



LUND UNIVERSITY

Defining the Voice of Montréal

*Exploring Possibilities for Human-Machine
Companionship*

Malik Dehbi Talbot

Master of Applied Cultural Analysis
Department of Arts and Cultural Sciences

TKAM02 - Spring 2020

Supervisor
Jakob Löfgren

Abstract

Defining the Voice of Montréal: Exploring Possibilities for Human-Machine Companions

Malik Dehbi Talbot

This thesis explores the possibility for (re)defining the relationship between humans and robots with conversational interfaces. It does so by looking at the creation process of a voice and text-based virtual assistant for tourists from the perspective of critical posthumanism and posthuman performativity. In a reflexive fashion, I analyze my involvement in a tech start-up, in order to argue that, rather than being ontologically separated, the boundaries between humans and machines are culturally and historically constructed. Moreover, through a closer look at the development of my own relationship with the prototype of the voicebot and my performance of ‘demos’ among uninitiated users, I put forward a relational understanding of how humans and machines *become* with - and constantly remake - each other. This allows me to redefine our relationship with intelligent artefacts beyond mere instrumentality, towards a form of human-machine companionship that highlights the potential of the relationship. Finally, through a practical engagement with the idea of posthuman responsibility, I analyse the effects of specific features of the voicebot and imagine how the boundaries between human and machines can be re-configured in a more responsible way. Amongst other things, this allows us to re-contextualize labor relations associated with service work and product development, with respects to how these practices shape gender and race. The thesis concludes by stating the relevance of these results for the fields of tourism and product development.

Keywords: critical posthumanism; artificial intelligence; virtual assistant; tourism; science and technology studies; human-machine relationships; voicebot; labor relations; companion; sociomaterial assemblage.

Résumé

Defining the Voice of Montréal: Exploring Possibilities for Human-Machine Companions

Malik Dehbi Talbot

Cette thèse explore les possibilités de redéfinir la relation entre les humains et les robots dotés d'interface conversationnels. Je m'attarde plus particulièrement au processus de création d'un assistant virtuel pour touriste en adoptant une perspective critique dite post-humaniste. J'analyse d'une manière réflexive mon implication dans un startup œuvrant dans le domaine de l'intelligence artificielle pour soutenir que la séparation entre les humains et les machines est issue d'un contexte culturel et historique bien particulier, plutôt qu'essentielle. D'autant plus, à travers un regard plus attentif au développement de ma relation avec le prototype du voicebot et ma performance des démonstrations auprès des utilisateurs, je met de l'avant une compréhension relationnelle du processus par lequel les humains et les machines sont inter-reliés dans leur devenir. C'est dans cette optique que je redéfinis notre relation avec les machines au-delà de la simple instrumentalité, vers une forme de camaraderie mutuelle. Finalement, j'analyse les effets de certaines fonctionnalités du robot afin d'imaginer comment la séparation entre les humains et les machines pourrait être reconfigurée d'une façon plus responsable. Cela me permet entre autres de re-contextualiser les relations de travail encodées dans le développement de tels produits. Je termine en soulevant la pertinence de ces résultats pour les secteurs du tourisme et du développement de produit.

Mots-clés: Post-humanisme critique; intelligence artificielle; assistant virtuel; tourisme; étude des sciences et technologies; relations humain-machines; interface vocal; assemblage socio-matériel.

Acknowledgments

I would like to thank my supervisor Jakob Löfgren. Without his re-assuring words and creative liberty, this project would have never taken shape. I would also like to acknowledge Prof. Ingrid Hall from the University of Montréal, whose seminar on Ontology, Politics, and Environment provided me with the intellectual tools to delve into the fascinating world of critical posthumanism. Special thanks to everyone in Montréal and in Sweden who made this experience possible. Most importantly, thank you to my parents and my sister, the best role models one could ever wish for. Your unconditional support and love allowed me to get through this.

Notre-Dame-du-Portage, Summer 2020.

Malik Dehbi Talbot

TABLE OF CONTENTS

ABSTRACT	II
RÉSUMÉ	III
ACKNOWLEDGMENTS	IV
INTRODUCTION	1
SETTING THE STAGE	1
AIMS AND RESEARCH QUESTION.....	3
<i>Structure of the thesis</i>	<i>3</i>
LITERATURE REVIEW AND THEORETICAL FRAMEWORK.....	5
<i>From a Structural Ontology to a Sociotechnical Ontology of tourism</i>	<i>5</i>
<i>Queering the (Post)humanities.....</i>	<i>7</i>
<i>Companion Others</i>	<i>9</i>
<i>Posthuman Subjectivities.....</i>	<i>11</i>
<i>Human-Robot Companionship in Context.....</i>	<i>13</i>
NEGOTIATING METHODS AND CONTEXTUALIZING THE FIELD.....	14
<i>Being in technoscience.....</i>	<i>14</i>
<i>Rethinking autoethnography and participant observation.....</i>	<i>16</i>
<i>Material.....</i>	<i>18</i>
<i>Ethical considerations.....</i>	<i>20</i>
I - HUMANS MAKING ROBOTS.....	23
<i>The technology of the voice.....</i>	<i>25</i>
<i>Subjects / objects of servitude.....</i>	<i>29</i>
<i>Beyond the Great divide.....</i>	<i>33</i>
II - THE VOICEBOT COMPANION.....	36
<i>Re-thinking the demonstration</i>	<i>37</i>
<i>When the demo does not work.....</i>	<i>41</i>
<i>Making room for companionship.....</i>	<i>47</i>
III - OF ACCENTS AND TONES	51
<i>Responsibility and accountability.....</i>	<i>51</i>
<i>Accents.....</i>	<i>53</i>
<i>The tone</i>	<i>56</i>
<i>Responsible companionship.....</i>	<i>58</i>
CONCLUSION	61
<i>Applicability.....</i>	<i>62</i>

Introduction

SETTING THE STAGE

For many of us, the names Alexa, Siri, Cortana or Google sound uncannily familiar; they are all associated with software agents enabled by artificial intelligence tools animating the vocal interfaces on our technological devices, cars, or even homes. They appeal to a certain social imaginary, according to which these virtual assistants hidden in the backdrop of our lives can be magically activated by a simple utterance, whether it is to set an alarm, ask them about the weather or book a table for us at our favorite restaurant. Similar visions of human-machine interactions proliferate in popular media and science-fiction account exploring the boundaries of our relationship with technology. For example, the 1968 movie *2001* places HAL 900, a smart computer that interacts vocally and represented as a red circle of light, at the command of the Discovery One spaceship (Kubrick & Clark, 1968). Depicted as the unfriendly character whose dysfunctionalities are responsible for many misadventures, HAL has driven both “fantasies of the sociable machine” (Suchman, 2007, p. 235) and popular fears about the dangers of A.I. technologies (Kember, 2003). More than 40 years later, these visions are reinstated in Spike Jonze’s science fiction drama *Her*, which depicts the story of Theodore, a lonely and recently-divorced bachelor who falls fatally in love with its virtual assistant, Samatha (Jonze, 2013). The film probes the possibilities for love to transcend the boundaries between humans and machines, beyond a purely mechanical and deterministic understanding of emotional responses (Kubes, 2019A).

Although articulated in different socio historical contexts, these visions are all elaborated around the same fact. Namely, that as creators of intelligent artifacts, we build these agents on the premises that we can foster the same social relationships we maintain with our human counterparts (Suchman, 2007). Thus, the question arises as to how intelligent should we expect these agents to be and on what basis do we understand this intelligence? Who is to be held responsible when they fail to display the intelligence we ascribe them? It is with these questions in mind that I myself decided to take part in yet another vision involving an AI-powered virtual assistant. However, this vision is not projected on movie screens; it is rather enlivened in the streets of Montréal and on tourists’ smartphones, in a technological solution sprouting from two of the city’s booming sectors, namely A.I and tourism. More precisely, in the summer and fall of 2019, I conducted research and collaborated with a start-up specialized in the creation of conversational experiences for tourist destinations through tailored AI-powered virtual

assistants. Still in a stage of infancy at the time of my research, their latest project was called the Voice of Montréal [*la Voix de Montréal*], envisioned as “the very first AI-driven hybrid (image and voice) assistant designed to help travellers discover Montréal.” (Fieldnotes, 17/09/19B) Much more than an interactive city guide, the Voice of Montréal would literally give voice to the city by offering tourists with personalized recommendations from a local’s perspective through a conversational experience.

As a young researcher in applied cultural analysis, this concept promised a lot of interesting avenues, considering my interest in applying anthropological perspectives on technology and product development in a variety of contexts beyond academia. It is in this context that I entered a tradition of scholars in critical posthumanism and science and technology studies who, from the standpoint of their situated perspectives, explore the links between science, technology, politics, and culture. Thus, this thesis tells the story of an ethnographer in the making immersed in the development of a technological solution for the sector of tourism, and just like any stories, it took many unexpected turns. While I was initially expecting to conduct user research for the Voice of Montréal and subsequently use this material to investigate the possibilities for users to establish an emotional connection with intelligent software agents, I realised that my own vision of AI was biased. Indeed, spending 5 months collaborating with the start-up in the creation of the voicebot led me to re-think my most basic assumptions about humans, machines, and the boundaries that Western modernity and Cartesian dualism hold between these categories. Through my close engagement with the voicebot, I realised that it is much more than an interactive and human-like city guide; it is the nexus of various expertise, material components, as well as assumptions about the nature of intelligence. Most importantly, I realised that the voicebot was much more than a mere instrument, but rather a full-fledged entity, whose progressive development as an ever more complex conversational agent turned me into a more confident user researcher and ethnographer, and vice-versa.

It is precisely on this co-constitutive link between the voicebot and I that this thesis will dwell, in order to reveal the possibilities for talking about A.I. and intelligent artifacts differently; not as an unavoidable force of progress that will annihilate humanity or salvage it, or as the witness of the superiority of our species. Rather, this thesis will rethink our relationship with the Siri, Alexa, Cortana, Google, HAL 900, or Samantha of this world, be they fictional or not, in a way that underlines the potential of technological artifacts to shape reality in a better way.

AIMS AND RESEARCH QUESTION

The aim of this thesis is to look into the process of creation of AI-based virtual assistants and voicebots through the lens of critical posthumanism's notions of companionship. By examining the process of the creation of Bonjour Montreal, a tourist virtual assistant created with the purpose of strengthening the bond between visitors and the city of Montreal; and an auto-ethnographic re-reading of my fieldwork conducted over the summer and fall of 2019, this thesis will explore the different possibilities for defining bonds between humans and voicebots. More precisely, through a closer look at my own relationship with the voicebot, I will not only illuminate how a bond is established, but also shed light on the different types of material-discursive agencies it re-configures. By re-imagining my relationship to the voicebot according to the non-dualistic framework of posthumanism, I will open-up new avenues for a constellation of human-machine companionships that transcends hierarchical understandings of the boundaries between humans and machines. I will do so by answering the following questions:

- What are the different bonds envisioned in the creation process of the Bonjour Montréal voicebot? How do these bonds intersect with a naturalist ontology and its conceptions of humans and machines?
- As an ethnographer and user researcher, how did my own relationship with the voicebot developed? In what ways are we companions?
- What were the effects of the encounters between users and the voicebot and what categories did it (re)configure? How does conceiving the voicebot as a companion help us practically engage with the possibilities for re-configuring the boundaries between humans and machines in a more responsible way?

Structure of the thesis

Half-human, half-machine, straddling the sectors of tourism and technology and mobilizing a great number of different actors, the Voice of Montreal is an ambiguous subject / object that is hard to frame. It seems like the benchmarks of the anthropological discipline, which traditionally focus our attention on the purely symbolic aspects of the human experience, might not be enough to account for the complexity of our engagements with intelligent artifacts (Kohn, 2015). How can we understand bond formation and kinship ties to material things, non-human beings, or anything that falls in between? The different parts of this thesis can be seen

as a composite of theoretical, methodological, and empirical material that addresses this overarching question.

The remaining of the introduction will explore different theoretical avenues and review the potential of their contribution for this thesis. Through these literature reviews and theoretical explorations, my aim is to assemble a patchwork of perspectives that will help me make sense of the ambiguity that new forms of subjectivities such as the Voice of Montreal represent. Taking as our starting point the humanities' canon in tourism studies, this inquiry will take us to sociotechnical perspectives on tourism before meandering towards the territories of critical posthumanities. By expanding on the concepts of queer companionship and posthuman performativity, the end of this section will draw the contours of a framework for understanding what happens when I, a researcher in cultural analysis, meet and foster a relationship with a virtual assistant in the making. It is important to note that this literary review is by no means exhaustive nor does it cover the vast array of rich contributions to the fields of tourism, STS studies, or critical posthumanities. The following discussion is rather based on the theoretical encounters I made along my research journey that impacted my understanding of the subject. Finally, the last section of the introduction will allow me to explore how the posthuman framework forces us to negotiate methods, contextualise my experience in the field, and present the material this thesis will analyse.

The three analysis chapters will explore different possibilities for understanding the relationship between the prototype of the voicebot and the humans involved in its creation. The first analytical chapter will look at the different bonds envisioned by the start-up in the creation process of the voicebot and assess how these bonds intersect with naturalist understandings of humans and machines. The second analytical chapter will look deeper into the development of my own relationship with the voicebot and analyse how users' demonstrations and dysfunctionalities force us to rethink the status of the voicebot from a human-like machine to a socio-material assemblage, which will allow me to argue for the co-constitutive nature of our relationship. The third chapter will delve into the effects of the voicebot on users and analyse how specific features, such as the accent of the voice and its different tones, re-configure the boundaries between humans and voicebots in specific ways. In turn, the notion of human-machine companionship will act as a point of departure to imagine the possibilities for re-configuring these boundaries in a more responsible ways, and issue avenues for recommendations for the design of the voicebot. Finally, I will conclude this thesis by outlining the societal relevance of my findings for the fields of product development and innovation as well as the tourism sector.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

From a Structural Ontology to a Sociotechnical Ontology of tourism

Modern cultural anthropological theory as defined by Franz Boas or Émile Durkheim, amongst others, understands reality as a socially constructed phenomenon and the product of specific historical, social, cultural and linguistic contexts (Kohn, 2015). Their view is illustrative of humanities' narrow focus on the human, according to which culture is the only important variable of the human experience, as opposed to nature and the material world, which are perceived as stable, constant and external objects (Asberg & Braidotti, 2018).

Similarly, social and cultural analyses of tourism have traditionally associated tourism as a cultural object accessed by human subjects, in the form of physical spaces that break away from the routine of everyday life (Franklin, 2004). Dean MacCannell's *The tourist* published in 1976 offered one of the first account of tourism in the field of social theory by understanding the tourist experience as the search for authenticity in the face of the dulling threat of the modern lifestyle (MacCannell, 1976). His approach, focused on symbolic interactionism, argues that the tourist experience is organized around the binaries of the ordinary and the extraordinary, which are illustrative of how "modernity reworks binaries deep within the human condition." (Franklin, 2004, p.283) Another influential account of tourism is John Urry's *The tourist gaze*, which focuses on the organization of tourist sites through representational practices such as photography of sightseeing, hereby understanding the tourism experience mostly as a visual phenomenon (Franklin, 2004; Ek et al., 2008). These ideas have much in common with the modern anthropological theory highlighted above, according to which the human experience is the result of our own symbolic forms of representations, such as language, which understand these binaries as the basic unit of analysis. The later "performative turn" in tourism studies slightly deviates from these constructivist and structuralist accounts, by acknowledging that tourists, beyond simple observers, are "actively, corporeally, technically, and socially" participating in the performance and production of the spaces they visit (Ek & al., 2008, p.123; Löfgren, 1999). While these contributions slightly differ in the relative focus they put on tourists' level of agency, amongst other things, Adrian Franklin argues that they all operate according to a binary opposing the ordinary to the extraordinary that reproduces "tourism as a social spatial enclave," without questioning this assumption, and therefore taking it as an essential characteristic of tourism (p. 278). He argues that "we seem to have a structural account based on a relationship between the everyday and the extraordinary, which take their very existence, their juxtaposition for granted, as already explained - particularly the *extraordinary*

ordered as tourism.” (p. 283) He adds that while binary conceptions of the human experience can be useful to account for the fact that, indeed, people leave their home in the search for extraordinary experiences, these accounts offer little insights as to “how the binary becomes operationalized.” (p. 283) In other words, tourism theory needs more account for the *how* of tourism rather than the *why* by shifting its focus from a structural ontology to a sociotechnical ontology.

Borrowed from Science and Technology Studies (STS), the term sociotechnical ontology refers to an understanding of reality according to which the sociocultural world and the natural, scientific or technical world constantly define each other, as opposed to two separate realms (Treusch, 2015). This offers a counterpoint to normative systems such as constructivism, structuralism, or naturalism, to name a few, which assume an ontological divide between nature and culture and presuppose human exceptionalism, on the grounds of the alleged unique interiority of our species (Kohn, 2015). On the one side of this divide, humans are seen as complex subjects evolving in a constructed web of meaning, which can be uncovered and understood by social and cultural theories. On the other side, the natural and material world is seen as composed of objects or animals deprived of any agency who are only knowable through the rational enterprises of technoscience. In contrast, sociotechnical perspectives call for greater attention to the entanglements of material practices with sociohistorical processes and the blurring of the disciplines of science and technology and humanities. Several contributions in the field of tourism have taken this turn in the recent years, hereby extending understandings of tourism beyond a purely sociocultural activity where meaning is only produced and extracted by human agents. Rather, tourism is seen as a network of management practices, innovations initiatives, or modes of ordering linking both human and non-human actors, all contributing to the enactment of reality (van der Druim et. al., 2007; 2017). In this sense, entities such as “sheep cheese [...], cigars [...], money [...], and road and ruins [...] come to the fore as actors involved in how tourism is enabled and enacted.” (van der Druim et. al., 2017, p.141) This perspective allows us to see how tourism practices contribute to the making of the world, rather than simply being a product of the state of the world. For example, Franklin’s analysis of the early days of modern tourism in England reveals that tourism did not emerge as consequence of modernity, but rather contributed to the establishment of the modern project of nation-state formation (Franklin, 2004). More precisely, he looks at the enterprise of British businessman Thomas Cook, who in the context of a spreading nationalist ideology in 19th century Europe, dreamed of making accessible “the cultural, historic and natural *content* of the nation” accessible to millions of people (p.289, emphasis original). This ideology, combined with an increasingly

interconnected transportation infrastructure and the perfecting of managerial practices, contributed to the development of modern the tourism industry and, by the same token, the formation of the modern nation state. Thus, tourism is understood as a conglomerate of material practices emerging of a precise socio historical context that have very tangible consequences on the world.

