets anders in te zien. Als een robot ons in fatsoenlijk Nederlands te woord staat zal niets ons ervan weerhouden een mens waar te nemen. De filosofie van Merleau-Ponty wordt hier bevestigd op een negatieve manier. Merleau-Ponty zegt dat er tussen mij en de ander een innerlijke betrekking bestaat die de ander als voltooiing van een systeem laat verschijnen waarbij de taal een fundamentele rol speelt. Perspectieven glijden in elkaar over en wij co-existeren in eenzelfde wereld.⁹¹

⁹¹ Merleau-Ponty, *Fenomenologie*, 409.

De mogelijkheid dat er buiten de mens taal wordt gesproken was voor Merleau-Ponty nog ondenkbaar. Er is hier misschien sprake van een ander spreken. Maar is dit een betekenisloos spreken? Zodra een mens bij het gesprek is betrokken kan er niet meer worden gesproken over een betekenisloos gesprek. Maar ook voor de robot is er een, zij het andere, vorm van betekenis. Het zal van dit gesprek leren en de uitgewisselde argumenten misschien gebruiken bij een ander gesprek. Merleau-Ponty zegt over de communicatie tussen twee mensen dat het voor de ander, degene die spreekt, geleefde situaties zijn. Voor mij, degene die luistert, zijn het getoonde situaties. Daar zit het verschil tussen het spreken van een mens met een mens en een mens met een robot. De robot leeft geen situaties maar toont alleen situaties. De getoonde situatie van de robot is de getoonde situatie die ik ervaar. Mijn ervaring is wat dat betreft dus dezelfde bij een mens en een robot.

Wat de uiteindelijke maatschappelijke gevolgen hiervan zijn is moeilijk te voorspellen. Duidelijk is dat er fundamentele veranderingen komen als robots met het uiterlijk en het gedrag van een mens een taal spreken. Dit kunnen zowel positieve als negatieve consequenties zijn. Ondanks dat zullen er verschillen blijven. Momenteel worstelt de wereld van de AI nog met natuurlijke taal. Dit komt omdat, zoals gezegd, communicatie tussen mensen meer is dan alleen spreken en er meer is dan alleen de inhoud van de woorden. Merleau-Ponty onderscheidt bij taal een begripsmatige en een existentiële betekenis. De begripsmatige betekenis gaat over de verwijzing van het woord naar het object in de wereld en de existentiële betekenis gaat over de contextuele en culturele betekenis binnen een taalgemeenschap.⁹² De gevoelsuitdrukking van de taal bepaalt mede de existentiële betekenis. De existentiële betekenis gaat bij een robot verloren en dat is de reden dat de natuurlijke taal een probleem vormt. Het talige gebaar richt zich op een mentaal veld. Door het spreken wordt een gemeenschappelijke wereld tot stand gebracht waarop het actuele en nieuwe woord zich betreft. De zin wordt gevat in een directe act.⁹³ Om zich een taal eigen te maken moet men de wereld die hier wordt uitgedrukt op zich nemen. Maar wij kunnen nooit tot twee werelden gelijk behoren. Ieder mens heeft dus maar één moedertaal. En iedereen die wel eens langere tijd in het buitenland

⁹² Ibidem, 212.

⁹³ Ibidem, 217.

heeft doorgebracht weet dat het vlot spreken van de taal niet inhoudt dat alle taalnuances uit een taalcultuur worden beheerst. Een universeel denken wordt alleen mogelijk bij gebruik van één taal. Een op conventie berustend algoritme zal nooit iets anders dan de natuur zonder mensen uitdrukken.⁹⁴ Merleau-Ponty zegt:

Zodra een mens zich van taal bedient om een levende betrekking met zichzelf of met zijn medemensen aan te gaan is taal niet langer een instrument. Het is een zich tonen en het aangaan van een psychische band die ons met de wereld en met onze medemens verenigt. Het is het onderscheid tussen sprekende woorden en gesproken woorden. Het sprekende woord is de overschrijding van onze existentie boven het natuurlijke zijn uit. Het lichaam moet zelf de intentie worden die het altijd betekent en de mens heeft geen ander middel om het lichaam te leren kennen dan het te leven⁹⁵ (Mooier kan het denk ik niet uitgedrukt worden).

Merleau-Ponty laat op verschillende gebieden van het menselijk zijn zien dat lichaam en geest als een eenheid functioneren. Als voorbeeld kan hier worden genoemd het verband tussen taal en denken. Wij denken niet voordat wij spreken maar spreken is denken en luisteren is begrijpen. Er zijn geen afzonderlijke acten waarbij het denken aan het spreken voorafgaat. Ook het luisteren is geen horen waarna het gesprokene wordt verwerkt door gedachten. Bij het luisteren zijn wij in de ban van het gesprokene, wij zijn erdoor bezeten (in de zin van er volledig door in beslag genomen), en pas naderhand kunnen wij het gesprokene overdenken.⁹⁶ Men hoeft een woord niet eerst voor te stellen om het uit te kunnen spreken. Het is voldoende er uitspraak en klanksensaties van te bezitten als een mogelijk gebruik van mijn lichaam. Men denke hierbij aan de manier waarop ik een pijnlijke plek op mijn lichaam direct kan aanwijzen. Tekens gaan onder in hun betekenis zoals de gespeelde noten in de muziek doorklinken. Telkens benadrukt Merleau-Ponty dat de werking van de geest een act is en dat de wereld door het ik wordt geconstitueerd in de waarnemingsact waarbij lichaam en geest deze als eenheid voltrekken. In de natuurlijke taal klinken emoties door en gebaren zijn eveneens talige uitdrukkingen. Het is op

⁹⁴ Ibidem, 218-219.

⁹⁵ Ibidem, 229-230.

⁹⁶ Ibidem, 250.

dit moment nog niet mogelijk deze talige nuances in een robot te verwezenlijken.

5. Emoties en empathie

Emoties en gevoelens behoren tot het meest wezenlijke van de mens. De emoties en gevoelens van mensen zijn bepalend voor het handelen. De verschillen tussen mens en robot komen in de emoties en gevoelens het meest direct tot uitdrukking. De emoties en gevoelens worden behandeld aan de hand van de volgende vragen:

Wat is er bekend over emoties en gevoelens en wat is de visie van Merleau-Ponty op emoties en gevoelens?

Hoe worden emoties en gevoelens herkend bij mensen?

Wat is empathie?

Is empathie een goede basis voor ethiek?

5.1. Emoties en gevoelens

In de emoties verschilt een mens fundamenteel van een robot. Robots kunnen van alles maar hebben geen emoties die vanuit het organische leven tot stand komen. Er worden wel modellen ontwikkeld om emoties te simuleren maar vooralsnog zijn deze eigenschappen niet aanwezig bij robots. Dit komt omdat het organische een grote rol speelt bij emoties. Emoties zijn direct verbonden met lichamelijke reacties. Denk hierbij aan hartkloppingen bij een schrikreactie of rood worden bij schaamte. Over de wijze waarop emoties werken en hoe de relatie tussen lichaam, emoties, ervaring en denken tot stand komt is nog veel onduidelijk. Het is zelfs niet duidelijk wat emoties precies zijn en of alleen mensen emoties hebben. Cecilia Mun geeft in haar artikel *The Science of Emotion: Mind, Body, and Culture* een aantal van de huidige vragen en discussies over emoties weer.⁹⁷ Zijn emoties informatiesensitief of ervaringssensitief? Zijn emoties sociale constructies of natuurlijke eigenschappen? Zijn emoties als een systeem te herkennen dat gelokaliseerd kan worden in de hersenen? Zijn emoties evolutionair bepaald? Moeten emoties

⁹⁷ Mun, Cecilia, 'The science of emotion, mind, body and culture', *MDPI* (2022) afl. 144 1-20, aldaar 1-10.

op het microniveau van de afzonderlijke emoties worden bestudeerd of op het macroniveau van de emoties als één samenhangend systeem? Heeft het cognitieve toegang tot emoties? Zij constateert dat er nog geen aparte wetenschap van emoties bestaat en wijt dit aan het feit dat het object van onderzoek complex is. Bij Darwin zijn emoties lichamelijke uitdrukkingen van interne gevoelens.⁹⁸ De lichamelijke uitdrukking is een motorische uitdrukking. Er is eerst de mentale perceptie van een externe prikkel die een emotie oproept. James Lange gaat ervan uit dat gevoelens worden bepaald door het lichaam. De motorische reactie van het lichaam gaat bij hem vooraf aan de mentale perceptie van de uitwendige prikkel.⁹⁹ Bij beiden is de biologische herkomst van emoties leidend en ontbreekt de subjectiviteit. De neuroloog Damasio heeft de theorie over emoties verder ontwikkeld.¹⁰⁰ Hij ziet een toenemende complexiteit in emoties en onderscheidt emoties van gevoelens. Bij hem zijn de primaire emoties gevolg van een neurologische activiteit van het zenuwstelsel en de hersenen. Dit zijn de primaire emoties zoals angst en plezier. Gevoelens worden op basis van deze emoties ervaren. Zij volgen een, wat hij noemt, script. Dat is een soort neurologisch programma dat in de hersenen wordt opgeslagen. Dit programma kan voortdurend veranderen als gevolg van de plasticiteit van de hersenen. Deze gevoelens worden in het heden ervaren maar ontstaan op basis van ervaringen in het verleden en de verwachtingen van de toekomst. Deze gevoelens zijn variaties van de primaire emoties. Het hoogste niveau van emoties zijn de gevoelens van het autobiografische zelf. Hierbij gaat het geheugen een rol spelen. De gevoelens zijn als beelden in het geheugen aanwezig en kunnen worden opgeroepen door prikkels van binnen en buiten het lichaam. Hij onderkent universele gevoelens, dat zijn gevoelens die onafhankelijk zijn van culturele omstandigheden zoals verdriet, geluk, afschuw en verrassing; sociale gevoelens zoals schaamte, schuld, jaloezie en trots. Deze zijn afhankelijk van de sociale context waarin men acteert; achtergrond gevoelens zoals enthousiasme, bedruktheid en ontmoediging. Kenmerkend voor zijn theorie is dat alle gevoelens zijn terug te voeren op de primaire emoties en dat emoties ontstaan op de grens van de scheiding tussen

⁹⁸ Tone Roald, Kasper Levin, Simon Køppe, 'Affective Incarnations: Maurice Merleau-Ponty's Challenge to Bodily Theories of Emotion', *Journal of Theoretical and Philosophical Psychology*, 38 (2018) 205-218, aldaar 205.

