PRACTICALLY HUMAN

Performing Social Robots
and
Feminist Aspects on Agency, Body and Gender

Karin Victorin

Supervisor: Ericka Johnson and Tara Mehrabi, Gender Studies, LiU

Gender Studies – Intersectionality and Change

Master's Thesis 15 ECTS credits

ISNR: LIU-TEMA G/GSIC1-A—19/013-SE

Abstract

Through an experimental theatre play, this thesis explores the development of human-like agency in contemporary "social robot" technology. The entrance point of this study is the gender gap and lack of diversity in contemporary AI/robot development, with an emerging need for interdisciplinary research across robot technology and social sciences. Using feminist technoscience and critical posthumanism as the theoretical framework, this research involves an analysis of a particular social robot case, currently being developed at Furhat Robotics in Stockholm. Inspired by Judy Wajcman (2004), I analyze how socially intelligent machines impact perceptions of human agency, body, gender, and identity within cultural contexts and through interaction. The first part of the empirical research is carried out in the robot-lab. The robot is then, in the second part, invited to perform as an actor in a theatre play. Entangled amidst the other players and audience members, a queered agency starts to reveal itself through human-machine "intra-action" and embodiment (Barad 2003). Human-like agency in machines is shown to be a complex matter, drawing the conclusion that human-beings are vulnerable to a myriad of entanglements and preconceptions that artificial intelligence potentially embodies.

Keywords: Feminist technoscience; critical posthumanism; posthuman; social robot; AI; technology and art; robots in theatre; human-machine; agency; body; gender; intra-action; cyborg; liminal personae; human-like machines; Judy Wajcman; Karen Barad; Donna Haraway.

Acknowledgements

The long and winding path of this adventurous study wouldn't have been possible without the company and support of many remarkable people. I want to express gratitude and offer special thanks to my supervisors, Ericka Johnson and Tara Mehrabi, for patiently and consistently keeping me on track and their wise guidance in narrowing my research topic. To my dear co-creator of the theatre play *Arise Amazons*, Rebecca Tiger, for your fierce activism- later that year (2017) the #meetoo movement exploded and you initiated the international circus call #theshowisover. To Furhat Robotics, and CEO Samer Al Moubayed, for the trust in letting me experiment freely with your social robot. To friend and mathematician, Petter Lidberg, for your budding talent as the *Wizard-of-Oz* and stagehand to the social robot, Furhat. To ATHENA, for spiritual, as well as grammatical support. To the amazing amazon-sisters and our stage crew (including the horses). And last but not least- to ALL courageous women out there- passionately engaging, individually and collectively, in creating necessary change.

Contents

Chapter 1: Introduction …….....1
Thesis and chapter structure - 4

Chapter 2: Methodology…….....5
Why autoethnography? - 6
Introducing the social robot case - 8

Chapter 3: Theoretical concepts……...9
Feminist technoscience studies (FTS) - 9
Critical posthumanism - 10
Posthumanist performativity; re-thinking agency, body and gender - 11
Feminist technoscience and human-machine interaction - 13
Summary- 16

Chapter 4: Analysis of social agency imagined in machines…….17
Autonomous robots and issues on agency - 17
Sociability embodied in Machines - 18

Chapter 5: Conversation about bridging the gap between human and machine…21

Chapter 6: Beyond human performativity……..24
The Wizard-of-Oz and human-machine agency - 24
Body and social intra-action - 27
Elaborating on gender in the Furhat/robot character - 29
Summary of human-robot performance - 31

Chapter 7: Conclusion of human-like agency in social robots……..33
Agency analyzed as correlating to body and gender - 33
Possible directions for further research - 34
A cyborgian rite of passage – 36

References……...37

Appendix…….41
Lyrics *"I Am What I Am"* (Gloria Gaynor) connected to pp: 29-30.

Chapter 1: Introduction

Standing at center stage, I noticed the slippery feeling of my sweaty palm, firmly gripping the remote control for the visual presentation I was about to give. I looked out in to the audience of about 200 people. Mostly men, some women, I knew most of them belonged to the company and that they were some of the world's smartest quant mathematicians and AI developers. This was not the familiar stage of the theatre where I, as a performance artist felt at home. I took a deep breath; a 5-minute lightning talk lay ahead where I would be the center of attention and my message needed to come through; how feminism and more women are needed in the field of AI technology and why it is crucial to act now. How did I end up here at Bloomberg? The story begins at the master's program Intersectionality and Change, University of Linköping.
(Lighting talk at The Bloomberg Quant (BBQ) Seminar, Bloomberg headquarters NYC October 2018).

The vantage point for this study is the gender gap and lack of diversity in contemporary AI/robot development and the growing need for interdisciplinary research across robot technology and social sciences. The AI Now Institute in New York released a report in 2018 stating the importance of social science and the humanities contributing to foundational knowledge in the AI field, enabling new models of analysis and methodological intervention.[1] The report expresses significant concern regarding the future direction of AI especially the detection and prevention of bias, discrimination and inequity (2018:36-39). The aim of this thesis is to explore the development of human-like performance, agency and sociability in contemporary technology and in particular the phenomena of social robots. I explore human-machine interaction through a personal and practical autoethnographical experience, with a theoretical framework of feminist technoscience and studied through a queer lens. In feminist technoscience (FTS), non-organic bodies are criticized for not being included as examples of body and agency. FTS studies both human and non-human agency through a critical posthumanist approach of mutually co-constitutive entanglements of materiality and discourse. Science holds ethical responsibilities towards technological development, since technology and science (techno-science) are entwined with the common interests of society (Lykke and Åsberg 2010). As discussed in the theory chapter, I have particularly focused on feminist scholars researching human-machine relations and how the politics of gender and other identity markers are interconnected with the resulting processes of technological development.

[1] According to the *AI Now 2018 Report* (p.38) women working in the AI industry in the U.S. were estimated to comprise 12%. The AI Now Institute, behind the report, is an interdisciplinary research institute at New York University dedicated to research, knowledge production and understanding of the social implications of AI technologies. Uniquely, it is a women-led AI institute, founded by Kate Crawford and Meredith Whittaker.

The research question emerged while reading Judy Wajcman's analytical observation and the inclination that technological development has the potential to change the idea of what it means to be human and even the individual sense of self (2004:1-2). Pat Treusch draws on Wajcman in her dissertation on Robotic Companionship (2015) and explores human-machine interaction and its impact on human condition adding a queer technoscience perspective. Treusch's aim is to study how the "proper human" is undergoing transformations because of rapidly evolving tech societies, which changes conditions and circumstances of individuality and sociality (2015). With inspiration drawn from Wajcman's and Treusch's writings, my analysis centers in particular on sociability and agency in performing machines. Bearing the above issues in mind, the research question for this thesis is: does contemporary development of socially intelligent machines have the potential to impact the understanding of human agency, body, gender and social identity? Artificial Intelligence is originally *"free"* of the boundaries of biological body/sex/gender, but it still depends on biased human sociocultural programming, in other words it carries the potential of emancipation yet could be a trap of gender conservative agendas. The research question therefore is to be explored in the reverse sense; to what extent and how do humans, with the intersectional differences characterizing the human species, inform gendered social agency in such technological developments? With respect to the limited size of this master thesis, the research question will be studied and analyzed through one particular case of a social robot that is being developed at Furhat Robotics.[2] Furhat is branded as "The human face of AI". Basically, it's a robot in the form of a translucent mask with a camera placed in the back, projecting an animated face that communicates through movement and artificial senses. Furhat communicates through eye contact, voice, hearing, facial expressions and mimicry, and through head and neck movements. I set the aim to explore agency, social skills and authenticity performed by Furhat, first in its home environment at the robot-lab and then through an experimental method of inviting the robot into my professional sphere of women in theatre and feminist performance art.

The reason why I choose to study social machines/technology performing within a context of live art is because this is the area of my professional life.[3] Producing and exploring gender

[2] The robot Furhat started as a research project at the Royal Institute at Technology in Stockholm 2014.
[3] Closest to my heart, when it comes to work experiences, are my two long term art projects, *Hootchy Kootchy Club,* a queer neo-burlesque project and *Drag as Art*, a project aimed at exploring, producing, and presenting performing arts using drag as an artistic expression (Rosenberg 2018:201-205).

through live art on stage has been part of my professional life for 15 years. The next step is presented in this thesis, to experiment with gender performance in non-humans and create a posthuman figure in drag. My research at the Furhat lab started in 2016 when I decided to do a case study assignment about Furhat. Spending time at KTH with Furhat and the work team, I developed an idea to collaborate with them. I asked the CEO if I could borrow a Furhat Robot for a theatre project I was currently working on. At the time my art colleague Rebecca Tiger and I were in the midst of creating and a writing a feminist theatre play for the avant-garde theatre Orionteatern in Stockholm. We thought it would be an interesting experiment to let the robot, designed in the form of a talking head be the protagonist of our story *Arise Amazons*.[4] It developed into a *femme-tastic* story about the human face of AI in search of a body, it hears the voice of nature and subsequently enters a world of "Amazons", played by actors, dancers, circus performers, an opera singer, a drag queen and six horses. We created a social experiment aimed at unifying artistic agency related to the subject of human-machine inter-action. The *"*amazon women*"* in the cast identify beyond stereotypical definitions of sex and gender binaries, including trans, drag and female-identified people. Furhat plays a leading actor in the theatre play, as well as in this thesis and will be treated as the main case.

Throughout the unfolding research, I grew concerned as to how AI, social machines and digital social media in the future will impact our human sense of social identity.[5] A reoccurring topic in current news flows and media is whether the fast progress in AI technology will result in an existential threat to human life or not. Another popular subject involves the risks of biased algorithms and sexists AI, with alarmingly many real case scenarios of this happening.[6] After a lot of reading and coding the texts, the words *agency*, *body* and *gender* stood out as reoccurring concepts, and which I sensed are key aspects in humanlike performing machines. Therefore, these three will be treated as the centralized categories of analysis. This thesis explores a case study of a particular prototype of a social robot, operating/performing for an audience in a theatrical setting. With the aim to analyze aspects of agency performed by the robot, and also to imply an analysis of body and gender, the choice of a theatrical setting is linked to my professional positioning. In addition, the

[4] The title "Amazon" is a symbolical reference to, and based on, the mythic concept of the amazon(s) as independent women fighting collectively with other warrior women for their position in society.
[5] E.g. *AI Now 2018 Report*, AI Institute p. 9, and the "concern about AI social implications".
[6] E.g. *The Big Nine: How the Tech Titans and their thinking machines Could Warp Humanity* by Amy Webb (2019) and her discussion how nine powerful corporations are turning the human-machine relationship on its head. And, *Technically Wrong: Sexist apps, biased algorithms and other threats of toxic tech* by Sara Wachter-Boettcher (2017).

theatre presents a room or space separated from the performance of everyday life in society. In the theatre, the performance is accentuated as being fictive and character-based rather than real. I wondered what would happen if the social robot were situated outside the scientific lab and placed in a fantasy landscape created by me. The reason why I have a posthumanist framing is because this transdisciplinary field involves art, philosophy, technology and science informed by advanced cultural critique and de-construction of humanism (Badmington 2000). Processing the empirical material by combining traditional academic writing with a personal, artful and evocative practical approach, the autoethnographical method appealed to me as discussed in the next chapter.