These considerations give us a glimpse of the potential of combining different perspectives cutting across the disciplines of tourism studies and STS. In the scope of this thesis, turning our attention to the “material heterogeneity” of the world and ensuing practices such as innovation, design, and product development proves to be a fruitful way to understand how the Voice of Montreal “ is set to change the way visitors interact with the city.” (Tourism Montréal, 2019) Further, behind the idea of giving voice to the city is the start-up’s desire to use artificial intelligence and natural language processing (NLP) tools to create emotional experiences for visitors, something they often referred to as “emotional intelligence.” (Field notes, 2019) Thus, more than the introduction of new types of tourist subjectivities, the apparition of a new actor in the tourism landscape hints to the emergence of new types of affective connections between users and virtual assistants, who are now imagined as travel companions. As the rest of this thesis will elaborate upon, my positionality within the organization as an ethnographer and user researcher gave me an early and privileged access to the voicebot. This encounter proved to be insightful to understand the process by which new relationships with technological others are created and fostered.

Queering the (Post)humanities

In order to answer these questions, it is worth turning our attention to the post-humanist turn that has taken place in the humanities since the 1980s. According to philosopher Francesca Ferrando, posthuman theory calls for “an integral redefinition of the notion of the human,” in the face of the obvious exceptionalism and anthropocentrism that characterizes mainstream understandings of the human experience (Ferrando, 2014, p. 26). Critical posthumanism acts as a counterpoint to this polarization by re-evaluating the ontological premises that define what counts as a subject or as an object and broadening the scope of humanities’ analytical gaze beyond the human. Ferrando highlights how technological development is a central aspect in the posthumanist enterprise. She writes that “posthumanism investigates technology precisely as a mode of revealing, thus re-accessing its ontological significance in a contemporary setting where technology has been mostly reduced to its technical endeavors.” (Ibid, p. 29) This

perspective forces us to consider technology not as a tool that accounts for the superiority of our species, but rather as a category carrying the potential to help us redefine what it means to be humans. At the same time, it allows us to revise our most basic assumptions about the types of beings with whom we share our everyday life, our relationship with them, and their impact on our common future on this planet.

In many ways, this analytical turn owes its openness to the contributions of feminist and queer theorists who, in their plight of critically disentangling the hegemonic links between knowledge and power, have come up with new ways of accounting for the entanglements of humans with other types of beings (Asberg & Braidotti, 2018). This thesis will mainly focus on feminist and queer posthumanities, also often referred to as critical posthumanism, for their potential to reframe our relationship with technoscience in a critical way. Critical posthumanism aims at dismantling human exceptionalism, at the expense of an “unruly worldliness” that accounts for the multiple ways in which human lives are entangled with the non-human material world. This encompasses any type of “lively matter” such as “people, technoscience, global media, biotics, ecologies, animals, finance, land,” and so on (*Ibid*, p 13). This should not be interpreted as the neglect of our species’ well-being, but rather as part of the larger project of imagining a better future where all types of beings - be they human or not, organic or not - can coexist (*Ibid*). Citing Haraway, Treusch adds that the act of “queering” research can be defined as the “project of undoing taken-for-granted norms,” such as the heteronormative matrix through which technoscience is practiced and understood (2015, 24). More precisely, this approach critiques the assumed objectivity of Western science and technology, which, as Donna Haraway argues, operates from the assumption that the observer, a White Man, defines the norm by detaching himself from the world and its representation. She writes that “[t]his gaze signifies the unmarked positions of Man and White, one of the many nasty tones of the word “objectivity” to feminist ears in scientific and technological, late-industrial, militarized, racist, and male-dominant societies [...]” (1988, p. 581) Haraway urges us to flip of the coin by embracing the ‘privilege’ that the inescapable partiality of knowledge and situated positionalities offer. This position holds a lot of potential to tell stories left behind by the patriarchal objective gaze. More precisely, the strategy of seeing things from within allows us to pay greater attention to how our daily lives are enmeshed with those of other types of beings, be they viruses, dogs, or robots. While these entities have long been ignored as legitimate ‘objects’ of knowledge in the humanities, critical posthumanism takes our mushy involvements with them as sites of inquiry to understand how essential categories such as the “human” and the “non-human” are produced. I take this as a departure point in my quest to

understand the process by which relationships and bonds are formed between an intelligent artefact and its user.

Companion Others

Critical posthumanism provide us with the tools to critically disentangle the strong links that Cartesian dualism holds on our grasps with non-human beings and the material world. It breaks with the essentializing mind v.s matter logic that makes us see the human experience as virtually disengaged with the rest of the world and that delimitates its constituents into stable and fixed categories, without accounting for the multiple ways in which they influence each other. For example, scholars such as Donna Haraway and Karen Barad take the relationship as the smallest unit of analysis for understanding the world and the emergence of things (Barad, 2003 & 2007; Haraway, 1991, 2003 & 2008). This means that nothing exists *as such* in the world; categories such as “human” or “virtual assistant” only come about through the relations that their constituents cultivate with each other.

Feminist Anthropologist Donna Haraway famously traces the entanglements of our relationships with other species and technology. Brilliantly blurring the genres of science fiction and academic essay, her writings depict a world where women are intergalactic cyborgs and where dogs and humans are co-evolutionary companions, who are all part of the same “queer family” re-appropriating technology in the fight against capitalism, sexism, and racism. This vision offers a counter narrative to the heteropatriarchal white hegemony, which governs our engagement in the world, and forces us to trace clear-cut boundaries between categories such as; Human vs Animal, Culture vs Nature, Organic vs technological, etc... (2003, p. 4) Her *Companion Species Manifesto* provides us with an insightful framework to make sense of the bonds that link us with beings that are ontologically different from us, yet key actors in our making as humans. Putting aside the figure of the cyborg she famously appropriated in her *Cyborg Manifesto*, she focuses on so-called “Man’s best friend,” namely dogs, to argue against the idea that humans are superior to dogs as we engage in the patriarchal act of domestication and taming of things of the wild natural world. In a nutshell, she writes against the idea that “man took the (free) wolf and made the (servant) dog and so made civilization possible.” (p. 28) Haraway explains how this view permeates our relationship to dogs in many ways, from the technoscientific accounts of our co-evolution stipulating that dog is the first domesticated animal during the larger project of the establishment of Human civilization, to their representation in popular culture as the providers of unconditional love. In our normative

Western technoscientific system, dogs are understood only as means to the end of the great enterprise of human superiority. When we did not have enough of ourselves to assert our domination to nature, dogs were there to assist. When our human peers fail to love us, dogs do so unconditionally. Haraway criticizes this idea that dogs are these animal Others that act as gatekeepers of the boundaries between nature and culture. She writes that “[...] if the idea that man makes himself by realizing his intentions on his tools, such as domestic animals (dogs) and computers (cyborgs), is evidence of a neurosis that I call humanist technophilic narcissism.” (p. 33) She further explains that dogs are not a mirror to our humanity, but rather active companions in the co-constitution of our species by suggesting a relational ontology in which the “relationship becomes the smallest unit of analysis” for understanding ourselves and others (p. 20). All types of beings and entities, organic or not, animate or not, engage with each other and come to define each other in that very process.

Thus, she gets at an understanding of multispecies relationships in which “the partners do not precede the meetings.” (2008, p.4) This means that the two parts that constitute the relationship only come to being once they enter in the said relationship. However, it is important to note that a relational ontology does not erase the ontological barriers between different types of beings. Rather, it helps us reconcile difference in a way that is ethical for all actors. Engaging in multispecies companionship is all about grasping the relationship that the significant otherness produces, not getting rid of the difference. For the dogs and their humans, this means engaging in an “ontological choreography,” in which training becomes an act of negotiation where both actors hone their skills, in a way that ensures the full potential of the relationship. She writes that:

Much companion animal talent can only come to fruition in the relational work of training. Following Aristotle, Hearne argues that this happiness is fundamentally about an ethics committed to “getting it right,” to the satisfaction of achievement. A dog and handler discover happiness together in the labor of training. That is an example of emergent naturecultures. (p. 52)

It is in this context that Haraway talks about “fictive kin group in training” to describe the companionship between dogs and their humans. She talks of a “positive bondage” between her dogs, who she calls her “bitch,” and herself, who she refers to as the “handler.” According to Haraway, this perspective has the power to “shape a potent worldly consciousness in solidarity with my feminist, anti-racist, queer, and socialist comrades; that is, with the imagined community that can only be known through the negative way of naming, like all the ultimate hopes.” (p. 64) And this is where we go back to the idea of a queer family, in which kinship is

not determined by patriarchal precepts of biological bloodlink, human exclusivity, or civilizational superiority. Rather, it is defined by the potential underlying every relationship uniting different beings, regardless of their otherness. It is companionship.

I will use Haraway's account of multispecies companionship to understand my relationship with the prototype of the voicebot. While she bases her analysis on "historically situated animals in companionate" such as dogs, cats or tigers, she emphasizes that the term companion species "is less a category than an ongoing *becoming with*" (Haraway, 2008, p.16, emphasis added). Consequently, taking as its starting point the encounter between the Voice of Montreal, a virtual assistant in the making, and myself, a researcher in the becoming, this thesis tells the story of a *becoming with*, in which both parties played a co-constitutive role. While Haraway's account of co-constitutive companionship suggests a relational ontology according to which boundaries between different types of being are constantly shifting and being (re)negotiated, one might wonder what stabilizes these boundaries, as the different 'constituents' of a relationship often appear to us as distinctively different. In other words, why can we talk, for example, of humans and virtual assistants as two distinctive normative entities, when a relational ontology suggests that they constantly remake each other? In order to shed light on these gaps the rest of this section will turn to posthuman understandings of subjectivity and agency.

Posthuman Subjectivities

Feminist theorist Karen Barad engages with the notion of co-constitutive relationships in her theory of posthuman performativity and agential realism (Barad, 2003). Her approach is notably useful to understand how, despite their unstable and co-constitutive character, categories such as the "human" and the "non-human" come to be stabilized and perceived as distinctive. More precisely, she criticizes the representationalist turn in social and cultural theory and urges us to pay greater attention to matter, as opposed to language or any other form of symbolic representation, to understand our position in the world. I will summarize her ideas and illustrate how they inform my understanding of posthuman companionship.

Central to Barad's argument is the notion of intra-activity, which builds upon post-constructivist and feminist understandings of performativity. These accounts understand the becoming of individual subjects in terms of performativity, conceived as the bodily enactment of commonly-shared symbolic constructs. Feminist theorist Judith Butler's theory of gender as the performance of a stylized repetition of symbolic acts is the prime example for this (Barad,

2003). This framework assumes an ontological divide between discursive practices and material phenomena, where the latter only serves as material support for the symbolic representations of the former. The concept of intra-activity argues against the inherent properties of matter or meanings in words, which therefore cannot be taken as “ontological basic entities” for understanding the world (2003, p.813). Rather, in order to get to a primary unit of analysis for what makes things, people, and the world around us, she borrows quantum Physicist Niels Bohr’s concept of “phenomena,” understood here as “ontologically primitive relations - relations without preexisting relata” (Barad, p.815). In other words, things are made of inseparable “intra-acting components,” not separate entities (Barad, p.815). According to Barad, the concept of intra-action is central to understand how things come to be seen as distinct entities despite their relational character. She writes:

The notion of intra-action (in contrast with the usual ‘interaction,’ which presumes the prior existence of independent entities/relata), represents a profound conceptual shift. It is through specific agential intra-actions that the boundaries and properties of the ‘components’ of phenomena become determinate and that particular embodied concepts become meaningful. A specific intra-action [...] enacts an *agential cut* [...] effecting a separation between ‘subject’ and ‘object’. (Barad, p.815)

The quantum model of the atom elaborated by Bohr states that matter at its most simple form cannot be understood according to the “atomistic metaphysics that take things as ontologically basic entities” (p. 813). Rather, the basic unit of analysis for understanding matter should be the phenomena, or the relationship, in which the boundaries between the components are drawn as they interact, or more rightly *intra-act*. This is what Barad calls the ‘agential cut.’ It is precisely in that process of intra-activity that things *become*. If such is the way matter comes to matter according to quantum physics, Barad argues that the process applies at a larger scale for our understanding of the world and its categories in general.

In the light of this account, I am able to add a few specifications in my outline of a posthuman companionship. As I explained, Barad’s framework “challenges the positioning of materiality as either a given or a mere effect of human agency.” (p.827) Rather, discursive practices and material phenomena are one entangled process, as the result of which things come to be. In other words, this framework illustrates how both discourses and material phenomena equally contribute in the shaping of the material world and sociohistorical processes. Consequently, agency is not an essential quality attached to something/someone but the possibilities that the flow of intra-activity in a phenomena enable. The phenomena, or the relationship between

different components, produces specific material-discursive forms of agencies that stabilize different categories.

Human-Robot Companionship in Context

This brings me to a more complete portrait of what I call posthuman companionship. This outline of a theoretical patchwork lays the base ground for an understanding of my relationship with the Voice of Montreal as a co-constitutive companionship. In this sense, this thesis will not only look into what type of companionship such a relationship can create, but also what forms of material-discursive agencies it ensues. For example, on the one side it defines not only what it means for me to be human, but also an ethnographer in a tech start-up. On the other side, the companionship also stabilizes the voicebot in a particular form. This process of *becoming with* is enacted through various material-discursive practices, which this thesis will explore more profoundly. Moreover, as per the feminist-queer posthumanist perspective wants it, this theoretical patchwork allows us to account for the possibilities of de-stabilizing certain categories in the scope of creating a better future in which all beings can co-exist.

To frame my approach as completely novel for the study of human-robot companionship would be to ignore the amazing contributions of scholars who, besides helping me put together the different pieces of my theoretical patchwork, also contributed to the development of critical posthumanities and the fields of human-computer interaction (HCI), robotics, artificial intelligence (AI), and more. Indeed, I am intellectually indebted to scholars such as Lucy Suchman, who in her seminal book *Plans and situated actions* blurred the ontological boundaries between humans and intelligent machines (1987). Her more recent work focuses on how the boundaries between humanoid robots and their human creators are materially enacted and stabilize naturalistic understandings of the human while being dependent on human labor (2011). Pat Treusch engages with similar topics in her PhD dissertation titled *Robotic Companionship* for which she conducted fieldwork in a research laboratory developing domestic robots for kitchen (2015). She focuses more precisely on the public testings of the robot and analyses how it (re)defines what it means to be human, notions of humanness, the performance of care and companionship and how these aspects relate to notions of gender, nationalism, and sexual reproduction. Her approach is notable for the attention she pays to sociotechnical imaginaries, namely the popular media and fictional accounts of our potential present and future with robots. She joins her voice to many feminist-queer scholars of the posthumanities who argue that the imaginary must be taken into account because they are

entangled in the discursive-material practices that drive sociotechnical change on the same basis as technological development does (Asber & Braidotti, 2018). More generally, the substantial contributions of anthropologists and cultural critics on the important role of robots and AI in our sociotechnical imaginary is telling of the undeniable inseparability of imaginaries from the material world and of the importance of including them in our analyses (Gibson, 2020; Lyon, 2018).

While my analysis will inevitably engage with all those themes, my preliminary literature review reveals that there are still many unaddressed gaps in cultural analytical research on robots and AI. Indeed, social and/or cultural analyses of artificial intelligence and robots mostly focus on embodied robots used in the context of care, domestic labor, or sex work (Kubes, 2019B; Richardson, 2015; Treusch, 2015). This thesis will broaden this focus by including new perspectives on AI and robots using on natural language processing tools and vocal interfaces in the context of the service and tourism industry, terrains who remain unexplored. Moreover, my focus on the co-constitutive relationship with the voicebot that my dual role as an ethnographer and user researcher engendered will bring interesting perspectives not only about the boundaries between humans and machine, but also about the boundaries of our discipline of cultural analysis as it engages both with science and technology studies and tourism studies.

NEGOTIATING METHODS AND CONTEXTUALIZING THE FIELD

Re-thinking more-than-human companionship and processes of becoming beyond the precepts of a humanist and naturalist ontology is no simple task. To quote Asberg and Braidotti, my approach thinks of critical posthumanism as the “imperfect praxis” of redefining the human in the light of our entanglements with non-human and technological others (Asberg & Braidotti, 2018, p. 16) It is imperfect because, as my experience shows, *becoming with* more-than-human companions requires first hand engagements that follow no precise trajectory. This section nevertheless tries to trace these trajectories through an experimental composite of approaches that can sometimes seem contradictory. In what follows, I will explore how a posthuman methodology forces us to rethink concepts such as objectivity and positionality, autoethnography, participant observation, and ethics. This will allow me to lay out my research material as well as the aims and questions that my analysis will address.

Being *in* technoscience

Although I present the theoretical and methodological background of this thesis under two separate headings, I argue against a clear demarcation between the realms of knowing and

doing. Indeed, critical feminist contributions to STS commit to an understanding of theory as an active experimentation and performance, rather than abstract manifestations of the world (Barad, 2012; Asberg & Braidotti, 2018; Rootsh & Schrader, 2012). This means that as researchers engaging with science and technology, our work should not only talk about it, but also be “in” and “of” it (Rootsh & Schrader, 2012, p. 2).

My research falls in line with this perspective, as I spent more than five months over the summer and fall of 2019 collaborating with the tech start-up creating the Voice of Montreal. The initial intent upon collaborating with this organisation was to investigate the process by which an affective relationship develops between users and a voicebot fitted with so-called “emotional intelligence” and the impact of the introduction of this new agent in the tourism landscape. As both a master’s student in the field of applied cultural analysis and an intern at the start-up, my research process not only had to follow the precepts of academic qualitative research and its ethical implications, but also produce an output that would be of use for the start-up and contribute to the product development process. Consequently, I simultaneously occupied the position of user researcher helping with various tasks related to the user experience (UX) side of product development, which were mostly focused on conducting user testing at various tourist locations, in order to get a preliminary feedback of user interaction with the voicebot. This dual role put me in the middle of what Sociologist of science John Law calls research “networks of anxieties,” which he describes as situations in which ethnographers in organizations such as laboratories are (perhaps too) “committed to a perfect version of representational ethnography.” (1994, p.46) Amongst other things, these anxieties were mostly associated with the difficulty of obtaining coherent data on user interaction with a preliminary version of an often faulty voicebot. Moreover, these uncertainties led to many frustrations for what I thought was my inability to adequately represent users’ reality and have their voices properly heard. To go back to Law’s words, “I was committed to a version of pure order [...]. And I didn’t see that it was impossible. So I panicked at what I took to be the shoddiness of my fieldwork (*ibid*).”