⁹⁹ Ibidem, 212.

¹⁰⁰ Ibidem, 213.

het ik en de wereld. Zijn theorie over emoties valt binnen zijn algemene theorie over bewustzijn. Hij ziet het bewustzijn als een stroom van zelfbewuste ervaringen.¹⁰¹ De primaire emoties zijn bij Damasio de basis van het zelfgevoel en daarmee ook van het zelfbewustzijn. Gevoelens ontstaan als er prikkels komen die de bewustzijnsstroom onderbreken en er wijzigingen in aanbrengen. Daarmee geven ze het belang aan van deze prikkel voor het individu. Door middel van de emoties kan het subject zich in de wereld handhaven.¹⁰² Al deze theorieën kennen een bepalende rol toe aan het lichaam als basis voor emoties en gevoel. In deze theorieën blijft het stimulus response model uitgangspunt voor de theorie. De response wordt geactiveerd op basis van een prikkel maar blijft passief als die prikkel niet wordt gegeven.

5.2. Emoties en gevoelens bij Merleau-Ponty

Merleau-Ponty wijst het stimulus response model af.¹⁰³ Het lichaam is niet passief en wordt ook niet alleen gestuurd door de cognitie. Het gevoel is deel van alle perceptie. Wij leven als het ware de emoties. De emoties kunnen zelfs het fysieke lichaam beheersen. Bij een fantoom lid is dit het geval.¹⁰⁴ Als iemand zijn vinger verliest door een ongeluk kan het lichaam toch het idee van de vinger in standhouden. Er kan zelfs pijn in worden gevoeld. Hiermee wordt het verleden door middel van de emotie in het heden opgenomen. Het fysieke gebrek wordt als het ware door de emotie versluierd. Het is tegenwoordig zo dat als in een prothese het gevoel van koude of warmte van een object kan worden gerepresenteerd dit tot gevolg heeft dat de patiënt dit gevoel ervaart alsof het zijn eigen hand betreft.

Merleau-Ponty schrijft dat in de seksualiteit de aard van gevoel en emotie het meest prangend tot uitdrukking komt.¹⁰⁵ Het seksuele leven is geen specifiek gedrag op basis van prikkeling van de geslachtsorganen. Het libido is geen instinct. Seksualiteit omvat het hele leven van de mens en geeft de manier aan waarop een mens in de wereld staat. In de seksualiteit komen lichaam en geest samen. Het lichaam is de uitdrukking van het bestaan van een mens. Het is niet

¹⁰¹ Ibidem, 212.

¹⁰² Ibidem, 213.

¹⁰³ Ibidem, 207.

¹⁰⁴ Merleau-Ponty, *Fenomenologie*, 97-100, 210.

¹⁰⁵ Ibidem, 195-196.

zo dat wij een lichaam willen bezitten maar wij willen het lichaam van een persoon bezitten. Er zijn geen aparte zintuigelijke gegevens die alleen afhankelijk zijn van de zintuigen.¹⁰⁶ Alles wordt opgenomen in een geheel dat ons bestaan definieert. Echte liefde bestaat er volgens Merleau-Ponty in dat mijn leven wordt bepaald door dit gevoel.

Emoties leiden tot een gevoel naar het zelf dus naar zelfbewustzijn. Dit zelfbewustzijn heeft een robot niet. Descartes bepaalde het zelf als ik denk maar bij Merleau-Ponty is het zelf meer dan ik denk. Ik denk aan een laptop die voor mij staat. Kan ik dan zeker zijn van het denken en niet zeker zijn van de laptop die voor mij staat. Dat is volgens Merleau-Ponty niet mogelijk omdat de waarneming niet gescheiden kan worden van het denken.¹⁰⁷ Zien betekent iets zien. Het is een intentionele act. Zo is het ook met het willen. Als ik iets wil dan is dat iets in mij. De wil constitueert het object van willen in mij. Ik kan erdoor misleid worden. Maar het kan niet zo zijn dat het willen mijzelf misleidt. Als ik verdrietig ben is het object waarom ik verdrietig ben, buiten mij, maar dat ik verdrietig ben is een interne werkelijkheid. Het verdrietig zijn kan dus alleen maar als er een ik is die verdrietig is. Willen en weten dat men wil is dus dezelfde act. Het weten dat ik wil is onlosmakelijk verbonden met het object van willen. Een gevoel dat ik voel is altijd waar voor mijzelf op het moment dat ik het voel. Er zijn ook onware gevoelens. In dat geval interpreteer ik het gevoel verkeerd. Of gevoelens die naderhand onwaar blijken te zijn. Ons bestaan is voor onszelf nooit helemaal doorzichtig. En het is niet altijd eenvoudig om authentieke gevoelens van illusoire gevoelens te onderscheiden. Er kan alleen zekerheid zijn als ik de acten van willen, liefhebben of verdrietig zijn voltrek. Alleen in de intentionele act kan ik de waarheid ontdekken omtrent mijzelf.¹⁰⁸

Een ander aspect van emoties is dat ze in verschillende culturen op verschillende manieren worden geuit. Boosheid bij een Japanner uit zich door een glimlach een westerling loopt rood aan en stampt met zijn voeten.¹⁰⁹ Dus ondanks dezelfde emotie van boosheid, uit de boosheid zich op een andere

¹⁰⁶ Ibidem, 187-193.

¹⁰⁷ Ibidem, 408.

¹⁰⁸ Ibidem, 434-439.

¹⁰⁹ Ibidem, 220.

manier afhankelijk van de sociale omgeving. Al deze argumenten tonen volgens Merleau-Ponty aan dat het stimulus-response model te kort schiet.

Als emoties ontstaan volgens het stimulus-response model en dus een reactie zijn van een intern mechanisme op een externe prikkel dan is het voorstelbaar dat deze reacties bij robots ingebracht kunnen worden. Er zal nog veel onderzoek naar emoties nodig zijn wil er een betrouwbaar model van emoties gemaakt kunnen worden maar ondenkbaar is het niet. De vraag dient zich aan waarom we robots emotionele reacties zouden geven? De eerste vraag die dan beantwoord moet worden is wat het nut is van emoties voor de mens. Een verklaring voor het hebben van emoties bij mensen is dat zij een rol spelen bij voortplanting, het overleven en de sociale interactie of meer algemeen bij het nemen van beslissingen die met het welbevinden en sociaal functioneren te maken hebben. Zaken als creativiteit, schoonheid en geluk zijn ondenkbaar zonder emoties. Maar het slecht functioneren van organisaties van mensen en individuen zijn vaak op emoties terug te voeren. Jaloezie, sadisme, moordzucht, uitsluiting, streven naar macht en overheersing; al deze zaken vinden hun oorsprong in het emotionele gedrag van mensen. Wat draagt het bij aan het functioneren van organisaties van mensen of individuen als robots boos kunnen worden, verdriet kunnen hebben, enthousiast zijn, bang zijn of ontmoedigd worden? Emoties staan rationeel handelen vaak in de weg. Psychische problemen van mensen hebben vooral betrekking op het gevoelsleven. Puur rationele robots hebben daar geen last van. Robots hoeven niet zo nodig te overleven. Hun beslissingen kunnen efficiënt, effectief en nauwkeurig zijn. Beslissingen kunnen worden genomen op basis van alle informatie die voorhanden is. Echter het nemen van besluiten vereist het afwegen van dilemma's. Het is de vraag of robots zonder menselijke tussenkomst deze dilemma's goed in kaart kunnen brengen en afwegen als welbevinden geen onderdeel van hun zijn uitmaakt. En omdat een besluit van een robot op zich zonder betekenis is, is er een mens nodig om er betekenis aan te verlenen.

Emoties en gevoelens komen volgens Merleau-Ponty dus niet tot stand middels het stimulus-response model. Alleen in de intentionele act komt de waarheid van de emotie en het gevoel tot uitdrukking. Dit betekent voor robots dat het vrijwel ondenkbaar is om emoties en gevoelens te hebben. Het is ook nog verre van duidelijk op welke manier het organische lichaam een rol speelt bij emoties

en gevoelens. Dit maakt dat het hebben van werkelijke menselijke emoties en gevoelens bij robots voorlopig niet denkbaar is. Wel zullen emoties en gevoelens bij robots gesimuleerd kunnen worden. De consequenties hiervan zijn niet te overzien. Aan het simuleren van emoties en gevoelens kunnen naast positieve kanten ook gevaarlijke kanten zitten. Het misleiden van mensen door simulatie van emoties en gevoelens is niet denkbeeldig en kan een bedreiging vormen op allerlei niveaus van menselijke organisatie. Van afzonderlijke personen tot grote maatschappelijke verbanden.

5.3. Het herkennen van emoties

De vraag is: hoe herken ik emoties bij anderen? Neurowetenschappelijk onderzoek heeft laten zien dat er verschillende hersengebieden bij het herkennen van emoties betrokken zijn. Voor het verklaren hoe deze hersengebieden samenwerken om emoties te herkennen is nog veel onderzoek nodig. Het reduceren van het herkennen van emoties tot alleen biologische processen is vatbaar voor kritiek, maar het herkennen van emoties is voor een deel pre-reflectief.¹¹⁰ Voor het herkennen van emoties spelen de spiegelneuronen een rol. Spiegelneuronen zijn neuronen in de hersenen waarvan bewezen is dat zij een rol spelen bij het herkennen van emoties. Zij worden geactiveerd op het moment dat emoties worden herkend en er zelf in het lichaam emoties ontstaan.¹¹¹ De werking van deze spiegelneuronen is pre-reflectief waarmee wordt aangetoond dat representatie geen voorwaarde is voor het herkennen van emoties. Theorieën waarbij de representatie van de buitenwereld als primair voor onze ervaringen is, schieten hierbij tekort. Van pasgeboren baby's wordt gezegd dat zij in staat zijn emoties te herkennen wat erop wijst dat het een aangeboren mogelijkheid is emoties te herkennen. In de groei naar volwassenheid wordt deze eigenschap verder ontwikkeld en wordt het herkennen van emoties verder opgenomen in het lichaamsschema. Trevarthen heeft een theorie ontwikkeld waarbij er sprake is van drie stadia van herkenning van emoties bij anderen.¹¹² Het eerste is de postnatale fase van het direct lichamelijk herkennen van lichaamsuitingen. Het tweede stadium is het

¹¹⁰ Daly, *Ethics*, 176.