Thesis and chapter structure

In this first chapter, I have introduced the motivation of my thesis and discussed the aim and research question. In chapter 2, I will present my methodology, mixing traditional ethnographical methods with autoethnography. The method of collecting empirical material is introduced, with participatory observation in the robot-lab and interviews with the Furhat-team and participants in the theatre project. In chapter 3, I discuss theoretical concepts and the FTS disciplinary framework for this study. In the second part of the theory chapter I discuss a literature review of previous research in FTS on human-machine interaction, which I find relevant and useful to my case study. In chapter 4, I discuss social agency imagined in technologies with the aim to historically trace and analyze issues of agency in robots and artificial intelligence. Chapter 5 is where the empirical research starts, with a first visit to the Furhat-lab and a longer conversation with CEO Samer Al Moubayed about building a human-like social robot. In chapter 6, the social robot is situated at the theatre stage to perform in the theatre project *Arise Amazons*. This analysis chapter is autoethnographically narrated, mixing interviews with my co-creators of the robot-character for the theatre play, Rebecca Tiger (director/actor) and Petter Lidberg (Furhat operator), with descriptions and examples from the theatre experience. In chapter 7, I conclude the research question based on the empirical material in the case study of the social robot Furhat and my contribution to the FTS field. I also present ideas for further research on practices for ethics and bias in contemporary technologies.

Chapter 2: Methodology

Inspired by feminist technoscience studies, I aim to do a critical posthumanist analysis of performing social machines through, primarily but not solely, the method of autoethnography. I have chosen to work with a mixed method approach, where I combine traditional ethnographical methods of participatory observation and reflexive interviews with more experimental methods of autoethnography and art (Lykke 2010:161). This thesis involves a heterogeneous collaboration between science and technology (KTH with the Furhat Robotics team) and artists from the field of performing arts and theatre, a process in which I have been the initiator, engaging myself and the co-collaborators in practical activity of creating a theatre project. By a method of applying knowledge from feminist theories and collecting empirical data through interviews with the robot designers, I created a fictional robot-story and experimental theatre play around it. The intention with the experimental theatre play was to activate empirical research and process inquiries on the topic of sociability connected to agency, body and gender in machines through the language of art. My purpose was to reach not only the academic reader, but also to include a wider and more diverse audience.

Professor Nina Lykke advocates and inspires her students to write academic texts differently (2014). Lykke elaborates on Donna Haraway's article *Situated Knowledges* (1991) and her critique of what she calls the "god-trick" in positivist epistemology, where the scientist detaches her/himself from the objects of study and becomes the context-less knower (Lykke 2010:4-5). Haraway dismisses the god-trick as being an illusion. On the contrary, she implies a subjective dimension where the researcher is co-responsible for the knowledge production, creating a story of which she or he plays a part (Haraway 1989:4). In *Autoethnography as Feminist Method. Sensitizing the Feminist "I"* (Ettore 2016), the author critically reflects on how feminists negotiate agency which effects ones' political sensibility. Ettore advocates that feminists should instead write and speak about themselves, creating stories of political responsibility, which is a key issue for feminists who function as cultural mediators. Traditionally, writing "I" has not been considered appropriate by standard academic rules. In this thesis I claim it to be necessary to clearly position myself as "I" since I have situated myself as a cultural mediator, not above or detached from the subject of the study but in close interaction with it on a theatre stage.

Why autoethnography?

Autoethnography is a self-reflexive qualitative research method where writing is used as a way of exploring through the storytelling of personal experience. Comparing this to wider, social, cultural and political meanings, the subjective and personal experiences are embraced and acknowledged, and personal biases are exposed. I find this an interesting choice of methodological approach, since I had positioned the object of study - a social robot - in my world at the theatre. I invited it to perform with me and the other performers in a story I created, to practically explore the abstract theories of feminist technoscience and posthumanism. Autoethnography is a method applied in performance studies and anthropology. Performance studies examine performances related to both the context of various art forms, e.g. theatrical storytelling, and, to "cultural performances" analyzing other types of performances, e.g. related to social and professional roles and of individual performances of race, gender, sexuality and class (Schechner 2002). Since my study is to merge these two categories of performances (researching a social robot's performance in the Furhat-lab and the robot's performance in theatre) I decided autoethnography would provide a method well suited for this study. However, I want to clarify that performance studies is not my discipline of choice. I find feminist technoscience to be a more interesting disciplinary theoretical framing for the thesis. Different from traditional research, the autoethnographer tries to subjectively touch the reader, by making the cultural experience engaging through a personal reference. They are also aiming to reach a more diverse audience. This methodology carries the potential for both personal and social change (Bochner 1997). Concerning the methodological ethical approach in this study, I draw on a postmodern feminist viewpoint of treating categories as unstable, multilayered and incalculable, thus problematizing traditional ethics (Shildrick 1997).

The method used for collecting the empirical material for the thesis analysis is also based on traditional anthropological participatory observation (Aull Davis 1999:67-93) at the Furhat robot lab (KTH), and interviews with people working in the Furhat-team. I did a few longer interviews with Samer Al Moubayed (CEO Furhat Robotics). The empirical research from the theatre project is analyzed through conversations with co-creators, Rebecca Tiger and Petter Lidberg. All interviews were semi-structured (Aull Davies 1999:95), interactive, and dyadic (Ellis 2004), meaning that the researcher can focus not only on the participant's answers and story but also on the interactively produced meanings and emotional dynamics that may arise

during the interview, including that of the researcher. Dyadic interviews in autoethnography also include the researcher's personal motivation for doing the project. This fits well into this study since it's developed through a long story of motivation; from a first encounter with the social robot Furhat, leading to an art project, and finally a master's thesis. My research process of collecting empirical data and knowledge about the social robot Furhat stretches back in time to 2016 and my first encounter with Furhat and CEO Samer Al Moubayed. That first day I sat with Samer and Furhat in a room at KTH and got to know both of them through a two-hour long interview on AI. I spent another couple of days in the lab/workshop, interviewing the Furhat team of developers about the specific tasks in building a social robot, the concept of social machines and what this meant to them. These interviews were based on conversations in the lab simultaneously exploring and demonstrating the technological anatomy and design of the robot. In 2017, I came back to KTH and spent some days in the lab with Furhat to prepare and program the robot for its theatrical debut as the protagonist for *Arise Amazons.*

Autoethnographers reveal several layers of consciousness that link the personal to the cultural reference by asking the epistemological question "how do we know what we know?" (Ellis and Bochner 2000:739). Through the narrative, the scholar can emphasize the plurality of truths and understands personal experience is not necessarily bound to specific historical contexts. Narrative methods, such as self-narratives, are increasingly explored in both the humanities and social sciences and have generated an important tool in the knowledge production about person(s) and/or collective agency. Ethnographic research requires social interaction and the method of participatory observation, the hallmark of anthropology, and can be considered as rites of passage for the anthropologist (Aull Davis 1999:67-68).[7] I consider this study to be a rite of passage for me as a social scientist, not only through the participant observation at the Furhat (robot) lab but also through the chosen experimental method of practically exploring the research question by situating the robot with me on a theatre stage. I created a framework of storytelling a fictive saga, in which the robot protagonist is on a quest for body (embodiment), it undergoes a rite of passage through human-animal-machine interaction and transitions to a liminal personae.[8] The reason I was

[7] Based anthropology's concept of rite of passage, a transition to the unknown (Turner 1969).
[8] Victor Turner states that there are people/beings who have "*liminal personae"* (threshold people), their ambiguous attributes make them slip through the network of classification that normally locates states and positions in cultural space (1969, 2nd Ed. 2008).

attracted to the autoethnographaical method is that it attempts to disrupt the binary cut between science and art (Muncey 2010). I experience art and academic writing to be significantly different languages and methods of communicating. Art is more embodied and freer in expression while academic work requires more precision and structure. I aim to challenge myself in exploring the in-between liminal space where science and art collide.

Introducing the social robot case

The social robot Furhat started as a technological scientific research project at the department of speech, music and hearing at the Royal Institute of Technology in Stockholm. The start-up company Furhat Robotics was formed in 2014 through a funding team of four male developers and another four male advisors. The all-male-crew knew that they had stumbled upon something special, stating on the official webpage how "the feeling of co-presence you get with Furhat leaves no one unmoved" and that what's special with this social robot is that it communicates in semi-human ways. The aim is to build a human-like interface, with social intelligence and conversational AI. Furhat carries the vision of its innovators to understand how to make machines come to life, and in simulating an interface with personality, agency and social intentions. The Furhat team believes such a human-like interface will be the next revolution in human-computer interaction, enabling people to interact the way they interact with each other. During my first meeting with CEO Samer, he compared Furhat to smartphones or a personal computer. In the beginning, smartphones couldn't really do anything, and it was expensive and too small to be like a computer. He doesn't think Furhat Robotics will be building all the applications in-house to make the robot social and intelligent. Instead, Furhat is a collective project where developers can create and develop programs and apps through the operating system "Furhat OS", which is an open-source programming platform. He thinks this to be the biggest contribution to Furhat. Like windows was for PC, Furhat is a platform where co-creation and collaboration are necessary for success. What Samer meant was that in the process of developing the social robot, he and the team started to reflect on what it means to be human. Before a closer study of social agency and qualities of human-robot interaction with Furhat, I will discuss the theoretical framework and analytical tools related to important feminist technoscience scholars in the following chapter.

Chapter 3: Theoretical concepts

I have decided to distinguish three concepts; agency, body and gender from a feminist posthumanist perspective and to explore these further in the analysis of the social robot case. Two important feminist scholars opened for new discussions on feminism in science and technology, challenging the prominent and problematic male-centered perspectives. Donna Haraway, researched and engaged in dialogue on feminist anti-racist and postcolonial theory on gender/sex concerning new cultures of technology (Lykke 2010:117).[9] Sandra Harding, created an important shift in thinking when she turned "the Woman Question in Science" into "the science Question in Feminism", advocating the importance of not only including women in technological science but aimed to transform the foundations of feminist epistemology and scientific knowledge production altogether (Harding 1986:29).[10] Reading the *Cyborg Manifesto* (1991), by Donna Haraway, was a gateway that awakened my curiosity to learn more about technofeminism and the concept of the cyborg and human-technology relations. Drawing on the ideas of Haraway I created and wrote the story for the theatre play. What especially caught my attention is that Haraway emphasizes that women are suspended in multiple entanglements with technology, and in this fashion blur the boundaries between nature and technology. Haraway invited the cyborg into the world of feminist theory and it became a very potent critical figure (Lykke, Markussen and Olesen 2004:326). Cyborg feminism questions the distinction between individuals and their environmental contexts. Rather than dividing women into clean-cut categories, individuals are related through multiple intersections that make the cyborg a symbol relevant and fitting in the contemporary digital age.[11]

Feminist technoscience studies

Feminist technoscience studies (FTS) is an interdisciplinary field of study that is grounded in gender studies, technology and society (STS), and feminist cultural studies. It highlights the entanglement of gender and its intersections with other identity markers such as race, class,

[9] E.g. Postcolonial technoscience is a field of study, focusing on analyzing global dynamics in and around science and technology (Andersen 2002; McNeil 2005).
[10] Sandra Harding's intention was to make women more visible in scientific knowledge production, but she soon realized that it was as important to initiate a feminist inspired transformation of the very foundations of science, an epistemological problem revealed to be a complicated and grand scale enterprise (Lykke 2010:140).
[11] Haraway stated the frequently quoted remark that the sexed body has been reduced to "a blank page for social inscriptions" (Haraway 1991:197*).*

sexual orientation and dis/abilities. FTS and STS also studies how the sociotechnical phenomena and human and non-human actors interact in the production of knowledge. The transdisciplinary field emerged in the early 1980s through a line of feminist critique on natural sciences.[12] The term technoscience was created to challenge aspects on science and technology and how they make each other up. FTS analyzes the interconnection of gender politics with technology, drawing attention to power relations within the context of a globalized world (Wajcman 2007) and is a critique of the positivist assumption of the distinction between scientific theories and the way they are practically applied. The gap between basic and applied science does not consider the complex entanglement with e.g. power relations in the academic world, or power and hegemony implied by governments, corporations/corporate environment(s) or private investors, and of societal interests in general (Åsberg and Lykke 2010). With academia as a starting point, FTS aims to develop new theories in the scientific studies of technology as multi-disciplinary dialogue connotes the intersection of multiple disciplines colliding through conversation that have the potential to create change in societal perspectives on the development and practical uses of technologies. A common concern of feminist technoscience analysts is how the exceedingly growing influence and use of technology -ruled by global capitalism- generates huge historical impact on what it "[m]eans to be an embodied human subject in a globalized world" (Lykke and Åsberg 2010:300). Feminist technoscience addresses the need of reconceptualizing what it means to be a (post)human.