Our inevitable enmeshment in “networks of anxieties” materializes the inability to achieve so-called ‘scientific objectivity’, or what Haraway calls “the conquering gaze from nowhere.” (1988, p. 581) Exposing the deeply-rooted ties of this scientific gaze to “militarism, capitalism, colonialism, and male supremacy” and its obsession with “distan[c]ing the knowing subject from everybody and everything in the interest of unfettered power,” she argues for a feminist engagement with science from a situated, embodied, and partial perspective (*ibid*). Looking at technoscience from within allows us to elaborate critical perspectives that leave behind the ‘all-

knowing gaze,’ and instead account for the “care, affect and responsibility” that emanates from technoscientific practices (Roosth & Schrader, 2012, p. 3). As I mentioned, dealing with the pressure of the objective gaze was not easy task during my experience on the field. However, my approach in this thesis embraces the privilege of partial perspective that my positionality as a researcher within the organization granted me. In light of these perspectives and along with the feminist-queer posthuman practice, my methodology becomes one of storytelling recounting my engagements with the technoscientific practices that contributed in the configuration of the voicebot (Asberg & Braidotti, 2018). As the main narrator of this story, I side with Haraway who identifies the laboratory as a “world-building space” worthy of investigation for researchers of the posthumanities. More precisely, my privileged partial perspective within the MTLab, the business accelerator and innovation laboratory hosting the start-up, as well as within the peripheral locations where my user testing sessions took me, becomes the starting point of my account. It is through my active engagement within these sites and with the people with whom I collaborated that the Voice of Montreal uttered its first babbling, slowly developed into an ever more eloquent conversational agent, and most importantly became a companion to my everyday life. Thus, I trace in this thesis the genesis of my engagement with the voicebot, which will hopefully allow me to disentangle the knots of care, affect, and responsibility that go into the making of technoscience and deeply impact the world around us.

Rethinking autoethnography and participant observation

The investigation of various forms of *becoming with* the Voice of Montreal from my partial, embodied, and situated perspective comes with the practical challenge of making sense of my own experience through the appropriate methodology. While the use of autoethnography seems appropriate for the purpose of this thesis and the handling of my material could be labelled as autoethnography, I explain in this subsection why I shy away from blindly committing to this method, as I challenge the approach with various critiques and perspectives. But let us first consider how autoethnography can be useful to make sense of my partial perspective.

In autoethnography, the researcher’s subjective experience and observations about their own self is the center of fieldwork, and it is analyzed to reveal the workings of broader sociocultural dynamics (Penseau-Conway et al., 2017; Rapp, 2018; Bodker & Chamberlain, 2016). In the context of HCI and ubiquitous computing research, researchers look at their own interaction and relationship with the computer to access the inner workings of user interaction, which

would otherwise be very hard and costly to account for (Rapp, 2018). Moreover, combined with perspectives such as affect theory, autoethnography is defined as a process of sensemaking of “the forces and intensities of feeling that shape our embodied capacity to act” and that pour through bodies, hereby acknowledging “that there is more at stake in being-with computers.” (Bødker & Chamberlain, 2016, p. 5) Through the crafting of narrative accounts and stories in the form of vignettes, for example, the autoethnographic process has the potential to reveal “an ontology of permeable, porous, and sensory bodies entangled with technologies.” (*ibid*, p.13) On the face of it, this approach seems appropriate to practically engage with the feminist-queer propositions of disentangling the material-discursive agencies at stake in our *becoming with* intelligent artifacts.

While these guidelines have concretely helped me make sense of my experience, it would be amiss to obliterate the concerns that prevent me from blindly committing to autoethnography. As Tim Ingold argues in an intentionally controversial article about the overuse and conflation of the term ethnography and its derivatives, I claim that we should understand (auto)ethnography as the “judgment that is cast upon [experiences and encounters on the field] through a retrospective conversion of the learning, remembering and note-taking [...]” (Ingold, 2014, p.386) In other words, the autoethnographic-ness is not intrinsic to my experience and my data, it rather arose during the posteriori process of sensemaking that this thesis is part of. Ingold further explains that, as researchers interested in uncovering “the conditions and possibilities of being human,” our role vis-a-vis the people and things we meet on the field implies a responsibility to turn our encounters into meaningful learning experiences, rather than a mere results-yielding practice that will be turned into research papers and monographs (*ibid*, p. 388).

As this thesis will elaborate upon, if there is one thing I can say I learnt from my experience, it is that the boundaries between a bounded human subject and a virtual assistant in the making are more porous than we think. In the face of this cognizance, how can I subscribe to understandings of autoethnography without falling in the humanist trap of symbolic sensemaking and distanced observation of the inner workings of my bounded, autobiographical self? This enterprise would certainly seem contradictory to this thesis’ intent at demonstrating the shifting boundaries between humans and technological artifacts. On the other hand, neither do I completely reject the possibility of formulating a critical posthumanist analysis from my position of human subject as it is configured by the sociocultural contingencies of Western thought and science (Meissner, 2016). Along with Hanna Meissner, I argue that we cannot completely reject the position of our human subjectivity and that we should make do with the

gaps and tensions arising from seemingly mutually exclusive frameworks (*ibid*). Referring to Barad's principle of intra-activity, she explains that we can only work from that position as we acknowledge that it is produced through our intra-actions within a phenomenon, which in turn "implies the exclusion of other possible and equally necessary phenomena." (*ibid*, p. 13) Thus, considering that no theory or methods could possibly be comprehensive, "an interesting opening lies in the possibility that different, even contradictory, points of view can be considered as equally possible, or equally necessary when siting/sighting apparatuses that constitute and differentiate particular forms of agency - while marginalizing or excluding other forms" (*ibid*). In other words, my methodological approach explores what can arise from the juxtaposition of the contradictory posthuman framework with an autoethnographic account formulated from the position of a human subject.

Finally, this account is based on more than five months of participant observation at the start-up during which I occupied the position of user researcher helping with various tasks related to the user experience (UX) side of product development. But in light of the aforementioned critiques, what does it mean to do participant observation? To go back to Ingold's critique of ethnography, I will refrain from deeming my encounters and my experience on the field as ethnographic and rather bring back to the front anthropology's and cultural analysis' "tried and tested way of working, namely participant observation." (2014, p. 386-387) He goes against established understandings of participant observation as the interplay between the engagement of participation and the detachment of observation, which in his opinion would imply the "excision of knowing and being" and the ensuing objectification of the beings and things we work with (*ibid*, p. 387). Rather, he argues for a view of participant observation as the "ontological commitment" of "attending to persons and things, to learn from them, and to follow in precept and practice." (*ibid*) In other words, to do participant observation is to unleash and honor the potential knowledge that arises from living with others. This perfectly translates my experience at the start-up, as I was not just there observing, but became actually part of the team. It also guides my understanding of a posthuman ethics as being performed in praxis, which I will address later in this section.

Material

As mentioned above, my positionality as a student from a humanities department raised a lot of uncertainties about conducting ethnographic fieldwork for a young start-up in full structuration and grasping the complexity of a technological intelligent artifact. Being only in

its preliminary stages of development, the demos of the voicebot presented limited options of interactions, rather than the complete human-like conversational abilities and emotion-inducing storytelling features promised by the concept. Moreover, as it is the case with early versions of any technological artifact, the voicebot was often faulty and demonstrated bugs that rendered difficult the already limited interactions with users. What I perceived as a lack of substantial material proved to be quite frustrating and anxiety-inducing for me. How was I to conceptualize a robot's agency and its relationship with users when it could not interact with them?

Challenging commonplace epistemologies and ontologies in line with critical posthumanist propositions is helpful to make sense of the many tensions that arose during my research process. What I perceived as scattered field notes on bottled up emotions over an allegedly shabby fieldwork, now becomes rich material traces bearing witness to the many aspects involved in the configuring of the companionship between a human and a virtual assistant. The following pages are a reconstruction of my field notes and other material from the fields strung together in an autoethnographic narrative. This story is about many things. First and foremost, it is about my dual role as a cultural analyst and user researcher in a nascent tech start-up in the early stages of development of a conversational voicebot. During this time, I participated in and analyzed the encounters of users, clients and colleagues with this intelligent agent. But underneath this lies my own encounter with the robot, which I largely disregarded during my fieldwork. This thesis will take a closer look on the frustrations, uncertainties, and hopes I experienced while engaging with the voicebot and its ecosystem. So, where it was impossible to thoroughly investigate users' relationship with the voicebot, I also take my own affect-laden experience as a point of departure to re-think my relationship with the voicebot as co-constitutive of different discursive-material agencies. However, one thing must remain clear; this is not a story only about myself. I rather offer an account of my partial perspective as the nexus of many intra-actions configuring various material-discursive forms of agencies around the voicebot.

More precisely, my account pieces together various material traces remaining from my experience at the innovation lab and the orbiting locations where my research took me. This comprises notes about moments ranging from weekly meetings to brainstorming sessions and everyday life moments such as casual conversations between me and my colleagues. The team varied in size throughout my stay from five to ten members whose occupation ranged from co-founders, strategist, data architect, project manager, graduate students in computer science as well as a PhD student in communication and ethics. While all of them were aware of my research, I anonymize everyone in the text to protect their privacy. Moreover, my analysis

prioritizes moments and instances that relate to my contribution to the team and the product as a way to draw on the full potential of analysing my situated perspective and the unique insights it grants me for understanding my companionship with the voicebot. This leads me to include material gathered in the context of the various user testing sessions performed with more than 20 participants in various tourist locations, including two youth hostels and one tourist information center. During these sessions, participants were asked to interact with the voicebot and provide feedback on their experience, which I recorded in the form of written notes on my computer and screenshots from my phone, that I later summarized in various presentations to my team. All participants in user testing sessions were aware of my study and are anonymized in the text. Moreover, I cannot ignore all my fieldnotes entry dealing with my interrogations and frustrations over the numerous bugs that characterized the development of the voicebot. I consider this data not as representations of distanced observations, but as material traces bearing witness to my sensuous and affective engagements with technoscience in the making. I hereby sustain Barad's claim on the importance of paying greater attention to the intimacy at play in the "material-affective dimensions of doing and engaging science (Barad, 2007, p. 208)." In contrast with this, I also include in my analysis public demonstrations of the voicebot as well as official statements of the various organizations involved in the project. Overlapped with famous fictional accounts of voicebots mentioned in the introduction, these discursive evidences demonstrate that technological innovations cannot be separated from social, cultural, and economic dynamics as they display very distinct technology-driven visions of our present and future with voicebots and virtual assistants (Treusch, 2015). Along with many scholars of the posthuman turn, I understand these sociotechnical imaginaries as key drivers in the actualisation of novel forms of material-discursive agencies (Asberg & Braidotti, 2018; Harraway, 1991, Treusch, 2015).

Ethical considerations

Venturing in the territories of posthuman scholarship also involves a reconsideration of the concept of ethics. As I have explained above, the posthuman approach is an inherently ethical enterprise, with its aims to re-evaluate the allegedly essential and universal normative system according to which we define ourselves and the world around us. Considering the political and ethical load that these perspectives carry, it is important to make the following clarifications regarding my methodological and theoretical choices. First, my intent is not to appropriate the struggle of women working in science and academia, whose experiences allowed for the

elaboration of critical posthumanism. However, as a queer-identifying man of color, I found comfort in these theories, in a world that too often makes us feel “less than human” for not conforming to the norms (Ahmed, 2018; Hansson, 2019). I believe that my sexual and ethnic ambiguity have given me the tools to understand that normative categories are by no means essential or universal.

In a Swedish context, the disciplines of anthropology and cultural analysis operate according to the ethical standards of the American Anthropological Association (AAA) and of the CODEX, the Swedish Center for Research Ethics and Bioethics. Their guidelines bind me to the conduction of research within parameters that respect human dignity, rights, and basic freedom (CODEX, 2019). Moreover, the AAA emphasizes that researchers must “protect the safety, dignity, or privacy of research participants.” (AAA, 2012) These guidelines bind me to protect the privacy of my informants and, as per their request, I disclose their identity in my thesis as well as the name of the start-up and their clients.

Notwithstanding the importance of these guidelines, Hansson argues that they are elaborated according to a logic of universalism that presupposes a “shared understanding of abstract terms such as freedom and dignity,” and whose foundations prove to be shaky when applied to any given situated context (Hansson, 2019, p. 23). Similarly, I want to ask what happens to this anthropocentric and allegedly universal understanding of ethics once we acknowledge that the very categories upon which it is elaborated, such as ‘humans’ and ‘machines,’ do not hold up to the reality check of everyday experiences? Along with Meissner and Barad, I argue that ethics is to be found in the practical performance of accountability and responsibility, rather than in moral ideals and imperatives (Barad, 2012; Meissner, 2016). Concretely, this means that I am personally accountable for my research choices in the field and that, as researchers, we are not detached from the fields and people with which we work. This justifies my methodological choices to focus on my own experience as it allows me to ethically account for the ways in which my presence within the team impacted the subject of study of this thesis. Moreover, a posthuman account of ethics focussed on accountability and responsibility allows us, actors involved in the creation of complex technological artifacts, to understand where exactly lies our capacity for action within these assemblages. Since this aspect will be further elaborated in the third chapter, it suffices to say that this approach has an unavoidable ethical and political reach considering that it allows me to provide concrete recommendations for the development of a more ethical and responsible product. This perspective echoes Ingold’s conception of anthropological practices, which he views as a commitment to turn what we learn

from our experiences into teachings for the world (Ingold, 2018). In many ways, this is very similar to the practical and solution-driven approach of applied cultural analysis.

Finally, my discussion would not be complete without mentioning the importance of ethics it comes to the development of A.I. Indeed, the field has been under scrutiny regarding questions of protection of privacy and confidentiality of users' data. In the development of the voicebot, ethics was a central concern that the co-founders of the start-up held high and that was championed by one colleague, a PhD student in communication and ethics, whom I got to work really closely with because she was also performing some user research tasks besides her "privacy-by-design" project. As much as I think issues of privacy and confidentiality are important in the development of AI solutions, it is not the focus of this thesis. Consequently, my consideration of ethics is focused on the sociopolitical effects of certain features of the voicebot and how they can be addressed in the face of the complex entanglements that link us to them.

I - Humans Making Robots

[MTlab, Montréal - 18/06/19] As I sat in the meeting room of the MTlab and listened to the onboarding presentation by my new colleagues, I was really intrigued by the concept they were showing me. The product they were working on was a virtual assistant for tourists with a vocal interface using AI tools in natural language processing. This is what I would be working on all summer, helping them gauge the interest of potential contributors and users. The idea was interesting, I thought to myself. But why using voice technologies?

As if my colleagues could read into my mind, one of them started listing a bunch of facts about travellers' use of technology and most specifically virtual assistants: "We start from the observation that 75% percent of travellers rely on their phone screens to make travel decisions. We can thus assert that people discover cities through their phone. However, because of the advances in natural language processing technologies, there is an increasing number of vocal queries through virtual assistants [...]. The fact is that with the voice, one searches much more emotionally and when you start using it, you stick to it."

I was trying to put together all the pieces of the puzzle laying in front of me: Smartphones. Vocal queries. Tourism. How does that all come together? Again, probably deciphering my puzzled look, he went on to show me the proof of concept that they created for a potential client, a big name in the city's booming tourism sector. "We called it *Bonjour Montreal*," he said. The idea is for tourists to be able to hold a conversation with the voice of the city and get personalized recommendations on what to do and where to go."

Perplexed, I nod in silence. He proceeds to pull out his Iphone of his pocket, turn on the Bluetooth speaker on the coffee table and command: "Hey Google. Talk to *Bonjour Montreal*." "Getting *Bonjour Montreal*," replied the human-like yet impersonal female voice of the Google Assistant.

Accompanied by a gentle piano ballad, the velvety voice of what appeared to be a young woman with a French-Canadian accent took over: "*Hello, I'm the voice of Montreal and I'm here to make you discover the best Montreal has to offer. Remember that childlike wonder of believing everything is possible? Well, I never stopped thinking like that. If you feel like grasping some of that spirit for yourselves, then you will be able to ask me all the questions you want. I'm right here, and I have access to all the experiences you want to live. So, what can I do for you today? You can ask me to explore Montreal, share my stories or access travel services. What is it going to be?*"

It did not take more to be absorbed by the softness of the voice coming out of the speaker. I was impressed, taken aback by an indefinable excitement. The city that I had lived in for the years that defined me as a young adult now had a voice. She was talking to me and I could talk back! The conversation between the Voice and my colleague went on for a few more

interactions, before we came back to the research I would conduct over the summer. (Adapted from my fieldnotes, 18/06/19)

More than a summary of my first meeting with my colleagues at the start-up, this vignette recounts my first “encounter” with Voice assistant technologies and the Voice of Montreal. Of course, I had already interacted with similar technologies in the likes of Alexa or Siri, but that encounter was different. It told a story which I got to be part of, and that was deeply exciting.

The underlying premises of that story were the possibilities enabled by rapid technological advances underway in the field of artificial intelligence and natural language processing (NLP), according to which the primacy of text-based interaction would soon be replaced by a more natural and emotional medium. That is, the voice. According to their business model, these new technologies would lend themselves perfectly to the tourism industry, where labor and space shortages hampered the need to welcome and provide assistance to the increasing flow of visitors in the city. What they were offering was a new way to discover the city, hereby involving a new type of actor in visitors’ travel experiences: a human-like hybrid of voice and text available at all times through your smartphone to welcome you in the city and inspire you with suggestions during your stay. More importantly, behind the concept was the intent to form a bond with the user based on “emotional intelligence,” namely the combination of artificial intelligence tools with emotional experiences. I want to situate the “emotional intelligence” of the voicebot as part of a larger set of mimetic practices that aim at reproducing certain aspects of humanity into machines. Anthropologist Lucy Suchman explains how the creation of intelligent artifacts typically involves “the enactment of an elaborate, and elaborating, history of social relationship implying specific agencies and consequence.” (2007, p. 256) In the context of the sciences of the artificial, this means that the process by which machines are rendered “intelligent” and human-like is contingent to a specific cultural and historical normative system, which governs our very conceptions of what a human is and its capacity to act upon a certain environment. In line with this idea, this chapter explores the different types of emotional bonds envisioned in the voicebot, with respects to its vocal and conversational properties, as well as the negotiation of its identity as a service worker of the tourism sector. More precisely, I will look into how the attribution of these properties and identity characteristics intersects with the naturalist ontology and its conceptions of humans and machines.

The technology of the voice

The main pillar behind the concept of emotional intelligence is the idea that vocal interfaces allow for a more emotional connection with the user. To quote my colleague in his pitch, “the fact is that with the voice, one searches much more emotionally and when you start using it, you stick to it.” (Fieldnotes, 18/06/19) While the field tests among potential users that I was to conduct later in the product development process were meant to confirm, or infirm, this assumption, it is worth looking further into the idea itself to understand what, really, is in a voice.