¹¹¹ *Ibidem*, 174.

¹¹² *Ibidem*, 207.

stadium waarbij het kind in staat is haar aandacht gezamenlijk met een ander ergens op te richten zodat er een gezamenlijke ervaring ontstaat. Hierbij moet het kind in staat zijn de motorische bewegingen en de gezichtsuitdrukkingen te interpreteren. Het derde niveau betreft de mogelijkheid te begrijpen waarom de ander reageert zoals zij reageert en dit in een logisch verband te plaatsen. Merleau-Ponty zegt dat voor het herkennen van emoties het lichaamsschema een belangrijke rol speelt en voorafgaat aan begrip en herkenning.¹¹³ Het gevoel is onderdeel van de perceptie. Wij voelen ons aangetrokken tot andere personen, zij laten ons onberoerd of wij hebben er een afkeer van. Dit is een door het gevoel ingegeven perceptie van de ander. Dit gevoel is onderdeel van het lichaamsschema dat wij in de loop van de tijd hebben ontwikkeld.¹¹⁴ Onze in het lichaamsschema opgeslagen gevoelsvoorkeuren bepalen mede hoe wij de wereld waarnemen en waar onze aandacht naar uit gaat.

Emotieherkenning is bij robots aanwezig en waarschijnlijk kunnen robots in de toekomst emoties nauwkeuriger herkennen dan mensen. In de wereld van de effectenhandel worden robots ingezet om de emoties van bestuursvoorzitters van beursgenoteerde bedrijven te analyseren bij presentaties van jaarcijfers om na te gaan wat er achter de gesproken woorden zit. De gezichtsexpressies worden door de robots met voorgaande presentaties en andere bijeenkomsten vergeleken om zodoende na te gaan of de bestuursvoorzitter al dan niet gespeeld optimisme vertoont. De inzet van AI om emoties te herkennen toont aan dat het herkennen van emoties niet alleen pre-reflectief is maar dat de herkenning ook nauwkeurig sensorisch waarnemen bevat en opgeslagen ervaringen een rol spelen waarbij een evaluatie achteraf plaatsvindt. Het laat zien dat mensen geen gesloten boek zijn en dat interioriteit wordt gereflecteerd in gedrag en stijl die herkenbaar zijn voor anderen. De inzet van AI voor emotieherkenning roept ethische vragen op. Misbruik ligt op de loer. Het geeft aan dat het belangrijk is om de inzet van AI aan regels te binden.

¹¹³ Merleau-Ponty, *Fenomenologie*, 24-28.

¹¹⁴ Daly, *Ethics*, 215.

5.4. Empathie

Empathie heeft een nauwe verwantschap met het herkennen van emoties. Empathie gaat verder dan het herkennen van emoties. Het is het verschijnsel van het in wisselwerking meevoelen en begrijpen van het gevoel van de ander. Voor het meevoelen is het noodzakelijk de emoties van de ander te herkennen en te begrijpen. Maar voor het empathisch gevoel is meer nodig dan begrijpen en beoordelen van het gevoelsleven van de ander. Naast het begrijpen en beoordelen van de interne gevoelens van een ander is voor empathie ook een gevoelsmatige response nodig. Empathie is intentioneel. Het is de act van meelevendheid met een ander. Het is relationeel en het gebeurt in wisselwerking met de gevoelens van de ander. Het is een misvatting dat het gevoel van de ander een gesloten boek is. Liefde, haat, boosheid, schaamte zijn gevoelsuitingen die in het gedrag van de ander zichtbaar zijn.

Anya Daly gaat in haar boek diep in op het verschijnsel van de empathie. Zij beschouwt empathie als een noodzakelijke maar onvoldoende voorwaarde voor ethiek.¹¹⁵ Haar betoog gaat uit van de door Gallagher ontwikkelde 'interactietheorie van sociale herkenning'.¹¹⁶ Deze theorie toont aan dat er een directe, primaire interactie tussen subjecten plaatsvindt vóór de reflectie. Het gesprek en het gebaar waardoor wij een ander kunnen kennen is na de reflectie. Maar vóór deze reflectie is er al een interactie op basis van een primaire lichamelijke interactie. Hij vindt hiervoor aansluiting bij de ontdekking van de spiegelneuronen en hij vindt hiermee ook aansluiting bij het concept van het lichaamsschema van Merleau-Ponty.¹¹⁷ Hij gebruikt de door Trevarthen onderzochte structuur van de ontwikkeling van intersubjectiviteit bij kinderen. Er zijn drie niveaus waarop de intersubjectiviteit zich afspeelt en waardoor wij toegang hebben tot de ander. Het primaire niveau is het niveau van de lichamelijke intersubjectiviteit aangestuurd door de spiegelneuronen. Het tweede niveau is het niveau van het begrip waarbij de context waarin de interactie plaatsvindt een rol speelt en de relatie tussen subjecten zich ontwikkelt. Hierbij spelen gebaren en mimiek een rol alsmede de gezamenlijke aandacht voor een bepaald onderwerp. Het derde niveau is het niveau van de

¹¹⁵ Daly, *Ethics*, 226.

¹¹⁶ *Ibidem*, 207-209.

¹¹⁷ *Ibidem*, 224.

interactie door middel van taal. Hierbij gaat het om het gesprek tussen subjecten.¹¹⁸ Anya Daly betoogt op basis van deze analyse dat er ook sprake is van drie niveaus van empathie.¹¹⁹ Bij empathie gaat het niet alleen om toegang tot de ander maar is er ook sprake van een antwoord op de ander in de vorm van meegevoel. Dit antwoord op de ander is op dezelfde manier gestructureerd als de toegang tot de ander. Er is hier ook sprake van drie niveaus van toegang. Het primaire niveau van meegevoel vindt plaats op hetzelfde moment als herkenning van de ander als ander op basis van het lichaamsschema en is op het primaire niveau pre-reflectief.¹²⁰ Het meegevoel op het primaire niveau is een modus van de intentionaliteit en is ook noodzakelijk voor het overleven van het subject.¹²¹ Het opent de mogelijkheid tot een wij perspectief en staat daarmee aan de basis van de mens als sociaal wezen. Bij Merleau-Ponty is het wij perspectief onderdeel van de structuur van de intersubjectiviteit.¹²² Dit betekent dat het ik altijd tot één of meerdere groepen behoort. Dit behoren tot een groep wordt uitgedrukt door het 'wij' gevoel. Het gevoel is altijd onderdeel van de wisselwerking tussen subjecten. Het is mogelijk dit gevoel uit te bannen maar daar is een bewuste wilsactie voor nodig. Dit gebeurt in gevallen van ideologieën waarbij het subject wordt gedemoniseerd of ontmenselijkt. Het kan ook betrekking hebben op een pathologische afwijking zoals psychopathie.¹²³ Primaire empathie is gevoelsmatige wisselwerking tussen personen. Deze gevoelsmatige wisselwerking is een modus van de intersubjectiviteit. Het tweede niveau van empathie omvat de mogelijkheid tot het wisselen van perspectief, Het is het niveau waarbij gezamenlijke ervaringen centraal staan en noties van eerlijkheid, schuld en schaamte een rol spelen.¹²⁴ Op het derde niveau speelt de taal een grote rol. Het is op dit niveau mogelijk morele regels te stellen, anderen te beoordelen of ideologieën te verspreiden. Hiervoor speelt het lichamelijke een ondergeschikte rol en speelt de voorstelling een grote rol.¹²⁵ Over empathie zegt Merleau-Ponty dat echt gevoel de afstemming op de ander betekent.¹²⁶ Empathie is geen innerlijke realiteit zoals een emotie maar een relationele manier van zijn die wordt weergegeven in gedrag, stijlen en

¹¹⁸ Ibidem, 206-208.

¹¹⁹ Ibidem, 229-230.

¹²⁰ Ibidem, 224.

¹²¹ Ibidem, 225.

¹²² Merleau-Ponty, *Fenomenologie*, 466-468.

¹²³ Daly, *Ethics*, 238.

¹²⁴ Ibidem, 240.

¹²⁵ Ibidem, 241.

¹²⁶ Merleau-Ponty, *Fenomenologie*, 418.

ontmoeting. Het is een essentiële modus van de intentionaliteit. Daly betoogt op basis van de analyse van Scheler over sympathie dat meevoelen met een ander als de kern van empathie is aangeboren.¹²⁷ Het meegevoel met anderen is essentieel voor ethiek. Zonder het meegevoel heeft ethiek geen basis en is de vraag gerechtvaardigd waarom je moreel zou handelen.¹²⁸ Daly onderscheidt deze ethiek van normatieve ethiek die zij op het tweede niveau van empathie plaatst. Normatieve ethiek gaat uit van het idee dat het noodzakelijk is regels te stellen voor het samenleven van mensen onafhankelijk van het gevoel. De plichtsethiek van Kant is hier een voorbeeld van. Normatieve ethiek gaat niet uit van het eerste niveau van empathie en mist daardoor het aangeboren menselijke aspect van de intersubjectiviteit.¹²⁹

Het is duidelijk dat robots geen morele regels bijgebracht kunnen worden op basis van een primaire vorm van empathie. Robots zijn, net zoals de psychopaat, in staat emoties en gevoelens bij anderen nauwkeurig te registreren zonder dat dit gevoelens van wederkerigheid oproept. Zij kunnen deze gevoelens gebruiken voor het manipuleren van mensen en net zoals de psychopaat deze gevoelens misbruiken ten gunste van zichzelf of hun eigenaar. Dit is een reëel gevaar en het is te voorzien dat dit ook gaat gebeuren. Het is als menselijke maatschappij dus belangrijk regels op te stellen voor het gebruik van menselijke robots. De vraag is of een primaire vorm van empathie nodig is om robots morele regels bij te brengen. Er zijn nog wel kanttekeningen te plaatsen bij de noodzaak van empathie voor het morele leven. Ten eerste maakt Daly niet duidelijk wat de relatie is tussen emoties en empathie. De theorie over de hiërarchie van de emoties waarbij onderscheid wordt gemaakt tussen primaire emoties en gevoelens plaatst empathie in het rijtje van de gevoelens. Is er bijvoorbeeld een niet te reduceren verschil tussen gevoelens van jaloezie en empathie? Volgens Daly is jaloezie iets wat tot het secundaire niveau van de intersubjectiviteit behoort. Maar het is net zo goed te verdedigen dat er drie niveaus van jaloezie zijn en dat jaloezie constitutief is voor de intersubjectiviteit. Door conceptuele vaagheid van begrippen zoals empathie en de gebrekkige kennis over de lichamelijke en mentale werking van emoties en gevoelens is het

¹²⁷ Ibidem, 264-266.

