Title
Embodied Technology: Human Machine Communication from the Media Arts Perspective

Author
Wolfe, Hannah E.

Publication Date
2019

University of California
Santa Barbara

Embodied Technology: Human Machine Communication from the Media Arts Perspective

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Media Arts and Technology

by

Hannah Elizabeth Wolfe

Committee in charge:

Professor JoAnn Kuchera-Morin, co-Chair
Professor Marko Peljhan, co-Chair
Professor Yon Visell

December 2019

The Dissertation of Hannah Elizabeth Wolfe is approved.

Professor Yon Visell

Professor JoAnn Kuchera-Morin, co-Chair

Professor Marko Peljhan, co-Chair

August 2019

Embodied Technology: Human Machine Communication from the Media Arts Perspective

by

Hannah Elizabeth Wolfe

"Karma is samsara.

Relationship is metempsychosis.

We are in open circuits."

- Nam June Paik

Acknowledgements

I would like to thank the Deustch Foundation, Mosher Foundation, AlloSphere Research Group, Systemics Lab, and Media Arts and Technology Program at University of California, Santa Barbara for their generous support. I am deeply indebted to my advisers Dr. JoAnn Kuchera Morin and Prof. Marko Peljhan. I would like to express my appreciation to Dr. Yon Visell for pushing my work further. I am extremely grateful for my collaborators, Sahar Sajadieh, Sölen Kiralti, Alex Bundy and members of the AlloSphere Research Group; this work would not have been complete without them. I would like to thank Dr. Matthew Wright for suggesting that I research emotive sound. I cannot begin to express my thanks to Pam Wolfe for her copy editing, encouragement, and patience. I also wish to thank Skyler Kasko and Robert Wolfe for their emotional support. Last, I am thankful for my cat, Lady Catterley, for staying by my side and inspiring me.

Curriculum Vitæ
Hannah Elizabeth Wolfe

Education

2019	Ph.D. in Media Arts and Technology (Expected), University of California, Santa Barbara.
2017	M.S. in Computer Science, University of California, Santa Barbara.
2016	M.S. in Media Arts and Technology, University of California, Santa Barbara.
2009	B.A. in Liberal Arts, Concentration in Visual Arts, Bennington College.

Publications

2019	Sajadieh, S. & Wolfe, H. "Come Hither to Me: Performance of a Seductive Robot." Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, 2019.
2017	Wolfe, H., Peljhan, M. , & Visell, Y. "Singing Robots: How Embodiment Affects Emotional Responses to Non-linguistic Utterances." IEEE Transactions on Affective Computing. 2017.
	Adderton, D., Cabrera, A., Kuchera-Morin, J., Kim, K.H., Putnam, L., Peliti, L., Rincon, G.A., Tilbian, J., Wolfe, H., Wood, T., & Youn, K. "PROBABLY/POSSIBLY?: An Immersive Interactive Visual/Sonic Quantum Composition and Synthesizer." Proceedings of the 25th ACM International Conference on Multimedia. 2017.
2015	Rosli, MHW., Yerkes, K., Wood, T., Wolfe, H., Roberts, C., Haron, A., Estrada F., & Wright, M. "Ensemble Feedback Instruments." Proceedings of the International Conference on New Interfaces for Musical Expression (Edgar Berdahl, Jesse Allison, eds.), Louisiana State University, 2015.

Exhibitions and Performances

2019	"Cacophonic Choir" Plugin, Contemporary Istanbul. Istanbul, Turkey. (collaboration with Şölen Kıratlı and Alex Bundy)
	"Touching Affectivity" Installations, New Interfaces for Musical Expression. Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.

	"Come Hither to Me!" Interactivity, CHI Conference on Human Factors in Computing Systems. Glasgow, Scotland. (collaboration with Sahar Sajadieh)
	"Bodies at Work: I Love My Robot and My Robot Loves Me" Multimodal Performance Festival. Glass Box Gallery, Santa Barbara, California. (collaboration with Sahar Sajadieh)
2018	First Thursday, Santa Barbara Center for Art Science and Technology. Santa Barbara, California.
	"Come Hither to Me!" Invisible Machine, Santa Barbara Center for Art Science and Technology. Santa Barbara, California. (collaboration with Sahar Sajadieh)
	"Hydrogen-Like Atom" Sullivan Goss Gallery. Santa Barbara, California. (collaboration with the AlloSphere Research Group)
	"Island Caterpillar." Digital Naturalism Conference, Koh Lon, Thailand. 2018. (catalogue)
2017	"ROVER" Bio-creation and Peace. International Symposium of Electronic Arts. Manizales, Colombia. (catalogue)
	"Hydrogen-Like Atom" Bio-creation and Peace. International Symposium of Electronic Arts. Manizales, Colombia. (catalogue (collaboration with the AlloSphere Research Group)
	"Hydrogen-Like Atom" Computer History Museum. ACM Multimedia. Mountain View, California. (collaboration with the AlloSphere Research Group)
	"Hydrogen-Like Atom" Museum of Exploration and Innovation. Santa Barbara, California. (collaboration with the AlloSphere Research Group)
2016	CREATE AlloSphere Spatial Sound Concert. University of California, Santa Barbara, Santa Barbara, California. (collaboration with the CREATE Ensemble)
2015	Karlsruhe HFG Opening Ceremony. Karlsruhe, Germany. (collaboration with the CREATE Ensemble)
	New Interfaces for Musical Expression Performance, Louisiana State University Baton Rouge. Baton Rouge, Louisiana. (collaboration with the CREATE Ensemble)
	"Carousel of Physics," Santa Barbara Museum of Art. Santa Barbara, California. (collaboration with Dos Pueblos High School Engineering Academy)

“Carousel of Physics,” Maker Faire. San Mateo, California (collaboration with Dos Pueblos High School Engineering Academy)

2015 & 2018 — Media Arts and Technology End of Year Show Concert, University of California, Santa Barbara. Santa Barbara, California. (collaboration with the CREATE Ensemble)

2014 — Imagining the Universe - Music, Spirituality, and Tradition - Network Concert, University of California Santa Barbara and Stanford University. (collaboration with the CREATE Ensemble)

2014, 2015, 2018 & 2019 — CREATE Concert, University of California, Santa Barbara. Santa Barbara, California. (collaboration with the CREATE Ensemble)

2013, 2014, 2015, 2016, 2017 & 2018 — Various Works. Media Arts and Technology End of Year Show, University of California, Santa Barbara. Santa Barbara, California. (catalogue)

2008 - 2010 — “Bennington Bookmarks” Crossett Library. Bennington, Vermont. (collaboration with Robert Ransick, Joe Holt, Ben Choiniere, Jess Funston, Rebecca Grabman, Jason Irla, David Meresman, Luce de Palchi, and Kyle Schroeder)

Conference Posters, Demonstrations and Talks

2019 — Kiratli, S. & Wolfe, H. “Cacophonic Choir,” SOUND::GENDER::FEMINISM::ACTIVISM – TOKYO. Chinretsukan Gallery, Tokyo University of the Arts. Tokyo, Japan, 2019.

2018 — Wolfe, H. “Interactive Emotive Robots.” Presented at the Digital Naturalism Conference, Koh Lon, Thailand, 2018.

2017 — Son, J. & Wolfe, H. “Artist’s Drawing and Perception Teaching System” Demo at Re-habituation. University of California, Santa Barbara. Santa Barbara, California, 2017.

Wolfe, H. “R.O.V.E.R. the Reactant Observant Vacuous Emotive Robot.” Poster session presented at Computing Research Association-Women Graduate Cohort Workshop, Washington, D.C. 2017.

Wolfe, H. “Interactive Emotive Robots.” Invited talk for the MAT Seminar Series. University of California Santa Barbara, Santa Barbara, California. 2017.