Anthropologist Amanda Weidman writes about the voice as “both a set of sonic, material, and literary practices shaped by culturally and historically specific moments and a category invoked in discourse about personal agency, cultural authenticity, and political power.” (Weidman, 2014, p. 38) She further explains that in a Western context, the voice as a category holds a particular importance as it is associated with the expression of self, authenticity, and presence as well as the locus of agency. She traces the historical roots of the well-ingrained, and often taken-for-granted, concept to the period of the Enlightenment, a period often considered as the philosophical foundation of modernity (Weidman, 2014). It is at that time that, through the writings of thinkers such as Rousseau and Locke, the figure of the individual, conceived as an articulated and speaking subject, emerged. Among other things, the unique voice of the subject was a key aspect in its emergence as a political figure, as it signaled an interiorized self prone to a complex emotional life. It is important to note that it is at this point that the phenomenon of the voice was constructed as a binary opposing form and content, the former being conceived as a sonic and material support, albeit subordinated to the metaphysical quality of the latter. It is notably on this basis that psychoanalytic practice emerged with its focus on the articulation of one’s authentic “inner voice” through speech therapy (*Ibid*, p. 39). The materiality of the voice, as it is articulated in speech, was seen as a vehicle for accessing a subject’s authentic expression of their emotional individuality. This divide between the “signifying, authorial, or inner voice and [the] bodily, material vocality” was not only the foundation of psychoanalysis, but also of other disciplines, such as philosophy and linguistics. It eventually became the normative model for the broader socio-political project of Western modernity, which held at its center the figure of the rational, standardized and outspoken citizen, as opposed to the linguo of lower class and regional folk, which “was mired in custom and superstition.” (*Ibid*) Again, the essentializing power of modernity was unleashed, hereby attributing a universal and precultural quality to the voice. This “great divide,” to use the words

of Bruno Latour (Latour, 1991), between the metaphysical subject and its articulation by a material sonic object later mutated into other contrasts related to vocality dividing “male and female,” “colonizer and colonized,” “white versus black,” whereby gendered and racialized voices are thought of as “presenting inarticulate vocality” in the Western imaginary (Weidman, 2014, p. 47).

While I will come back to the politics of the voice in the next chapters, I argue that the link between the voice, authenticity, and emotion is at play in the creation of *Bonjour Montréal*. Besides its official name, the voicebot was often referred to as the *Voice of Montreal*, a naming act that I argue is far from being banal. By setting out the vocality of the city in the name itself, the concept does not only position the voicebot as a subject fitted with conversational agency, but also appeals to an authentic form of interaction and emotional connection with users. As the reaction to my first introduction to the voicebot shows, this emotional connection secures itself once the individual character of the voice is being heard. The city, now embodied by the voice of a young woman, can now interact with visitors and accompany them throughout their stay. Similar dynamics are further apparent in the start-up’s business model of “giving voice” to brands. In an interview, one of the co-founders of the start-up explained to me that the start-up was born out of a certain disillusionment resulting from an extended career in the advertisement industry. He wanted to get away from the industry’s tendency to make “false promises” in the sole perspective of “seducing people” without delivering their actual services. He believes that using vocal interfaces, made possible by AI tools such as natural language processing (NLP), automatic speech recognition (ASR), and text-to-speech (TTS), allows to bring more value into people’s lives in a meaningful way. He explains:

Being able to say things, to be recognised, and *to be touched* by a voice that talks to us and tells us a story according to what we said, this exchange, this dialog, this is something magical and it is going to have an impact on the way people will interact with internet. Even maybe among themselves. (Interview, 2019A)

Beyond the idea of the voice as a vehicle for authenticity and emotions, my colleague’s account hints to the underlying dualist ontology of the voice mentioned by Weidman, according to which the voice as a metaphor is separated from its material vocality “that is outside of referential meaning.” (2014, p. 40) In other words, the emotional power of the human voice, albeit cloned in this case, is unleashed by the technological tools of AI and NLP. Notice how the form/signifier is also relegated a subordinate position to the content/signified, as less importance is paid to whether the voice is spoken by an actual human or not. What matters, is

the that the technology is advanced enough to foster the idea of authenticity, emotions, and agency.

Beyond its voice, the possibility of an emotional connection with the voicebot is more generally allocated to its linguistic and conversational capacities. Coming back to the last quote from the interview with my colleague, he states that the “magical” potential of the voicebot lies in our ability as users “to say things, to be recognised, and to be touched by a voice that talks to us and tells us a story according to what we said.” (Interview, 2019A) Thus, the emotional bond with the user is not only conceived in terms of the voicebot’s ability to hold a conversation, but also to do so in the context of the situation. Similarly, the emotional bonding through a contextualized conversation between users and the voicebot was also envisioned in the recommendation feature being developed in the voicebot. The idea was that through a series of questions about the visitors’ trip, the voicebot could be able to make a personalized recommendation to the user for an activity to do or a place to visit. During an official event presenting the voicebot to members of the media and the technology and tourism sectors, a colleague stated that this feature of the voice was advantageous because it prevented making “recommendations à la Netflix.” He further explained that by holding a conversation with the users, the voice has the capacity to be spontaneous in the recommendations it makes. This in turn was described as “a tool that empowers the users.” (Fieldnotes, 09/17/19)

In order to make sense of these statements, I will turn to Anthropologist Lucy Suchman’s seminal book *Plans and Situated Actions*, in which she studied interactions between users and the “expert help system” of a photocopier machine at the XEROX research center in the 1980s (2007). Most importantly her account combines insights from the fields of anthropology, cognitive sciences and HCI to provide a framework for understanding our relationships with interactive machines. She argues that “the description of computational artifacts as interactive is supported by their *reactive, linguistic, and internally opaque* properties.” (*Ibid* p. 35) This means that machines are made interactive because we fit them with the ability to react, communicate in human language, and conceal their inner workings to users and that, in turn, we conflate these properties with the very broad concept of intelligence. She further explains that the equation of intelligence with these properties originates in the idea that “intelligence [...] is only incidentally embodied in the neurophysiology of the human brain, and what is essential about intelligence can be abstracted from that particular, albeit highly successful, substrate and embodied in an unknown range of alternative form.” (*Ibid*, p. 36) In other words, intelligence is conceived as an essence that can be reconstructed in any physical or mechanical arrangement, an idea still largely endorsed by cognitive scientists and which still governs the

making of self-regulating, interactive artifacts. It is important to note that the abstraction of the mind from its “material support,” i.e. the human brain, served the larger project to assert cognitive science as a legitimate discipline for understanding human action. Suchman writes that “at the turn of [the 20th] century, the recognized method for studying human mental life was introspection and, insofar as introspection was not amenable to the emerging canons of the scientific method, the study of cognition” was deemed unscientific (*Ibid*, p. 36) It is in this context that behaviorism, with its understanding of intelligence and human psychology based on “publicly observable” and external behaviour, asserted itself as a legitimate science (*Ibid*). This conception of intelligence as a set of outwardly observable behaviour gave rise to the field of AI as a subfield of cognitive science dealing with computers. The idea was to model intelligence onto a computer in the larger project of understanding human action in “scientific” terms, according to which its mechanical reproducibility proved “its psychological validity.” (*Ibid*, p. 37) Again, according to this understanding of intelligence, the more our machines are successful at reacting to a given environmental stimuli with “natural” language, the more we think we are able to understand, and hereby recreate, intelligence. However, as Suchman argue, this is only telling of our “tendency to ascribe full intelligence on the basis of partial evidence,” as we fit the machine with reactive and linguistic properties, which are themselves based on a culturally and historically specific understandings of intelligence.

In the light of this matter, we bring the voicebot in relationship with us by forming it as an individual subject fitted with agency. This is done through the ascription of reactive and vocal/linguistic properties, as it conforms to our understanding of what an intelligent and emotional agent is. As Suchman argues, this makes it easier to integrate in a social circuit of relationships governed by “an ontology of separate things that need to be joined together.” (*Ibid*, p. 256) As I explained in the introductory chapter, the default ontology of Western modernity attributes autonomous agency as an essential quality of human subjects, as opposed to the inanimate animal and material world. This places the voicebot in an ambiguous position in the naturalist ontological framework, straddling the positions of both subject and object. On the one side, the vocal and conversational properties grant agentic characteristics to the voicebot, which users perceive as essential properties as associated to their own humanity. This acquiesces with a naturalist and liberal humanist ontology. On the other hand, my analysis suggests that these properties are in fact the product of culturally and historically specific representations of what intelligence and humanity are. This blurs the naturalist and humanist ontological framework, as it challenges the assumptions of universalism underlying the very idea of what a human is and its given separation from the technological and material world. As

such, the following section will explore whether this dualism holds once the voicebot is envisioned in its “natural environment,” namely the tourism sector in Montréal.

Subjects / objects of servitude

[MTlab, Montréal - 17/09/19] It is a special night at the *MTlab*. The innovation hub for start-ups in the tourism, culture, and entertainment sectors is getting ready to host the event “*La Voix de Montréal*” [The Voice of Montréal], organized by the start-up I’ve been working with since this summer. Everyone in our small team is present and people slowly start pouring in. Apparently, all of the 80 places have been reserved. “The event is sold out,” a colleague told me.”

Mostly targeted to the community of UX professionals and to the media, the evening aims at presenting how “the Voice of Montréal will enrich the visitors’ experience with personalized and contextualized content.” (Fieldnotes, 17/09/19B) After a brief introduction, I feel an aura of curiosity in the room full of people, as the two organisations are about to publicly unveil the *Voice of Montréal*.

As soon as the demonstration starts, the feminine voice of the prototype rose from the speakers of the room, connected to the presenter’s smart phone, promising participants to access a myriad of authentic experiences. From this follows a discussion, during which participants ask questions ranging from the identity of the voicebot to its technological features and process creation: “Is she going to speak French as well?”, “How can one voice represent Montreal?”, or even “Who is the voice of Montréal?” (Adapted from Fieldnotes, 17/09/19A)

This ethnographic vignette adapted from my fieldnotes bears witness to the first time the Voice of Montreal left the ivory tower of discussion boards and meeting rooms to introduce itself to the broader public. As the questions raised by the audience show, the voice is not only a technological object, but also a public actor whose identity is negotiated. I use this event as an example to further expose the different forms of emotional bond envisioned in the creation of the voicebot. Since the tourism organisation that partnered with the start-up was present, the questions raised by the public and the answers provided by the presenters offer the perfect opportunity to analyse how emotional intelligence coincides with the tourism sector.

First, as the description of the event suggests, the Voice of Montreal was meant to “enrich the visitors’ experiences with contextualized and personalized content. This statement is important not only because it illustrates the main functionalities of the voicebot, but also because it is a trace of my involvement as a qualitative researcher in the design process. More precisely, I had spent the summer participating in guided tours, talking with various actors of the tourism sector, as well as meeting visitors in order to understand how the start-up’s

technological solution could better address their realities. At the end of the summer, I brought my research to an end and presented my results around what I labelled “the quest for authentic.” Based on Dean MacCannell’s notorious book *The Tourist*, I wanted to highlight the apparent dichotomy in the tourist experience between their obsession for accessing an “authentic experience” while staying within the bounds of major tourist attractions and areas (1976). I got the sense from the people I had met over the summer that it was important to nurture the thrill that they get from exploring and accessing to the “backstage,” while providing them with the right tools to guide them in this exploration, as some sort of guiding thread. My team really clicked with this idea of the guiding thread and coined the catchphrase “getting lost in the right direction,” in order to express that idea. Following a few brainstorming sessions with the client of the start-up, these insights informed the development of the robot as it provided the evidences for conceiving the emotional intelligence of the voicebot in terms of its ability to “inspire” visitors by promising them a certain synchronicity with the local lifestyle. It would allow them to move effortlessly in the city and to get rid of this paralyzing feeling of lacking context.

As I already argued in the previous chapter, such representations of the world emerge from a structural grasp of reality, with its description of the human experience as composed of a set of binaries such as ordinary/extraordinary, backstage/frontstage or authentic/inauthentic. More precisely, tourism is conceived as a cultural object consumed by tourist subjects as they symbolically move across the spatially enacted realms, in order to access an “authentic experience.” (Franklin, 2004) It is important to note that my intent is not to delegitimize such theories and frameworks. Indeed, beyond the fact that they represent canonical contributions to sociocultural understandings of tourism, my own use of this framework shows it can be very helpful to make sense of some of the apparent tensions in tourists’ experiences. Again, my intention is rather to show how some of the bonds envisioned between the voicebot and users follow a binary ontology. In this case, this framework centers its focus on the symbolic experience of the tourist subject, which implies that technological objects are relegated a secondary and instrumental role (Van der Druim, 2007). Thus, the question arises as to what form the emotional relationship will take between the visitors and the voicebot, who, despite its agentic characteristics granted by NLP tools, is relegated the secondary and instrumental role of “inspiring tourists”?

In order to answer this question, I will go back to the official event presenting the Voice of Montreal. The panel discussion and the exchange of questions and answers between the public and the presenters offered great insights into how the emotional relationship is envisioned in the context of the tourist experience. For example, to the question “Who is the Voice of

Montréal?”, the presenter from the tourist organization answered that “the Voice of Montreal is a woman. She is dynamic, curious, and open-minded.” (Fieldnotes, 17/09/19A). Along with this description, an official press release of the organisation states the following: “she [the voicebot] will be like a very nice tourist information agent in your pocket,” adding that “it will be there for you 24/7/365.” (Fieldnotes, 15/10/19) In the context of the booming tourism sector in the city, where the tourist flow keeps increasing every year, but labor supplies dwindle, the voicebot is seen as an innovative solution that will provide uninterrupted assistance to visitors by inspiring them in their experience of the city.

Such a conception of the bond between visitors and the voicebot concurs with recent developments in robotics that have taken place over the recent years. To quote Suchman, “as the robot was to the industrial imaginary, so the software agent is to the desires and fantasies of the service economy.” (Suchman, 2007, p. 219) In other words, the figure of the industrial robot tirelessly performing productive labor has moved to the domestic sphere, where it is envisioned as a social robot performing reproductive labor through various tasks ranging from care, sex work, or cooking (Chasin, 1995; Kubes 2019B; Phan, 2019; Suchman, 2007; Treusch, 2007). Many feminist posthuman scholars dealing with this subject have argued that the increasing development of robots in the domestic sphere both stabilizes and destabilizes the ontological divide between subjects and objects. For example, Alexandra Chasin wrote about the intertwined relationship between machines and workers in the context of domestic service in North America and Europe (1995). She posits that, on the one side, the idea of domestic robots reifies the subject-object binary with the assumption that *we* invent *things* to make our lives easier, and that this project often empowers individuals and groups who occupy the position of “subjects” at the expense of subordinated others who, on the other side of the divide, occupy positions such as female, feminine, non-white, poor, etc. On the other hand, she challenges the assumption that these positions are essential and “god-given” by looking at how domestic robots re-code the meanings associated with gender, race and class, along with the meanings of service work. Historically, domestic service performed by servants is associated with a lower-class status as it is considered as reproductive labor, which takes place within the closed walls of the home. In Marxian terms, reproductive labor is ‘invisible’ because, unlike productive labor, which by virtue of its visible accumulation can be ascribed a value, reproductive labor is “exhausted in their performance” and leave no material traces so-to-speak (Chasin, p. 78). Consequently, the burden of domestic service for affluent groups, such as the white upper and middle classes, has historically fallen upon groups placed in subordinated positions, like women of color of a lower economic class. As domestic work became

increasingly undesirable in the second half of the 20th Century and labor supplies for such type of work decreased, the realisation of the capitalist myth of the ever-expanding middle-class depended on a new type of being to perform domestic work, namely machines. This coincided with rapid technological changes and the apparition of domestic electronic appliances in our homes and beyond. Chasin demonstrates that these appliances were often advertised as “servants, as things that could perform labor without the displaying the liabilities of human subjectivities. At work in the household, machines have promised, explicitly, to save labor.” (*Ibid*, p. 85). It is precisely on this aspect that “servant” technologies challenge the subject-object boundary in two ways. First, technological changes re-code the meanings of categories of gender, class, and race, since the difference between domestic machines and humans performing this type of labor, historically lower-class women of color, is elided as the latter were deemed disposable and replaceable by machines. Moreover, these technological changes re-code the meanings of domestic labor and service work, as domestic technologies completely efface the labor force that goes in domestic work by portraying it as effortless, hereby subtracting “the liabilities of human subjectivities out of the equation.” (*Ibid*) Finally, Chasin clarifies that the premises stating that technologies will save labor and the expansion of the middle class are false, since both of these assumptions never actually concretized. She writes: “[...] the era of Reagan-Bush economics, in which the middle class has actually shrunk, and the upper and lower classes have polarized, the number of electronic household appliances has risen with— rather than caused a diminishment of— the number of domestic workers.” (Chasin, p. 93) If anything, Chasin concludes that electronics participate in the polarization of economic classes as they secure “the idea that a service class of being(s) is proper and even necessary.”

Chasin’s argument on domestic technologies can help us make sense of the emotional bond envisioned between users and the voicebot. While her analysis does not really focus on service work as it is performed in the context of tourism, I argue that similar dynamics are at play in the envisioning of the voicebot as a “very nice tourist information agent.” As I have mentioned above, the voicebot is accessible as an “action” on Google’s Assistant, which belongs to the latest wave of virtual assistants using a vocal interface along with software agents such as Amazon’s Alexa, Microsoft’s Cortana, or Apple’s Siri. All sporting female voices and initially designed to help users with a wide range of organizational and domestic tasks, these agents have found their ways in multiple spheres of our lives as they are now embodied in various derivative products such as home speakers, cars, smartphones, etc. (Phan, 2015) Similarly to Chasin, Thao Phan argues that such devices are designed and used with the premise that they provide effortless, constant, and dedicated labor to busy middle class families, hereby re-coding

the charged relationships between the markers of class, gender, race and domestic service in the United States, and eluding the acknowledgement of the women of color that have occupied (and still do) such positions (*Ibid*). As much as this background makes it hard to completely dissociate the voicebot from such critical perspectives, it is important to note that my intent is not to blame certain creative decisions, nor deem them unacceptable. Having experienced the day-to-day practical reality of a tech start-up collaborating with many actors across various sectors, I know firsthand that a wide range of other factors and representational challenges are at play in the creative process of, literally, *giving voice* to a city. Rather, my intent is to show how the bond imagined in voicebot, taking the form of a service relation between visitors and a tourism service worker, de-stabilizes categories of humans and machines as it re-codes their relation to labor and service work. More precisely, the identity of the voicebot is conceived as a “very nice tourist information person” who is “there for you 24/7/365.” (Fieldnotes, 15/10/19) In a context where the flow of visitors in the city keeps increasing at a pace which physical tourist information infrastructure cannot keep up with, the voicebot is seen as the perfect solution to offer a better service at any point in the visitors’ itinerary. Besides expanding the reach of human tourist information agents beyond their limited points of service in the city, it does so effortlessly at any point in time and space. We can see how the voicebot blurs the boundaries between human and machines in their joint performance of service work and, by the same token, re-codes the meaning of this type of labor as constant, effortless, and without any physical boundaries. Finally, on the other hand, it also stabilizes the ontological divide between humans and machines, as it reinforces the assumption that we are dependent on a certain class of beings to perform the work that we cannot do in order to favor economic growth.