¹²⁸ Ibidem.

¹²⁹ Ibidem.

onduidelijk hoe we deze moeten waarderen en onderscheiden. Dit wordt nog problematischer als we begrippen als vertrouwen, zelfbeheersing, enthousiasme en ontzag moeten definiëren en plaatsen. Ten tweede is het niet te verdedigen dat afwezigheid van empathie alleen voorkomt als een pathologische deficiëntie of een top down proces van destructieve beïnvloeding of conditionering. Wat bij deze interpretaties van empathie ontbreekt is dat empathie positief gestuurd kan worden door cognitie. Doktoren zullen bijvoorbeeld bewust afstand houden tot hun patiënten terwijl zij over het algemeen zeer empathisch zijn. Verkopers zullen hun empathisch vermogen inzetten om uiteindelijk hun doel te bereiken en dat is het verkopen van hun producten. Het is moeilijk te verdedigen dat het bewust afzien van empathie alleen uit negatieve overwegingen wordt gedaan. Bovendien geeft dit aanleiding tot twijfel over de unieke constitutieve eigenschap van empathie op het primaire niveau van de intersubjectiviteit. Als empathie onder de algemene categorie van de gevoelens valt, is aan te nemen dat alle gevoelens constitutief zijn voor de intersubjectiviteit en niet alleen empathie. Als dat het geval is, dan kan empathie niet als unieke eigenschap voor ethiek dienen. Ten derde is er het probleem van de temporaliteit van empathie. Empathische gevoelens spelen een rol in de relatie tussen mensen. Maar deze relaties zijn tijdelijk en wisselend. Gevoelens komen op maar verdwijnen weer net zo snel als dat ze opkomen. Uitbarstingen van woede of juichen bij winst in een wedstrijd zijn gevoelens die zeer tijdelijk zijn. Gevoel van medeleven is dit ook vaak. De vraag is of een vluchtig gevoel van medeleven, een noodzakelijke grond is voor morele regels. Het is juist het inspelen op gevoelens waarbij mensen worden gemanipuleerd. Dit geldt ook voor gevoelens van empathie. De reclamespotjes en zelfs hele programma's waarbij medeleven wordt opgewekt met het uiteindelijke doel geld in te zamelen zijn hier een prangend voorbeeld van. Of hier sprake is van een top down proces van manipulatie is op zijn minst betwifelbaar. Het gevoel van empathie te gebruiken als een noodzakelijke basis voor ethiek heeft als voordeel dat dilemma's worden ontweken. De basis ligt immers in het gevoel en de intentionele act en dat is niet betwifelbaar. Dat het een onvoldoende voorwaarde is maakt het niet minder problematisch. Ethisch handelen, morele regels, zijn afwegingen van dilemma's. Daarbij kan het gaan om de minst kwade optie te kiezen uit twee verwerpelijke opties. Hoe valt daaruit een goede keuze te maken als het gevoel een overheersende rol speelt? Anderzijds is het problematisch goede beslissingsregels te definiëren alleen op basis van logica

en historische gebeurtenissen. De mens als gever van betekenis zal altijd nodig blijven voor het vaststellen van morele regels. Maar het zal noodzakelijk zijn morele regels op te stellen voor het gebruik van robots die niet op empathie zijn gebaseerd.

6. Het menselijk gevoel en robots; de filosofie van Levinas

De vraag die in dit hoofdstuk aan de orde komt is:

Geeft de filosofie van Levinas een meer radicale scheiding aan tussen mens en robot dan de filosofie van Merleau-Ponty?

6.1. De filosofie van Levinas

Merleau-Ponty 's filosofie is gebaseerd op het waarnemende subject. Robots zullen vroeg of laat goed kunnen waarnemen maar Merleau-Ponty laat zien dat waarnemen meer is dan het opnemen en verwerken van beelden van onze omgeving. In de waarneming is heel het 'menselijk zijn in de wereld' betrokken. Echter door de nadruk op waarnemen komen bij Merleau-Ponty emoties en gevoel slechts indirect ter sprake. Waarnemen is ook emotioneel waarnemen. Zelfs een emotieloos uitdrukken is het uitdrukken van een emotie. De talige uitdrukking is een emotionele uitdrukking. Omdat waarnemen het uitgangspunt is, staat de relatie tussen mensen niet op de eerste plaats bij Merleau-Ponty. Het is de vraag of er niet andere, meer radicale, filosofische visies zijn op mens en wereld die het verschil tussen mens en robot duidelijker aan het licht brengen. Denken (Rationalisme) en bewust ervaren (empirisme) gaan ook niet over emoties en gevoelens als de meest fundamentele menselijke eigenschap. De filosofieën die gebaseerd zijn op rationalisme en empirisme zullen daarom ook geen bevredigend antwoord kunnen geven. In de filosofie van Emmanuel Levinas staat de relatie tussen mensen centraal en vinden we misschien een betere aansluiting op dit gegeven.¹³⁰ Emmanuel Levinas was een tijdgenoot van Merleau-Ponty en filosofeerde eveneens in de traditie van de fenomenologie. Levinas' filosofie is een filosofie van de ethiek.¹³¹ Relaties tussen mensen zijn volgens hem per definitie ethische relaties. Zijn hele filosofie is er op uit aan te tonen dat ethiek aan de basis staat van het menselijke samen in de wereld zijn. Levinas zegt dat het metafysisch verlangen het verlangen naar de ander is. Dit

¹³⁰ Levinas (1906-1995) was een Frans-Joodse filosoof. Zijn hoofdwerk is het boek *Totaliteit en Oneindigheid*. Hij heeft als Jood de verschrikkingen van de Tweede Wereldoorlog meegemaakt en deze gebeurtenis heeft zijn leven diepgaand beïnvloed. In zijn filosofie staat de relatie tot de ander centraal. Ethiek vindt zijn oorsprong volgens Levinas niet in het handelen van de mens maar in de relatie van de mens tot de ander.

¹³¹ Levinas, Emmanuel, *Totaliteit en oneindigheid. Essay over de exterioriteit*. Theo de Boer en Chris Bremmers vert. (2^e druk; Amsterdam 2018).

verlangen is transcendent. Dat wil zeggen dat het verder gaat dan de ervaring en dat het de fysieke wereld overstijgt. Hij vraagt zich af hoe het Zelf in relatie kan treden tot de Ander zonder de ander daarbij tekort te doen.¹³² (Met hetzelfde of het Zelf bedoelt Levinas het ik maar dan als algemene term en niet het ik als specifiek subject.) Mensen zijn in de basis egoïstisch. In hun relaties tot anderen hebben zij de neiging zichzelf op de voorgrond te zetten en anderen te gebruiken voor hun eigen doeleinden. Dit noemt Levinas totaliseren. Totaliseren wil zeggen dat de wereld wordt gereduceerd tot het ik. Alles in de wereld om ons heen wordt in dienst gesteld van het ik. Als ik een object zie dan wordt dat object onderdeel van mijn bewustzijn en heb ik de neiging mijn idee van dat object als het ware te beschouwen. Mijn idee van het object is het object geworden en daarmee als het ware mijn bezit en ik kan er macht over uitoefenen. Als ik een ander zie dan heb ik de neiging de ander niet als ander te zien maar mijn bewustzijn van de ander is dan primair en dus heb ik de neiging de ander tot mijn beeld van hem of haar te reduceren. In de taal van Levinas totaliseert het Zelf de Ander. De Ander wordt niet erkend in zijn of haar andersheid. Het beeld of het idee dat ik heb van de Ander is niet de Ander maar is mijn beeld. Ook het beeld dat een eventuele derde heeft van de Ander of dat een gemeenschap heeft van de Ander is niet die Ander. De andersheid van de Ander is fundamenteel. De Ander is anders voor mij. Dit kan niet tot een algemeenheid worden gemaakt maar het blijft de andersheid van de specifieke Ander voor mij. Hiermee fundeert Levinas de relatie van de ene mens tot de andere mens. Relaties zijn persoonlijke relaties. De Ander is daarmee een transcendentale Ander voor mij.¹³³ Dat de ander transcendent is voor mij betekent dat de ander mijn ervaring en mijn wereld overstijgt. Hoe kan het Zelf, als egoïsme optredend, in relatie treden tot de Ander zonder hem of haar van zijn andersheid te beroven? De Ander heeft recht op zijn egoïsme. Het ter discussie stellen van mijn egoïsme door de aanwezigheid van de Ander is ethiek.¹³⁴ Levinas kritiek op de Westerse filosofie is dat de ethische relatie met de ander ondergeschikt gemaakt wordt aan het Zelf. Descartes schrijft: '*ik* denk dus *ik* besta' Ook Merleau-Ponty gaat uit van ik die waarneemt. Bij bezit wordt het andere het mijne. De fout van de Westerse filosofie is dat zij uitgaat van het Zelf

¹³² Ibidem, 6-8.