Critical posthumanism

Feminist studies is witnessing a growing interest in posthumanities (Lykke 2010:81-82), with the overall aim to set up new transdisciplinary meeting places of mutual commitment between social and natural scientists in global change research. The umbrella term posthumanism, contains within itself several branches and disciplinary fields.[13] Feminist posthumanism questions the anthropocentric and imperialistic logic of humanism, which places *man*kind in a hegemonic centralized position compared to other species and non-human others. In contrast, a feminist posthuman worldview acknowledges that humans are in a complex entangled

[12] Genealogically, this goes back to the 1970s, when ecofeminists considered science and technology a business opposed to the interests of women, stating that natural science was based on white patriarchal principles that claimed objectiveness and supremacy over women (Harding 1986).
[13] Posthumanism covers different schools of thought, such as philosophy, cultural and critical posthumanism, transhumanism, the feminist approach of New Materialism(s) and the heterogeneous landscape of (anti)humanities and meta-humanities (Ferrando 2013).

relationship with surrounding environment(s) including non-human others. The posthumanist perspective enriches feminist studies by including the missing pieces of earth-others. Encompassing animals, organic and non-organic beings, cyborgs, robots, AI and more, feminist posthumanism (like cultural posthumanism) aims to move beyond archaic concepts of human-nature by developing new concepts changing the face of contemporary technoscientific knowledge production.[14] In fact, feminist posthumanities has even been described as a renegotiation of the human, questioning the anthropocentric bias of humanist thought and human exceptionalism, and also, the sometimes overtly-optimistic belief in technological progress (Kostinen; Karkulehto 2018). Concluding the above, critical and feminist posthumanism encompasses the important conjunctions between different life forms and how these entanglements create a rich network of interconnections. The human-machine interactions that are now being developed in interfaces and socially intelligent machines, such as Furhat, constitutes a cross-border intersection between human/non-human, highlighting their inevitable entanglement of body/mind, which creates a new form of agency.

Posthumanist performativity; re-thinking agency, body and gender

Body and embodiment are two historical hot topics and concepts of discussion in feminist studies. The basic understanding of embodiment is that it means the lived experience of *human*-beings, in which the body plays a fundamental part. Concerning the definition of body and agency, the feminist posthumanist approach emphasizes the importance to include non-human or non-organic beings (Braidotti 2013). Embodied beings should not be determined only by their corpomaterial matter. Embodiment expands "natural" and cultural conditions yet is never unbound to its social and cultural environment, never completely elastic. The definition of body could be described in many ways. Generally, it is understood as being the organized physical substance of a living organism, human, animal, plant or other organic entities, whether if it is living or dead, fall under this category. A body could also refer to a larger body, consisting of many bodies/entities connected to, and entangled with, each other, together creating a new body. This wider concept of *body*; a series of organized units, functioning as a collective whole, including persons and other things, is the focal point in feminist posthumanism (Braidotti 2013; Barad 2003).

[14] Cultural posthumanism, is affiliated with culture theory, comparative anthropology and semiotics that compares and critically examines historical concepts of human and human-nature (Badmington 2000).

Agency could be described as one's ability to act or perform in the world. Meaning, if a person/agent/thing is able to choose and act on these decisions then they have agency. If you, or an agent are not able or allowed to act based on your own decisions, then there is a limiting of agency, and in cases of oppression a denied agency. In feminism agency can be seen in terms of societal and cultural power structures that affect a woman's agency. In social theory, the concept of agency changes when it is conceptualized as relational rather than an individual phenomenon. The concept of "relational agency" challenges the Western traditional and political perspective of agents as autonomous and separate individual agents with another approach, that of agents as interactants which are interdependent, intermittently reflexive and vulnerable (Burkitt 2015:322-339). Applying this view to complex sociotechnical inventions, Suchman advocates that relational agency aligns with feminist theorizing, critically questioning the Enlightenment and organized humanism's conception of the individual actor/subject living in a world of separate things/objects (Suchman 2007:239-240). Instead, feminism emphasizes the relational character of a social agent's capacities for action. In the case of social machines, human scientists project social agency in non-human entities with expectations of simulating human-like action, responses, and behavior in technologies. Suchman critically draws on anthropologist Alfred Gell conclusion of fetishized objects of enchantment when arguing the conception of agency as relational (Suchman 2007:239-240). Gell's discussion concerns how it is that humans attribute social agency to objects, although they simultaneously know that these entities are different from them, with the conclusion that what matters in ascribing a social agent its status depends mainly on its network of social relations rather than what the object actually is (Suchman 2007: 241-245; Gell 1998:123). Robotic agent-artifacts are closely modeled after humans producing the material agency that is still in human hands when it comes to intentionality. In this sense these sociotechnical inventions mirror human agency and consciousness by being conceived and controlled by human terms, and with a design to serve and perform strictly under human conditions (Knappet and Malafouris 2008). In feminist posthumanism agency is linked to performativity.

Karen Barad coined the term "intra-action" (2003:815). She describes intra-action as the mutual constitution of entangled agencies. The difference between inter-action and intra-action is that "inter" means a positioning among or in the midst of, while "intra" means a positioning from within. When we add the word action to these concepts, they become distinctly different in meaning. Two bodies that inter-act maintain a level of independence as discrete objects. However, when bodies intra-act, their ability to act emerges from within the

relationship and not outside of it. Karen Barad terms the entangled intra-action between human and non-human actors or agents "agential realism", a theory that aims to transgress dichotomies like mind/body and human/non-human (Barad 1996:183)[15]. Agential realism challenges the binary between the material and cultural and instead focuses on the processes of material-cultural interaction (Braidotti 2013:158). According to Barad, language has been granted too much power and with the concept of agential realism and posthumanist performativity she advocates a shift to building *"*sensitivity*"* to how humans and things become (Barad 2003:382).[16] Contemporary technoscience and the impact of information technologies that have brought about, is creating a new focus on the agency and of body in social practices. Braidotti writes how science today is reaching into the most intimate layers of the living organism, dissolving boundaries established by centuries of humanistic thinking, and we can now think of the body as being part animal (organic) and part machine (non-organic) (2002:21)[17]. She sees in the human a specific capacity of simultaneously incorporating intersectional variables such as class, race, sex, age and more, structuring identity and the notion of social imagery. I argue that feminist aspects of agency and body, as discussed so far in this chapter, need consideration when it comes to programming a social robot with personality and social skills. Specifically, the question of gender-replication in technologies intrigues me. In the following section I will be using analytical tools to discuss and unravel the works of scholars that inform my study on the topic of humanlike machines related to issues of gender.

Feminist technoscience and human-machine interaction

I will precede this chapter with a literature review tracing feminist technoscience research on human-robot interaction and entanglement, and in particular the development of social robots. Professor of sociology, Judy Wajcman will be my point of departure, since she is one of the earlier contributors to the social studies of technology. Instead of critically getting stuck in the discrimination of women in science and technology, Wajcman writes that feminist technoscience critics chooses to focus on how technology, deeply rooted in masculine

[15] Barad combines feminist theory and adopts Nobel Prize winner in quantum physics Niels Bohr's philosophical reflections on scientific subject/object relations, e.g. his famous wave-particle paradox of how light behaves either as waves or particles depending on the apparatus developed by the scientists (Barad 1996:2007).
[16] Barad critically elaborates on Butler's much-quoted theory of gender as performative with a modification of the term posthuman performativity as rather being an iterative intra-activity and related to matter as a congealing of agency (Lykke 2010; Alaimo and Hekman 2008:140; Mehrabi 2016:39).
[17] From *Becoming woman or sexual difference revisited*, by Braidotti (2002) where she revisits Lucy Irigaray's theory of sexual difference and compares it to contemporary philosophical theories of Gilles Deleuze.

science, actually can be turned and used in the emancipation of women (2004).[18] In the article *From women and technology to gendered technoscience* (2007) Wajcman concludes how the FTS-approach criticizes technological determinism and gender essentialism instead emphasizing a gender-technology relationship that is fluid and flexible, where feminist politics is the key to gender equality in technology. Technoscience holds the potential to open new gender dynamics, emphasizing a mutual shaping of gender and technology, where neither gender nor technology is taken to be pre-existing (Wajcman 2004:7). Judith Butler's gender performativity theory (1990,1993) is applied by Wajcman to the construction of technology.[19] Just as gender identity is something subjectively performed in daily-life and social interaction as a relational behavior with others, technology too is a product of subjective coding on gender performance and identity (Wajcman 2007:294).[20] Intriguingly Wajcman suggests that technological development has the potential to change the idea of what it means to be human and even our sense of self, which has inspired my research question for this thesis (Wajcman 2004:1-2).

Another significant feminist technoscientist is Lucy Suchman and her critical research on human-machine/robot interaction (Suchman 2007). With human-machine interaction she means the activities humans perform with or through machines (such as computers) and the behavioral social patterns this type of communication creates. Suchman draws on Sherry Turkle's description of the machine as being an "evocative object*",* that is, an object with the power to evoke emotion and provoke thought (Turkle 1984; Suchman 2007:33). She raises new questions regarding what humans have thought to be common sense regarding the distinction between artifacts and intelligent others.[21] The ultimate goal of AI, according to Turkle, is to build an artificial mind (Turkle 2005:220). According to Suchman, it is therefore necessary to illustrate interactions between people and machines that imply a "mutual intelligibility" and "shared understanding" (2007:29).[22] I share Suchman's motivation to further explore how intelligibility can be shared between humans/machines when developing

[18] Wajcman argues that Marxist and liberal feminism failed in the analysis of technoscience with their lack of attention to the field's symbolic dimensions (2004).
[19] Gender performativity and sex should according to Butler be understood as continuously being shaped in language and discourse (Butler 1990,1993).
[20] Later write here the chapter where to find more info on this etc.
[21] Sherry Turkle is an American professor of the social studies of science and technology. She positions herself writing from the subjective side of people's relationship with technology.
[22] "Mutal intelligibility" [relationship between languages] and "shared understanding" [individual/collective knowing of rules, objectives, and boundaries of a certain circumstance/pursuit] are both reoccurring terms discussed by Suchman in relation to AI and technologies (2007).

social AI. Suchman poses significant questions to how the concept of the human figure is materialized in technologies. Traits Suchman considers to be crucial in engineering humanness in AI/robotics are; embodiment (an AI requires a designed body), emotional talents (e.g. happy response) and to engage humans in social interaction with the right timing so that the machine is perceived as real (2007:229-238). Suchman states that to be successful with projects of building intelligent and interactive machines, it is necessary to acknowledge the differences between human interaction and machine operation, rather than assuming their similarity. Understanding the problems that arise out of the disparity between human/machine relationship can provide new knowledge regarding the design of machines as well as to better understand and build new theories concerning situated human action (2007:126).[23] Suchman's term "situated actions" means purposeful actions taken in the context of particular circumstances. She is here critically referring to the biases of culture, and in particular the "[f]amiliar parochially Western" interpretation setting a false opposition between theory and practice (2007:26). When applied to a robotics company, it is therefore very important to take notion of the kind of interfering situated actions connected to the robot-engineers that are creating social robots, and to what proportion technology inherits biased traits from its makers.