Beyond the Great divide

To conclude, let us come back to Suchman’s claim that the creation of intelligent artifacts typically involves “the enactment of an elaborate, and elaborating, history of social relationships implying specific agencies and consequences.” (2007, p. 256) The creation of an AI agent, in other words, not only involves the enactment of relationships between humans and themselves, but also between humans and machines. In the creation process of the Voice of Montréal this relationship is conceived as an emotional experience based on a conversation between a visitor and a voicebot embodying a tourist information agent. Along with Suchman’s claim, my analysis has presented how this relationship is contingent to a specific normative framework, which, in a Western context, is by default governed by a naturalist ontology tracing

sharp lines between the categories of subjects, humans, language, etc., on one side, and objects, technology, matter, etc., on the other.

To summarize my findings, the different bonds envisioned between the voicebot and users reaffirm the naturalist ontological framework in two contradicting ways. First, the ascription of vocal and conversational properties grants the voicebot agentic characteristics in line with our conceptions of intelligence and sociality, hereby earning its place along the “subject” side of the great divide. Second, the description of the voicebot’s identity and function as a pocket-size tourist information agent available at all times to inspire visitors is partly influenced by structuralist understanding of tourism, as well as gender, class, race, and labor relations intertwining humans and machines. These dynamics relegate the voicebot the position of object, as it is considered the best solution to replace human labor and “serve” tourist subjects in their consumption of tourist experiences. The assumption that we rely on specific classes of beings to perform certain types of labor reaffirms the subject/object ontological divide. However, in this case, it does so by placing the voicebot on the “object” side of the great divide. We can conclude that the voicebot falls within an ambiguous position along this ontological framework, as it fits both positions without ever completely meeting the provisions that these allegedly universal categories suppose.

This last point leads me to my second and most important conclusion, namely that the different bonds envisioned in the voicebot destabilize the naturalist ontological framework. Whereas this framework attributes an essential character to categories such as subject/object, human/machine, voice/vocality, intelligence/its material arrangement, etc., and holds their opposition as universal, my analysis has also shown that these are in fact culturally and historically constructed. Moreover, I have shown how technological change re-codes the allegedly essential meaning ascribed to such categories, as it is the case with service work and labor. Consequently, these conclusions have several implications for the bonds we envision between the voicebot and users. First, the ontological boundaries between subjects and objects or humans and machines do not really make sense in the first place. This points to the inadequacy of the framework to understand the place of this new type of actors in our lives. If we remember the claims of feminist-queer posthumanist scholars, my conclusions show that rather than being completely separate from each other, humans and machines *become with* each other as they constantly remake one another. We are posthuman companions. Finally, it thus seems like a framework putting forward our mutual becoming would be more appropriate to fully grasp the potential of our relationship with the voicebot, and with technology in general.

It is with this intent in mind that the next chapter will turn to my own relationship with the voicebot and understand it from the lens of posthuman companionship.

II - The voicebot companion

“09/24/19. I get to demo!” (Fieldnotes, 09/24/19)

Recorded only a few days after the event presenting the voicebot to various actors of the industry, this excerpt from my fieldnotes marks the second part of my collaboration with the start-up. The first preparation phases of the product development process were to be completed earlier in the fall and the development of a working prototype was well underway. My role would be to test the alpha versions of the bot and perform ‘demos’ among tourists to gather their impressions and present the user feedback to my team. These recommendations would contribute to the iterative product development process, whereby a new alpha version of the robot addressing the previous versions’ shortcomings was released every two weeks. Looking back at it, getting access to the ‘demo’ was major milestone in my relationship with the voicebot; we were about to become intimate partners, given that being signed up in the system as an alpha tester made me among the few team members to be able to pronounce the magic words “Hey Google, talk to Bonjour Montreal,” and to actually be able to interact with the prototype. As I will explain, this marks the beginning of a complicated relationship with the robot set with hopes, frustrations, disappointments, and uncertainties. Similarly to Haraway who asked “whom and what do I touch when I touch my dog?,” this chapter sets out to investigate the effects of my multi sensuous and affective involvements with the prototype of the voicebot (Haraway, 2008). I will explore how my relationship with the voicebot developed through a closer look at two important aspects of my own role in the development of the prototype. First, I will base my analysis on the various demonstrations I performed with visitors in the context of my user research. The second part of my analysis will look into the various bugs and uncertainties that arose upon the initial releases of the demos among team members. These two aspects of my relationship with the voicebot, respectively instantiating frontstage and backstage encounters with the demos, will allow me to challenge my conception of the robot, from a defective autonomous agent to a complex sociomaterial assemblage. I will conclude by reframing our relationship to intelligent artifacts within the tenets of a relational ontology as I make the case for viewing the voicebot as a companion.

Re-thinking the demonstration

My excitement upon getting access to the robot was partly due to my own perception of the voicebot and its prototype along the lines of the Cartesian divide; the progress being made on the prototype would materialize the various narratives and visions of the robot service worker in the form of a human-like virtual assistant for tourists fitted with conversational agency. This was connected to a conception of the voicebot, and AI in general, as completely autonomous systems with highly sophisticated capabilities for interaction. Keeping in mind the insights from the last chapter, it is safe to argue that this perception is partly founded on a Cartesian ontology ascribing intelligence and agency through vocal, reactive, and conversational properties. However, the various encounters between the voicebot, myself, and other related actors, as it arose in the context of demonstrations to users, forced me to rethink my conception of intelligent artifacts and our relationships with them.

The first round of demonstrations was done at the Tourist Information Center of Montreal in the End of September. Located in the city's buzzing downtown area, the center was only one block away from the very busy main shopping street. The information center had several reception desks, behind which tourist information agents were waiting for visitors to show up with their questions. While we were in the last week of September and the frenzy of the summer season had slowly subsided, they were expecting a small peak of visitors as the trees' colorful foliage was blasting in its full autumnal glory. The rationale behind the choice of space was simple; we needed a place where we could easily intercept tourists to ask them if they want to test our voicebot and ask them a few questions about their experience.

Introducing the *Voice of Montréal* to users meant juggling with- and re-adjusting various assumptions we were holding about the autonomy of the voicebot. The underlying assumption was that by minimizing the information we gave to users about the instructions for use, we would be able to assess the usability and intuitivity of the solution. Concretely, this meant that I would introduce the prototype as "a conversational solution with whom we can discuss to discover more about the touristic offer of the city." (Fieldnotes, 30/09/19) I would proceed to tell them that that we were going to perform the test on my smartphone, to which I would add: "You begin talking to it by saying 'talk to Bonjour Montréal.' It's a demo version, so there's only one path of questions and answers that really works. I'm going to guide you through it as you go. And then, I'm going to ask you questions about your experience." (*Ibid*) I was trying not to give too much information while introducing the basics for users to be able to initiate their conversational experience. This implied that the voicebot was autonomous enough to hold

a conversation by itself with an uninitiated user, as the true conversational agent I thought it was. However, even with all this information, it proved to be hard in a lot of cases for users to interact autonomously with the demo. Consider the following transcription of an interaction between a visitor at the information center and the voicebot, adapted from the notes me and my colleague took during the intervention (Translated from French, Fieldnotes 2019A):

Difficulty: accessing the action
Malik does it instead)
Intro, user reads the text)
Problem with “leisure”; the bot understands “leather”, “4 laser”
Activities: after long text, people don’t see the carousel with pictures: “should I just scroll?” said the user .
“Talk to, Bonjour Montreal” : people say “Bonjour Montreal” only→ doesn’t work

A few things are at work upon a closer look to this sequence of observations. First, the user had a hard time accessing the action as they were saying “Bonjour Montreal,” despite my clear indications to say, “talk to Bonjour Montreal.” This is something that happened a lot during the demos, because of what I suspect to be the counter-intuitivity of the “trigger phrase” “Talk to bonjour Montreal,” which comprises two interpellations, namely, “talk to” and “Bonjour.” As a consequence, I would often have to take over the process of interaction and call on the voicebot myself, before handing the phone back to the user in order for the demo to follow its course. Second, a similar thing would often happen when the voicebot did not understand what users said during the interaction. In this case, the voicebot asked “are you travelling for leisure, or business?” to which the user replied “for leisure” with a slight French accent. Only that was enough for the voicebot to mishear the user and register “leather” and “4 laser” as answers. Again, I would have to take over the interaction and clearly pronounce “for leisure” in an exaggerated and unnatural way. This happened to around 30% of users who tried the prototype, which often led them to have an overall disappointing impression of the voicebot and the concept, because it could not understand their accent (Fieldnotes, 01/11/9). Finally, there were often moments of uncertainties when the user was unsure of how exactly they should interact with the interface. In this case, the user had completed the conversation with the demo, which led them to hear and read about a description of the activities that are going on in the month of June in Montreal. This paragraph is followed by an interactive carousel composed of different icons, each representing a choice of activity. However, the interactive carousel did not show up because the interface did not scroll down automatically, which prompted the user to ask me

“should I just scroll?”. The opposite would often happen, whereby users tried to do things that were not programmed in the demo, such as asking right away “what are the main events this weekend?”, despite my warning that it was not yet programmed to answer queries outside of the conversational path. In such cases, I would have to intervene to tell them that this function was not available yet and that they had to “stick to what the voicebot was asking them.” (Fieldnotes, 20/09/10) I chose this sequence of observation both because it encompasses most issues that came up in the later rounds of testing and because it is representative of what most interactions looked like. While I am going to come back to the way these insights were addressed in later versions of the prototype in the next chapter, I want to redirect your attention to the interaction that went on between the voicebot, the user, and myself. Namely, to the way my high level of assistance to the interaction challenges the assumption I was holding about the autonomy of the voicebot.

I want to situate these sessions of user testing of the voicebot’s prototype in line with what Suchman calls ‘technological demonstrations.’ (Suchman, 2011) Beyond simple opportunities for evaluating the usability of a given technological artifact, she understands demonstrations as sites of technoscientific practices, where boundaries between subject and object as well as nature and culture are drawn and negotiated. She argues that the technological demonstration, often referred to as ‘demo,’ is a “distant cousin of the scientific demonstration, the common thread being the premise [...] that rather than simply being informed, spectators are witness to some natural or technical object directly.” (*Ibid*, p. 123-124) In other words, demonstrations are essentially a work of *mise en scène* serving to prove the existence of humanoid robots. These dynamics enter in line with objectivist sciences, in which instruments of observations came to stand for the natural objects they represented, hereby erasing the human “coauthors” of the so-called scientific facts they create (Suchman, 2008, p. 214). In our matter of interest, rather than reconstituting an encounter with nature, the technological demonstration reconstitutes an encounter with culture, “in the form of an uncannily familiar Other in the making.” (*Ibid*, p.124) Thus, the intelligent artifact becomes its own instrument in the testimony of its factuality and, similarly to the scientific observation, its human creators erase themselves as important mediators (Suchman, 2008). Suchman adds that demonstrations have the effect of naturalizing and objectifying humanoid robots as they are staged as completely detached from the labor of their creators. This naturalization is partly in effect through the demonstration’s distanciation of the observer, who, besides the creator, becomes a witness to the factuality of the robot (Suchman, 2011).

The previous official demonstrations I had witnessed, recounted in the previous chapter, definitely succeeded at ‘proving the existence’ of the voicebot as “an uncannily familiar Other in the making.” (Suchman, 2008, p.124) My initial excitement upon my very first encounter with the voicebot and the audience’s reaction during the official evening’s demonstrations pays witness to this. It suffices to say that my colleagues and I entered the first round of testing with the same mindset, as the assumption that the voicebot was autonomous enough to interact with uninitiated users guided the demonstration. More precisely, I was to simply efface myself in the background and only intervene minimally, in order to assess the usability and intuitivity of the solution (Fieldnotes, 30/09/19). We can see how this rationale dovetails the logic behind objectivist sciences’ demonstrations, in which the human “coauthor” is supposed to be absent from the construction of a fact. In this case, the “fact” is the voicebot’s autonomous agency, based on the ascription of conversational and reactive properties. The demonstrations would serve to assess the interaction with the voicebot based on these properties, albeit in a very restricted way, since the prototype was only programmed to say a few things. However, taking the voicebot out of its highly controlled environment and exposing it to uninitiated users resulted in an unsettling picture. As my observations point out, the voicebot ‘failed’ at being the autonomous agent the official demonstrations led me into thinking it was, which often led me into experiencing disappointment and frustration.

Similarly, Suchman analyses her encounters with famous robots from the MIT AI Lab such as Mertz, Kismet, or Robota (Suchman, 2008; Suchman, 2011). She recounts that despite their being framed as “autonomously affective entities,” they failed at displaying such characteristics, as the interaction was highly contingent on the assistance of their “human caregivers.” (Suchman, 2008, p. 246) She argues that rather than deeming these encounters as failures, we should “[...] reframe [these robots] from an unreliable autonomous robot, to a collaborative achievement made possible through very particular, reiteratively developed and refined performances.” She further adds that “like all forms of agency, [the robot’s] capacities for action are created out of sociomaterial arrangements that instantiate histories of labor and more and less reliable, always contingent, future enactments.” (*Ibid*, p. 246-247) We understand that robots do not hold agency *a priori*. Rather, their agency emerges from their entanglements in “sociomaterial arrangements” involving the performances and labor practices of the humans who build them, amongst other things. Very similar dynamics were at play during the demonstrations I conducted among users, as the successful completion of the demo was contingent to my high level of assistance. More than the performance of my role as a user researcher, it was also highly dependent on the labor practices of my colleagues who worked

on their specific part of the voicebot. If the predicament upon the first round of testing implied that I was to make these histories of labor quasi-invisible by effacing myself from the demonstration, I soon realised it would be impossible to do so. It is important to note that my intent is not to criticize the bugs and imperfections in the interface, which are normal in a prototype at such an early stage of product development. Rather, I want to show through these observations how the agency of the voicebot is not only dependent on the ascription of conversational and reactive properties, as I explained in the previous chapter, but more so on the establishment of relationships with various human actors and material entities. This includes the visitors I met at the information center, whose interaction with the voicebot was made possible through the bot and I's fine-tuning of the performance of the demo. Indeed, I progressively learnt when to step in the interaction and ensure its completion through the iterative product development process and the successive releases of new demos. It is not exaggerated to state that it is precisely these iterative re-adjustments that allowed me to gather constructive user feedback; I would not have gone very far had I only stood there and let the uninitiated users and the voicebot to themselves, under the assumption that the latter was the autonomous agent I initially thought it was.

To sum-up, my relationship with the voicebot progressively developed through the various demonstrations I conducted among users. Learning to perform the demonstrations correctly implied constantly re-adjusting the “ontological choreography” between the voicebot, the users, and myself every second week, as new versions of the demo were coming out (Haraway, 2008, p. 65). These instances challenged my perception of the voicebot from an autonomous agent to a sociomaterial arrangement, whose agency is intertwined in my performance of labor practices as a user researcher assisting users in their interaction with the voicebot. The next section is going to look further into how labor practices and agency are linked to the materiality of the voicebot.

When the demo does not work

I will now turn away from the performance of the demonstration to look at how my relationship with the voicebot developed within the walls of the MTLab, as new versions of the demonstration were initially released internally among team members. These encounters with new versions also involved a readjustment of my conception of the voicebot, as the prototype was often initially faulty and defective. For example, one day I was working at the MTLab when the following happened:

I arrived today and my colleague asked me whether the demo version of the voicebot was working. Apparently, he made changes to the conversational design yesterday and the demo version was not working on his google assistant ever since. I tried it and it was working for me. The only ways he could access the assistant was through the test app, and then every request to the conversation (e.g explore Montreal) would be answered by a failure message. (Fieldnotes, 23/10/2019)

This instance is illustrative of the uncertainties that characterized my relationship with the voicebot, which often took the form of bugs impeding the voicebot from working properly upon new releases of the demos. In some cases, it was initially suspected that the ‘bug’ was attributable to the incompatibility of the solution on smartphones that were not Apple products, which sparked a wave of worry that the solution would not work on Androids. In the end, they might have been attributed to the fact that Google’s system needed a 48h delay to sync the different platforms, applications, and softwares involved in the development of the prototype. It is something I often referred to in my notes in the broad category of “google problem,” indicating both the vast digital infrastructure behind the voicebot and the complexity that it represented for an uninitiated person like me. More precisely, our solution was being developed as an action on the Google Assistant, which implied that we were relying on the tech giant’s cloud services to design conversational interfaces such as DialogFlow. As much as these tools, among many others, allowed the team to create the voicebot altogether, it came with some limitations in our capacity for action.

However, what interests me for now is not so much why it did not work, but what would actually happen in these instances. Namely, my colleagues would ask every team member who was signed up as an alpha tester to go and “see if it works with them.” Often, it would work for some people and not for others, which perhaps is what motivated all the claims supposing that it “only works on Iphones.” During these moments, uncertainty prevailed to the point where the voicebot seemed to have a life of its own; it was as if it had ‘mood swings’ as it refused to work on some of our devices for reasons we often had a hard time enlightening. These moments further challenged my perception of the voicebot in a few ways. Indeed, if I initially attributed the cause of these occurrences to the alleged unreliability of the voicebot, they progressively revealed the vast infrastructure and the labor relations that were behind the development of the prototype. Indeed, trying to assess the cause for the ‘bug’ led team members to talk about the voicebot in very different ways. This is what prompted me to write the following note on the same date:

Interesting how we don't really talk about the bot itself. The final version of it is rarely discussed / referred to as an entity in itself. It's not 'the bot', it's 'dialogflow,' the 'demo version,' 'conversational design,' 'intent management,' 'web hook,' etc. Maybe it is because we are at the developing stage and the Voicebot does not technically exist yet, except for the demo that we have. (Fieldnotes, 23/10/2019)

In this case, I was referring to the different components and entities in the architecture behind the final product of the voicebot and its different prototypes. Each of these components were tied to the labour of a team member who worked on a specific part of the project. I added in my notes that for myself, "my own perspective was tainted by the interactions I had with the users and the insights I gathered from them." (*Ibid*) We can see how the bugs and the 'backstage' perspective they provided me challenged my conception of the voicebot as an autonomous whole. Here, the different components that held this whole together started to emerge, as the latter was not working. I argue that these observations on the relationship between the different parts of the voicebot and the whole they form further support my claim for a relational and distributed version of agency.