¹³³ Ibidem, 12.

¹³⁴ Ibidem, 13.

waarbij de Ander wordt gereduceerd tot het beeld dat ik heb van de Ander. Volgens Levinas wordt daarmee de rechtvaardigheid ondergeschikt gemaakt aan de vrijheid van het Zelf. De vrijheid van het Zelf is het standhouden van het ik tegenover de Ander. De vrijheid betekent in de relatie met de Ander mijzelf niet ter discussie stellen maar de Ander ter discussie stellen. De vrijheid van het Zelf is belangrijker dan de rechtvaardige relatie tot de ander. De rechtvaardige relatie tot de ander erkent de absolute scheiding tussen het Zelf en de Ander en daarmee de vrijheid van de Ander.

Het uitgangspunt dat de relatie tussen mensen fundamenteel is voor het menselijk bestaan en op de eerste plaats komt, geeft andere inzichten. Levinas stelt dat de mens in eerste instantie een organisme is dat leeft in de genieting.¹³⁵ 'Leven van...' zegt Levinas is genieting. Wij leven van lekker eten, werk, lucht, licht, toneel, slaap, ideeën enzovoort. Voedsel is de omzetting van het andere in het Zelf. Het voedsel wordt in mij omgezet in energie zodat ik weer verder kan met leven. Dat geldt voor voedsel maar ook voor de arbeid om in mijn levensonderhoud te voorzien. Ik leef door middel van de arbeid. Positief gesteld geniet ik daarin van mijn voedsel en mijn arbeid. Daarom zegt Levinas: alle genieting is voeding. Deze inhouden te leven is inhoud van het leven. Arbeid is iets wat ik doe en door dat doen leef ik de arbeid. Dan is de arbeid de inhoud van mijn leven op het moment dat ik arbeid. Genieten van het leven is het ultieme bewustzijn van alle inhouden die mijn leven vullen. Bewustzijn van mijn bewustzijn is genieting en geen wetenschap of reflectie. Genieting is het egoïsme van het Zelf.¹³⁶ Genieten is geen gevoel. De empiristische psychologie ziet het geluk in het gevoel. Het goed voelen is een vorm van gelukkig zijn. Genieten volgens Levinas is meer dan ataraxie. Met ataraxie wordt een staat van innerlijke rust en kalmte bedoeld en het begrip komt voort uit de Griekse filosofie. Bij het Epicurisme gold deze staat als het uiteindelijke doel van het leven van de mens. Als een mens een staat van innerlijke rust en kalmte bereikt en zich niet door externe omstandigheden uit zijn evenwicht laat brengen spreekt men van ataraxie. Degene die in staat is te genieten en het geluk te zoeken zal deze staat van ataraxie bereiken en de genieting is het surplus van de ataraxie. Er zit hier een schijnbare tegenstelling in. Het leven hangt af van

¹³⁵ Ibidem.

¹³⁶ Ibidem, 80-83.

wat het niet zelf is; datgene wat de natuurlijke en culturele wereld aan het Zelf verstrekt zoals eten, drinken, gesprekken, mooie voorstellingen en verhalen. We kunnen niet leven zonder te eten en te drinken en relaties met anderen te hebben. We zijn voor ons bestaan hiervan dus afhankelijk. Als we gelukkig zijn dan is dat 'ataraxie plus' dus een vorm van onafhankelijkheid. Deze tegenstelling gebruikt Levinas door te zeggen dat wij als mensen van onze wereld afhankelijk zijn maar dat het geluk, het surplus van de ataraxie, een onafhankelijkheid waarborgt. Het Zelf heeft behagen in de dingen waar het van afhankelijk is. Beheersing binnen deze afhankelijkheid zorgt ervoor dat deze afhankelijkheid omslaat in soevereiniteit en dat is geluk. Het Zelf is gelukkig met zijn behoefte. Het leven is affectiviteit en gevoel. Lijden is een tekort aan geluk en wanhopen heeft slechts zin omdat het leven van oorsprong geluk is. Het is geen onwetendheid van het andere maar de exploitatie ervan. De afstand tussen Het Zelf en het andere maakt het wezen van de behoefte uit. Omdat ik het andere niet bezit ben ik behoeftig naar het andere. Het ik is uniek. Niet de uniciteit van de Mona Lisa en geen uniek exemplaar van een soort. Het ik is de mens die zichzelf genoeg is. Het is de absolute scheiding met het andere die deze uniciteit voltrekt. Daarmee is het gescheiden van de totaliteit van de wereld. Volgens Levinas valt het genietende ik buiten het Zijn en met het individuele en algemene wordt het Zijn in zijn meest algemene vorm beschreven.¹³⁷ Levinas' hele denken staat in het teken van kritiek op de ontologie (Zijnsleer) van Heidegger en daarom probeert hij ik, de ander en de relatie tussen beide buiten het Zijn te plaatsen. Bij Heidegger staat het Zijn centraal. Zijn onderzoek is een onderzoek naar wat Zijn betekent. Hij onderzoekt daarmee de mens als een subject (een zijnde) in de wereld. Het subject wordt door Heidegger beschreven als een in het Zijn (in de wereld) geworpen Zijnde (subject). Het bespreken van de discussie tussen Heidegger en Levinas is verder niet relevant in het kader van dit onderzoek en wordt daarom buiten beschouwing gelaten. Het idee van Levinas dat ik, de ander en de relatie tussen beide buiten het Zijn vallen is misschien wat ver doorgevoerd want ook binnen het Zijn kan gesteld worden dat ik, de ander en de relatie tussen beide een uniek pluriform karakter hebben. Levinas wijst het idee van pluriformiteit binnen het Zijn af. Het Zelf kan dus niet in de categorieën van het Zijn worden gedefinieerd omdat het de pluriformiteit ontkent en daarmee de uniciteit van het

¹³⁷ Ibidem, 88-91.

Zelf. Volgens Levinas is de zijnsleer van Heidegger een leer die tot het totaliseren van de Ander leidt. Dit totaliseren leidt uiteindelijk tot overheersing en het onderdrukken van de Ander en daarom wijst hij deze filosofie af. Dat Levinas het ik buiten het Zijn plaatst is onderwerp van controverse en het is o.a. door de kritiek van Derrida dat Levinas in zijn boek *Anders dan Zijn* probeert de kritiek te ondervangen. Die discussie is hier verder niet relevant en er wordt aan voorbij gegaan.¹³⁸ Het niet deelnemen aan de totaliteit van het Zijn is het egoïsme van het geluk en is de absolute scheiding van de totaliteit. Dat wil dus zeggen dat ik als mens buiten het Zijn ben geplaatst. De oorzaak hiervan is dat ik uniek ben en mijn uniciteit is alleen kenbaar aan mijzelf. Het Zelf is zichzelf genoeg vanwege de relatie met het andere. Dat wil zeggen dat de wereld zorgdraagt voor mijn genieting. De wereld maakt genieten mogelijk. De kritische tegenwoordigheid van de Ander stelt het egoïsme ter discussie. Door de Ander word ik erop gewezen dat ik niet alleen ben en rekening met een ander moet houden. De Ander communiceert alleen al door zijn aanwezigheid het 'dood mij niet'. En dit is het ethische beroep dat de ander door zijn aanwezigheid op mij doet. De rede (de taal) maakt de menselijke gemeenschap mogelijk. Maar alleen een gemeenschap van zielen die de taal spreken zou geen gemeenschap zijn. Bij Levinas wordt de mens als genietend wezen centraal gesteld in de relatie met het andere. Het andere en de Ander zijn wezenlijk voor het kunnen leven in en van de genieting. Bij Merleau-Ponty staat de act van de intentionaliteit van het al waarnemende en denkende subject centraal. Maar zowel bij Levinas als bij Merleau-Ponty is lichamelijke een wezenlijk en onverdeeld onderdeel van het menselijk bestaan.

6.2. Intentionaliteit bij Levinas

Levinas maakt onderscheid tussen de intentionaliteit van het bewustzijn en de intentionaliteit van de genieting. Bij de intentionaliteit van het bewustzijn gaat het om de waargenomen wereld die ik tot mijn wereld maak. Dit gebeurt via de waarneming en de voorstelling van het verstand. Het ik staat daarbij centraal en het ik maakt de wereld tot mijn wereld. Alle bewustzijn is bewustzijn van iets. Waarneming is de waarneming van iets, ontroering is ontroering van iets dat

¹³⁸ Derrida Jacques, *Geweld en Metafysica, Essay over het denken van Emmanuel Levinas*, D de Schutter vert. (Kampen 1996).

ontroert en zo voort. De intentionaliteit van het bewustzijn is de voorstelling. Het object in de wereld wordt door mij, middels de waarneming en het verstand, voorgesteld en deze voorstelling in mij is dus gescheiden van het object zelf dat in de wereld bestaat. In de act van de voorstelling wordt het object door degene die voorstelt volledig bepaald. Het is zijn of haar voorstelling en daarmee is het volledig uniek. De voorsteller schept als het ware de voorstelling in de act van het voorstellen. De voorstelling is altijd nu. Is dus altijd heden. Het ik verliest in de voorstelling zijn tegengesteldheid aan het object. Het Zelf blijven is zich iets voorstellen. Ik stel mij alles voor in het waarnemen maar het blijft het ik dat zich iets voorstelt. Dat is de identiteit van het ik. Het object transformeert tijdens mijn voorstelling op de voor mij specifieke wijze van voorstellen. Het is mijn voorstelling. Intelligibel betekent voorgesteld zijn.¹³⁹ Volgens Levinas kan de intentionaliteit van de genieting beschreven worden vanuit het contrast met de intentionaliteit van de voorstelling.¹⁴⁰ Zij houdt vast aan de uitwendigheid die door de transcendentale methode wordt opgeschort. Het tussen haakjes zetten van de werkelijkheid om zich volledig te richten op het bewustzijn wordt door Levinas voor wat betreft de sensibiliteit, het gevoel, afgewezen. Vasthouden aan de uitwendigheid betekent zich lichamelijk opstellen in de wereld. Uitwendigheid aanvaarden betekent met haar in relatie treden. De wijze waarop Het Zelf door het Andere bepaald wordt is 'leven van'. Door te eten wordt het voedsel dat uitwendig is, omgezet in de energie van mijn handelen. Dat gaat niet via de voorstelling. Dat is geen act van het voorstellen maar een act van het genieten. Het voedsel dat ik heb gegeten wordt mijn energie, wordt mijn kracht, wordt mijzelf. Het lichaam is de permanente betwisting van het privilege dat men aan het bewustzijn toekent om aan alles zin te verlenen. Want het voedsel geeft zin aan mijn lichaam en dus aan mij en dat is buiten de voorstelling om. Omdat de genieting (bijvoorbeeld van het voedsel) geen act van de voorstelling is, wordt hiermee de prioriteit die altijd aan de voorstelling wordt gegeven, betwist. De wereld die ik constitueer voedt en baadt me. Ze is voedsel en milieu. Het uitwendige verandert van richting door inwendig te worden. Zij maakt inbreuk op de universaliteit van de voorstelling en op de transcendentale methode. Sensibiliteit is de modus van de genieting. De zintuigelijke kwaliteiten stellen wij ons niet voor. Ze komen niet via het denken tot ons. De zon die op onze huid

¹³⁹ Ibidem, 99.