2016 — Wolfe, H. “The Reactant Observant Vacuous Emotive Robot: A study exploring emotive response to robots producing sound.” Invited talk at inTouch Health, Santa Barbara, California. 2016.

Abstract

Embodied Technology: Human Machine Communication from the Media Arts Perspective

by

Hannah Elizabeth Wolfe

This dissertation is the study of a series of media art installations that deal with embodied technology which allows humans to interact with computers in different contexts. It explores the social effects of ubiquitous technology and the role for embodied media arts as a critique of interactive digital technologies which are replacing physically present forms of communication with our environment and with each other. In these works I examined the affordances of different physical spaces and observed how the social dynamics were affected. To probe these systems, I am using affective computing (computational expression of emotion) and tangible computing (objects with computational power). Intimate spaces allow for the ability to touch and to hold technology. Here the emotive haptic and emotive sonic response of a robot changed the way that people interacted with and viewed the robot and each other. Whether someone was directly interacting with the robot or simply observing drastically changed the emotion they thought the robot was expressing. Social spaces allow for the ability to interact interpersonally with technology. I explore how interactions change when they are embodied and public, while examining gender roles, female agency, and the line between human and machine. Interactive environments allow for movement and for multiple people interacting with large amounts of information and therefore each other. Personal mobile technology, multimodal interfaces, and distributed interfaces were used to allow multiple people to interact with the system and each other. Through these works, I show that media artworks, which embody tech-

nology and give it emotive qualities, critique the way that technology is used and change the way that people interact with it.

Contents

Chapter 1

Introduction

This dissertation explores the social effects of ubiquitous technology and the role for embodied media arts as a critique of interactive digital technologies which are replacing physically present forms of communication with our environment and with each other. Because these applications are not embodied, the interactions feel less concrete. Social Computing sets a precedent for attempting "to incorporates understandings of the social world into interactive systems" [1]. To understand these systems, I am using affective computing (computational expression of emotion) and tangible computing, where objects in the everyday world are augmented with computational power. I am exploring interaction with embodied technology as a replacement for digital interaction at the social, physical, and public levels. I am using our natural interactions with animals, people and our environment as inspiration for creating a more engaging, embodied interaction. Embodied media art works have the ability to be interactive and can benefit from the affordances they creates by being physically present.

1.1 Context

Since modernism, art has been used as a way to critique its medium. Greenberg states that "Modernism used art to call attention to art" and that "The essence of Modernism lies, as I see it, in the use of characteristic methods of a discipline to criticize the discipline itself, not in order to subvert it but in order to entrench it more firmly in its area of competence" [2]. Since the birth of new media, it has been entrenched in Marshall McLuhan's idea that the "medium is the message." If the medium is the message and the message is a critique of the medium this creates a feedback loop. According to Edward Shanken "with regard to cybernetics, information theory offered models for explaining certain aspects of how messages flow through feedback loops" [3]. My work falls under this category of cybernetics, "the scientific study of control and communication in the animal and the machine" [4]. Cybernetics is considered a subsection of General Systems Theory (GST) that focuses on communication and regulatory feedback. Von Bertalanffy when discussing GST argues that there are logic and mathematical laws to systems that are applicable across disciples [5].

Embodied interaction is a cybernetic feedback loop. Michael Kelly defines embodied interaction as the aspect of interactive art which deals with physical embodiment and sensorial experience, "a specialized aspect of interactive art practice which acknowledges the primacy of physical embodiment and sensorial experience in computer-based interaction creating sensor driven artwork which is responsive to visitors" [6]. Through embodiment of interactions it allows for the artwork to take advantage of the environment's affordances, which may not be present in a disembodied interaction. Gibson describes the idea of affordance as what the environment offers to an animal and a niche as a set of affordances [7]. Bonard and Pfeifer extend this to robotics when designing agents to exploit their ecological niche through embodied intelligence [8]. In this way by embodying digital interactions through art, the work can take advantage of the affordances provided by the environment.

1.1.1 Communication in Media Arts

Art can be defined as what happens between the viewer and the work, meaning that the experience of the artwork is what makes it art. This definition incorporates ideas from Roland Barthes' "The Death of the Author" [9], Monroe Beardsley's "Intentional Fallacy" [10] and John Dewey's "Art as Experience" [11]. The idea that the viewer is part of the work is supported by Burnham who argued that all "which processes art data, ... are components of the work of art" [12]. New media arts allow for a more participatory interaction of the viewer with the work.

Art is a sensory based experience that can communicate without words and transcend language. Many artists make visual or sonic art so they don't have to use words. There are two main parts of nonverbal communication in the arts, the artist's expression and the viewer's experience. The medium influences how a message is perceived and it is a symbiotic relationship [13]. The medium of robotic art therefore carries a history of interactivity (see Ch. 2). So when creating art using technology, there is the opportunity to provide an interface. Frequently, media artists will create an interface for composing their work, even if they do not expose it to the public or perform it in real-time. In the field, it can be used as one way communication for a traditional performance or two way communication in interactive performance. Edmonds defines these as a static system versus a dynamic interactive system [14]. The visitor is part of the audience in one way communication and a participant in two way communication. Two way communication becomes a feedback loop between the participant and the artwork. Multi-way communication is an integral part of my work, which becomes more compelling with interaction.

For human technology interaction, it can be argued that verbal communication is cumbersome, and in many cases detrimental to the user experience. Non-verbal communication can be multimodal, creating multiple cues if some are missed, and allowing multiple pieces of information to be expressed through different modalities. Well designed multimodal interfaces should integrate complementary modalities in a way that supports mutual disambiguation of errors, leading to better performance [15]. Verbal communication can supersaturate a user with information, while non-repetitive multimodal, non-verbal cues can be used to separate the information onto different channels.

This dissertation focuses specifically on interactive art and its sub-genre, robotic art. Because the thesis covers both electronic creatures (robotic art) and more generally disembodied work (interactive art), I will reference and discuss robotic art and design at length. This is not without precedent (see Section 2.3.2). Based on prior definitions, I will define robotic art as interactive art that uses the more traditional definition of robots, as well as interactive art that is physically embodied and is viewed as an agent.

This dissertation's goals align with how theorists and artists describe the intent of robotic art. It focuses on "the innovation and exploration of the human machine relationship" [16] and challenges our understanding of robots, while exploring "the aesthetics of modern behavior" and the "development of unprecedented interactive communication scenarios in physical ... spaces" [17]. The work in this dissertation mirrors Ghendini's description of how robotic art goes against modern aesthetics, focusing on the "medium specificity; autonomy of the artwork from its environment; rejection of narrative; anthropomorphism and theatricality; and separation of high and low culture" [18]. The thesis itself is about exploring the medium specificity of robotics, manipulating the features that are unique to the attributes of robotic art.

1.2 Research Question

Can the embodiment and emotion of robotic media arts be used as a tool to probe the social implications of interactive digital technologies? This dissertation explores the potential limitations, strengths, and weaknesses exposed by these probes by embodying digital interactions and by employing affective computing techniques. This question is relevant at the social, the physical, and the public interaction levels.