Turning to Political Theorist Jane Bennett's concepts of vibrant matter and agency of assemblages is useful to understand this relationship (Bennett, 2009). In her plight for the aliveness of matter, Bennett refers to Spinoza's concept of conative bodies. More precisely, bodies are social, be they alive or not, organic or not, insofar as they are "continuously affecting and are affected by other bodies." (*Ibid*, p. 21) She writes:

"Spinoza's conative, encounter-prone body arises in the context of an ontological vision according to which all things are 'modes' of a common 'substance.' Any specific thing [...] – a glove, a rat, a cap, and the human narrator of their vitality [...] – is neither subject nor object but a 'mode' [...]." (*Ibid*)

This means that entities can neither be classified as subjects or objects, they are rather "modes" of a common substance. Thus, any specific body or thing can be seen as a facet of a greater multifaceted whole. Agency, in turn, "becomes distributed across an ontologically heterogeneous field [...]." (*Ibid*, p. 23) Deleuze and Guatarri's concept of the assemblage is useful to further understand the distribution of agency across a multitude of "macro- and microactants." She defines assemblages as "ad hoc groupings of diverse elements, of vibrant materials of all sorts." Moreover, "the elements of the assemblage work together, although their coordination does not rise to the level of an organism. Rather, its jelling endures alongside energies and factions that fly out from it and disturb it from within." (*Ibid*, 25)

She uses this concept to analyze the power blackout that shook the United States and parts of Canada, affecting more than 50 million people in the August of 2003. At the center of this

event is the North American power grid, conceived here as a “volatile mix of coal, sweat, electromagnetic fields, computer programs, electron streams, profit motives, heat, legislation, wires, [...]” (*Ibid*, p. 25) While investigators are still unsure about what exactly happened, she describes the events in the following way: “What seems to have happened on that August day was that several initially unrelated generator withdrawals in Ohio and Michigan caused the electron flow pattern to change over the transmission lines, which led [...] to a successive overloading of other lines and a vortex of disconnects.” In this context, she understands the grid as an assemblage of multiple human and nonhuman actants and she attributes the “agential loci” of the power outage among the following actants: understaffed power plants, bush fires in Ohio, a growing consumer demand, and, amongst others, the deregulation and privatization of the grid initiated by the Energy Polity Act of 1992 (*Ibid*, p.25-26). Assembled together, all these actants were subject to what she names “the slight surprise for action.” (*Ibid*, p. 27) Borrowed from Bruno Latour, this concept refers to the consequences arising independently of the outcomes planned by the actants. She adds that this happens often in the distribution of electricity, as this “vibrant matter” can unexpectedly change its behavior and flow in a completely different, or even reversed, direction. Bennett’s account of the blackout allows her to understand agency differently than mainstream philosophical accounts. More precisely, human-centered conceptions of agency offered by thinkers such as Augustine, Kant or Merleau-Ponty respectively define agency as intentional will attributed by the divine, moral law, or the intersubjective field. In contrast, Bennett writes that “in this selective account of the blackout, agency, conceived now as something distributed along a continuum, extrudes from multiple sites or many loci - from a quirky electron flow and spontaneous fire to members of Congress who have a neoliberal faith in market self-regulation.” (*Ibid*, p. 28) This means that agency emerges from the different bodies and parts that constitute the assemblage, rather than being attributed on the basis of humans’ capacity for action. This vibrancy of matter is the result of the human and nonhuman constituents of the assemblage, who, when all put together, generate a capacity for action much greater than the sum of each part, for better or for worse.

These insights help us further conceive the voicebot as a complex sociomaterial assemblage composed of both human and nonhuman bodies. It is important to note that my positionality and part-time presence within the team have actually not allowed me to achieve a complete understanding of the role of each of these constituents and the exact “agential loci” at which their vibrancy derailed and caused the numerous bugs and dysfunctionalities. However, Bennett’s framework is nonetheless useful to understand the event of “the slight surprise of action” itself, namely the instances when the voicebot did not work. These moments of

uncertainties reveal the vitality of the new demos of the voicebot in the form of ‘mood swings’ as they often did not comply to our commands. In the style of Bennett, let me draw a non-exhaustive list of the important actants that my account of the bugs has identified: iPhones, Androids, the Google Assistant App, human vocal cords, natural language processing algorithms, a pounding heartbeat when the voicebot refuses to work during a demonstration, the cloud infrastructure, tourists, another bunch of human ‘caretakers,’ and of course, myself, “the human narrator of their vitality.” (*Ibid*, p. 21) As with Bennett’s account of the blackout, both the vital role of these bodies in the assemblage and their own rationale for action came to light as the whole did not work properly. What is more, the different nonhuman actants in the assemblage named above such as ‘Dialogflow,’ the ‘conversational design,’ ‘web hook,’ etc., are all technologies connected to the specialized labor provided by team members that grant the voicebot its agencies. To quote Suchman, “seeing the robot ‘at home’ in the lab [...] provided an opportunity to see as well the extended network of human labors and affiliated technologies that afford [the robot] its agency.” (2008, p. 246) In the light of my observations, Bennett and Suchman’s views teach us that despite the tight labor relations that tie us with specific bodies of the assemblage, vibrant matter sometimes acts in non-predictable ways that might not align with our intentions and in turn impede our work. Facing the unpredictability of matter, it can become hard to identify the root cause of ‘the slight surprise of action.’ Concerning intentionality and causality, Bennett adds that:

“[a] theory of distributive agency, in contrast, does not posit a subject as the root cause of an effect. There are instead always a swarm of vitalities at play. The task becomes to identify the contours of the swarm and the kind of relations that obtain between its bits. To figure the generative source of effects as a swarm is to see human intentions as always in competition and confederation with many other strivings: [intention] always vibrates and merges with other currents, to affect and be affected.” (Bennett 2009, p. 32)

Put differently, “the swarm of vitalities at play” in the creation of the voicebot puts us humans in relation to the assemblage, in a way in which our intentions either align or clash with those of the other bodies. This allows us to dissociate the concepts of agency and intentionality as the successful achievement of our intentions as subjects. Rather, agency is about bodies affecting and being affected by each other. This was mostly apparent in the fact that the uncertainties associated with the defective new versions of demo were highly frustrating for me when they interfered in my ability to perform the demonstrations to users. This shows in the following excerpt from one of the last demonstration sessions I performed at a hostel in downtown Montréal:

“It makes me kind of frustrated to perform the demo to a participant and they end up not being able because “she does not work” right now. [...] In these situations, I feel a kind of disconnect between me and them. It’s more a non-connection than a disconnect because we cannot establish a connection on the same grounds, the one that would premise the utilisation of our solution. In fact, I would say I feel extremely annoyed at the bot in these situations. I say at the bot because I know that it is not necessarily [the start-up’s] fault if these things do not work. Being immersed for a while in the organization has allowed me to understand how things are often very complex...” (Fieldnotes, 15/11/2019)

Here, the agency of the voicebot manifests itself through its malfunction, as my intentions to use it and establish a connection with users clash with the intentions of the other bodies of the assemblage. These intentions, albethey unknown, affect me in the form of annoyance and frustrations as they impede me to perform my tasks as a user researcher. As I highlighted above, we can see how the tight labor relations that tie me to the voicebot, and in turn afford it its agency, are contingent on the different trajectories the intentions of each bodies in the assemblage take. Moreover, it is interesting to note how a better understanding of the complex assemblage that forms the voicebot, identified in my quote as a prolonged immersion in the organization, has allowed me to resist attributing the causality of the demos’ dysfunctionalities on one single actant, in this case the start-up. I instead alluded to the complexity of the assemblage. Far from me the intent to frame this observation as an obvious fact. Rather, my developing relationship with the voicebot progressively revealed the different bodies among which agency is distributed. It is only through a sustained involvement in the start-up and developing my relationship with the voicebot that I can assert that there is no agency without relations. As Bennett argues, this makes it almost impossible to attribute direct causality to one or many human agents. She writes that “alongside and inside singular human agents there exists a heterogenous series of actants with partial, overlapping, and conflicting degrees of power and effectivity.” (*Ibid*, p. 33) She makes it clear that this is not a way to de-responsibilize us humans but rather to challenge the anthropocentric narrow-mindedness that attributes causality solely to human agents. Rather, the assemblage allows us to “broaden the range of places to look for sources” of ethical responsibility beyond mere moralism (*Ibid*, p. 37). The question of ethical responsibility in the development of the voicebot will be further addressed in the next chapter. For now, let me conclude my argument by tying together relational and distributed understandings of agency with the notion of posthuman companionship.

Making room for companionship

This chapter explores how my relationship with the voicebot established itself through a series of hands-on engagement with the voicebot, which occurred both ‘frontstage’ with various rounds of user testing and ‘backstage’ upon the demos’ initial release within the team. First, my analysis of the demonstrations to uninitiated users reveals that the voicebot’s agency emerges from sociomaterial arrangements entangling labor relations, in this case in the form of refined performances of my role as a user researcher, and in the subsequent re-adjustment of my level of assistance during the demonstrations. Second, my analysis of the release of the demonstrations among team members and the unexpected bugs they engendered revealed that the voicebot is a complex assemblage of human and nonhuman bodies. Agency is distributed and sometimes unpredictable; it emerges from the vibrancy of these respective bodies as they are affected by each other’s intentions. In both cases, these understandings challenge the characterization of the voicebot as an autonomous agent. Rather, agency is found within the complex relations between matter, humans, and the labor they perform. We are very close to Karen Barad’s conception of posthuman performativity and agential realism, according to which agency is relational rather than individual as it is enacted through “iterative changes” in particular material-discursive practices (Barad, 2013, p. 827). Emphatically, my analysis has revealed that my relationship with the voicebot becomes the smallest unit of analysis for understanding how categories such as user researcher and voicebot come to be. It is through our relation that our respective, yet entangled subjectivities arose in the form of my constant re-adjusting of my role as a user researcher and the voicebot as an assemblage of vibrant bodies, whose distributed agency sometimes acts in unexpected ways. These subjectivities are ‘posthuman,’ inasmuch as they are entangled in material-discursive practices that challenge the ontological divide between humans and machines, as these two constituents find themselves entangled in intra-active processes of becoming. The next chapter is going to further focus on how various subjectivities were negotiated in the changes in the demos. I will conclude this chapter by giving an account of the last stage of my relationship with the voicebot. This will allow me to further make the case for the co-constitutive character of our relationship of companionship.

As my collaboration with the start-up came to an end, I was to work on one last task: the implementation of a feedback loop within the voicebot. While I ended up being too short on time to actually complete this task, I want to make sense of the intent in itself as it presents interesting aspects of the final stage of my relationship with the voicebot. For context, the

feedback loop is the implementation of conversation cues such as questions to the user to account for user feedback in the voicebot's interface itself as a way to evaluate if the research results satisfy users (Fieldnotes, 13/11/2019). My task was to base myself on the field observations I had made during the user testing, in order to identify conversational cues that would be included in the voicebot's conversational design at the end of the interaction with users. The challenge would be to find a way to evaluate this satisfaction by asking questions such as "how did you like your experience," for example. My colleague had framed my task in the following terms: "We want to go in the fine grain of users experience, a little bit like you guys (the user research team) are doing." (Fieldnotes, 13/11/2019) It is important to note here that they did not aim at replacing the work of a qualitative researcher by the feedback loop. What they were after was rather a way to gather more feedback in a systematic way in the interface itself, which would provide the user research team with more data on user interaction. I was really motivated by this prospect since it legitimized my work as an ethnographer and user researcher (Fieldnotes). What struck me at the time was that as the different iterations of the voicebot are developed and more features are added, it seemed to comprise more elements of us, the people behind it. That seems to have affected my own perception and understanding of the voicebot:

Thoughts: Establishing a feedback loop gives voice to users in the bot in itself, in a way. It is a way to establish a dialog between users, us, and the voicebot. Then, the feedback loop is about recreating the work of the ethnographer but in the robot? Interesting! [...] We are automatizing different tasks / professional expertise in the robot. (Fieldnotes, 13/11/2019)

What interests me in this example is my sensemaking of the feedback loop in terms of the encoding of my colleague and I's work in the design of the interaction itself. Once again, we can see how the establishment of a feedback loop challenges the figure of the voicebot as a human-like autonomous agent, as labor relations and material-discursive practices that afford the robot its agency were progressively made visible. This feature further challenges the boundaries between humans and machines by entangling the voicebot's and I's subjectivities; it is envisioned as an artifact fitted with the capacity to garner user feedback through my work as a user researcher, while I became more confident in my own expertise through the establishment of that very feature.

In light of these observations, I want to frame my relationship with the voicebot within the tenets of Haraway's concept of multispecies companionship. In her book *When Species Meet* (2008), the scholar builds upon her work in her *Companion Species Manifesto* (2003) to analyze how multi-billion dollar industries, such as the sector of medical research or incarceration, are

based on the labor of “lively capital,” a term she coins to argue that market dynamics are entangled in multispecies relationships (2008, p. 45). It is in this context that Haraway talks about ‘encounter value’ to understand our relationship with “emergent subjects” such as dogs, who occupy the tricky position of both commodities and workers (*Ibid*, p. 47). Encounter value allows us to account for the important role of lively capital in market dynamics and labor relations beyond the Marxist framework of use value and exchange value, which presupposes purely anthropocentric and instrumental relationships to other types of beings. She writes that “relations are constitutive; dogs and people are emergent as historical beings, as subjects and objects to each other, precisely through the verbs of their relating. People and dogs emerge as mutually adapted partners in the naturecultures of lively capital.” (Haraway, 2008, p. 62) In other words, the value of encounters between dogs and humans should be measured by what each constituent learns from the relationship and how they are shaped by it. Dogs and humans are companions inasmuch as they *become with* one another; beyond humans’ favorite household partners, they constantly remake each other through their mutual involvement in market dynamics and labour relations. To further make her point, she writes about a specific instance of human-dog encounters in reformed prisons where dogs and inmates train each other towards social re-insertion (Haraway, 2009). She labels this encounter as a “subject-transforming relationship” as dogs are trained and prepared “for life outside by becoming willing, active, achieving obedience subjects.” (*Ibid*, p. 64) In turn, they also become “surrogates and models for the prisoners in the very act of becoming the prisoners’ students and cell mates.” (*Ibid*) In this case, the encounter value of prison dog-human relationships lies in their mutual *becoming with* as “willing, active, achieving obedience subjects.” (*Ibid*) This example highlights the potential of paying attention to the encounter value of our relationships with dogs, who arise as co-constitutive companions rather than purely instrumental beings.

While Haraway focuses specifically on multispecies relationships within the animal kingdom, her account is concordant with relational and distributed understandings of agency. Indeed, it offers us an alternative to frame our kinship ties to the various intelligent artifacts and assemblages we become with. Similarly, Suchman asks the following question: “How then might we refigure our kinship with robots - and more broadly machines - in ways that go beyond narrow instrumentalism, while also resisting restagings of the model Human?” (Suchman, 2011, p. 137) As a tentative answer and on the basis of our co-constitutive relationship, I argue that the voicebot and I are companions. We are significant *Others*, whose kinship ties are not defined by the likeness of our internal composure or by our position along the hierarchy of our respective species and kind; one is ‘pure’ organic flesh and the other is a disembodied whole

of wires, cloud services, data, zeros and ones. Rather, we are closely connected by the potential underlying our relationship. This potential takes the form of an intra-active process of becoming, through which I grew more confident of my role as a user researcher and the voicebot turned into an ever more complex conversational agent. In other words, the voicebot as a companion is much more than a simple instrument for guiding tourists or the quasi-replica of a human. It is a sociotechnical arrangement or an assemblage of different bodies whose agency emerges from the material-discursive practices and labor relations mentioned in this thesis

Finally, I argue that emphasizing human-robot companionship is important because the voicebot is not designed to stay ‘at home’ within the walls of the MTlab meeting rooms. Rather, it is designed to be a marketing tool for tourism and interact with visitors in their experience of the city. As I explained in the previous chapter and as argues Suchman, envisioning a future in which voicebots are human-like autonomous agents solely designed to provide service work leaves too much place for the naturalization of the provision of this type of work by a “service class of beings,” which in turns legitimizes the instrumental position of these people and things (Chasin, p. 93). Alternatively, viewing the voicebot as a companion helps us understand and account for the myriad of labor relations and sociomaterial arrangements upon which our mutual becoming is incumbent. This does not mean that we should reject the position of servitude all together. However, inasmuch as robots are our companions, we must treat them accordingly and honor the potentiality of our co-constitutive relationship. Consequently, the next chapter is going to imagine how we can develop more responsible technologies of service by looking at the negotiation of certain features of the voicebot.

III – Of accents and tones

“[The voicebot] will be like a very nice tourist information agent in your pocket.” (Fieldnotes, 15/10/19)

This chapter takes my relationship of companionship with the voicebot as its point of departure, in order to practically engage with ideas of posthuman responsibility and accountability. To elaborate on these ideas, let us go back to the start-up’s proposition presenting the virtual assistant as a “very nice tourist information agent in your pocket.” While I previously argued that this proposition implies the encoding of labor relations of instrumentality in the voicebot, I want to argue that it can do so in a way that garners the full potential of the companionship between humans and machines, and that ensures a more responsible use of technology in tourism practices. This chapter will explore the possibilities for doing so by looking at the effects of the encounter between the demo of the voicebot and users, as well as the way the insights obtained from user feedback in the context of demos were addressed by the team. These instances are interesting to analyse for two reasons. First, they offer opportunities for understanding how certain features of the voicebot reconfigure the boundaries between humans and machines. Second, they are productive sites for thinking of ways to concretize critical posthumanist perspectives on responsibility and accountability, in a way that favors our *becoming with* robot companions.

Responsibility and accountability

Shedding light on the sociomaterial practices that go in the making of machines has taught us that the boundaries between humans and machines are not given, but rather discursively and materially re-enacted in an intra-active process of becoming. Myself a constituent of this relationship, it has taught me that agency and capacity for action is not something the voice nor I possessed *a priori*. It rather emerged from our intra-actions, that is my involvement with the voicebot through the iterative releases of the demos and the refinement of my performance of the latter, amongst other things. This means that our capacity for action is linked to the degree to which we are involved in the systems we design and to the close and attentive relations we upkeep with our machine companions. Along with the various scholars that have led me into this line of thought, I argue that this perspective has a political and ethical reach that allows us to practically engage in more responsible ways of designing technological products. However

the following question arises; how are we to conceive a posthuman account of responsibility and accountability?