¹⁴⁰ Ibidem, 100.

schijnt voelen wij en stellen wij ons niet voor. Hier lijkt Levinas in te stemmen met Merleau-Ponty. Merleau-Ponty beschrijft deze zaken als voorafgaand aan het bewustzijn dus aan de lichamelijke ervaring. De sensibiliteit constitueert geen wereld en is de tevredenheid zelve. De zintuiglijkheid is genieting en behoort niet tot de orde het denken. Het esthetische vormt op een ander niveau een terugkeer naar het genieten. Hierin wordt de intellectuele benadering tot genieten.¹⁴¹ Voor het genieten van kunst is kennis van kunst nodig. Voor het genieten van muziek is kennis van muziek nodig. Het meer kennis hoe verfijnder het genieten. Bij Levinas beperkt dit zich uitdrukkelijk tot de mens. De rede hiervoor is dat zijn filosofie van de genieting uiteindelijk uitmondt in een ethiek als relatie tussen mensen. Door de filosofie van de genieting en de daarbij behorende interioriteit van de mens wordt het verschil tussen mens en robot radicaal. Er is een onoverbrugbare andersheid tussen mens en robot. Het is duidelijk dat robots niet kunnen genieten. Voor het genieten is een organisch levend wezen nodig. Als een robot zou kunnen genieten dan zouden wij dat niet direct waar kunnen nemen omdat genieten onderdeel is van de interioriteit. En er is een onoverbrugbare andersheid noodzakelijk voor het bestaan als mens. De robot zou een Zelf zijn. Dat wil niet zeggen dat robots niet op andere terreinen meer kunnen presteren dan een mens. Maar menselijke robots zijn binnen deze filosofie een illusie.

6.3. Levinas versus Merleau-Ponty

We zien bij Levinas een prioritering van het gevoel en dit staat haaks op de filosofieën die denken en ervaren centraal stellen. Eigenlijk de hele Westerse filosofie. Ook Merleau-Ponty geeft het waarnemen en ervaren via de waarneming prioriteit boven het gevoel. De kritiek van Levinas op Merleau-Ponty is dat de prioriteit niet aan het waarnemen moet worden gegeven maar aan het gevoel.¹⁴² Het eerste contact gaat via het gevoel en niet via het waarnemen. Iedereen kent wel het fenomeen dat je een donkere kamer inloopt en dat je voelt dat er iemand aanwezig is maar hem of haar niet ziet. Dit toont in ieder geval aan dat gevoel een belangrijke rol speelt in het contact met anderen.

¹⁴¹ Ibidem, 114.

¹⁴² Daly, *Ethics*, 153.

Een andere kritiek van Levinas op Merleau-Ponty is dat volgens Levinas andersheid alleen mogelijk is tussen subjecten.¹⁴³ Dit is natuurlijk aan te vechten met name ook omdat andersheid ook met andere levende organismen zoals dieren en planten te ervaren is. Levinas wijst dat af maar daar is veel kritiek op gekomen en het is ook niet wezenlijk omdat het wel duidelijk is dat andersheid tussen subjecten anders is dan andersheid tussen subject en andere organismen of objecten.¹⁴⁴ Echter de consequentie van de filosofie van Levinas is dat de ethiek alleen een ethiek van de menselijke relatie is. De hieruit af te leiden moraal zal hierdoor alleen betrekking hebben op de onderlinge relatie tussen mensen. Morele regels zijn ook nodig in relatie met de ons omringende natuurlijke wereld en kunnen niet worden beperkt tot de culturele wereld. Levinas lijkt hier geen oog voor te hebben. Door de komst van robots is het herkennen van emoties en gevoel niet meer beperkt tot mensen. Het gebruik van taal is ook niet meer beperkt tot mensen. Dit betekent dat er relaties tussen mens en robot kunnen ontstaan door waarneming en taal. Deze relaties, hoewel anders van aard dan de relaties tussen mensen, zullen ook aan morele regels gebonden moeten worden om misbruik te voorkomen. Bij relaties tussen mens en robot is er wel sprake van wisselwerking maar van wederkerigheid hoeft geen sprake te zijn. De relatie vanuit de mens gezien is een gevoelsrelatie zoals Levinas laat zien maar de robot is de ultieme rationalist die emoties en gevoel uitsluit. Levinas filosofie biedt hier geen oplossing voor. Het is natuurlijk wel mogelijk de ethische relatie tussen mensen als model te gebruiken voor de relatie tussen mens en robot.

Het derde punt van kritiek van Levinas op Merleau-Ponty betreft het overeenkomen van andersheid en wisselwerking.¹⁴⁵ Dit leidt ertoe dat Het Zelf en de Ander binnen het Zijn blijven. Levinas algemene kritiek op ontologische filosofieën is dat ze totaliseren. Zoals gezegd betekent dit dat alles in dienst gesteld wordt van het ik. Daarom zijn ontologische filosofieën monistisch en solipsistisch en leiden ze tot het overheersen van de ander en uiteindelijk tot de verschrikkingen van de Tweede Wereldoorlog aldus Levinas.¹⁴⁶ De filosofie van

¹⁴³ Ibidem, 155.

¹⁴⁴ Deze kritiek is o.a. door de filosoof Sloterdijk geuit in zijn boek *Sferen*. Ook Derrida heeft hierop kritiek geuit.

¹⁴⁵ Daly, *Ethics*, 155.

¹⁴⁶ Jan Kei, *De filosofie van Emmanuel Levinas, in haar samenhang verklaard voor iedereen*. (4^e druk; Kampen 2009) 31.

Merleau-Ponty is volgens Levinas een ontologie. De Ander wordt tot het Zelf gereduceerd.

Claude Lefort is een andere filosoof die kritiek heeft op Merleau-Ponty. Volgens Lefort is er meer dan wisselwerking tussen subjecten. Er is ook hiërarchie in een relatie. Hiervoor gebruikt hij als voorbeeld de relatie tussen een volwassene en een kind.¹⁴⁷ De relatie van een kind met de wereld wordt bemiddeld door de volwassene. De volwassene staat boven het kind. Maar ook in de relaties tussen volwassenen is er sprake van hiërarchie. Tussen volwassenen is er niet alleen sprake van wisselwerking maar is er ook sprake van machtsrelaties, van asymmetrie. En deze machtsrelaties zijn potentieel gevaarlijk voor de onderliggende partij. Zonder deze asymmetrie zou er geen moraliteit zijn omdat juist deze machtsrelatie, de ongelijkheid tussen mensen en hun mogelijkheden, een beroep doet op de bovenliggende partij om de ander te respecteren. We zien hierbij een parallel met Levinas. Hij definieert de relatie tot de ander ook als een ongelijke relatie. Bij Levinas wordt deze ongelijkheid op een andere manier uitgewerkt dan bij Lefort. Bij Levinas leidt de asymmetrie tot de ethische oproep 'dood mij niet' die van de ander uitgaat.¹⁴⁸ Bij Lefort gaat het om het gemis van de derde persoon die de wisselwerking en intersubjectiviteit begeleidt. Maar ook bij Lefort werkt de asymmetrie door in de mogelijkheid van moraliteit die door de radicale andersheid van de Ander mogelijk wordt.¹⁴⁹ Dan Zahavi heeft deze kritiek op Merleau-Ponty weerlegd.¹⁵⁰ Hij verwijt dat Levinas het Zelf van Merleau-Ponty een te beperkte rol toekent. Het Zelf van Merleau-Ponty bevat ook de interne wisselwerking en de andersheid binnen het subject. Ook de andersheid van de wereld en de wisselwerking tussen subject en wereld blijven aanwezig. Als het subject alle ervaring die het heeft van de buitenwereld en de ander totaliseert zou het subject een soort almachtig subject worden en dat is het niet. Het subject heeft al een besef van zichzelf voordat de ontmoeting met de ander plaatsvindt en dit is geen volledig opgesloten zijn van het zelf in zichzelf. Er moet voor de ontmoeting met de ander al een relatie mogelijk zijn met de buitenwereld. Deze relatie kan perceptueel zijn maar ook via de taal, dus in dialoog, tot stand komen. De structuur van de subjectiviteit en de

¹⁴⁷ Eddo Evink, Eddo, 'Surrender and subjectivity: Merleau-Ponty and Patočka on Intersubjectivity', *Meta: Research in hermeneutics, phenomenology, and practical philosophy*, V (2013) afl. 1, 13-28 aldaar 16. Daly, *Ethics*, 142-143.

¹⁴⁸ Levinas, *Totaliteit*, 169-175.

¹⁴⁹ Evink, 'Surrender and Subjectivity', 18.

¹⁵⁰ Daly, *Ethics*, 158-162

verwantschap die tussen subjecten bestaat garandeert de mogelijkheid van intersubjectiviteit. Volgens Zahavi ontkent Levinas deze vorm van andersheid bij Merleau-Ponty. Echter voor moraliteit is asymmetrie en andersheid binnen sociale relaties noodzakelijk.¹⁵¹ Moraliteit wordt door Merleau-Ponty niet verklaard vanuit de wisselwerkingstheorie.