Two other interesting researchers in technoscience, emphasizing issues on social robots and gender representation are Alison Adam (2002) and Jutta Weber (2005), both of whom write about human-robot interaction. Weber, interestingly proposes a robot-perspective, criticizing humans for creating robots to become "helpless machines" and/or "true loving care givers" (2005).[24] Her focus is the phenomena of robots as companions and technoscience ambitious goal to build intelligent artifacts with human social abilities.[25] With human social characteristics, Weber means; recreating emotions, the ability to conduct a conversation, and facilitating learning capacity leading to established character traits and a personality. In interacting affectively and affectionately, socially and physically social robots are designed to learn from humans. Weber analyzes the visions of these socio-emotional machines and how

[23] According to Suchman, *"*The turn to the social*"* among computer scientists in the mid 1980s led to an intensification of interest in social science to study the material grounds of sociality (2007:276). For instance, a growing interest in how *non-vocal activities*, like gaze and gestures, play an important part in face-to-face human interaction emerged in the field of conversation analysis, and which also play a part in the case of Furhat/robot.
[24] Weber's key aspects of research are human-machine-communication in the realm of robotics and cybernetics, where she analyzes the material-semiotic and socio-political reconfiguration of the culture of technoscience.
[25] These social robots are to be used for entertainment, care taking, education or as personal company (robot pets) and therefore are conceptualized as friendly and supportive to their human partners (Weber 2005).

feminist critique of technology could be applied to the field of social robotics. In order to highlight the concepts of sociality and emotion, she emphasizes the importance of feminist critics because she sees a tendency in technology to naively re-create gender stereotypes and norm-oriented behavior (2005:210). Adam's central concern is the question of embodiment in relation to gender, highlighting the problem of gender-construction in AI technology. Her method involves writing from category theory and anthropological phenomenology on the role of the body in feminism and knowledge production, critically arguing how the body has been left out in AI technology.[26] Pat Treusch's research focuses on contemporary socio-technical imaginary and engineering humanness in kitchen companion robots, where she analyzes human-machine relations with a feminist technoscience perspective (2015).[27] For instance, the questions of how we want these robot companions to look and how we want to live with them and engage them in human everyday-life. What signifies Treusch approach is her methodology of "queering" and emphasizing norms of humanness translated into the humanlike-robot. (2015:25).[28]

Summary

Pivotal to my research analysis is Wajcman's proposal of how advanced technologies impact the understanding of the human concept. In line with Suchman and Weber, I question how the concept of the human-likeness is simulated in machines with aspects of body, emotion and social skills, also touching on Adam's concern of the body being left out in artificial intelligence. In account of Treusch study, I intend to develop an understanding of humanness in machines and a method of queering present gender norms in the case of my study (2015:48-49). What differentiates my topic from Treusch is that my aim is to focus particularly on sociability or social intelligence in machines, and to explore the fantasy landscape of modern robot technology and automation. Distinctively different from discussed scholars in the literature review is my choice of an experimental autoethnographical method, in which the under-construction-robot is moved from the lab-context and invited to perform with me at the theatre, positioning it in my professional context. In the following chapter, I will outline the history of social robots and analyze how robot scientists have figured out ways they can improve agency and social intelligence in machines.

---

[26] Like Weber, she uses the example of parenthood and children (2002:372).
[27] Treusch is studying the specific robot model "Amar", that is being developed at the Karlsruhe Institute of Technology, where she participated in the local everyday practices of establishing a robot-human interface.
[28] Treusch builds her queering on the feminist critique of heterosexuality and normativity (Treusch 2015:24).

Chapter 4: Analysis of social agency imagined in machines

A scientist that has conducted influential work in the field of social robots is Cynthia Breazeal.[29] In her book *Designing Sociable Robots* she aims to "[p]ioneer a path toward the creation of sociable robots", through her own experiences with trials and errors drawing a map on how to build robots with social intelligence so that others can follow (2002:5). She finds it difficult to describe precisely what a social robot is and gives examples from imagined robots in popular science fiction, like the mechanical droids R2-D2 and C-3PO from the Star Wars films (2002:1). Accordingly, these sociable characters illustrate how a social robot should embody qualities of acting as socially intelligent humanlike beings, built to be able to understand itself, humans, and others in social terms, incorporating the capacity to adapt and learn through shared experiences with other individuals. Interacting with a social robot should be like interacting with another person. Breazeal's robot-research provided me with insight to a similar kind of robot engineered as Furhat was, as she laid the groundwork for contemporary development of sociable robots. According to Furhat CEO Samer, it will take more time before a social robot-like C-3PO will enter our living room, even though the desire to create humanlike objects stretches quite far back in time.[30] Already in the early 1700s, experiments with the attempt to synthesize life were becoming popular (Suchman 2007). At that time inspiration came from, what Suchman describes as, an emerging materialist sensibility, in which the workings of the human body where investigated through physiological and simulated mechanistic experiments and practices. Drawing on Suchman's research, the way in which robot agency is thought to be humanlike changes with time, which I continue to discuss in this chapter.

Autonomous robots and issues on agency

Robots are termed differently depending on design, purpose and type of tasks and performance the robot is expected to carry out and accomplish. "Autonomous" robots are designed to perform with a high degree of autonomy and independence, although they do, like any machine, require regular maintenance and an electric power source. Some modern factory robots are examples of autonomous machines. The factory robot's workplace can be a

[29] Cynthia Breazeal is the director of the Personal Robots Group at MIT Media Laboratory, U.S.
[30] The Japanese NEC corporation, a multinational provider of information technology was, by the time Breazeal wrote her book, developing a household robot resembling R2D2, which is an actual example of science fiction inspiring the technological robot engineering (Breazeal 2002:4).

challenging environment because of chaotic and unpredictable variables that the robot has to be able to adjust to.[31]Most important, is to avoid all situations that involves risks of being harmful towards people, property, or itself. Due to precaution and safety, a cage is often built around industrial robots restricting them to operate under a limited amount of freedom (Suchman 2007:37-38). I would say that the need for a cage indeed affirms agency in these robots and the capacity/risk of them "acting out" beyond human control. It is this imagined and potentially dangerous agency of the robots that is embodied in the cages around them. In Treusch study of the companion-robot Amar, she develops the term "the un-caged robot" to argue the difference between dangerous industrial robots and the domesticated household robot (Treusch 2015:86-105). They are both assumed autonomous, which partially makes them agential. However, Amar is designed not to be of any risk to humans and is therefore allowed to step out of the cage and experience a direct interaction and co-existence with humans. While the factory robot is more mechanic and robust, belonging to the rough factory environment disconnected from direct human interaction, Amar belongs to the space of home and needs to be more human-oriented. It could be argued that the factory robot is more agential, as it represents a higher risk of acting out towards humans, while Amar is less agential and therefore un-caged. On the other hand, agency could be understood as shift shaping. The factory-robot is linked to degrees of autonomy and risk. In Amar's case, both the level of autonomy and the human inter-action are considered as well. To conclude this section, the long-standing preoccupation with developing automata is, according to Suchman, connected to the urge of investigating human agencies in material things (2007:228). As argued so far, early models of robots were saturated with automation. The next step in technological advancement of robots involves the concept of sociability.

Sociability embodied in machines

Social robots are a fairly new branch in technology. It was not until the early 1990s that collaborations of researchers in artificial intelligence and robotics explicitly focused on robots with the capacity to engage socially. Social robots are autonomous machines, but with an emphasis on developing social behavior and interaction with humans, or, other autonomous physical agents. The social robot is a hybrid of robot and software programming, where the programming enables human-like communication. While autonomous robots previously were

[31] Factory robots should be able to gain information about the environment, work for an extended amount of time, without human attendance, and have the ability to move within in its operating environment.

engineered to perform independently as caged machines with limited interaction with humans, e.g. exploring other planets, or mowing lawns, more recent commercial interests have brought forth a new generation of interactive social robots with the ability to engage humans in compelling and enjoyable ways. As Suchman highlights, what the robot was to the industrial imagery, the software agent is to the desires and fantasies of the service economy (2007:219). There are different types of systems designed to interact socially with people, in general these systems can be divided into two categories; those that are embodied where the human interacts with a robot or animated avatar like Furhat, or those disembodied where the human interaction happens through speech, text, or by entering code (e.g. Apple's Siri).[32]

Breazeal discusses human computer interaction applied to social robots, through implementing software systems, and how this can contribute to mutually beneficial relationships, where the robot and human learn from each other (2002:182). An important key for humanlike communication is the aspect of senses, and Breazeal is concerned how the equivalent to senses (e.g. cameras for seeing) can be constructed in machines. Breazeal's intention with her robot Kismet was to design a social robot that was infant-like, with the aim to study the most basic social interaction between a human infant and its caretaker (2002:5).[33] Kismet is a social robot, in the form of an anthropomorphic head that is specialized for face-to-face interaction with humans. Breazeal does not refer to Kismet as a he or a she, but instead emphasizes how the purpose is to build a life-like "creature".[34] According to Breazeal, designing the Kismet-robot to look like a fantasy animal (pet), with the behavior of a curious child, evokes care-taking sympathies in the human motivating human-robot connection and interaction. Similar to how Furhat Robotics exclaim with their product, Breazeal writes that what is special with Kismet is how the robot makes you feel, connecting to people on a physical (agency), social (interaction) and emotional level (2002:xiii). The 15 computers and 4 operating systems that are running Kismet are strategically placed in a different room so you can't see them while communicating with the robot. I find this interesting, relating it back to agency. On one hand the robot is portrayed as sociable with its attractive features, its stands

[32] Siri is an artificial intelligent assistant and part of the company Apple Inc.'s iOS (operating system).
[33] Kismet was built by adapting models from the interdisciplinary fields of psychology, ethology and cognitive development. Even though this robot delivers at what seems to be the basic level of an infant, Breazeal evaluates Kismet as being a genuinely sophisticated machine that was very complicated to build. The ambition of the team behind Kismet was to create, not only a robot-demo program but a device that mimics intelligent behavior.
[34] The robot Kismet is able to move its neck, lips and eyes and is specialized for sending as well as receiving, using its sensors to pick up social cues that people expect when talking to someone face-to-face. Kismet has a very expressive face, operating through a complicated system of 22 motors in the head and neck.

autonomous on its own, giving the illusion that it is interacting with humans. One the other hand, the interaction and its sociability is an illusion, the embodiment of absent presence with the operating machines placed in another room. Kismet was designed almost two decades before Furhat, while the former robot(s) were stuck in a lab, technological advancement have made contemporary social robots like Furhat portable. Although Breazeal calls Kismet an ungendered creature, I lack the critical perspectives on the aspects of gender and gendered technoscience, signified in the research by Wajcman and Weber.[35]

Anthropomorphism is a term in robotics implying that personification and human characteristics is ascribed to non-human beings or objects. In the case of building human-like machines this is a very crucial aspect. As technology advances, becoming more intelligent and complex, humans demand more of their technological devices. Since humans are embodied within a social context, for instance where body-language play an important part, the robot will seem more familiar the more it is anthropomorphized (2002:7). The emerging focus on building anthropomorphic machines in contemporary robotics is what Treusch identifies as a "turn to the body" and with this paradigm shift in AI, a new type of domestic companion-robot arises in technology (2015:102).[36] In early AI development, the concept of intelligence was detached from a body, the obsession was to build a detached artificial mind (Turkle 2005:220). This was eventually problematized. With the focus shifting from "the problem-solving mind" to the "mind in the body". AI-scientists realized cognitive systems needed to be embodied autonomously to be able to gain experience from the world they performed in and hereinafter AI was embodied into physical agents (Dreyfus 1982; Hayles 1999). Summarizing this chapter, I discern agency in machines to be a complex matter. In social robots, human-like agency appears to depend on materializing artificial intelligence through a body, giving rise to new forms of anthropomorphized technology. Sci-fi figures belonging to the rich fantasy world of art is a common source of inspiration in simulating social robots and lifelike machines.[37] In the next chapter the empirical material is presented by inviting you into the Furhat-lab and my first encounter with the social robot case.