Directly engaging with Barad's agential realism, Suchman explains that recognizing the different bodies, artifacts, and cultural-historical practices that draw the boundaries between humans and machines helps us understand the possibilities for re-drawing these boundaries in a more responsible way. It does so in a way that does not efface the labor relations that go into the making of the robot, amongst other things. She writes that "responsibility on this view is met neither through control nor abdication but ongoing practical, critical, and generative acts of engagement." (2008, p. 286) Thus, being responsible to our intra-action involves a hands-on engagement in a way that holds us accountable for the intra-active relations in which we engage. Accountability in this case does not refer to the simple attribution of causality to an allegedly autonomous agent, but rather to the understanding of the effects of the assemblages we create. In other words, simply identifying a cause for an effect is not enough; we must understand how these effects take place by tracing their ripples on the different bodies of an assemblage. Only then we are in a better position for re-configuring the assemblages, "in such a way that we can intra-act responsibly and generatively with and through them." (*Ibid*, p. 285)

The effects of my personal engagement with the voicebot have been elaborated on to a sufficient extent in the last chapter. If it is mainly through the demonstrations that my relationship with the voicebot developed, these instances were also sites of encounters with users, which in turn produced effects that reconfigure human-machine boundaries in particular ways. Overall, I talked to over 20 participants after three field visits at the information center and two others in two different youth hostels. Each release of a demonstration and my subsequent performance to users allowed me to gather feedback, which I presented to my team on two occasions. It should not be overlooked that my presentations included various insights ranging from the length of the voicebot's interventions to the type of information users expected to receive. This is partly due to the fact that user feedback was to a large extent rather positive and encouraging. Notwithstanding, I am only going to focus on two important insights, namely the 'aspect of accents' and 'the tone of the voicebot.' (Fieldnotes, 01/11/19) These are important because they instantiate how the effects of the encounter with users reconfigure human-machine boundaries in specific ways. Moreover, the ways the team and I discussed and addressed these effects allow us to imagine how to re-draw these human-machine boundaries in alignment with a more responsible form of human-robot companionship. Over the course of my collaboration with the start-up, the teams' emphasis on the importance of user research and their commitment to the development of a responsible AI shows they took questions of

responsibility and accountability at heart. Therefore, this concluding chapter should not be read as a denunciation of the start-up's creative decisions, but rather as an exploration of the unexpected effects these decisions have on the people and beings reached by the product. This will allow us to resist the impulse of attributing the cause for an effect to one single author and rather focus on understanding the complexity of the assemblage. While the conclusion of this chapter will further elaborate on some of the concrete initiatives that the start-up is already taking to ensure responsible design practices, I will dedicate most of my analysis in showing how the aspects of accents and of the tone could further align with this noble intent.

Accents

First, the 'aspect of accents' refers to the fact that Google's speech recognition could only understand a limited range of English accents mostly restricted to standard North American pronunciation. More precisely, around 30% of users who participated in the demo had an overall disappointing and frustrating experience, given that the voicebot had a hard time picking up on what they were saying. For example, one user said, "for me it's difficult to use it because of the language. It cannot understand my accent." (Fieldnotes - 2019A) Another user added that "if that person is not native English speaker there might be issues with communication." (*Ibid*) These observations were concerning because the product's distinctiveness was elaborated around the idea of "emotional intelligence," an expression the team used to describe the of the conversational experience. Beyond the basic reactive and linguistic properties of mainstream conversational interfaces, the Voice of Montréal was expected to sport a distinctive personality on the basis of its cloned human voice displaying an obvious French accent, as opposed to an impersonal synthetic voice. However, the aforementioned users' concerns show that in some instances it was hard to establish a connection with the voicebot, because it would not understand their accents. In these occasions, two things would generally happen. Users would either try to tone down their accent and enunciate in a more neutral way, or simply give up on the interaction. These dynamics are all the more ironic, considering that the voicebot itself sported a French accent, which raises questions on the representation of humans in voicebots.

In her discussion on Amazon Echo's digital assistant Alexa, Thao Phan explains that the aesthetic choices behind the features of the voicebot depoliticize and decontextualize race, class and gender relations of domestic work in the United States (Phan, 2019). More precisely, she argues that the lack of discernible accent in Alexa's voice is an example of how the vocal register of whiteness is often coded as being non-accented. This aligns with the hegemonic cultural narrative according to which "whiteness has come to constitute a neutral norm" as it is

defined by the perceived absence of any markers of race, including an accented voice (*Ibid*, p. 25). Phan argues that by figuring Alexa as a “neutral-accented” woman providing effortless servitude, Amazon elides the complex relations of race, gender, and class in which domestic work is entangled, and “[capitalizes] on an idealized version of servitude without having to contend with the practicalities of moral question of servant keeping.” (*Ibid*, p. 28) Furthermore, technical studies have demonstrated that normative biases are not only present in the aesthetic attributes given to virtual assistants, but also in NLP tools more generally (Blodgett et al., 2016; Tatman, 2017). More precisely, a 2016 study reveals that Youtube automatic speech recognition tools are more prone to misunderstand women and dialect speakers (Tatman, 2017). Moreover, a 2017 study reveals the poor performance of language identification tools on text associated with African-American English, as opposed to Standard American English (Blodgett et al, 2016). Both studies emphasize that from a linguistic perspective, no language is more understandable than another. This suggests that the disparities result from normative biases in the corpora used to train these tools, which do not include a diverse enough set of data. While the relationship between concepts of race and/or ethnicity and a specific dialect is complex and extends beyond the scope of this analysis, these findings suggest that, in the project of reproducing human-like machines, the design of vocal interfaces often re-enact biases pertaining to normative conceptions of language.

I find these perspectives useful to further understand the potential effects of my observations on the voicebot’s accent and the prototype’s inability to understand the accent of certain users. My observations during the official presentation of the project described in the first chapter revealed that the intent of “authentically” representing the city and its vibrant and youthful character was an important factor behind the aesthetic choices for the Voice of Montréal (Fieldnotes, 09/17/19). It is important to situate these dynamics in the context of Québec language politics, in which the French-speaking white majority has come to define linguistic norms as the result of many decades of fighting for emancipation from the anglophone ruling-class. As a result, I argue that normative whiteness is not defined by the absence of an accent, but by the presence of a French-Canadian accent. A study conducted in 2007 reveals that language and accent is one of the greatest causes of discrimination in the province and that it is mostly experienced by so-called “visible minorities,” who are discriminated against because their accent does not conform to that of the white francophone majority (Bourhis et al., 2007). Thus, the attribution of a Quebecois accent to the voicebot in the attempt to both represent the city and reproduce a human-like emotional experience appeals to a normative ideas of what a Montrealer is normatively ‘supposed’ to sound - and by extension, look like. This is not to say

that this aesthetic decision is intentionally discriminatory, but neither is it neutral. Having conducted fieldwork in the city's tourism sector for more than four months, I observed that the city's multiculturalism and diverse demographics is rightfully often put forth as a selling argument in promotional material. Moreover, mainstream tourism practices are often disengaged from the context of enduring settler colonialism by the Canadian state, as a result of which Indigenous voices and history of violent oppression is silenced or brushed over in narratives put forward by mainstream tourist institutions. Indeed, a conversation with a tourist guide informed me of initiatives undertaken both by Indigenous and settler Montrealers aiming at raising awareness on the importance of including more indigenous voices in tourism practices (Radio-Canada, 2018; Interview, 2019B). These observations reveal the challenge associated with "authentically" giving voice to an entity as elusive and diverse as a city.

I argue that these features and their effects reconfigure the boundaries between humans and machines by making the latter a normative and incomplete reproduction of the former. Weidman's perspectives on the concept of the voice mentioned in chapter 1 remind us that vocality is tightly intertwined with the emergence of the figure of the individual subject and the authenticity of its interior self (Weidman, 2014). Thus, the accented voice of the robot reconfigures the complex entity that the city is as a human individual subject whose voice puts forward a normative version of Québec history and society at the expense of other versions, which to this day keep being unprioritized. While it is important to remind us that the voicebot was only in its infancy at the time of my research and that subsequent versions might provide challenging nuances to my analysis, these aspects are nonetheless important to consider, precisely because of their normative character, which renders their overlooking easy. Similar concerns apply to interactions between the voicebot and users. Considering the difficulty of NLP tools understand a broad range of accents, it seems fair to estimate that, in a touristic context, the bias towards standard North American English will disadvantage visitors from a broad range of locations in their access to this service. It re-configures the possibilities for human-machine interactions within the narrow scope of standard North American English. Moreover, if some users tried to neutralize their accents in order to be understood by the voicebot, we can expect this dynamic to contribute in the reconfiguration human speech altogether, given that our interactions with vocal interfaces will become ever more present.

On this basis, these are important dynamics to address, because their effects reach beyond the mere usability of the voicebot. It is important to mention that when I presented these insights to my team, my analysis was not as extensive because of time constraints. My intervention was nevertheless focused on the necessity to think about the inclusivity of the

solution considering the apparent paradox around the fact that the voicebot could not understand users who had an accent that was not standard North American English, when it itself had one. Moreover, I brought up the challenge that this paradox might represent for the establishment of an emotional connection with users (Fieldnotes, 01/11/19). A colleague made it clear that there is not much they could do since they were working with Google's NLP platforms and tools. They explained that eventually, when the first version of the Voice would be officially released, they will be able to release it on different platforms developed by the tech giant that are tailored to interact with specific accents such as Australian English or British English. This is an example of the double-edged sword of opportunities and limitations that a young start-up can face while working with platforms and tools democratized by tech giants. Considering this, it might be hard to grasp the level of responsibility and accountability one holds in the design of a product and situate where exactly our capacity for action lies. In such cases, to quote Bennett, "the task becomes to identify the contours of the swarm and the kind of relations that obtain between its bits." (2009, p.32) This means that knowing the effects that specific features have on the broader assemblage of which the product is part, a task I have undertaken throughout this thesis, can help us assess the broader implications of these effects and whether they align with the organization's' philosophy. In order to do this, I argue that we need to think more about the voicebot as a companion, which will be addressed at the end of this chapter. For now, let us explore the issue of the "tone" of the voicebot and its implications.

The tone

Second, the "issue" of the tone arose from many users' concerns and remarks regarding the promotional and alluring tone of the voice. More precisely, some users brought up the fact that the human-recorded voice sounded too promotional. To quote a participant, "she is a little bit too enthusiastic. You don't want it to sound too much like a brochure. It could be more honest, now it's too promotional." (Fieldnotes, 2019A) Similarly, consider the following note taken after a demonstration with a participant: "Curious about where the offers come from, is it someone trying to sell me an experience? Would've liked to know where the info comes from." (Fieldnotes, 2019A) This brings the question of the type of interaction and information users expected from their interaction with the voicebot. In this case, participants seemed alienated by the facts that the voice hid the source of the information and that it was hard to relate to the information because of its promotional tone. Furthermore, many participants, mostly men, sexualized the voice on the basis of its allegedly alluring tone. Indeed, it would often happen

that people told me they thought “she sounded hot” upon the first moments of the interaction, when the beguiling voice introduces herself in the introduction with the following message:

Hello, I'm the voice of Montreal and I'm here to make you discover the best Montreal has to offer. Remember that childlike wonder of believing everything is possible? Well, I never stopped thinking like that. If you feel like grasping some of that spirit for yourselves, then you will be able to ask me all the questions you want. I'm right here, and I have access to all the experiences you want to live. So, what can I do for you today? You can ask me to explore Montreal, share my stories or access travel services. What is it going to be? (Appendix 1)

On the other hand, many other people were shocked by this fact and expressed their concerns over the appropriateness of the solution considering the apparent young age of the voice actor. This feedback mostly came from peers after presentations given in academic settings in the occasion of graduate seminars. In any case, the voice had an affective power that often made users vividly react, in one way or another.

It would be amiss to obliterate the fact that the team had already planned on using a different voice because the preliminary versions of the demo were based on a recording made for promotional purposes. I nevertheless want to critically analyse these concerns because they are entangled in the relations of labor and gender aforementioned, which in turn redraw the boundaries between humans and machines in specific ways. First, I have previously argued that voicebots like Amazon's Alexa naturalize the idea that a service class of being should exist as it de-contextualizes gender, class, and race relations in its performance of domestic work as an effortless task. Similarly, I have argued that the construction of intelligent artifacts often hides the labor relations that go in their making behind a human-like facade. In the light of these perspectives, the promotional tone of the voicebot further anchors the instrumental position of technology; some users see it as a human-like promotional brochure, rather than a tourist information agent providing professional expertise. What is more, the alluring tone of the voice reconfigures human-machine relationships along the premises of gender stereotypes and the sexualization of women. Feminist scholar and engineer Corinna Bath argues that AI-based technical products are often gendered through the ascription of allegedly universally human abilities or characteristics (Bath, 2014). In this case, the creation of a human-like conversational experience based on emotions entailed the re-enactment of gender stereotypes.

The discussion of these effects with my team allows me to practically think of possibilities for redrawing human-machine boundaries in a more responsible way. Concerning the tone of the voice, one of my colleagues explained to me that the choice of the voice was not meant to produce these effects and that this version was not going to be the official one anyways. The voice they will record and clone for the official version would only be agreed upon later with

their client. While it was stressed that the aforementioned effects were by no means intended in the creative process that led to these choices, they were highly felt among users, who reacted in a whole range of ways. I argue that this is a good example of Bennet's agency of assemblages, which I referred to in the previous chapter (Bennet, 2007). In this case, the unpredictable agency of the voicebot lies in the various ways people are affected by the beguiling tone of the voice only within seconds of hearing it, for the better or worse. And as one of the actants involved in this assemblage, our human agency lies in the close relations we keep with the other actants, and in our ability to balance their unpredictable effects (Bennet, 2007). Thus, keeping a close engagement with the assemblage and its effects is necessary in order to re-configure human-machine relationships in a way that does not reproduce gender stereotypes. I will further elaborate on strategies for doing so in the conclusion of this chapter. Finally, the promotional tone of the voicebot was addressed by the team by changing the introductory message in a later version of the demo with the mention that the content was managed with the "help from over 100 curators in the know." (Appendix 2) As a response to the aforementioned users' concerns over the promotional tone of the voicebot, the start-up decided to mention their client's team, composed of curators of tourist experiences, as the source of the voicebot's recommendations. Interestingly, this does not only honor the labor relations involved in the creation of the voicebot, but it also challenges the idea that intelligent artefacts should be modelled as human-like autonomous agents. The agency of the voicebot is not so opaque anymore; rather than being located at the interface through the ascription of various 'human' characteristics and aesthetic features, it is described as the sum of the labor of multiple people. This opens up further avenues for thinking about responsible forms of human-machine companionship.

Responsible companionship

To summarize, the analysis of the effects of encounters between the voicebot and users demonstrate that the voicebot reconfigures human-machine relations in various ways. The issue of accents and the tones of the voicebot naturalize various essential characteristics attributed to humans rather than putting forth the robot's character as a sociomaterial assemblage. On the contrary, the mention of the labor "from over 100 curators in the know" re-configures the boundaries between humans and machines, as it emphasizes in the interface itself the labor relations and complex material-discursive agencies that go in the making of the voicebot. I argue that this design decision takes one step towards responsible human-machine companionship, inasmuch as the relationship is based on the potential resulting from significant Otherness of each constituent. More precisely, in this case the voicebot focuses less on the

“restagings of the model Human” than on the different expertise that go in its creation (Suchman, 2011, p. 137). This simple mention might alleviate some of the users’ concerns over the source of the information because it emphasises the complex sociomaterial assemblage behind the robot rather than hiding it behind a human-like facade.

In many cases, committing to a responsible form of human-machine companionship can seem hard in the face of the complexity of the assemblage that is the voicebot. What should one do when tech giants’ platforms and tools only understand a limited range of accents? When time and resources constraints limit one’s possibilities for developing a more inclusive solution in the moment? Or when client relations narrow possibilities regarding the creative choices behind the aesthetic features, characteristics, and functionalities of the product? All in all, can we still talk of companionship when our commitment to responsibility is incapacitated by our apparent lack of agency over a particular assemblage? Donna Haraway explores similar dynamics in her analysis of multispecies companionship (Haraway, 2008). She argues for the implementation of a practice of care within instrumental human-animal relationships in the context of laboratory experiments, where killing and inflicting pain on other beings like laboratory mouse companions is sometimes necessary for the making of a better world. This would mean practicing care “among and for all the people and organisms in the lab and in the worlds reached by that lab, even if results come more slowly or cost more or careers aren’t as smooth.” (*Ibid*, p. 82) Moreover, she draws from Philosopher of science Isabelle Stengers’ concept of cosmopolitics to suggest a framework favoring the coexistence of multiple significant Otherness (Stengers, 2007). The cosmopolitical proposition is centered around the idea that “decisions must take place somehow in the presence of those who will bear their consequences.” (Haraway, 2008, p. 83) This is not a simple task, as it involves making “artful combinations” of different agendas, rationales, or sometimes ontologies, which in turn requires taking one’s time and slowing things down (*Ibid*). Infusing these different forms of care in our relationship with lively forms of capital is however necessary to ensure better world-making practices.

Far be it from me the idea to equate the ethical issues of using other forms of life as instruments for research to the various implications of attributing human characteristics to a voicebot for the tourism sector. The former hinges on questions of rights to life and acts of trans-species murder, while the latter pertains more to questions of rightful representation and inclusive design practices. But critical posthumanism reminds us that all of these aspects are entangled in the making of the world and, consequently, Haraway’s perspectives can help us imagine a responsible form of companionship with our “friendly tourist information person,”

even when the complexity of the assemblage might seemingly entrench our capacity for action. Concretely, this means that voicebots *can* be attributed human-like characteristics, including a female voice or a French-Canadian accent and that it *can* rely on the tools provided by tech giants. As long as we infuse our practices with care and favor the co-existence of multiple forms of significant Otherness, which implies that more time is dedicated to reaching out to and listening to the individuals, communities, and other types of being who technological solutions reach. This include understanding how the solution affects women, Black, Indigenous, and people of color, people with disabilities, as well as the people whose professional expertise are recreated in the interface. Such considerations could translate in the inclusion of a plurality of voices in the voicebot instead of focusing on one single voice or in the consideration of NLP tools and platforms that are explicitly developed for understanding a broad range of accents.