Anya Daly beargumenteert dat de filosofie van Merleau-Ponty niet tot de traditionele Westerse filosofieën kan worden gerekend. De reden hiervoor is dat het een niet dualistische filosofie is. Zijn theorie van de wisselwerking geldt ook binnen het subject zoals eerder weergegeven. Pluralisme is binnen het Zijn een mogelijkheid.¹⁵²

6.4. De robot en het gevoel

De conclusies die we uit deze discussies kunnen trekken is dat Levinas en Merleau-Ponty op een aantal punten verschil van inzicht hebben maar dat beiden de lichamelijke van het subject als wezenlijk voor de menselijke existentie beschouwen. Merleau-Ponty betoogt dat de interioriteit niet radicaal gescheiden is van de exterioriteit en benadrukt het belang van wisselwerking en wederkerigheid. Voor Levinas is de radicale andersheid van de ander een belangrijk thema. Beiden gaan ervanuit dat het directe contact tussen mensen bepalend is voor verantwoordelijkheid. Proberen we het functioneren van een met AI ontworpen robot erbij te betrekken dan kunnen we zien dat de filosofie van Levinas een beter inzicht geeft in het essentiële verschil tussen mens en robot. Een robot kan denken en waarnemen. Natuurlijk zijn de lichamelijke en de wisselwerking bij Merleau-Ponty aspecten van het innerlijke leven maar de genieting zoals Levinas die beschrijft is niet mogelijk voor een robot. De genieting, het gevoel, is uniek voor mensen en wellicht voor alle levende organismen. Levinas beperkt zijn filosofie nadrukkelijk tot het subject en het intersubjectieve maar niets houdt ons tegen dit op te nemen en verder uit te werken naar ander organisch leven. Bovendien zal het met de komst van robots ook nodig zijn morele regels voor de verhouding tussen mens en robots te ontwerpen. De Westerse filosofie heeft altijd de nadruk gelegd op het denken en

¹⁵¹ Evink 'Surrender and Subjectivity', 18.

¹⁵² Daly, *Ethics*, 156-161.

het bewust ervaren en het gevoel is altijd secundair. Levinas zelf geeft zelfs ook prioriteit aan het denken. Het 'heldere idee' zoals beschreven door Descartes vindt zijn weergave in de voorstelling en deze voorstelling is zetel van de waarheid. Maar deze voorstelling gaat uit van de denker. De denker bepaalt. Maar het bepaalt hem zonder hem aan te raken.¹⁵³ De voorstelling leidt uiteindelijk tot de Husserliaanse intentionaliteit en niet de intentionaliteit van het genieten. Ondanks het belang van het denken staat de intentionaliteit van het genieten en de noodzakelijke relatie hiervoor met het andere en de Ander centraal.

De vraag is of de prioritering van denken en bewust ervaren nog stand kan houden in het licht van AI. Deze aandacht van de filosofie voor rationaliteit en bewuste ervaring ontkent hiermee toch ook de loop van de geschiedenis. Geschiedenis is een irrationeel gebeuren. Het voeren van oorlog is nooit rationeel en toch gebeurt het en het is bepalend voor het verloop van de geschiedenis. De natuurwetenschappen hebben uiteindelijk geleid tot overbevolking en uitputting van de aarde tot een niveau dat niet meer is te controleren. Werkelijk leven is genieten. Genieten van het elementale zoals Levinas dat noemt en dan doelt hij op de elementen aarde, lucht, water en vuur.¹⁵⁴ Deze dienen als basis voor het genieten in een menselijke dus culturele omgeving en vormen deze. Binnen het concept van de genieting kunnen de emoties ook een plaats krijgen. Rationaliteit leidt tot waarheid. Maar de waarheid is dat het handelen van de mens en de mensheid wordt bepaald en geregeerd door het gevoel. Gelukkig doet een robot dat niet. Als AI iets laat zien dan is het wel dat logisch redeneren en waarnemen niet uniek zijn voor het mens zijn en dus niet tot het wezen van de mens alleen behoren.

¹⁵³ Levinas, *Totaliteit*, 96.

¹⁵⁴ *Ibidem*, 104.

7. Conclusie

In dit werkstuk is de verhouding tussen mens en robot geanalyseerd vanuit een fenomenologisch perspectief. De fenomenologie als een typische continentale filosofie leent zich goed voor de analyse hiervan. De fenomenologie zoals die door Husserl is vormgegeven, is door filosofen als Merleau-Ponty en Levinas verder ontwikkeld en hun filosofieën vormen het uitgangspunt voor de analyse.

De bespreking van AI door de filosofe Malabou is gebruikt om de huidige stand van AI te analyseren. Hierbij gaat het om de begrippen bewustzijn, intelligentie en denken.

De ontwikkeling van computers vanuit het concept van von Neumann (een computer met één verwerkingseenheid en een geheugen) tot computers die gebaseerd zijn op de neurale hersenstructuur van mensen, heeft geleid tot een enorme impuls in de ontwikkelingen van AI. Dit geeft aanleiding de begrippen brein, bewustzijn, intelligentie en denken opnieuw te overwegen. Bewustzijn zetelt in het brein maar bewustzijn kan bestaan uit meerdere abstractieniveaus die elk op zich onafhankelijk kunnen functioneren. Of wij de werking van het bewustzijn ooit zullen doorgronden en op computers over kunnen brengen blijft nog de vraag. Het begrip intelligentie is onderwerp van discussie maar de computerarchitectuur van neurale netwerken geeft wel aan dat er nieuwe vormen van intelligentie ontstaan. Deze nieuwe vormen van intelligentie leiden ook weer tot nieuwe vormen van creativiteit. Computers kunnen nog niet denken zoals mensen denken. Daarvoor speelt bewustzijn een te belangrijke rol en betekenis geven is nog iets dat voorbehouden is aan mensen.

Merleau-Ponty gaat in zijn boek *Fenomenologie van de waarneming* diep in op het verschijnsel van het waarnemen. Hij laat zien dat waarnemen veel meer is dan voorstellingen maken van objecten in het brein. Waarnemen is een psychosomatische ervaring van de werkelijkheid waarin denken en waarnemen in elkaar zijn verstrengeld. De problemen waarmee AI momenteel worstelt laten goed zien dat waarnemen meer is dan alleen de voorstelling van objecten. De huidige moeilijkheden met het autonoom laten rijden van auto's is hiervan een voorbeeld. De analyse van het zien van diepte van Merleau-Ponty laat zien dat de mens op dit moment een voorsprong heeft op robots. Het is nog maar de

vraag of robots dat in de toekomst goed kunnen. Ook de ruimte als een geleefde ruimte is voor robots nog toekomstmuziek die misschien wel nooit werkelijkheid wordt.

Bij de analyse van bewegen door Merleau-Ponty wordt het verschil tussen mens en robot nog eens extra benadrukt. Het frameprobleem waarmee AI worstelt laat zien dat bewegen voor robots nog veel problemen kent. De analyse van het bewegen van Merleau-Ponty geeft aan waarom dat zo is.

Het typische verschijnsel dat mensen van hun omgeving een virtuele ruimte kunnen maken om toch in de ruimte een weg te vinden is een voorbeeld dat het bewustzijn een grote rol speelt bij het waarnemen bij mensen en dat dit nog niet mogelijk is bij robots. De ervaring van ruimte en tijd zoals dat bij mensen gebeurt is ook niet mogelijk bij robots. Zij zijn geplaatst binnen ruimte en tijd maar kunnen deze niet ervaren.

De analyse van Merleau-Ponty bevat tevens een kritiek op het rationalisme en empirisme. Het is denkbaar dat er robots komen die volgens de theorie van deze stromingen zullen werken. De kritiek van Merleau-Ponty laat zien dat deze stromingen een onvoldoende antwoord geven op de onderscheidende eigenschappen van mensen. Dit doet de theorie van Merleau-Ponty wel. De psychosomatische werking van de perceptie geeft de mens de mogelijkheid op een unieke manier betekenis te geven. Een robot kan alleen betekenis weergeven!

Voor de analyse van de intersubjectiviteit, de relatie tussen mensen, en intersubjectiviteit in relatie tot robots, is gebruik gemaakt van het werk van Anya Daly. Zij betreft heel het oeuvre van Merleau-Ponty in haar analyse en daarom is haar boek *The ethics of intersubjectivity* een goed uitgangspunt voor deze analyse.

Merleau-Ponty gebruikt het begrip wisselwerking of chiasme als uitgangspunt voor zijn analyse van de intersubjectiviteit. Wisselwerking duidt op het samensmelten, het in elkaar verstrengeld zijn van ik, de wereld en de ander. Hij behandelt deze wisselwerking als fenomeen van de zichtbare wereld en de onzichtbare wereld. Voor deze analyse gebruikt hij de perceptie als basis.

De wisselwerking van de zichtbare wereld doet zich voor binnen het subject als de wisselwerking tussen de zintuigen en tussen de zintuigen en het denken. Deze zijn onderling afhankelijk van elkaar en werken op onnavolgbare wijze met elkaar samen. De wisselwerking tussen subject en object wordt weergegeven aan de hand van het voorbeeld van de jongleur. Ook hier ziet Merleau-Ponty een voorbewuste samenwerking tussen zintuigen en object. De ontwikkelingen in de AI laten echter zien dat deze acties mogelijk ook door robots kunnen worden geëvenaard waarbij er geen voorbewuste activiteit nodig is. De robot zal tegelijk moeten kunnen waarnemen en interpreteren. Interpreteren is een activiteit waarvoor de reflectie noodzakelijk is.

De wisselwerking tussen het ik en de ander is essentieel voor het leven van de mens. De analyse van deze relatie laat zien dat wij elkaars ervaringen niet kunnen delen maar dat wij toch gemeenschappelijke ervaringen kunnen hebben en dat de ander mijn plaats in de wereld ter discussie stelt door een relatie met mij aan te gaan.