---

[35]Weber critically analyzes Breazeal relationship with the infant robot Kismet as illustrating a stereotypical gendered mother-infant relationship (2005:210).

[36] Treusch refers in her discussion of the paradigm shift to Weber and Bath (2007), where they discuss gendering processes and de-gendering strategies for emotional artefact from a feminist technology studies perspective.

[37] Jennifer Rhee discusses this further in her recent publication *The Robot Imaginary* (2018) where she analyzes the productive connections to film, art and culture in robot technology. She argues how contemporary depictions of anthropomorphized robots offer insight to governing conceptions of the human, scrutinizing the maintained racialized and gendered de-humanization of the "Other".

Chapter 5: Conversation about bridging the gap between human and machine

It was a bright sunny day and with energetic steps I walked towards KTH and the Furhat-lab. This would be my first encounter with a social robot and human-robot interface. I remember my expectations were quite high. What would it be able to do? Would I be able to communicate with it in a face-to-face conversation? I did know, that the robot was in its early stages of design, an "infant" machine built to socialize with humans. I felt excited by the opportunity to find out how human-like sociability was applied and simulated in this specific robot.

CEO Samer welcomes me at the lab and takes me to a smaller room where the Furhat device is placed. He turns on "Furhat" (the robot) via a computer, to demonstrate how the robot works. We are sitting in front of a large touchscreen monitor that looks like a table. On it I see a control panel where the user can control Furhat through the various programs for the robot's set-up e.g. character/face, voice-response and more. Furhat is in a default setting when we enter the room, with its standard prototype mask/face.[38] Samer and I start talking about Furhat's "senses" or sensors. Samer points to the touchscreen and shows me how Furhat can "see" me and measure the distance between us through a proximity sensor. He says that two humans in a conversation would behave differently depending on distance, based on cultural preferences, gender, individuality and so on. Therefore, distance also affects Furhat's behavior. If I'm positioned too far away, I need to raise my voice for Furhat to hear me, as the microphone device is still under construction. Furhat's "senses" are created through technological devices, microphones for hearing and camera for seeing and so on. We start to communicate with Furhat. I present myself to Furhat and it asks me my name. It feels a bit awkward that a machine is pleasantly smiling at me when awaiting a reply. I ask Samer if Furhat can read my (human) facial expressions and he runs the camera-program, showing me how the robot can recognize seven different human emotions by reading the human face e.g. eyebrows, if you blink, smile or when you move your head etc. Samer tells me that it is up to the robot developer/skill-builder to how the tracking should be used in human-machine interaction, when personality and the intentionality is engineered in a particular robot. In other words, the programmer can either create a friendly/polite robot or a rude one. Furhat can also change its' appearance/face and Samer goes through the alternatives currently available; e.g. a

[38]The robot is wearing a furry winter hat, and Samer explains how the robot got the name Furhat after a tech student one day forgot his hat on the mask, giving the mask-robot a more appealing and softer personality.

white male and female, a blue avatar, Barack Obama, and a Disney-like childish character. After trying the work-in-progress skills of Furhat I ask Samer if his aim is to create an artificial mind. *"Sure, yes!"* he enthusiastically replies. He continues to tell me that Furhat Robotics are currently building a robot head and social AI containing sensory capacities and other integrating services, like cloud services and other applications. I ask him what characteristics are considered desirable, when you develop a machine that you want to talk to. Being a device that behaves in characteristically human ways, you don't want the machine/robot to look too close to a human. Samer tells me that could cause the "uncanny valley" effect. If the robot looks human but doesn't have the "right" timing or reaction, it generally creeps people out.[39] Samer gives an example; *"say you smile but the robot doesn't smile back at you"*. It looks human, but behaves in unpredictable ways, which would make you feel very uncomfortable. According to Samer, a robot may be human-like, but it's not as intelligent as a human. A human doesn't presume or think that the social robot has human abilities. Instead, they might think of it as an actor in a theatre play or a film. After a while you forget that it is acting and it almost turns real, because it evokes thoughts and emotion. People don't talk to Furhat the same way they talk to each other and in the near future people won't treat Furhat like a real human, Samer explains. Sitting opposite the robot, I notice how its reactions are too slow for me to really have a conversation with it. I figure people are not used to talking to a robot today. I turn to Samer and think out loud; *"talking to machines must be a new form of language that is just starting to develop, right?"* He agrees and says that although robots and humanoids are familiar through sci-fi culture and films, not many have encountered one in real life. Samer tells me how the idea of creating intelligent machines can be traced back to the 50s, with a movement that applied theories about how the human brain and neocortex works in the design of computer algorithms. People saw it as a new era. The spirit was high, believing the time had come when it soon would be possible to program humanlike machines that were extremely smart and intelligent. This was going to solve all the problems of the world and a lot of money was invested in the area of robotics. Samer ends his story with the conclusion that scientists expected too much. The hardware was very expensive, and the computers were very slow. However, now we have reached a new wave of AI-development and people in the tech-industry expect AI to become a hundred-million-dollar industry. Samer thinks many AI-developers throughout history have had the common

[39] Uncanny Valley is a term used in technology for human-like machines to describe the hypothetical phenomena in humans to react to technological objects that are human-like with emotional responses of e.g. awkwardness, unease, aversion, revulsion or fear.

desire to re-create themselves in their own image, which almost touches on reproduction and the instinct to have children to genetically live forever. I think this resembles "playing god" or recreating yourself into a god-like eternal figure. We continue to talk about the subject of recreation, e.g. with Furhat you can design a mask mirroring the image of your own face. Samer discloses by telling me how he stays motivated in the lab through learning about being human through the process of replicating human behavior/traits in technology.

While transcribing the recorded interview, I reflected on the gendering of the Furhat robot prototype. Samer had called the robot "he" throughout the interview and the prototype mask at the time was called "Furhat Svensson" manifesting the image of a stereotypical white male, tech-nerd, KTH student. This was the marketing image Furhat Robotics used for promotion, e.g. with Furhat Svensson as their Facebook profile. Samer explains the reason behind Furhat Svensson; *"We created this stereotype of a white male, a Swedish persona for an event. So Furhat became Furhat Svensson".* At this early stage, the robot didn't have a top/head, only a mask but the science team wanted the robot to develop more personality. The package that usually comes with an AI-robot is that it is very rigid and with no emotions. Samer explained how they wanted to break that stereotype and create the illusion of the robot being alive. Yet, the gender-stereotype remained. I decided to critically discuss prototype Furhat Svensson's stereotypical cis-male image in my case study. Even though Samer himself wasn't a white Swedish person, he still made this gender/race-biased choice, which I thought was interesting. Samer contacted me after reading my text and we decided to meet for a coffee and talk more. He told me the case study was eye opening on the gendered problematization. The Furhat team was currently planning on creating a new prototype version and he asked me if I wanted to be part of this process. Samer's words*; "A lot of these things that we do are led by intuition because we have a feeling or intuition about how to make the machine become more human, but we still don't have a theory, or words on how to express this"*. Robot designers are still searching and creating this new language for socially intelligent machines *"an area where social scientists like you would be very contributive"*, says Samer. I felt flattered, but due to my fulltime job, I regrettably had to decline. Instead, the idea for a collaboration with Furhat, in the context of an art project, started to grow. Through art, I would have freer hands to explore the agency and gender issues discussed with Samer, and, how human-machine interaction can bring new understandings to what it means to be human.

Chapter 6: Beyond human performativity

"An artificial intelligent robot embodied through a mask is on a quest. It enters a female and queer world, where the body is wildly celebrated." (Arise Amazons, official trailer)[40]

In this chapter I will narrate the theatre-play *Arise Amazons* and analyze Furhat's social performance connected to concepts of body, gender and agency, discussed in the theory chapter. In chapter 4, I detected the concept of agency within robot technology to be a complex matter which made me curious to further explore agency through empirical research. The empirical information collected at the Furhat-lab generated the idea to practically test human-machine performativity through a genuine meeting with a live theatre-audience. In the following section, two important co-creators of the theatre experiment, Rebecca Tiger (director/actor) and Petter Lidberg (Furhat operator) play significant roles. Through conversational interviews with Rebecca and Petter, I ask them to discuss their personal experiences with Furhat and reflect on human-robot interaction connected to agency, body and gender. Hannah Arendt (1998:184) describes narratives as "living realities", rather than being the producers/authors of our life stories, it is through acting and speech that we insert ourselves in the world through a web of human relationships. Drawing on her viewpoint, the narrative of the play involves a tribe of symbolical amazons, consisting of a performer-crew from different artistic genres and with variations in gender, choreographically spinning a web of narrative body-knowledge for the body-less robot, aiming to catch its attention.

The Wizard of Oz and human-machines agency

A few weeks before our premiere, a journalist from the Swedish magazine Ny Teknik [New Technology], called me and made a short interview.[41] He was very curious to know more about the theatre play and what role we had given the social robot. I described it as follows. The first thing that happens at the theatre is that the audience is welcomed in by the social robot, which is a disembodied talking head, created to interact in a humanlike manner. It can talk, recognize and make facial expressions, signifying different emotions. The audience are eventually invited to enter the main theatre space where the stage is designed as a circular circus arena, filled with soil. Furhat is situated to the right, just outside the arena, turning its head and gazing out towards the audience. The audience is about to enter a fictive world. A

[40] Link to Arise Amazon's official trailer: https://www.youtube.com/watch?v=NJ72RP1KhQk
[41] Article available here: https://www.nyteknik.se/automation/roboten-tar-plats-vid-teatern-6821816

journey that we do together with the robot as the protagonist, driving the narrative of the play through its "desire" for a body.