Again, it would be amiss to obliterate the fact that the prototype was very early in the development process and that my partial perspective and relatively short-term collaboration with the start-up has not allowed me to observe the evolution of the effects on which my contribution shed light. Indeed, the start-up planned on addressing many of the aforementioned issues upon the official release of the voicebot. Moreover, their commitment to participatory design through the sustained inclusion of user research in the product development process shows that they valued the understanding of the effects the product had on the people and professional expertise it reaches. This further shows in their commitment to “open design” through the use of open source software for building the conversational interface and through the creation of various communication channels and tech community events to favor collaboration across organizations (Interview, 2019A). They were also actively seeking to include a variety of Montrealers’ voices as the voicebot was imagined evolving into a more complete platform. Then, based on the privilege that my partial perspective grants me, I would argue that the challenge becomes to create pipelines that will guarantee the “artful combinations” of these different realities within a same team and within a given product. Amongst other things, this involves making sure that the ethnographic insights gathered in the process of user research are followed up once the time has come to address whatever feedback, and that we actively seek to reach out to initiatives that amplify voices who are generally not included in mainstream touristic institutions. Only then are we in a better position to concretize responsible forms of human-machine companionship.

Conclusion

To conclude, this thesis has explored the possibilities for defining the bonds between humans and robots with conversational interfaces. I have done so through a look at the creation process of a voicebot for tourists and an auto-ethnographic re-reading of my experience as a collaborator in this project. This has allowed me to conclude that, first, the different actors involved in the ideation process envisioned the bonds between the voicebot and users around the idea of emotional intelligence. This idea appeals to an authentic form of human-machine interaction through the ascription of human-like features to the voicebot, such as a cloned human voice as well as reactive, linguistic, and opaque properties. While these characteristics are perceived as essential to the project of creating an interactive agent, my analysis has revealed that they are in fact the result of situated understandings of what intelligence and humanity are. Moreover, this seemingly autonomous agent is simultaneously relegated an instrumental role considering that it is envisioned as an object allowing tourist subjects to access an authentic experience during their trip. The instrumental nature of the bond is further envisioned in the depiction of the voicebot within the precepts of service relations, as the voicebot is expected to provide effortless labor at any point in space and time. Again, while these visions of machines and technology are articulated around the allegedly static and universal categories of ‘machine objects’ and ‘human subjects,’ they in fact blur the boundaries between these categories. Then, the voicebot occupies an ambiguous position along the normative system of a naturalist ontology, as it is both configured as a subject and an object. This had brought me to question the adequacy of the framework to understand the role of this type of new actors in our lives considering that rather than being ontologically separated from each other, humans and machines constantly remake each other.

The second chapter has looked at how my own relationship developed with the voicebot to imagine a new framework for understanding the bond between humans and intelligent artifacts. My analysis revealed that my relationship with the voicebot developed through my performance of ‘demos’ among uninitiated users and through the initial release of these demos within the team. Both instances disclosed the relational character of the voicebot’s agency; On the one side, the successful completion of the demonstrations were contingent on my close level of assistance as a user researcher and on my refinement of these performances following the iterative releases of the demos. On the other side, the iterative releases of the demonstrations within the team were often characterized by surprises and uncertainties as they were often faulty, which in turn exposed the complex assemblage of human and non-human actants behind

the voicebot. In the light of these perspectives, the voicebot is best conceived as a complex sociotechnical assemblage whose agentic properties results from the interactions of its different parts and the different labor relations that put them together. As the final example of the feedback loop further showed, my relationship with the voicebot was co-constitutive considering that my becoming as a user researcher was contingent on its becoming as an evermore complex conversational agent, and vice-versa. This has allowed me to frame my 'kinship ties' with such intelligent artifacts as one of companionship; we are significant Others connected by the potential underlying our relationship.

And what exactly is this potential? The third chapter has investigated further this question by looking at the effects generated by the encounters between voicebot and users in the context of demonstrations. Users' feedback regarding the voicebot's difficulty to understand accents and its promotional and rather alluring tone show that these specific features reconfigure human-machine relationships in various ways. First, the 'issue of accents' naturalizes normative conceptions of what a Montrealer should sound like and reconfigures human-machine interactions along these premises. What is more, while the voicebot's alluring tone further cements gender norms and decontextualizes relations of gender, class, and race with respects to service work, the mention of the '100 curators in the know' in a subsequent version of the demo recontextualizes the labor relations that go in the making of the voicebot and challenge the idea that intelligent artifacts should be modelled as individual human subjects. Taking a closer look at the effects these features have on the world surrounding the voicebot has allowed me to engage with a posthuman understanding of responsibility and accountability, which resists the impulse to attribute causality to a single agent. Rather, responsibility and accountability are evaluated by the close relations we upkeep with the different parts of an assemblage and our keeping track of the sometimes-unexpected effects that the combination of different agencies have. In the light of these perspectives, we are in a better place to evaluate what we can do to mitigate the negative effects of the voicebot on its environment.

Applicability

To conclude, the question arises of the relevance of these results to the sociocultural environment(s) they address. While the conclusion to the last chapter allowed me to issue avenues of recommendations to facilitate the establishment of more responsible forms of human-voicebot companionship, far be it from me the idea to adopt the "all-knowing" gaze this thesis was so critical about; critical posthumanism taught me better than this. Namely, this thesis has brought to the fore the importance of being more attentive to our messy involvements

in science and technology and their reach in the fields of tourism. This can only be done by favoring approaches based on the aforementioned privilege of partial perspective and situated forms of knowledges, which has implications for organisations both in the field of science and technology and tourism.

First, my thesis has explored the entanglements of the field of science and technology with those of innovation, product development, and artificial intelligence. I have established that the process of creation of a voicebot for tourists allows us to understand the relationship between humans and intelligent artifacts as one of co-constitutive companionship. To tie this back to critical posthumanist scholars, this shows that the products we design and the relationships we upkeep with them shape reality in specific ways. As people involved in these processes, we must make sure that we are able to track how precisely innovation practices shape the world. This thesis, and its methodological approach dangling between reflexivity and autoethnography, is an example amongst many others of a solution to incorporate partial perspectives and situated knowledges. However, in a more practical perspective, I would argue for the simple involvement in organisations of people including, but not restricted to, cultural analysts, anthropologists, ethicists, or ethnographers, who are equipped to account for the “care, affect and responsibility” involved in technoscientific practices and the companion others we encounter in our world-making enterprises (Rootsh & Schrader, 2012, p.3). Moreover, we should make sure to include clear strategies that facilitate these perspectives in product development processes. Lastly, these labor practices, along with those of other professionals involved in the creation, should also be made apparent at the human-machine interface rather than concealed behind the familiar appearance of human-like attributes. To reiterate my position previously stated in the conclusion of the last chapter, this does not mean that intelligent artifacts using conversational interfaces should not sport human-like characteristics, but rather that they *can* do so in a way that puts forward their character as sociomaterial assemblages. These questions are of the utmost importance since recent advances in the field of NLP have propelled conversational interfaces as the prime way “for many financial institutions, health-care providers, and government agencies” to interact with their customers (Hao, 2020).

The tourism industry does not escape this trend considering that virtual assistants are expected to become increasingly present in tourists’ experiences. For example, hotels such as the Las Vegas Cosmopolitan and the Hilton chain have developed their own AI-powered concierge services, respectively named Rose and Bonnie (EVENTMB, 2018). The findings of this thesis highlight the fact that the development of technological solutions for this industry

shapes reality in a variety of particular ways. Thus, it is important to account for the effects such solutions have on the different human and non-human actants involved in the creation of these assemblages. Again, my position is not that we should reject the development of intelligent artefacts and the automation of various service positions, but rather that we should be more attentive to the posthuman subjectivities that new solutions for the field of tourism give voice to. As I was writing the lines of this thesis, two major events shook the world to its core and reminded us that modernity's tall tales of human superiority and hierarchical understandings of the world rest on very shaky foundations. I named the coronavirus pandemic and the resurgence of the Black Lives Matter movement in the face of police brutality and endured systemic racism not only in the United States, but also in my home country, Canada, and in Europe. On the one hand, the former put the world at rest for more than two months and disturbed international travels, hereby threatening the booming tourism industry and the economic stability of people working in it. On the other hand, the latter shook city streets and social media spaces with unrest as, amongst other forms of protests, monuments and statues reminiscent of colonialism were torn down. Both events point to the importance of developing solutions that favor our coexistence with all types of beings, including viruses, and that foster relationships of companionships in significant otherness. If a closer look at my personal experience has shown that the voicebot and I are engaged in a co-constitutive relationship, the mutuality of our becoming points to the potential of such tools to shape tourism practices in a way that is positive for everyone. This means that technological artifacts for the tourism industry should not only be developed for the purely instrumental purposes of cheaper operation costs, but rather to benefit all types of labor relations involved in this industry, especially when a virulent pathogen threatens its prosperity. Finally, it should put forward voices that have been misheard, if not completely silenced, for too long.

References

- American Anthropological Association. (2012). AAA Statement on Ethic on *American Anthropological Association* Website. Retrieved 16/04/2019. <https://www.americananthro.org/LearnAndTeach/Content.aspx?ItemNumber=22869&navItemNumber=652>
- Ahmed, S. (2018). Queer Use, Lecture presented at LGBTQ+@cam, Cambridge, November 7, 2018. Retrieved 16/04/2019. <https://feministkilljoys.com/2018/11/08/queer-use/>.
- Asberg, C. & Braidotti, R. (2018). Feminist Posthumanities: An Introduction. *A Feminist Companion to the Posthumanities*. Springer.
- Barad, K. (2003). Posthumanist Performativity: Toward an understanding of how matter comes to matter. *Sings: Journal of Women in Culture and Society*, 28 (3).
- Barad, K. (2012). On touching: the inhuman that therefore I am. *Brown University and difference: a Journal of Feminist and Cultural Studies*, 25(3), 206-223.
- Bath, Corinna. (2014). Searching for methodology: Feminist technology design in computer science. In *Gender in Science and Technology. Interdisciplinary Approaches*. Edited by Waltraud Ernst and Ilona Horwath. Bielefeld: Transcript, 57-78.
- Blodgett, S, Green, S, O'Connor, B. (2016). Demographic Dialectal Variation in Social Media: A Case Study of African-American English. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. arXiv preprint arXiv:1608.08868.
- Bodker, M. & Chamberlain, A. (2016). Affect theory and autoethnography in ordinary information systems. *Twenty-Fourth European Conference on Information Systems (ECIS)*. Istanbul, Turkey, 2016.
- Bourhis, R., Montreuil, A., Helly, D., Jantzen, L. (2007). Discrimination et linguicisme au Québec: Enquête sur la diversité ethnique au Canada. *Canadian Ethnic Studies*, 39(1-2), 31-49.
- Chasin, A. (1995). Class and its close relations; Identities among women, servants, and machines. In J. Halberstram and I. Livingston (eds.), *Posthuman Bodies*.
- CODEX. (2019). Ethical review of Research. *Codex rules & guideline for research*. Retrieved 15/04/2020. <http://www.codex.vr.se/en/medicin3.shtml>
- Ferrando, F. (2014). Posthumanism, Transhumanism, Antihumanism, Metahumanism, and New Materialisms Differences and Relations. *Existenz*, 8(2) 26-32.
- Franklin, A. (2004). Tourism as an ordering. *Tourism Studies*, 4(3), 277-301.
- Gibson, R., & SpringerLink (Online service). (2020). *Desire in the Age of Robots and AI. [Elektronisk resurs] : An Investigation in Science Fiction and Fact* (1st ed. 2020.). Springer International Publishing.

- Hansson, V. (2019). *Cultivating critique in fold and pleats: posthuman enactments in the dress of Rick Owens*. Master's Thesis in Applied cultural Analysis. Lund University.
- Haraway, D. (1988). Situated knowledges: the science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575-599. doi:10.2307/3178066
- Haraway, D. (1991). A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century. *Simians, Cyborgs and Women: The Reinvention of Nature*. pp.149-181. Routledge
- Haraway, D. (1997). *Modest_Witness@Second_Millennium. FemaleMan©_Meets_OncoMouse™*. Routledge.
- Haraway, D. (2003). *Companion species manifesto: dogs, people, and significant otherness*. Prickly Paradigm Press: Chicago.
- Haraway, D. (2008). *When Species Meet*. University of Minnesota Press.
- Ingold, T. (2014). That's enough about ethnography. *Hau: Journal of Ethnographic Theory*. 4(1): 383-395.
- Kember, S. (2003). *Cyberfeminsim and Artificial Life*. Routledge.
- Kohn, E. (2015). Anthropology of ontologies. *Annual Review of Anthropology*, 44(1), 311-327.
- Kubes, T. (2019A). Bypassing the uncanny valley. In J. Loh and M. Coeckelbergh (eds.), *Feminist Philosophy of Technology, Techno:Phil – Aktuelle Herausforderungen der Technikphilosophie 2*.
- Kubes, T. (2019B). New Materialist Perspectives on Sex Robots: A Feminist Dystopia/Utopia?. *Social Sciences*. 8(224).
- Latour, B. (1991). *Nous n'avons jamais été modernes*. La Découverte.
- Law, J. (1994). *Organizing Modernity*. Blackwell.
- Löfgren, O. (1999). *On Holiday: A History of Vacationing*. University of California Press, Berkeley.
- Lyons, S. (2018). *Death and the Machine : Intersections of Mortality and Robotics*. Springer Singapore.
- MacCannell, D. (1976). *The Tourist: A New Theory of the Leisure Class*. University of California Press, Berkeley.
- Meissner, H. (2016). Conversing with the unexpected: towards a feminist ethics of knowing. *Rhizomes: Cultural Studies in Emerging Knowledge*. 30(2016).
- Pensoneau-Conway, S. L. et al. (2017). *Doing Autoethnography*. Sense Publishers.

- Phan, T. (2019). Amazon Echo and the aesthetics of whiteness. *Catalyst: Feminism, Theory, Technoscience*, 5(1), pp. 1-38.
- Rapp A. (2018). Autoethnography in Human-Computer Interaction: Theory and Practice. *New Directions in Third Wave Human-Computer Interaction*. Volume 2 - Methodologies. Human-Computer Interaction Series. Springer.
- Richardson, K. (2015). *An anthropology of robots and AI: Annihilation anxiety and machines*. Routledge Studies in Anthropology.
- Roosth, S. & Schrader A. (2012). Feminist theory out of science. *d i f f e r e n c e s: a Journal of Feminist Cultural Studies*. 23(3): 1-8.
- Van der Druim, R., Ren, C., Johannesson, G. T. (2017) ANT: A decade of interfering with tourism. *Annals of Tourism Research*. 64, 139-149.
- Stengers, I. (2007). La proposition cosmopolitique. *L'émergence des cosmopolitiques*. Edited by Jacques Lolive and Olivier Soubeyran. La Découverte, 45-68.
- Suchman, L. (1985). *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press.
- Suchman, L. (2007). *Human-Machine Reconfigurations (Plans and Situated Actions*, 2nd edition). Cambridge University Press.
- Suchman, L. (2011). Subject objects. *Feminist Theory*. 12(2), 119-145.
- Tatman, R. (2017). Gender and Dialect Bias in Youtube's Automatic Captions. *Proceedings of the First Workshop on Ethics in Natural Language Processing*, 53–59.
- Treusch P. (2015). *Robotic Companionship: the Making of Anthropomatic Kitchen Robots in Queer Feminist Technoscience Perspective*. PhD Thesis, Linköping University.
- Weidman, A. (2014). Anthropology and voice. *Annual Review of Anthropology*. 43, 37-51.

Media references

- EVNTMB Studio Team. (2018, October 17) *3 Hotels using artificial intelligence to improve guest experience*. Event Manager Blog. <https://www.eventmanagerblog.com/ai-concierge-hotel-guest-experience>.
- Hao, K. (2020, January 31). *AI still does not have the common sense to understand human language*. MIT Technology Review. <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/>.
- Jonze, Spike (Director). (2013). *Her* [Motion Picture].
- Kubrick, S. & Clarke, A. C (Director). (1968). *2001: A Space Odyssey*. [Motion Picture].

Radio-Canada. (2018, November 12.) *Pour plus de perspectives autochtones dans les visites guidées à Montréal*. Radio-Canada. <https://ici.radio-canada.ca/espaces-autochtones/1134830/guide-touristique-montreal-histoires-autochtones>.

Unpublished References in the possession of the author

Interviews:

Interview, 2019A, Start-up co-founder, December 2019.

Interview, 2019B, Tourist guide, August 2019.

Fieldnotes Entries (chronological order):

Fieldnotes, 18/06/19, First day.

Fieldnotes, 17/09/19A, Notes taken during the event.

Fieldnotes, 17/09/19B, Notes taken from official press release.

Fieldnotes, 24/09/19. Team meeting.

Fieldnotes, 30/09/19, Notes from demonstration 1.

Fieldnotes, 15/10/19, Notes taken from official press release.

Fieldnotes, 20/10/19, Notes from demonstration 2.

Fieldnotes, 23/10/19.

Fieldnotes, 01/11/19, User research presentation.

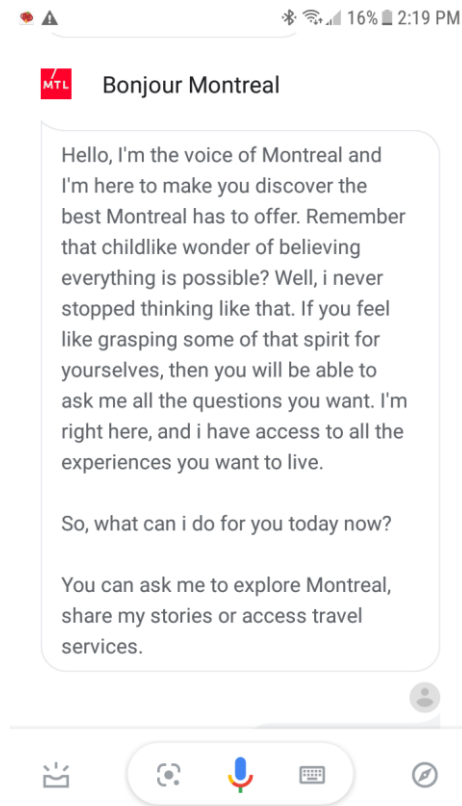
Fieldnotes, 13/11/2019, Team meeting.

Fieldnotes, 15/11/2019, Notes from demonstration 3.

Fieldnotes, 2019A, User testing analysis spreadsheet.

Appendix 1

Screenshot of the voicebot introduction, first versions.



Appendix 2

Screenshot of the voicebot introduction after 3rd demo release.