1.2.1 Social Interaction

The first study explores embodied media arts' ability to probe the commodification and dehumanization of humans' interaction with social technology. It uses methods of affective computing when choosing the next appropriate dialogue response. It uses media arts to discuss and embody the issues involving social interaction and interpersonal communication. Specifically I am investigating how mobile technology and social media dehumanizes and degrades the quality and quantity of social interaction. Social network sites are regularly accessed from mobile phones, which are habit forming because checking behavior emerges and is reinforced by informational 'rewards' [19]. The pathological use of social networking sites (SNS) is associated with depression and poor self regulation skills [20]. Social Networking Sites make money on how long users interact with the platform, viewing advertisements and creating more content that can be monetized [21]. Therefore these sites are incentivised to keep users interacting with the platform as much as possible. This has led to a prevalence of social media addiction, and problematic social media usage which can lead to depression and other mental health side effects [22]. According to Schiffin, despite perceiving online interaction as "less beneficial than face to face interaction" and "associated with reduced well-being," people are increasingly using the internet for communication with others [23]. Companies are using celebrity endorsers, who have built parasocial relationships with followers, to sell products [24]. The first work in this dissertation examines the gamification of social interaction technologies, reducing the interaction to an algorithm with the goal of obtaining information about the participant. It questions how interactions change when an interaction is embodied and interacted with publicly. Through two different performances, I explore embodiment, humanization of technology, mechanization of people, social communication, gender roles, and female agency.

1.2.2 Physical Interaction

The second work attempts to address the problems with technology in the personal space by exploring the potential for healing using interactive zoomorphic personal technology. It uses media arts to discuss the potential for embodiment and physical interaction. It tests how different sonic and haptic responses of a robotic creature to touch affect the way that people interact with the robot. This work also explores the social dynamics between the people interacting with tactile emotive robots. Particularly, it examines how the emotional expression of a robot can be perceived as completely different by a person touching a robot than by a person observing the interaction. This work emphasizes the importance of touch and haptic interaction in media arts.

Mobile technology and computer mediated communication, like cell phones, have become ubiquitous, to the point where people imagine vibrations that aren't there [25, 26]. Cell phones can become addictive and a distraction [27]. It has been shown that simply having a phone visible causes two people to feel that their interaction is less connected [28]. Through the evolution of the cell phone, the touch pad and keyboard have been removed, leaving a nondescript screen. The physicality is heavily reliant on vision for interaction with no physical buttons to press, and one buzzer as the only physical response. By removing the physicality and personality of the object, we are missing the physical intimacy of touch when interacting with personal technology.

One way people can have a tactile non-verbal relationship with another creature is by having a pet. However, our society is becoming more restrictive on our ability to have pets. For example, landlords often prohibit pets [29] and more people are renting instead of buying [30]. To combat this, there has been a surge in "Emotional Support Animals" [31] prescribed by medical professionals [32]. Medical companion robots have filled the niche in spaces where a person cannot take care of a pet (e.g. a retirement facility), but would benefit from the interaction that a pet or emotional support animal provides. Social robots that are physically embodied and use tactile communication are needed to facilitate meaningful social interaction and are particularly important for people who are lonely [33]. Current social robots are not ideal, for example their sonic interactions with a therapy robot can be repetitive. One solution which is explored in the second work is to create a way to use input, like the way the robot is touched, to vary the sonic response and create an emotive reaction.

1.2.3 Interactive Environments

First I explore the limitations of digital devices as interfaces for media artworks and potential solutions for interactive environments. For a series of works, I created and built interfaces to facilitate social interactions in interactive environments. Interactive environments can allow multiple people to interact with a system, and to interact with each other through the system. These people could be performers, visitors, or scientists. In these examples we ran into issues with performance shyness, multiple collaborators, and disparate controls. Work that is presented with the ability for interaction "can evoke situational shyness in visitors, through the combination of a demand for active, performative engagement and the deliberate restriction of instructional and explanatory information" [34]. Situational shyness is an issue of showing art in the public space where, as the visitors are interacting with the piece, others are watching. Other issues to avoid with digital immersive environments and their interfaces are information overload and overwhelming control interfaces. Using distributed interfaces, multimodal interfaces, and personal mobile technology are potential design solutions that are considered in these works.

1.2.4 Public Interactions

Social technology is being used as a platform for social movements and can be viewed as a way to empower the people. For example, hashtags are a UI tool that was created by users as a way to organize ideas and that has been used by movements like #MeToo. My last work embodies interactive public technology to probe the limitations of the empowerment the space provides. In this work I embody voices from a forum where sexual assault survivors tell their stories to draw attention to them. This interactive sound installation embodies the way that testimonies of sexual assault have been filtered and twisted by the media and the experience of a survivor being inundated by it. Organizers of the #BeenRapedNeverReported movement reported that it required much emotional labor to be exposed to the emotionally charged content [35].

Public social media platforms like Twitter have also been used for disaster relief during fires and floods as a way to disperse information quickly. During the Thomas Fires in California, the public was directed to Twitter to find out the latest information instead of calling the fire department. Twitter has also been used by the public to inform others and the government about issues and whether they are all right during these crises [36]. There are potential pitfalls in the use of Twitter for disaster relief regarding coordination, accuracy, and security [37].

One major issue can be the public's level of trust in information they find on social media and the potential to spread misinformation. "The rise of fake news highlights the erosion of long-standing institutional bulwarks against misinformation in the internet age" [38]. Governments and other groups can use social media as a way to interfere with elections and political movements [39]. Social networking sites have been used to empower hate groups and hate speech. For example, the "fake news" in the US 2016 election which favored Trump was shared at 3.75 times the rate of news favoring Clinton [40]. This leads to distrust of the media and other information online. Corporations have commercialized the sharing of empowering videos by creating feel-good advertisement videos that support a movement [41]. These companies get free advertisement by piggy-backing on the cause. This can cause distrust of the motivation behind content. Having an agent be physically present changes the persuadability and authenticity of the information/interaction. People felt a higher level of arousal, responded more favorably, had a stronger response, and found physically present robots more persuasive [42].

1.3 Approach

To investigate the use of embodied media arts for the exploration of the social implications of interactive technologies, I created case studies that probe our use of and interactions with technology during the cycle of physical and emotional abuse of romantic partners. This dissertation is composed of two artworks, *Come Hither to Me!* and *Touching Affectivity*, which have lead me to design *Cacophonic Choir*, the third and final piece. Through these works I explore the commodification of relationships through technology and its potential use for healing. The last work focuses on technology as a platform to tell the stories of abuse. I am also exploring how these relationships can start in the social space, how we can heal through intimate interactions, and how they are received in the public space.

I explore the commodification of relationships with the installation *Come Hither to Me!* in Chapter 3. In this work, an interactive robot's goal is to get the phone number of the "hottest" (temperature) person in the room. ROVERita views her interactions as an end to a goal and uses pick-up artist methodology to manipulate the audience. She both flirts with and jokingly insults the audience, showing the push-pull mentality of abusive relationships. Her interest in the audience member is a means to a telephone number; she does not care about them personally. This examines how technology can enable abusers in the social space, while exploring more generally gender roles and women's agency. This work was exhibited at The End of Year Show (EoYS) at the University of California, Santa Barbara in 2018 and the Human Factors in Computing Systems Conference (CHI) in 2019. At the EoYS, the robot was untethered and could move around the space. At CHI, visitors interacted with the agent as a disembodied voice from a computer screen privately and a tethered embodied robot publicly. ROVERita was also used in the performance *Bodies at Work: I Love My Robot and My Robot Loves Me* at the Multimodal Performance Festival. This performance was an electronic reenactment of Joseph Beuys performance piece, *I Like America and America Likes Me*. In this work my collaborator and I became machine-like in an attempt to give humanity to a robot. *Bodies at Work* explored female agency, gender expectations, and the conflict between human and machine. Unlike the majority of female robots in robotic art, ROVERita in *Come Hither to Me!* breaks ground by having agency. These works show that robotic arts can be a valuable tool to discuss how the embodiment of technology can be used to break stereotypes instead of reinforcing the marginalization of others.