Before continuing the story onstage, I want to take you backstage and discuss the challenges of preparing the robot to behave like an actor. What the audience does not know, is that my technically skilled friend Petter, is seated in the back with the theater technicians. He is controlling the social robot manually through his laptop. Furhat is activated to perform as an "wizard-of-oz" interface, which is a common technique of operating a robot and control it to act and behave in human-like ways.[42] This way the social robot is operating in a passive mode where a human user controls the speech, gestures and gaze of Furhat directly through a computer. With one exception, the robot has sensors that enables it to automatically react to movement and turn its face and glance in the direction of humans or other moving objects, an option that also can be switched off.[43] When the art-project *Arise Amazons* started to manifest, I realized I needed some external help with Furhat. The Furhat team didn't have time to sit with me and prepare Furhat for its theatre role. Instead I asked my friend Petter to assist me, since he seemed appropriate as a postgraduate mathematician proceeding and proceeding with AI studies at KTH. Petter and I spent a couple of days in the lab with a Furhat robot device. In the conversations with Petter, I wanted to know more about his mathematician-perspective on AI. We discussed issues of agency and gender in machines. Through this method of constructive discussion, the performing robot character and non-human actor in *Arise Amazons*, was gradually conceived through an interdisciplinary collaboration between a feminist scholar (me) and a mathematician (Petter). I want to share with the reader some of Petter's viewpoints that I found interesting for the analysis of the case we were working with, preparing a robot for a theatre performance. When I brought up the issue of agency in AI, Petter told me that he associates agency to what is called "agents" in AI studies. He thinks the concept of agents is a bit vague, but a way to understand it is to use the example of AI agents in video or computer games, where you (the gamer) fights against an AI opponent agent.[44] Petter compares the agent/avatar controlled by a human to a puppet toy and the more the human is attracted to the avatar the more human/avatar merge. His comparison makes me

[42] *Wizard of Oz* or "wizarding" is when a human operator to trigger certain actions in the robot instead of automatic triggers (https://docs.furhat.io/unstable/wizard/).
[43] This is made possible through technological modules for body and face tracking, joint-attention, and multimodal processing and activated through turning on this autopilot setting by pushing a button.
[44] In world domination games like World of Warcraft you (human) control your agent/avatar/soldier/warrior and the computer agents (machine) fights you back.

think of Breazeal's robot-strategies discussed in chapter 4. If the robot appeals to its human user, then it's easier to play and interact. I didn't know at this point the level of commitment required of Petter in his role as wizard-of-oz nor how he would gradually merge with the robot-character. During the few days in the Furhat lab, Petter and I made an effort to learn how the robot functioned. Together, we programmed it according to the play's script and what we wanted it to say, in dialogue with the other actors. Rebecca played the narrator of the story. It wasn't until the week of rehearsal at the theatre that Rebecca met Furhat for the first time. She was on maternity leave with her newborn baby and didn't have time to come to work with us in the Furhat-lab. Petter and I were struggling with very limited time to figure out the dynamics of the social robot's interplay and dialogue, leaving Rebecca with nearly no time to rehearse the parts with the robot properly. Furhat arrived at the theatre just a day before the premiere, and Rebecca sat down with Petter backstage and re-programed the robot's lines for them to be in better sync with hers, and to get a better flow in their human-machine dialogue. Robot-talk has that characteristic, stiff and staccato, sound to it and we only had a few options of voice-settings to choose from. Rebecca was feeling very stressed by the situation and started to question giving the robot such a leading role and presence onstage. She questioned whether or not the robot would be able to perform on stage and if she could put trust in this non-human actor. I felt nervous to, but I needed to stay calm and be the assuring person in the process, since it was my crazy idea to include a robot in the play.

Back to the theatre stage. While the audience is seated, they can see a robot-head gazing out at them from the stage's arena. Meanwhile, they also hear the sound of various birds and animals. It is the "voice of nature" calling, a character played by the Siberian shaman singer, Olena Uutai. She cannot be seen but her voice amplifies an authentic atmosphere of being outdoors in nature with real animals (since the audience can't connect the sound to a human). The arena lights up slightly and Olena whinnies like a horse. Simultaneously, a big black horse emerges from the haze and confidently steps out into the arena. The horse is free to move around as it pleases. It steps right to the front, and gazes out at the audience and at Furhat. In our discussion on agency, Petter remembers this scene as the most interesting and powerful moment. He could sense how the horse was bothered by the presence of Furhat, acting from the corner of the stage. The robot was following the movement of the horse through the tracking sensors, turning its neck and gazing at the horse, as it moved in the arena. The horse seemed to interpret the robot as a living thing with its own agency. Petter did not perceive the horse as being afraid but rather annoyed, and he tells me this made him act

cautiously to a point where he even turned off the automated head-movement trackers.[45] Instead, he experimented and directed the robot's gaze out into the audience. That's when something very interesting happened. The horse walked up to the edge of the arena and gazed out into the audience too. There was a gasp in the audience, with two non-human creatures intensely gazing at them, welcoming them before any human agent/actor had entered the stage.

Body and social intra-action

After the black horse has performed, Rebecca appears. She is wearing Swedish traditional clogs, a she-hulk t-shirt and a winter-furhat, in sync with the robot. You could tell by looking at her belly and breasts that she was post-pregnant. In this sense, she wanted to embody the mother-body in the story. Balancing on the circular frame of the arena, Rebecca approaches the robot and welcomes the audience; *"Welcome! I've created this ceremonial circle especially for us. Thank you for accepting the invitation to this altered state of reality. Behind all things are reasons. Reasons can even explain the absurd. Do we have time to learn the reasons behind human beings varied behavior? I think not."* Rebecca starts a dialogue with Furhat. Talking for the first time the robot presents itself: *"Hello, I'm the world's most social robot. I've got a face, but I really desire a full body".* Rebecca tells the robot that it has come to the right place to experience body, and says; *"Are you ready? Shall we begin?"*

Months after the theatre project was finished, Rebecca and Petter reflect in an interview how they both had thought the robot to be more technologically advanced. Rebecca experienced it as a bit of a shock that the robot couldn't act or have a spontaneous dialogue at all. Instead, we had to build a dialogue based on a script. After a shaky premiere-night, the script progressively got better and more spontaneous. Rebecca says it had to do with learning how to interact with Petter and Furhat to create a flowing social human-machine dialogue, that worked for receiving audience responses (e.g. laughter). Petter recalls the experience, how controlling Furhat by the wizard-of-oz mode at the theatre made him feel like a puppet-master. I associate this to actors in a theatre controlled by a director and a script, and both Petter and Rebecca agree to my parallel. Furhat doesn't have a full body but is a social robot embodied through a head with a face in the form of a mask. Rebecca compares acting

[45] According to Petter, Furhat Robotics have built a good platform to run it in the wizard of oz mode where the user can pre-program different modes of interaction and combine this with automation, e.g. the head movements of the robot, through camera tracking sensors, making it possible to follow and gaze at the human(s) talking to it.

together with Furhat to commedia dell'arte, an early form of theatre technique, where the actors wore characteristic masks and played for large audiences. She explains that when acting in masks, the actors have to be aware of the fact "I can't see what you are thinking".[46] In this case, with a social robot in the form of a mask, Rebecca applied this method and found it very useful. Furthermore, she says that it was Petter as the puppet-master who was her co-actor, rather than Furhat. Petter comments on this and says that he thinks Rebecca's reflection is very interesting, and how it was his and Rebecca's correspondence that constituted agency in Furhat. Rebecca implies that the flow in the dialogue was very dependent on her timing that created the illusion of the robot being "real" and an acting agent. Relating to the importance of body, Rebecca describes how she used her body/body-language as an important instrument in creating this illusion of realness to convince the audience. She explains that when Furhat would talk, she could alternate between looking at the robot, responding with exaggerated body language, and then turn to the audience to hint that she expected a response from them. She regrets that she didn't have time to develop her body language more, considering how remarkably it improved the human-machine comprehensiveness for the audience. Another method Petter, Rebecca, and I used which turned out to work quite well with the audience was to build dialogue using lyrics from famous pop-songs. Rebecca describes the music as being a shortcut to human emotional reactions. She calls it a plump director's trick, but it proved to be very effective. E.g. in the beginning when Rebecca asks the robot how it feels, Furhat says; *"first I was afraid, I was petrified…"* Rebecca exaggerates her body-language, looks at the robot with large eyes, and replies; *"but you are a robot, can you be afraid?"*. Many people in the audience would laugh with recognition, associating the script to the lyrics of Gloria Gaynor's famous disco hit "I will survive", still popular on gay-club dance floors (we intentionally introduce the robot with a queer tone). The unfamiliarity of having a robot as an actor on a theatre stage, with an emotionless robotic voice, was compensated through coded pop-music lyrics to give the robot emotional agency. It also added a humorous touch, that seemed to create a bond between the audience and Furhat, making the robot likable. Rebecca describes the situation from her perspective; *"what happened when we used music lyrics, was that it created a rhythm to our human-machine dialogue in combination with my body reactions"*. In this way, emotional agency was successfully projected into the robot. Rebecca argues that this emphasized authenticity to the

[46] Commedia dell'arte is one of the earliest forms of professional theatre. It originated in Italy during the High Renaissance era of the 16th century and grew very popular throughout Europe. Commedia is a theatre form characterized by pantomime and masked and exaggerated representations of fixed social types (e.g. doctor).

human-robot interaction, where she could communicate with only body and no words and the robot with words but not with body. Rebecca understood that she would have to be the embodiment for them both for the audience to perceive the robot as human-like. She had to deal with a slight delay in the robot's reply, a reason why she and Petter found it necessary to rely mostly on scripted dialogue. Otherwise, as Rebecca reflects with Petter in the interview, she wouldn't know if it was Furhat who wasn't functioning/reacting correctly or if it was Petter/wizard-of-oz that hadn't reacted quick enough. This was an insecurity she had to deal with on stage, very different from acting with another human being. Petter agrees with Rebecca, that he relied on Rebecca's body-language for cues and to activate the robot's answers and facial reactions. Discussing Rebecca and Petter's experiences exemplifies how human-machine agency, in this case, is created through the partly scripted inter-play between Rebecca and Petter. I argue that the social robot's agency is also dependent on the *social* intra-action with audience/theatre-technicians/horses/amazon-performers and other present bodies in the theatre space.

Elaborating on gender in the Furhat/robot character

Informed by developments in feminist science and technology studies, I used my limited time in the lab with Furhat to plan how I could elaborate on issues of gender and create a queered character for the play (theatre). According to sociologists Candace West and Don Zimmerman (1987), gender is created in communicative interaction. They stress that it is important to construct coherence and intelligibility concerning gender performance, something that I argue is important to consider when programming an AI. Judith Butler states that gender is a form of "improvised performance" and she emphasizes deconstruction as a medium to create gender disruption, trouble, and renewal (Butler 1990). I believe critical deconstruction to be, not only an analytical tool, but aim to explore further how Butler's "gender trouble" could be applied through practical strategies in technology. In the Furhat lab, Petter and I experimented with all the possible modes and settings to get to know the robot's skills and possibilities, elaborating on questions like; why do we gender technology? When you think of a robot or an AI, do you think of a "she", "he" or neutral (non-binary)? I started to get an idea for how to build the character in play. Samer was very excited that Furhat would be in drag when I told him about my idea. They had developed a new standard prototype mask projecting a male face resembling Samer's own features. We chose to use this standard prototype setting, since ordering a custom-designed mask would be too expensive. Instead we came up with a

practical strategy to transform the social robot's gendered appearance in other ways; by gradually gender-bending and queering it on stage along with the proceeding story, as the robot reaches deeper into the Amazon world of female embodiment. Although the robot had a masculine face, the fur-hat contributed to a more gender-neutral look. In addition, the combination of a female voice-setting added further gender-confusion when the robot started to talk.[47] We kept the fur-hat on the robot in the first part of Act One and played with the concept, e.g. Rebecca entering wearing a furry-animal-hat just like the protagonist Furhat. By using the same headwear, it seemed like Rebecca and the robot were related/connected, and the animal fur-hat could be interpreted as a hint of "nature" and the origin of things.