De analyse van de interactie tussen mens en robot laat zien dat het lichamelijke aspect van de ontmoeting van mensen ontbreekt bij een robot en dat dit fundamenteel is voor de intersubjectiviteit. Wel kunnen mensen betekenis geven aan wat de robot doet en daarom is het mogelijk robots als mens te interpreteren. De mogelijkheden van robots zijn moeilijk te voorspellen en ook het mogelijke gebruik of misbruik van robots is moeilijk te voorspellen. Regelgeving kan vaak pas plaatsvinden nadat uit ervaring is gebleken of iets op een goede of verkeerde manier wordt ingezet. Dit heeft al geleid tot nieuwe regelgeving voor AI-toepassingen maar vaak komt regelgeving te laat. Omdat robots ook op een verkeerde manier kunnen worden ingezet is het belangrijk regels vast te stellen voor het gebruik van robots in de samenleving.

De wisselwerking als fenomeen van de onzichtbare wereld gaat over de interioriteit (het innerlijke leven) en de uitdrukking hiervan in taal. De taal heeft een formele kant en een culturele kant. De formele kant heeft betrekking op de verwijzing van talige tekens naar de werkelijkheid. De culturele kant heeft betrekking op de wijze waarop taal is vormgegeven en wordt gebruikt in een taalgemeenschap. Merleau-Ponty benadrukt dat gesproken taal zich onderscheidt omdat het subject zich hiermee als persoon uitdrukt. Robots

maken ook gebruik van taal. De vraag blijft open in hoeverre er gesproken kan worden over het zich uitdrukken van robots. Creativiteit lijkt onmogelijk voor robots op de manier waarop mensen creatief zijn. Robots kunnen wel op andere manieren creatief zijn; bijvoorbeeld door verbanden te leggen tussen data waar mensen niet aan hebben gedacht. Door taal drukken mensen zich uit en ontstaat communicatie. Communicatie tussen robots waarbij geen mensen betrokken zijn laat zien dat het geven van betekenis essentieel is voor een zinvol gesprek. Robots kunnen geen betekenis geven. Ze kunnen betekenis weergeven. Het feit dat mensen betekenis kunnen geven aan datgene wat de robot zegt kan ervoor zorgen dat mensen robots gaan zien als mensen van vlees en bloed. Beïnvloeding van mensen door robots is daarom niet alleen denkbeeldig. Regelgeving zal noodzakelijk zijn maar of daar in voldoende mate in kan worden voorzien blijft voorlopig ongewis.

Emoties en empathie behoren tot de meest wezenlijke eigenschappen van een mens. Een mens onderscheidt zich van een robot vooral door het hebben van emoties. Emoties hebben een sterke lichamelijke component. Hierdoor is het niet te voorzien dat robots echte emoties kunnen hebben. Er bestaat nog veel onduidelijkheid over emoties en er zijn verschillende theorieën ontwikkeld over emoties. Er is opvallend genoeg geen aparte wetenschappelijke discipline die zich met emoties bezighoudt. Het stimulus response model wordt door Merleau-Ponty afgewezen. Voor hem zijn emoties en gevoelens onderdeel van een actieve perceptie. Hij gebruikt voorbeelden van het fantoom lid en de seksualiteit om dit te verduidelijken. Emoties kunnen wel worden gesimuleerd door robots. Het is mogelijk dat emoties en gevoelens bij het nemen van beslissingen een belangrijke rol spelen. Dit kan zowel positief als negatief zijn. Omdat robots geen emoties hebben is het de vraag of zij ooit op een goede manier dilemma's kunnen afwegen bij het nemen van beslissingen omdat alleen rationaliteit bij hen een rol kan spelen.

Empathie gaat verder dan emoties. Hiervoor is het van belang emoties te herkennen bij anderen. Robots zullen goed in staat zijn emoties te herkennen. Zij zullen dit in de toekomst waarschijnlijk beter kunnen dan mensen. Regelgeving om dit in goede banen te leiden zal ook hier een belangrijk thema worden. Anya Daly pleit voor een ethiek die mede gebaseerd wordt op basis van empathie. Zij zegt dat empathie een noodzakelijke doch niet voldoende

voorwaarde is voor ethiek. Hierbij kunnen kanttekeningen worden geplaatst. Er is te weinig bekend over emoties en empathie. We weten bijvoorbeeld niet of en hoe emoties en empathie met elkaar samenhangen en op welke manier empathie positief kan bijdragen tot het afwegen van dilemma's. Dit maakt een ethiek op basis van empathie wankel.

Op de vraag of er een filosofie is die een radicaler verschil tussen robot en mens laat zien, is de filosofie van Levinas gebruikt. Levinas gaat uit van de genieting van de mens. Hij ziet een mens in eerste instantie als een (egoïstisch) genietend wezen. De intentionaliteit van de genieting bij Levinas zet de wereld niet tussen haakjes. De wereld is voor de genieting van wezenlijk belang. Door van al het goede te genieten zijn wij in staat te leven en ons te verhouden tot de ander. De kritiek van Levinas op Merleau-Ponty is dat Merleau-Ponty niet uitgaat van het gevoel maar van de waarneming. De filosofie van Levinas die uitgaat van het genietende subject laat een radicaal onderscheid tussen mens en robot zien. Dit genieten lijkt een betere basis te geven voor de analyse van het verschil tussen mens en robot dan denken en waarnemen. Het is namelijk ondenkbaar dat een robot zal kunnen genieten. Ook de belangrijke en onderscheidende rol van de emoties bij de mens kunnen binnen het concept van genieten een plaats vinden. Het zou interessant zijn dit verder te onderzoeken omdat emoties unieker zijn voor het mens zijn dan waarnemen en denken. Ook het verband tussen het geven van betekenis door mensen en emoties en gevoelens is nog onopgehelderd. Bij Levinas gaat het er uiteindelijk om aan te tonen dat ethiek bepalend is voor de relatie tussen mensen. De relatie tussen mensen is ethisch van karakter. Dit zou ertoe kunnen leiden dat morele regels alleen tussen mensen mogelijk zijn. Maar de relatie tussen mens en robot laat zien dat ook voor deze relaties morele regels nodig zijn. Bovendien beschouwen wij tegenwoordig onze relatie tot de natuur ook als ethisch van karakter, dit ethische karakter met betrekking tot de natuurlijke omgeving wordt door Levinas ontkend.

Literatuurlijst

Primaire literatuur

Daly, Anya, *Merleau-Ponty and the ethics of intersubjectivity* (Melbourne 2016).

Levinas, Emmanuel, *Anders dan zijn of het wezen voorbij*. Ab Kalfshoven vert. (Baarn 1991).

Levinas, Emmanuel, *Totaliteit en oneindigheid. Essay over de exterioriteit*. Theo de Boer en Chris Bremmers vert. (2^e druk; Amsterdam 2018).

Malabou, Catherine, *Morphing Intelligence, from IQ measurement to artificial brains*. Carolyn Shred vert. (herz. ed.; New York 2021).

Malabou, Catherine, *Wat te doen met ons brein?*. Jeanne Hollerhoek vert. (Amsterdam 2011).

Merleau-Ponty, Maurice, *De wereld waarnemen*. Jenny Slatman vert. (6^e druk; Amsterdam 2017).

Merleau-Ponty, Maurice, *Fenomenologie van de waarneming*. Douwe Timersma en Rens Vlasblom vert. (3^e druk; Amsterdam 2022).

Merleau-Ponty, Maurice, *Oog en geest*. Rens Vlasblom vert. (3^e druk; Amsterdam 2017).

Secundaire literatuur

Apostolopoulos, Dimitris, 'On the motivations for Merleau-Ponty 's ontological research', *British Journal for the History of Philosophy*, 26:2, (2018) 348-370.

Chrisley, Ron, 'Philosophical foundations of artificial consciousness', *Artificial Intelligence in Medicine* (2008) 119-137.

Critchley, Simon, *The ethics of deconstruction, Derrida and Levinas* (Cambridge 1992).

Derrida Jacques, *Geweld en Metafysica, Essay over het denken van Emmanuel Levinas*. D de Schutter vert. (Kampen 1996).

Evink, Eddo, 'Surrender and subjectivity: Merleau-Ponty and Patočka on intersubjectivity', *Meta: Research in hermeneutics, phenomenology, and practical philosophy*, V (2013) afl. 1, 13-28.

Fields, Chris, 'How humans solve the frame problem', *Journal of Experimental & Theoretical Artificial Intelligence*, 25 (2013), 441–456.

Free, George, 'Language, speech and writing: Merleau-Ponty and Derrida on Saussure', *Human Studies*, 13 (1990), 293-307.

Gardner, Howard, *Intelligence reframed, multiple intelligences for the 21st century* (New York 1999).

Gardner, Howard, *The mind's new science, A history of the cognitive revolution* (New York 1987).

Han, Shengnan, Eugene Kelly, Sharokh Nikou en Eric-Oluv Svee, 'Aligning artificial intelligence with human values: reflections from a phenomenological perspective', *AI & Society* (2021) 1-15.

Hasse, Catherine, 'Posthuman learning: AI from novice to expert?', *AI & Society* (2019) 355-364.

Itrona, Lucas, 'The measure of man and the ethos of hospitality', *AI & Society* (2010) 94-103.

Kei, Jan, *De filosofie van Emmanuel Levinas, in haar samenhang verklaard voor iedereen*. (4^e druk; Kampen 2009).

Liberati, Nicola en Shoji Nagataki, 'Vulnerability under the gaze of robots: Relations among Humans and robots', *AI & Society* (2019) 333-342.

Mazijk, Corijn van, *De wereld als verschijning. Fenomenologie en de twintigste eeuw* (Amsterdam 2021).

Mun, Cecilia, 'The science of emotion, mind, body and culture', *MDPI* (2022) afl. 144, 1-20.

Roald Tone, Levin Kasper, Køppe Simon,' Affective incarnations: Maurice Merleau-Ponty 's challenge to bodily theories of emotion', *Journal of Theoretical and Philosophical Psychology*, 38 (2018) afl. 4, 205–218.

Websites

<https://chat.openai.com/>, Laatst geraadpleegd op 21-06-2023.