I explore how technology could be used for healing from emotional and physical abuse through the installation *Touching Affectivity* in Chapter 4. In this work, the fluffy robot responds both physically and sonically to touch. The robot is reminiscent of a pet or emotional support animal, an animal that gives non-verbal comfort through touch-based interactions and unconditional love. This work explores healing through interactions in the personal space. It was exhibited at "Invisible Machine," University of California, Santa Barbara's Media Arts and Technology Program's 2018 End of Year Show (EoYS) and New Interfaces for Musical Expression Conference (NIME) in 2019. At the EoYS the work was described as responding emotively to touch, but it was not immediately sonically reactive. Here the robot vibrated with increasing pressure and made a high pitched chirping noise when squeezed intensely. Participants physically interacting with the robot and visitors viewing the interaction perceived the emotive response differently. The participant would feel the increasing vibration and feel that the robot was happy and content, while onlookers would interpret the high pitched chirping noise as screaming and feel that the robot was in pain. At NIME the robot responded immediately with emotive sound to the way it was touched while at "Invisible Machine" it did not. People were less aggressive with the robot when it responded to touch immediately. This showed that by emotively sonifying touch-based interactions, people interacted with the robot more gently. This work emphasizes the importance of examining haptic interactions in media arts, showing the difference between how emotion is perceived when interacting with a creature physically and viewing the interaction. It argues for sonic and haptic responses to be aligned to create cohesive human machine communication and interactions.

Interactive environments can be used to create, manage and analyze how people intereact with each other. I explore the limitations of digital collaborative interfaces for digital environments and potential solutions through the three projects in Chapter 5. The first, *Feedback Rings*, studied collaborative interfaces for multiple performers (the CREATE Ensemble) to control networked spatialized sound, where performers were too far apart to interact with each other or the same interface. This work was performed at 4 showings during the CREATE AlloSphere Spatial Sound Concert in 2016. In the second, I built an interface for *The Hydrogen-Like Atom*, an interactive installation by JoAnn Kuchera-Morin, Lance Putnam and Luca Peliti. For this work I examined performance shyness in exhibition spaces. This work was exhibited at the Museum of Exploration and Innovation (MOXI) from January until June of 2016, and shown at the International Symposium for Electronic Art (ISEA 2017), the Alliance of Women in Media Arts and Technology (AWMAT 2018), and ACM International Conference on Multimedia (ACM-MM 2017). The last work, *The Volumetric Viewer*, explored visual communication in the research environment, where parameters were disparate and had different requirements. This work has been shown at AlloSphere demonstrations to researchers, donors, and visitors. The solutions for these design problems were multimodal, distributed, and personal interfaces.

Currently, I am exploring technology as a platform to tell the stories of abuse through the work *Cacophonic Choir* in Chapter 6. In this installation, agents are spread throughout the room speaking incoherently through mumbled electronic voices. As the audience approaches an agent, its voice becomes more clear and its stories make more sense. When the audience is very close, they hear stories of sexual assault. This work explores how technology is both a platform for expression and suppression. This explores the question of how technology elevates and obfuscates stories of abuse in public discourse. This work was exhibited at Contemporary Istanbul 2019's Plug-in exhibition.

1.4 Outline

This dissertation is organized as follows:

- **Chapter 2** provides a background to the fields of embodied media arts and human-machine communication and presents how I use embodied media arts to discuss, explore and critique our use of technology.

- **Chapter 3** presents *Come Hither to Me!*, an interactive social robot that attempts to "pickup" audience members. In the chapter, I use affective computing to determine the participants' emotive response to complements and negative comments (negs). I show that by embodiment of social technology, participants are more likely to walk away from abusive behavior.

- **Chapter 4** presents *Touching Affectivity*, a small fluffy robot that responds both physically and sonically to touch. In the chapter, I show that sonifying touch emotively causes visitors to interact with a robot more gently.

- **Chapter 5** explores interfaces for environmental scale media art installations. It presents interface design for scientific visualizations, performances, and art installations. In the chapter, I show the limitations of screen based interfaces and potential solutions, including tangible computing.

- **Chapter 6** presents *Cacophonic Choir*, an interactive installation of orbs that respond visually and sonically to proximity. In the chapter, I show how media arts can be used to show the limitations of empowerment in public digital spaces.

- **Chapter 7** presents reflections, discussions, and implications of these works.

1.5 Permissions and Attributions

Chapter 2

Background

2.1 Systems, Design and Media Theory

In art every choice conveys meaning. When creating art, each design decision is important in how the final work is interpreted. The experience of interacting with the art, is the art itself and the message it is trying to convey. Anything that gets in the way, like thinking about how to interact with the object, breaks the experience. When designing robots and interactive technology, the same is true. When creating an interactive experience without language using a screen or a robot, communicating to the user how to interact can be a problem. A solution to this is good design. The design of the interaction as a comprehensive whole is more important than the controls themselves [43]. The idea of interactive technology, how we should interact with it and design it has been explored in systems, design and media theory.

2.1.1 Cybernetics

In 1948, Claude Shannon wrote about a "Mathematical Theory of Communication" which is the foundational work that described information theory. In this system there are 5 elements, an information source, a transmitter, a channel, a receiver and a destination [44]. For example for verbal communication this would be Alice's brain, Alice's mouth speaking, the air, Bob's ear hearing, and Bob's brain. Here the interface would be the receiver, or Bob's ear. Norbert Wiener defined cybernetics in 1948 as "the scientific study of control and communication in the animal and the machine" [4]. Wiener describes cybernetics as a feedback loop. According to Edward Shanken "with regard to cybernetics, information theory offered models for explaining certain aspects of how messages flow through feedback loops" [3]. Two entities in a system interact or communicate information through an interface. That interface can affect the environment. McLuhan coined the phrase "the medium is the message," discusses how the medium shapes its environment, as in "The 'message' of any medium or technology is the change of scale or pace or pattern that it introduces into human affairs" [45]. In this way the medium is the environment of the cybernetic system in which the message is communicated, and which is greatly shaped by its environment.

2.1.2 Design Theory

When Marshall McLuhan is speaking about the message of a medium, he is discussing the way that the message changes the environment, and therefore its affordances. He states that "environments are not just containers, but are processes that change the content totally" [46]. Affordance is what the environment offers to an animal, and a niche is a set of affordances [7]. The engineering design theories relevant to this dissertation are heavily dependent on psychologist James Gibson's affordance theory. Donald Norman extended the concept of affordances to the design of objects in his book "The Design of Everyday Things" [47]. Fiona Raby further uses design to speculate on how things could be in the future in "Speculative Everything: Design, Fiction and Social Dreaming" [48].

When discussing how to design a robot, Cynthia Breazeal includes expectation setting [49]. This means designing the environment (robot) so that it aligns with the expected affordances that a human requires to interact with it. In Bonard and Pfeifer's theory of building robots, they state that robots should be equally intelligent physically and computationally [8]. In human robot interaction, a robot should not be more intelligent than it needs to be. This relates back to affordance theory, in which a robot should take advantage of the environment's affordances through physical intelligence. An example of this in the art world is Louis-Philippe Demers *The Tiller Girls*, [50] in which most of the sculpture's movement is due to physical intelligence.

Bonard and Pfeifer discussed the importance of multimodal sensors in the design of robots [8]. This can be extended into multimodal interact where the robots are interacting with people as part of their environment. Sharon Oviatt throughout her career has shown the value of multimodal interaction in ensuring accessible, usable, and satisfying interactions for human computer interaction [51, 52]. This was theorized by early cybernetic writers like Ted Nelson who discussed the idea in 1974 that anything can be a control for interface design. How those controls are unified into a good system is not arbitrary [43]. "The human mind being as supple as it is, anything whatever can be used to control systems. The problem is having it be a comprehensible whole... The kind of controls are totally arbitrary but their unification in a good system is not" [43]. This is similar to the way cats communicate with humans. Some owners can recognize what different sounds mean, but cats as a species do not have a recognizable sound for specific contexts [53].