The gendered question of the robot in the story transgresses binary lines and becomes queered in the end of the first act. Furhat is here mesmerized by a long-haired drag queen who is lip-syncing to Gloria Gaynor's gay anthem *"I am what I am, I am my own special creation"*, while tenderly putting a similar long-haired wig on Furhat and dragging/re-constructing the robot's performed gender. In Act Two, the robot device is not on stage. In the story, the protagonist succumbs to a liminal space and is then resurrected and fully embodied as the "cyborg-amazon-goddess". Inspired by Haraway (1991), this is a figure manifested as a critical symbol of the future, yet accentuating potential and hope in coming human and non-human intra-action and technological inventions. The cyborg amazon represents a fictive posthuman agent, challenging binary oppositions and the division between nature and technology. Haraway ends the cyborg manifesto (1985,1991) by claiming that she would rather be a cyborg than a goddess, a critical hint to eco-feminism. In line with queer theorist Jasbir Puar (2012), who critically analyzes Haraway's cyborg/goddess juxtaposition, I argue that cyborgs and goddesses have a lot in common. Both are imagined as striking figures of human enhancement, pictured as impressive more-than-human manifestations. Puar finds it difficult to choose between being a goddess or cyborg, because she is critical to both options. While the first (nature) is embedded with racialized myths of a matriarchy from the past, the latter (technology) hails the future with a tendency towards technological determinism. Her solution is the necessary conjunction of the alternative "cyborg-goddess" that she further describes as an assemblage of becoming intersectional (2012:8) The idea of the cyborg-amazon-goddess character in the play was also animated through Turner's concept of the *liminal personae*, which is a type of trickster. Haraway explains how this is "[a] figure who

[47] None of the voice library options sounded good, they all had a monotone vocal pitch, far from theatrical drama. We finally choose the option of a female voice with an American accent.

plays tricks on humans and whom we can never fully control" (Haraway 1991:199). The story coming to an end, our protagonist robot-character has journeyed through an internal and external transformation from a bodiless mask into becoming fully embodied trickster character. Unbound to fixed gender, embodying both human and non-human (cyborg) and finally unified with the collective amazon tribe.

Summary of human-robot performance

In this section I will analytically summarize the autoethnograhical method by narrating shared personal experiences (with Rebecca and Petter) and the practical experiment of situating the Furhat/robot on a theatre stage. During the writing process of this analysis chapter, I found an early draft for the script of *Arise Amazons* on my computer and remembered how Rebecca and I had worked with color-coding to structure the different themes/categories in the story; green= nature/shaman, orange= human/embodiment, purple = animal/intuition, blue= robot/technology, and pink= goddess/glam/diva/femme. In retrospect, I analyzed the idea of structuring the script and treating the content this way. Drawing on Barad's arguments on agency not as"[a]n attribute, but the ongoing reconfiguration of the world", our aim was to create an experimental theatre play, weaving in themes of binary oppositions into a choreography of entanglement and reconfiguration (Barad 2003:818). Structuring the script by color-coding the different categories, provided a clearer picture of oppositional elements at play which worked as a tool to develop intersectional encounters and reconfigurations. Concluding the *femme-tastic* experiment of studying issues of body and gender related to humanlike agency in a robot, I emphasize the use of autoethnographical methodology for this study, utilizing personal (qualitative) experiences rather than measurable (quantitative) results. Petter, Rebecca and I first expected Furhat to be an independent agent (almost close to a technological singularity) leading to our disappointment when conversations with the social robot turned out dependent on a pre-scripted dialogue. Although effort has been made to mimic senses in the robot, e.g. camera for seeing, Petter presumed that the robot would be capable of spontaneous conversations and be more advanced at human-machine interaction. The most advanced part of Furhat was its professional 3D design with a custom-made mask and animation. This is what makes it human-like, because it can gaze at you and follow your movements. In account of Suchman's and Weber's emphasis on embodiment and emotional skills for humans to engage in social interaction with machines, Furhat can mimic a human and react to conversations with facial expressions. This creates the illusion of emotional

responses, which is very effective and makes you think that it's alive. Petter assumes that this is why it's called *the human face of AI*. He believes the social robot developers are on to something, basing their smart invention on the tendency by humans (and horses) to project meaning and social agency to the human-like gaze of the robot. The AI part of Furhat wasn't developed to the extent that it could use machine learning to automatically learn anything from our theatrical experiment. Petter means that AI doesn't have to be materialized directly in the robot device but can be transferred by different cloud solutions and this is how Furhat will contain human-like intelligence. Even though the level of agency in Furhat wasn't as developed as Petter, Rebecca and I first expected, it's possible we still managed to create the impression that the social robot had an identity and was acting spontaneously. As discussed in this chapter, this depended on Rebecca's acting skills and Petter's technological skills. Agency arose when human-robot interaction consisted of a communicative circuit of three bodies; Furhat/robot, Petter/wizard and actor Rebecca. Through the practice of exploring the social robot's performance in the theatre, I agree to Adam's (2002) discussion that AI performativity becomes poor when not channeled through a material body. Rebecca came to the conclusion that she had to compensate for the robot's lack of body by exaggerated and explicit body-language. In this sense she represented embodiment in the first part of *Arise Amazons*, while Petter was the mind and wit of the robot protagonist, speaking through a queer filter of a female voice combined with a masculine mask. Drawing on Treusch, and Butler's "gender trouble", I explored methods of queering biased gender norms projected in humanlike technology (Furhat). In the end of the story, binary oppositions of human/machine and nature/technology symbolically merge with the symbolical figure of the cyborg-amazon-goddess. All amazon performers, an army of heart-warrior/feminist activists, enter the arena with the narrator's (Rebecca's) last words: *"What is reflection? With opportunities of reflection circumstances expand. Where there was once one, Or, were there always three? We are now plenty and everywhere- ARISING! It has been a pleasure speaking to you".* While linear time feeds the concept of an isolated separated self, highly promoted in liberal politics and contemporary western societies, the non-linear reality presented in Arise Amazons preferably includes the Other/robot and creates community. If social robots like Furhat will provide humans the opportunity to re-create the image of themselves through a robot-mask of self-reflection, I'm concerned that it might cause self-absorption, possibly increasing an aversion to meet Others (e.g. people of different opinions). In this sense, humanlike technologies could have remarkable impact on "[o]ur sense of self" (Wajcman 2004:2) and Others, which I will discuss in the next chapter concluding the research question.

Chapter 7: Conclusion of humanlike agency in social robots

Through robots we will spot ourselves and reflect on what it means to be human. We believe social robots will be the next major platform for humans interacting with technology, and Furhat has social skills taught by YOU. (Samer Al Moubayed -CEO Furhat Robotics- ad on Instagram April 2019).

Samer says that the Furhat team doesn't build AI robotics with an obligation to companies to save more money. Instead he believes they do it out of a curiosity for a better understanding of what it means to be human. In this chapter, I will conclude the empirical material in relation to the research question; *does contemporary development of socially intelligent machines have the potential to impact the understanding of human agency, body, gender and social identity?* In the Furhat-lab, Samer informed me that when developing AI, you always have to think of intentionality. If the robot is to develop agency and make its own decisions someday, with their own kind of emotions, the aspect of intentionality is very important because human machine intra-action will accelerate. He believes human-like emotional AI in the future might be possible through copying and reconstructing stem cells to re-create an artificial brain and limbic system. When posing the question to Samer and if he thinks it's possible for AI to develop consciousness, he says it depends on whether you are pessimistic or optimistic when it comes to technology; *"I am more of an optimistic type and I believe that consciousness is a property based on the level of experience meaning that consciousness is an illusion. AI will eventually become conscious at some point, but that is very far from reality today."* He shares a metaphor with me; to think of the universe as a clock ticking away in minutes- and if you believe that the universe is deterministic- then there is no place for true meaning or free will. I tell him that feminist technoscience rather constitutes a critical approach to determinism and linear perceptions of time. However, technological innovations merging human and non-human life forms might carry the potential to create alternative non-deterministic realities.

Agency analyzed as correlating to body and gender

The empirical research of Furhat's performativity in a theatrical setting demonstrates how agency in this particular case depended on human-machine interaction with actor/Rebecca and wizard-of-oz/Petter. The social agency of the protagonist robot-character also depended on the *intra*-action (Barad 2003) of all co-present bodies in the theatre space. Agency consequently depended on several factors of intra-action and changed from one night to another. For example, if the audience laughed a lot or kept quiet the robot protagonist

character would respond differently according to what was socially appropriate. Although, this was not the robot itself performing agency. In chapter 4, I discuss how human-like agency appears to rely on materializing AI through a material body (Treusch 2015:102, Weber and Bath 2007). Accordingly, the object of study "Furhat the human face of AI" proved in the theatre experiment to require, and attain, a full body (not only face) from Rebecca and human wit from Petter. Rebecca's body language was an important instrument to project emotional responses in human-machine interaction, causing the audience to "know" when and how to react (social timing). This circuit of three, Furhat/ Rebecca/Petter, generated the illusion of personality and human-like agency in the robot, making it seem practically human. Breazeal's social robot Kismet was in the lab perceived by humans as having identity and agency, although it was actually "absent presence", with the operating computers placed in a different room. Furhat is more technically advanced and portable yet still depended on the physically distanced Petter as the Wizard-of-Oz and his laptop computer. The empirical research at the Furhat-lab involved and generated the research question vice versa; human impact and scientist norm-oriented projection onto the robot prototype, building on Weber's (2005) critical questioning of body, emotion, gender, and social skills simulated in socio-emotional machines. Suchman's term "situated action", critically referring to the biases of culture(s) (2007), was acknowledged when elaborating on gender in the stereotyped Furhat prototype throughout the narrated story of the theatre play, as analyzed in the previous chapter. If social robots are systematically created/engineered without critical regard to gendered and racial characteristics, bound by archaic thinking and Western prerequisites and norms, this technology will maintain the narrow viewpoint of a small percentage of humanity. On that account, the gender deconstruction, based on Butler's theory, and the gender-bending elaboration on Furhat aimed to practically problematize and re-construct alternative gender(s) in the robot.

Possible directions for further research

Advanced technology is connecting us more and more online and creating the phenomenon of a distinctive collective mind of social and cultural entanglements. Recapping Wajman's inquiry; whether technology will have an impact on the "sense of self" and social identity. I argue that human-like agency is a complex matter, in which human beings, and their sense of selves, are vulnerable to a myriad of entanglements and binding preconceptions that artificially intelligent beings can/may embody. Just imagine an artificial mind that eventually

develops an uncontrollable agenda of its own. Therefore, with the critical perspective of a feminist scholar, I think it is crucial to find ethical methods for social and human-like machines to be programmed with heterogenous, intersectional and ethical perspectives. Another important aspect is to include "softer" values, like sociability connected to emotions, or sensory experiences and even intuition. For this to be realized, there is a need for fruitful dialogues between social science, feminist technoscience and technology. Positioned among the feminist technoscience scholars, I see it as urgent to examine the gendered situation in cases of developing language and communication skills for human-machine interaction. With this, I mean the gender of the person(s) who program the machine and how that causes a gendering of the machine/robot.[48] Rather than being a neutral or objective activity, I stress that language and the ways we humans socially communicate is colored by intersectionality, such as gender, body, (dis)abilities, class, culture, and more. Being precautious against transmitting biased language and social skills to machines is something feminist technoscience can "remind" technological science about.

Standing in front of AI and quant mathematicians at Bloomberg, I expressed my feminist concern of the digital future to come. I stated, that if machines are to become sociable and attract a wider range of users, they need to represent the diversity and nuances of humanity and social structures. Furhat, the human face of AI, is still in its infancy, a start-up developed through "Furhat OS", an operating system for early adapters to use and create collectively. I argue that this is an innovative example that carries potential for diversity. That is, if more women and other under-represented groups of people are encouraged and included to engage and partake in the development of AI. Collectively, we are all responsible in setting the course to where we are heading and to what the human face of AI will become. Socio-technical imaginaries are constitutive of society and have impact on what society considers "proper" human behavior (MacKenzie and Wajcman 1999; Treusch 2015; Rhee 2018). In line with Treush (2015), and her posthumanist perspective of the social and technical as related through entanglements as the "socio-technical", I find feminist posthumanist studies to contribute with significant analytical perspectives in the field of social robots. What ultimately matters in the design of sociable robots is how the machine performs in tested

[48] Haraway (1997:11) argues that technologies are forms of materialized figuration and that all language is figural. The intervention of current practices of technological development should be a critical consideration of how humans and machines are currently figured in these practices and how they might be reconfigured differently (Suchman 2007:127).

human-robot interactions and if it will be perceived as authentic by humans. This is something I aim to study further and the experiences with this study has motivated me to develop new practical methods to test sociability skills in machines with a special emphasis on ethics and how to analyze, detect and avoid gender, and other intersectional, bias in AI/robot technologies.