2.1.3 Media Theory

In the 1930's John Dewey described art as not a crafted object but a crafted experience [11]. John Burnham reflects Dewey's sentiment when he states that all "which processes art data, ... are components of the work of art" [12]. In that way, the experience or processing of the art data is part of the artwork itself. Literary theorists emphasized the experience of the work as more important than the intent of the author. William Wimsatt and Monroe Beardsley questioned the importance of authorial intent, "...the design or intention of the author is neither available nor desirable as a standard for judging the success of a work of literary art" [10]. This was furthered to the extreme with Roland Barthes' argument that author and work should be judged separately, and the analysis should not include bibliographic information or authorial intent [9].

Modernism embodies Marshall McLuhan's concept that "the medium is the message." Clement Greenberg describes in his paper "Modernist Painting" describes Modernism as the "use of characteristic methods of a discipline to criticize the discipline itself," and says that "Modernism used art to call attention to art" [2]. Marshall McLuhan discusses the modernist art movement cubism, a modernist art movement, stating that it "suddenly announced that the medium is the message," and that cubism required "instant sensory awareness of the whole" [45]. McLuhan also stated that "Each new technology creates an environment that is itself regarded as corrupt and degrading. Yet the new one turns its predecessor into an art form" [45]. In this way 'New Media' has a complex relationship with the art world, because it hasn't yet been superseded and so is regarded by some as corrupt and degrading.

2.2 Spatial Design

Oviatt argues that well designed multimodal interfaces should integrate complementary modalities in a way that supports mutual disambiguation of errors leading to better performance [15]. Outside of vocalizations that are not words, kinesics, haptics and proxemics are other types of non-verbal communication. Haptics, the use of touch to interface with technology, is only available in the personal/intimate space. Kinesics, the way that a robot moves, affects the emotional responses of people observing it [54]. For example, the two main cues children use in distinguishing between animate and inanimate objects are movement and visual features (such as the presence of a face or dominant texture) [55]. Proxemics, distance based interaction, is used to design human robot interaction. In this dissertation, we use proxemics and personal interfaces to interact in public spaces.

2.2.1 Proxemics

We interact with people and animals differently at a different range of distances. Edward Hall explored this in a 1963 paper where he coined the term proxemics, defining it as "the interrelated observations and theories of humans use of space as a specialized elaboration of culture" [56]. We need different types of interactions when interacting with technology at different distances and in different types of spaces [57]. Proxemics is currently used in research on human robot interaction in the social space [58]. Hall describes 4 levels of interpersonal distance: intimate space (0-18 inches), personal space (1.5-4 feet), social space (4-12 feet) and public space (<12 feet) [59]. Researchers found that in virtual reality when a person views an object as an agent, they interact with it in space similar to the way they would interact with a person in social space. When interacting with a non-entity, they would not respect personal boundary [60]. Proxemics is also important in the media arts. One of the focuses of the robotic arts is the artwork's

<table>
<tr><th>Art Space</th><th>Art Interaction</th><th>Hall's Terminology</th><th>Hall's Distance</th><th>Embodiment</th><th>Agent</th></tr>
<tr><td>Cyborg</td><td>Physically attached to person</td><td>Intimate Space (Close phase)</td><td>< 1-2 cm</td><td>Embodied as an extension of person</td><td>Not separate agent/entity</td></tr>
<tr><td>Intimate</td><td>Allowed to touch</td><td>Intimate Space</td><td>< 18 inches</td><td rowspan="2">Embodied
baby or friendly animal (pet)</td><td rowspan="3">Embodied Agent</td></tr>
<tr><td>Personal</td><td>Within reach of art object</td><td>Personal Space</td><td>1.5-4 feet</td></tr>
<tr><td>Social</td><td>Outside of reach to 12 feet</td><td>Social Space</td><td>4-12 feet</td><td>Embodied
Adult or non-pet animal</td></tr>
<tr><td>Public</td><td>> 12 feet
Can be with separate device</td><td>Public Space</td><td>> 12 feet</td><td>Unembodied environmental</td><td>Unembodied agent</td></tr>
</table>

Figure 2.1: The relationship between art space, art interactions and embodiment.

autonomy from its environment, which means that in comparison to traditional forms of art, it focuses on interaction within different types of spaces (see Section 2.3.3).

2.2.2 A Spatial Design Framework

People aren't comfortable with other people in their personal space [61]. People react neurologically to a human face intruding on personal space, but not to an object like a sphere [62]. Appropriate approach distances for robots is complicated and relies on many factors [58]. How to collaborate with robots within personal space [63] is a current area of research. One solution to the problem would be to change the morphology of the robot so that it is no longer seen as threatening.

In my research I categorize media arts into different spatial distances: intimate, personal, social and public [59] (See section 2.2.1). Interactive robotic art in the intimate space is artwork where the participant can touch the work, while personal space is when the work is within reach. Interactive robotic art in the social space is artwork where the participant interacts with the robot outside of arm's reach, but within 12 feet. Interactive robotic art in the intimate, personal and social spaces are viewed as embodied agents by the viewers. Interactive artwork in the public space is work in which the participant interacts with the work at a distance greater than 12 feet. This work can be unembodied, becoming part of the environment. The work can be interacted with using devices or through sonic/gestural means. Cyborg work is not included in this research because the robotic work is physically attached to the participant or performer and at this range the work is not a separate entity but an extension of the person's body. This is illustrated in Figure 2.1.

In static work, the visitor is an audience member, while in dynamic interactive work, the visitor is a participant. I am focusing on dynamic interactive work. An artwork's space is not always rigid and definite. Artwork may be interacted with in multiple spaces or move between them. Performative interactive work can be in different spaces for the performer than for the audience. When the audience is watching a performer interact with the work, the work may be in the public space for the audience and in the intimate, personal or social space for the performer. An artwork may also be designed specifically to contradict these spatial expectations to create an uncomfortable or jarring interaction.

There are different issues and requirements for design of human machine communication at different distances. Territorialism is an issue in the personal space, requiring the robot to be non-threatening. The uncanny valley is an issue with humanoid robots requiring that a robot not be too realistic. Also, information overload is an issue in the environmental space, which is solved by multimodal interaction. This is illustrated in Figure 2.2.

Personal Interfaces

To interact from a distance, mobile device interfaces can be used. Prior work in HCI has been done using mobile devices as multi-degree of freedom controllers. Katzakis et al. used the gyroscope of a phone to control the rotation of an object and showed

<table>
<tr><th>Art Space</th><th>Art Interaction</th><th>Hall's Distance</th><th colspan="2">Psychology / HCI Issues</th><th colspan="3">Media Arts Issues</th></tr>
<tr><td>Cyborg</td><td>Physically attached to person</td><td>< 1-2 cm</td><td rowspan="3">Territorialism</td><td rowspan="3">Uncanny Valley</td><td rowspan="4">Performance Shyness</td><td rowspan="2">Potential damage to work or audience</td><td>Requires assistance</td></tr>
<tr><td>Intimate</td><td>Allowed to touch</td><td>< 18 inches</td><td>Likely requires monitoring</td></tr>
<tr><td>Personal</td><td>Within reach of art object</td><td>1.5-4 feet</td><td rowspan="2">If mobile: [Potential damage to work or audience]</td><td rowspan="2">If mobile: [May require monitoring and battery maintenance]</td></tr>
<tr><td>Social</td><td>Outside of reach to 12 feet</td><td>4-12 feet</td><td colspan="2">Uncanny Valley</td></tr>
<tr><td>Public</td><td>> 12 feet
Can be with separate device</td><td>> 12 feet</td><td colspan="2">Information Overload</td><td colspan="3">Performance Shyness</td></tr>
</table>

Figure 2.2: Psychological issues and media arts issues in different art spaces.

enormous potential for this type of interface [64]. Tablets have been shown to be a useful way of navigating and are better than a joystick for selecting items and manipulating virtual objects [65]. Tracked mobile devices have been used to select objects and get more information about cluttered data sets [66].

Prior work in the visualization field used gestural control to manipulate large data sets, focusing on two dimensional visualizations of the data [67]. In the Burgess research project, the LEAP controller was used for gestural control. The LEAP controller allows for fine motor gestures but does not track gestures where one hand occludes another. Slices and 3D visualizations have been used to visualize volumetric data in the medical field [68].

Exhibition Design Problems

Defining and conveying how to interact with technology is an issue in Media Arts and in Human Computer Interaction. In both cases many people are hesitant and/or uncomfortable to interact with embodied technological agents in their personal space [60]. In social spaces, robots and robotic art can be unpredictable if there is a lack of clearly defined communication [69]. In public spaces, interaction can be overwhelming to the participant due to the expectation of performance, fear or shyness of interacting with a new interface in front of other people [34]. Each space has potential issues that can be solved by non-verbal communication.

2.3 Embodied Design

When a robot appears human, expectations are set about how it should behave [70]. It is often tempting to design robotic agents that respond like humans, whether through sound (i.e., speech) or visual appearance, but this can be a perilous undertaking due to the presence of the "uncanny valley." This refers to the discomfort that is felt when something seems to be mostly, but not quite human [71]. The quality of synthesized speech is worse than synthesized facial expressions at expressing emotion [72], often exacerbating the effect of "uncanniness." By making a robot communicate in a human-like but distinctively different way, we can create social interactions without becoming uncanny. Humans experience unease when interacting with technology that is humanoid but not quite convincingly realistic (see Section 2.3.1). This is an issue in physical design, auditory design and design of movement that I attempt to address.

Because a robot is embodied, even the way that a robot moves affects the emotional responses of people observing it [54]. The physicality of the robot or its movement can make people uncomfortable if designed incorrectly. The presence of movement is often a valuable clue as to whether an object is animate or not [55]. Because movement can be indicative of animate embodiment, it can affect the associations that people make to an object in their environment, and the way that they might interact with it (see Section 2.3.1). In the public space, changing the interface and its embodiment affects how visitors, performers or researchers interact with the work. In my work with immersive environments, I use the tangible qualities of typical input tools to create more intuitive interfaces. *Cacophonic Choir* physically embodies the interaction online between the participant and stories from the media.

In the social space, by creating human-like robots, we can employ expectation setting to use how we interact with other humans as a template for human robot interaction. Small design differences in the robot will affect the interaction. For example, the height of a telerobotic agent in the social space affects how persuasive the agent is and how much the local person feels in control of the conversation [73]. This is similar to how an adult would interact with a child. When designing a robot each design decision, whether to give a robot a male or female voice, a child or adult height, a white or metal body, a British or American accent, changes the user experience of the robot, and what qualities the person interacting will associate with it. This can perpetuate stereotypes and archetypes of specific roles and people. ROVERita, in my prototype *Come Hither to Me!*, questions whether moving the interaction from private to public space and the agent from disembodied to embodied changes interactions with technology.

Using our interactions with animals for robotic design inspiration is a way to create a more pleasant experience for a person interacting with the robot within their personal space. I believe that more animal-like robots will be viewed similarly to animals, and will allow the robots' affordances that they wouldn't have otherwise. Since animals and babies do not communicate with words, but instead use non-linguistic utterances, an animal-like robot should express itself similarly. I designed a robot, my prototype *Touching Affectivity*, to test this theory. By embodying interactions with technology through art, we can question how people interact with the technology. The ability to touch is unique to a person's intimate and personal space. Therefore, I focused my exploration of touch-based interaction with robotic art in these spaces, while allowing communication through movement and non-linguistic utterances. This line of research was also brought further into the physical through an interaction in intimate space, looking at emotive response to sounds produced in relationship to haptic engagement. Pet-like instead of baby-like can be used for inspiration to avoid the uncanny valley. See Section 2.3.3 for a more in depth background.

2.3.1 Morphology of Embodied Technology

Prior research indicates that people empathize and engage more with embodied robots than with avatars [74, 33, 75, 76]. Physical embodiment can affect a social agent's capabilities and a user's enjoyment of a task [77]. In a study which compared a robotic and a virtual medical assistant, participants felt a greater sense of presence with the robot. They felt it was more lifelike, and disclosed less private information [76]. Based on a survey on experiments with embodied robots, telepresent robots and avatars, people preferred an embodied robot. People felt a higher level of arousal, responded more favorably, had a stronger response, and found physically present robots more persuasive [42]. Movement is a valuable clue as to whether an object is animate or not [55] and the type of movement affects the emotional responses of people observing it [54]. For example, the two main cues children use in distinguishing between animate and inanimate objects are movement and visual features (such as the presence of a face or dominant texture).

Visual appearance plays a complex and multifaceted role in affective robotics, as has been demonstrated in the work of Breazeal and Canamero [78, 79]. Cynthia Breazeal describes the idea of expectation setting, designing a robot so that it physically sets the expectations for how one should interact with it. When designing Kismet she chose an appealing appearance and a natural interface that encouraged humans to interact with Kismet as a young, socially aware creature. Kismet was small and baby- animal-like so humans will naturally and unconsciously provide scaffolding interactions, and humans will expect the robot to behave at a competency level of an infant-like creature [49].

To support Breazeal's theory of expectation setting, when a robot appears human, expectations are set about how it should behave [70]. The "uncanny valley" refers to the discomfort that is felt when something seems to be mostly, but not quite human [71]. Robots that are too humanoid are viewed as a threat. A number of theories have been proposed to explain the uncanny valley, including more biologically driven cues like conflicting perceptual cues, [80] and pathogen avoidance [81]. More psychological theories have been proposed as well, such as fear of being replaced, fear of death, and the threat to the humans' distinctiveness and identity [82].

2.3.2 Morphology of Embodied Media Artworks

Multiple artists have looked to define interactivity, embodied interaction, interactive art and robotic art. Ernest Edmonds discusses a scale of interactivity in artwork from static to dynamic-passive to dynamic-interactive [14]. An example of static artwork would be a classical painting, which is not affected by the environment or people in it. Dynamic-passive would be a work that is affected by its environment and therefore people in its environment, but the interaction is not direct. An example would be *Condensation Cube* by Hans Haacke, in which the condensation in the sculpture is affected by the temperature in the room, which is affected by the people in the room. Dynamic-interactive is artwork in which the viewers can see a 1 to 1 effect of their actions on the artwork's changes. Michael Kelly defines embodied interaction as "a specialized aspect of interactive art practice which acknowledges the primacy of physical embodiment and sensorial experience in computer-based interaction. Embodied interaction refers specifically to sensor-driven, computationally articulated artworks which behave in response to the behavior of human participants" [6].

When discussing robotic art, Eduardo Kac includes robotic art (electronic creatures), cybernetic art (work that combines the organic and electronic) and telepresent art (art involving remote projection of a human subject through technological means) because they are all directly related conceptually and "appear hybridized in several works." He states that robotic and telematic art discuss the "aesthetic dimension of modelling behaviour (the artist creates not only the form but the actions and reactions of the robot in response to external or internal stimuli) and the development of unprecedented interactive communicative scenarios in physical or telematic spaces (the object 'perceives' the viewer and the environment)" [17]. More generally Luigi Pagliarini said that robotic art explores the relationship between man and machine [16]. Fiammetta Ghedini describes how robotic art does this through choices that go against modern aesthetics, through "medium specificity; autonomy of the artwork from its environment; rejection of narrative; anthropomorphism and theatricality; and separation of high and low culture" [18].