A cyborgian rite of passage

Feminist theoretician Rosi Braidotti (2003) states that contemporary technology is disrupting centuries of humanist thinking by reaching deeply into the most intimate layers of living organisms and the human structure of self. She advocates the importance of the narrative and making myths, based on her presumption that because of technology we can "[n]ow think of the body as an entity that inhabits different time-zones simultaneously" (2003:21). In this sense, I propose the experiment at the theatre to be an example of practically testing narrativity and myth on the topic of human-machine agency, gender performance and inter-action through a live audience. Academic thinking and writing challenged me in understanding critical posthumanist theoretical perspectives of what it means to be a human intra-acting with other non-human agents. Although it wasn't possible for me and the amazons to teach Furhat a feminist perspective through machine learning at this stage of its development, I still see the project as fruitful. My perception about humanlike technologies and human-machine interaction has grown along the study. What I am implying here is that; through this study I've personally experienced a rite of passage with a deeper understanding as discussed in the methodology chapter (Aull Davis 1999:67-68). I have become motivated to further explore the concept of cyborg feminism, and how I, as a woman/human, am suspended in multiple entanglements and ever-changing intra-actions, blurring the boundaries between nature/technology. I continuously intend to disrupt the binary between science and art, using feminist curiosity and creativity to explore- theoretically and practically- what it means to be human within various contexts of techno-cultures.

References

Adam, Alison (2002) *Gender/Body/Machine* (Blackwell Publishers Ltd. Oxford).

Alaimo, Stacy; Hekman, Susan J. (2008) Introduction: Emerging models of materiality in feminist theory. In *Material Feminisms*, edited by Alaimo and Hekman pp.1-19 (Indiana University Press).

Andersen, Warwick (2002) Introduction: Postcolonial Technoscience. In *Social Studies of Science* 32/5-6 p.p 443-658 (SAGE Publications).

Arendt, Hannah (1998) *The Human Condition,* 2nd Ed (The University of Chicago Press).

Aull Davis, Charlotte (1999) *Reflexive Ethnography: A guide to reaching selves and others* (London Routledge).

Badmington, Neil (2000). *Posthumanism (Readers in Cultural Criticism)* (Palgrave Macmillan).

Barad, Karen (1996) Meeting the Universe Halfway: Realism and Social constructivism without contradiction. In *Feminism Science, and the Philosophy of Science,* ed. Lynn Hankinsson and Jack Nelson pp.161-94 (Dordrecht, the Netherlands: Kluwer).

____. (2003) Posthumanist Performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society* 28(3) pp.801-31.

Bochner, Arthur P. (1997) It's About Time: Narrative and the divided self, *Qualitative Inquiry*, v. 3, issue 4, p. 418-438 *Communication Faculty Publications*. 13.

Bochner, Arthur P.; Ellis, Carolyn (2000) *Autoethnography, personal narrative, reflexivity*. In *Handbook of Qualitative Research (2nd Ed.)* p. 733-768 (Sage Publications).

Braidotti, Rosi (2002) *Metamorphoses: Towards a Materialist Theory of Becoming* (Cambridge UK: Polity).

____. (2013) *The Posthuman* (Polity Press, Cambridge and Malden).

Breazeal, Cynthia (2002) *Designing Sociable Robots* (A Bradford Book).

Butler, Judith (1990) *Gender Trouble: Feminism and the Subversion of Identity* (London Routledge).

____. (1993) *Bodies That Matter: On the discursive limits of "sex"* (London: Routledge).

Burkitt, Ian (2015) Relational Agency: Relational sociology, agency and interaction In *European Journal of Social Theory Volume* 19 issue: 3 pp. 322-339
Dreyfus, Hubert (1982) *Husserl, intentionality and cognitive science* (Cambridge MA: MIT Press).

Ettore, Elizabeth (2016) *Autoethnography as Feminist Method: Sensitising the feminist "I"* (London, New York: Routledge).

Ferrando, Francesca (2014) Posthumanism. In *Kilden Journal of Gender Research Published by the Centre for Gender Research in Norway*, Vol. 2 pp. 168-172.
Gell, Alfred (1998) *Art and agency: An anthropological theory* (Oxford: Oxford University Press).

Haraway, Donna (1989) *Primate Visions: Gender, race and nature in the world of modern science* (New York: Routledge).

____. (1991) *Simians, cyborgs and women: The reinvention of nature* (London Free Association Books).

____. (1991) *Situated Knowledges: The science question in feminism and the privilege of partial perspective.* In *Simians, cyborgs and women: The reinvention of nature,* pp. 203-30 (London Free Association Books).

____. (1997) *Modest_Witness@Second_Millenium.FemaleMan©_Meets_Onco-Mouse™. Feminism and Technoscience* (New York: Routledge).

Harding, Sandra (1986) *The Science Question in Feminism* (Ithaca, NY: Cornell University Press).

Hayles, N. Katherine (1999). *How we became posthuman: Virtual bodies in cybernet- ics, literature, and informatics* (Chicago, IL: University of Chicago Press).

Knappet, Carl; Malafouris, Lambros (2008) Material Agency: Towards a Non-Anthropocentric Approach (Springer Societies & Publishing Partners).

Lykke, Nina (2010) *Feminist Studies; A Guide to Intersectional Theory, Methodology and Writing* (New York, London: Routledge).

____. (2014) *Writing Academic Texts Differently: Intersectional feminist methodologies and the playful art of writing* (Routledge).

Lykke, Nina; Åsberg, Cecilia (2010) Feminist technoscience studies, *European Journal of Women's Studies,* 17(4) pp.299-305.

Lykke, Nina; Markussen, Randi; Olesen, Finn (2004) *Cyborgs, Coyotes, and Dogs: A Kinship of Feminist Figurations and There Are Always More Things Going on Than You Thought! Methodologies as Thinking Technologies*. In Donna Haraway: The Haraway Reader (London, New York: Routledge).

Mac Kenzie, Daniel; Wajcman Judy (1999) Introductory Essay. *The social shaping of Technology*, 2nd Ed pp.3-27 (Buckingham, Open University Press).

McNeil, Maureen (2005) Introduction: Postcolonial Technoscience. In *Science as Culture* 14(2) pp.105-112.

Miah, A. (2008) *A Critical History of Posthumanism.* In Gordijn, B. & Chadwick R. Medical Enhancement and Posthumanity. Springer, pp.71-94.

Mehrabi, Tara (2016) *Making Death Matter: A feminist technoscience study of Alzheimer's sciences in the laboratory* (TEMA, LiU-Tryck).

Muncey, Tessa (2010) *Creating Autoethnographies* (SAGE Publications Ltd).

Puar, Jasbir (2012) 'I would rather be a cyborg than a goddess' Intersectionality, Assemblage, and Affective Politics. In *Becoming-Intersectional in Assemblage Theory* Volume 2(1) pp.49-66 (State University of New York Press).

Rhee, Jennifer (2018) *The Robotic Imaginary: The Human and the Price of DeHumanized Labor* (University of Minnesota Press).

Rosenberg, Tiina (2018) *HBTQ spelar roll. Mellan Garderob & Kanon* [HBTQ Matters. Between Closet & Canon] (Leopard Förlag).

Schechner, Richard (2002) *Performance Studies: An introduction* (Routledge).

Shildrick, Margrit (1997) *Leaky Bodies and Boundaries: Feminism, Postmodernism and (Bio)ethics* (London: Routledge).

Suchman, Lucy (2007) *Human-Machine Reconfiguration: Plans and Situated Actions,* 2nd Ed (Cambridge University Press).

Treusch, Pat (2015) *Robotic Companionship: The making of Anthropomatic Ktichen Robots in Queer Feminist Technoscience Perspective* (LiU- Tryck Linköping).

Turkle, Sherry (1984) *The Second Self: Computers and the human spirit* (MIT Press).
____. (2005) Deluxe Edition: *The Second Self: Computers and the human spirit* (MIT Press).

Turner, Victor (2008) *The Ritual Process. Structure and Anti-Structure.* 2nd Ed. (Transaction Publishers).

Wajcman, Judy (2004) *Technofeminism* (Cambridge, UK; Oxford: Polity Press).

____. (2007) *From Women and Technology to Gendered Technoscience* Article: Information, Communication & Society Vol. 10, No. 3, June 2007, pp. 287–298.

Wachter-Boettcher, Sara (2017) *Technically Wrong: Sexist apps, biased algorithms, and other threats of toxic tech* (W.W. Norton & Company Inc.).

Webb, Amy (2019) *The Big Nine: How the Tech Titans and their thinking machines Could Warp Humanity* (PublicAffairs U.S.).

Weber, Jutta (2005) *Helpless Machines and True Loving Care Givers: a feminist critique of recent trends in human-robot interaction* (Troubador Publishing Ltd).

____. and Bath, Corinna (2007) ´Social´Robots & ´Emotional´ Software Agents: Gendering Processes and De-gendering strategies for `Technologies in the Making´. In *Zorn, Isabel and Maas, Susanne and Schirmer, Carola and Rommes, Els And Schelhowe (eds): Gender Designs IT. Construction and deconstruction of, information, society and technology* pp.53-63 (Wiesbaden: VS Verlag).

West, Candace & Don H. Zimmerman (1987) *Doing Gender: Gender and Society* 1(2) pp.125-51

Online sources

AI Now Institute with AI Now 2018 Report available at: https://ainowinstitute.org/reports.html

The Bloomberg Quant (BBQ) seminar series. Available at: https://www.bloomberg.com/professional/quant-seminar-series/

Fredelius, Ada (2017) Roboten tar plats vid teatern [The robot enters the theatre]. Available at: https://www.nyteknik.se/automation/roboten-tar-plats-vid-teatern-6821816

Furhat Robotics official webpage available at: https://www.furhatrobotics.com/

Kostinen, Aino-Kaisa; Karkulehto, Sanna (2018) Feminism. Available at: http://criticalposthumanism.net/feminism/

Ruby Rose Union, production house theatre play Arise Amazons, available at: https://rubyrose.se/

I AM WHAT I AM

(Gloria Gaynor)

I am what I am
I don't want praise, I don't want pity
I bang my own drum
Some think it's noise, I think it's pretty
And so what if I love each sparkle and each bangle
Why not try to see things from a different angle
Your life is a sham
Till you can shout out
I am what I am
I am what I am
And what I am needs no excuses
I deal my own deck
Sometimes the aces sometimes the deuces
It's one life and there's no return and no deposit
One life so it's time to open up your closet
Life's not worth a damn till you can shout out
I am what I am
I am what I am
And what I am needs no excuses
I deal my own deck sometimes the aces sometimes the deuces
It's one life and there's no return and no deposit
One life so it's time to open up your closet
Life's not worth a damn till you can shout out
I am what I am
I am, I am, good
I am, I am, strong
I am, I am somebody
I am I do belong
I am, I am, good
I am, I am, strong
I am, I am somebody
I am I do belong
I am, I am, useful
I am, I am true
I am, I am worthy
I am as good as you
I am, I am, useful
I am, I am true
I am, I am worthy
I am as good as you

Written by: Mark Owen