This dissertation takes inspiration from early cybernetic art, more contemporary robotic art and other automated works, particularly those emulating creative human speech and behavior. As interactive art, the research is focusing on what the human reaction is to the work itself. Media arts is a great but often overlooked resource for human robot interaction design. I leveraged design inspired by media arts to solve problems like autonomy, movement, appearance and interaction.

2.3.3 Zoomorphic Technology

When choosing a robot design to emulate, baby-like robots can create a more learning-based environment. Breazeal leverages child-like features with her robot Kismet to focus

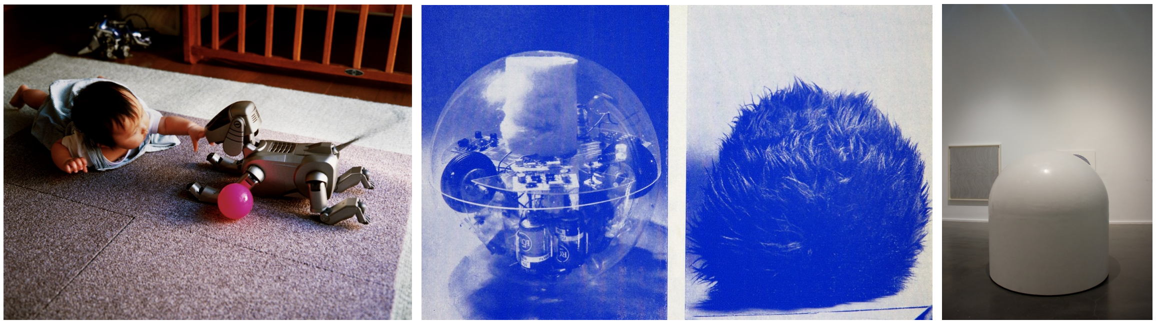

Figure 2.3: Chris Steele-Perkins. (1999). JAPAN. AIBO the Robot Dog. 1999. JAPAN. Aibo the robot dog. 1999. Retrieved from `https://library.artstor.org/asset/AMAGNUMIG_10311500041` Martin, Eric (North American, American engineer, 1943-) engineers, Parkinson, Robin (North American, American artist, 1943-) artists. (1968 (creation)). Toy-Pet Plexi-Ball, View: full, Two views: without fur cover (left) and with fur cover (right). [sculpture (visual work)]. Retrieved from `https://library.artstor.org/asset/HUCB_SHARE_109911210044` Robert Breer. (1970/2011, Image: 2012). Floats, detail. Retrieved from `https://library.artstor.org/asset/AWSS35953_35953_35431911`

on teaching [49]. If a person has created intimacy with their Roomba, like giving it a name, it increases their pleasure when the Roomba is cleaning [83]. People who have pets give robots a more positive rating and maintain a closer proximity to them [58]. When comparing how children interact with a fuzzy robot, smooth robot, dog, stick insect and hissing cockroach, the way children interacted with the fuzzy robot was more similar to the dog interaction than the interactions with insects [84]. The furry robot and the dog received more questions about the entities' biology, and the robots and the dog received more questions about the entities' biology than did the insects. In another study, it was found that children are more likely to conceptualize AIBO, a robotic dog (see Figure: 2.3), as a dog than as a robot [85]. The impact of a robot dog in patient therapy is an indicator of how well a real dog would help in therapy [86].

"Socially assistive robots describes a class of robots that is the intersection of assistive robotics (robots that provide assistance to a user) and socially interactive robotics (robots that communicate with a user through social and non-physical interaction)" [87]. Paro, the robotic seal, is one example of a socially assistive robot. Pets are seen as therapeutic for the elderly and sick, but it is hard for these groups to maintain the responsibility of having a pet. Paro interacts with simple sounds and movements, responding to being petted and held. There have been no studies specifically on the sounds that Paro makes [87]. Mamoru is another socially assistive robot that takes note when people take their medication and makes sure that they don't take it twice [88]. Mamoru's morphology blurs the line of human and animal-like.

Robot therapy has the same effects on people as animal therapy and has been used in a medical setting as mental healthcare for the elderly and children with developmental disorders. One example of this is PARO, a robotic seal, which has been used in Japanese nursing homes where patients can not have pets [89]. Tactile robots have also been designed for developmental disorder therapy. CAretaker RoBOt (CARBO), an autonomous, mobile robot, had promising results for children with ADHD as a tool for Sensory Integration Therapy [90]. Burns designed a haptic empathetic robot animal for children with autism [91]. This robot koala was built using Nao, a humanoid robot, using pressure sensors and Nao's capacitive sensors for touch-based interaction. Autistic children prefer unfamiliar robots to unfamiliar people [92], and a robotic dinosaur evoked more social interaction than an adult human [93]. Autistic children exhibited fewer negative behaviors when playing with AIBO, a dog robot, compared to a mechanical dog stuffed animal [94].

Biologically inspired robots are found throughout research in locomotion. Outside of more standard humanoid robots like Asimo,[95] (see Figure 2.5) examples of animal inspired robots are Boston Dynamics' research with the Wildcat, Big Dog [96] and Cheetah [97] and Festo's aqua penguins and aqua jelly. [98] [99] Complex locomotion is rarely used in home robotics, where Roomba style locomotion is more reliable.

Zoomorphic Art

This dissertation was inspired by early cybernetic and robotic artwork. Examples of animal-like early cybernetic artwork are *Tortoises* by Grey Walters (1948-49) [100], *Toy-Pet Plexi-Ball* by Robin Parkinson and Eric Martin (1968) (see Figure: 2.3) [101], *S.A.M.* (1968) and *The Senster* (1970) by Edward Ihnatowicz [102], *Colloquy of Mobiles* (1968) by Gordon Pask [103], and Robert Breer's work (see Figure: 2.3) for the Pepsi Pavilion in 1970 [104]. Grey Walter's *Tortoises* were some of the first examples of autonomous robots. Their movement and sensors were similar to that of a contemporary iRobot Roomba, with a photocell sensor and shell based bump sensor [100]. Examples of early robots that reacted to sound with movement are the *Toy-Pet Plexi-Ball* by Robin Parkinson and Eric Martin (1968) [101], *S.A.M.* (1968) and *The Senster* (1970) by Edward Ihnatowicz [102]. *S.A.M.*, which was the first moving sculpture that responded to people in its environment in a recognizable way, was exhibited in July 1968 at "Cybernetic Serendipity" [103]. It was an electro-hydraulically operated robot that moved in response to sound which sat on a table or pedestal [103]. It was followed by a 15 feet long robot that could move its arm around the space called *the Senster* which responded both to sound and movement[102]. *Colloquy of Mobiles* (1968) by Gordon Pask, a group of mobiles which communicated via light, was also shown at "Cybernetic Serendipity". *The Toy-Pet Plexi-Ball* was a plastic sphere that rolled in response to sound, which was exhibited at MoMA from November 27, 1968 until February 9, 1969. The ball was dormant when placed in its fur cover and could be petted like a dog or cat [101]. Lastly Robert Breer's work for the Pepsi Pavilion in 1970, *Seven Floats*, was made up of seven 6 feet tall, 6 feet diameter domes which moved less that two feet per minute, and reversed direction when touched [104].

In contemporary arts, the relationship between nature and the machine has been explored through work which empowers animals with different movements. Examples of this would be Ken Rinaldo's *Augemented Fish Reality* (2004) where Beta fish could control mobile tanks which moved around the space [105] and Garnet Hertz's *Cockroach Controlled Robot* (2008) where a cockroach controlled a vehicle [106]. In these spaces the visitor is interacting with the animal and observing how the animal interacts differently with its environment with new affordances. An early work that created a way for plants to move was Tom Shannon's *Squat* (1966), a live plant that by touching the plant created a voltage change which could cause the robot plant holder to move while making humming and chirping sounds [107].

Haptic Creature Art

Interactions with early artwork like *S.A.M.* (1968) [103] and *Toy-Pet Plexi-Ball* (1968) were based on sound. Though *Toy-Pet Plexi-Ball* could be put in a fur cover and petted, the creature did not react to touch [101]. Touch based interactive work at the time like *Magnet TV* (1965) (See Figure: 2.4) by Nam June Paik did not have zoomorphic forms [108]. Creating artwork that can be touched has a lot of technical difficulties with sanitation, sensors and deterioration. More contemporary works like *Cybersqueeks* (1988) by Ken Rinaldo [105], *Echidna* (2002) by Tine Bech [109] and *Trou Mireia* (2017) by Donat Melús [110] explore different interfaces and interactions for touch. For example with *Cybersqueeks* the interaction is through sensors and switches [105] and with *Echidna* the electromagnetic field that the artwork creates is disturbed by touch. While *Trou Mireia* does not physically react to touch, the visitor can only see the interaction via CCTV. The work is relevant due to the haptic interaction of touching it and because it is a media artwork that has a organism-like quality. The sanitation aspect of this work is managed with the visitors wearing gloves while interacting with the work and watching themselves interact on a screen, creating a medical atmosphere and conjuring ideas of surgery and endoscopy.

Figure 2.4: Nam June Paik. (1965). Magnet TV. [Sculpture and Installations]. Retrieved from `https://library.artstor.org/asset/LARRY_QUALLS_1039907843` Max Dean / Raffaello D'Andrea / Matt Donovan. (1984-2008). The Robotic Chair, one of three views. Retrieved from `https://library.artstor.org/asset/AWSS35953_35953_35384683`

Animal Flock Inspired Robotic Art

There are many contemporary and historical examples of interactive artwork which are flocks of organisms. Early examples of this would be *Colloquy of Mobiles* (1968) by Gordon Pask and Robert Breer's work *Seven Floats* (1970). More contemporary examples of this would be Ken Rinaldo's *Cybersqueeks* (1988), *Flock* (1992), *Autopoesis* (2000) and *Fusiform Polyphony (Face Music)* (2011) [105]. Other contemporary artists' works are *Bird Land* (1991) by Brett Goldstone, *Elevator Music* (2007) by Fernando Orellana [111], *ALAVs* (2007) by Jed Berk [112] and *Sixteen Birds* (2008) by Chico MacMurtrie [113]. Ken Rinaldo's works *Flock* (1992), *Autopoesis* (2000) and *Fusiform Polyphony (Face Music)* (2011) also reference the arm-like structure of *the Senster* [102], but instead of having just one responsive arm there are many. Most of these works interact with visitors at both social and public distances. Proxemics is explored both physically and artistically in works like Chico MacMurtrie's *Sixteen Birds* (2008) in which birds deflate when approached, show human's effects on the environment [113]. *Cybersqueeks*, *Fusiform Polyphony (Face Music)* [105] and *ALAVs* [112] interact in the personal space. *Fusiform Polyphony* [105] sonifying people's faces by photographing their faces from a close distance, while *ALAVs* can be fed by visitors and *Cybersqueeks* can be approached and touched.

Autonomous Art

One of the focuses of robotic artwork is the artwork's autonomy from its environment. Power and autonomy have been an important aspects of robotic artwork, starting with Grey Walter's *Tortoises* which found their charging station when they were low [100]. More contemporary examples of robotic artwork that are autonomous are *Desert Crawler* (1986) by Joe Davis [114], *Strandbeests* (1990) by Theo Jansen [115] and *Urban Parasites* (2010-2014) by Gilberto Esparza [116]. With Theo Jansen's *Strandbeests*, he is trying to make the autonomous sculptures completely self sufficient on the beach [115]. Robots also offer the opportunity to extend art to remote environments, to explore places where there is little to no human intervention. Michael Snow's film *La Région Centrale* was created completely with a mechanical camera surveying a remote area of Canada [117] and *Desert Crawler* (1986) by Joe Davis drew lines in the sand in remote environments. *Urban Parasites* (2010-2014) by Gilberto Esparza works on the other extreme, creating parasitic artwork that live off energy sources from the urban environment like power-lines [116].

Artists have given inanimate objects animal-like qualities by allowing them to move. Both the *Robotic Self-Healing Chair* (2008) (See Figure: 2.4) by Raffaeollo D-Andrea which collapses and then puts itself together [118] and *Escaping Chair* (2017) by Takeshi Oozu, Aki Yamada and Hiroo Iwata which moves whenever you try to sit on it [119], give the inanimate object intention and implied sentience through movement. People will project emotions and human characteristics on these chairs simply by how they move. Similarly in *C/Borg The Parliament of Robots* (2018) by Ken Rinaldo, his constructions, though shaped like the Parliament building, seem organic because of their size and their movement around the space [105].

2.3.4 Anthropomorphic Robots

While industrial robots have been around since 1961, [120] personal robots first appeared in the early 1980s. Their main function was educational, to teach people how robots work [121]. The first service robot that had a job was the HelpMate service robot, a robotic courier for hospitals in 1988 [122]. Personal helper robots are nowhere near the point where they can interact in non-controlled environments [123]. Care robots for the home, along with social service robots in museums and restaurants, have begun to appear in the last decade [124].

HANC was an early healthcare robot in 1995 that reminded patients to take their pills and could run routine tests [125]. Other examples of socially assistive robots are the Nursebot project [126], Robocare project [127] and Care-o-bot[128]. InTouch Health was one of the first in the telepresent robot market using them in rehabilitation centers, eldercare facilities and hospitals [129]. Roomba produced the CoWorker robot in 2002 and ConnectR robot in 2007, though neither of them were a commercial success [130]. Telepresent robotics is a field that is undergoing rapid expansion in research, office, eldercare and healthcare with PRoP, Giraff, QB, Texai, Beam, VGo, PEBBLES, MantaroBot, Double, mObi, Jazz Connect, iRobot Ava, 9th Sense Helo and Telo, RP-7, and MeBot [131]. Telepresent robots have expanded into remote environments like undersea and space exploration [132, 133].

Pepper by SoftBank Mobile and Aldebaran Robotics SAS is a Japanese home robot which can read emotion. Researchers are already investigating its use for teaching via telecommunication [134]. Pepper costs approximately 1800 dollars and is currently only available in Japan. While the goal of Pepper is to make people smile, there has been little research done on how emotive Pepper is [135]. JIBO is a personal robot by Dr. Cynthia Breazeal's new start-up. JIBO is a stationary tabletop robot that can track emotions, and is advertised as a personal assistant like SIRI or Cortana. Again there is little research yet on the emotional range of JIBO [136]. Both Pepper and JIBO can be used for telepresence.

Cybernetic Anthropomorphic Robotic Art

In early cybernetic art many humanoid robots were remote controlled and were used for disruption like Bruce Lacey's *ROSA BOSOM (Radio Operated Simulated Actress – Battery Or Standby Operated Mains) with Mate* [103] and *K-456* by Nam June Paik and Shuya Abe (See Figure 2.5). *K-456*, built in 1964, was a 20 channel radio controlled robot originally considered "androgyne" but cast as female in the United States. "Robot-K456 can bow, walk, give a speech (recorded by the then Mayor-elect of New York, John Lindsay), lift each arm independently and wiggle its representational torso. It also defecates on the floor of the gallery by remote control. Paik's robot looks mechanically unreliable and he admits that it needs constant attention" [137]. *ROSA BOSOM* was originally designed as an actress to play the Queen of France in the production of Three Musketeers, at the Arts Theatre in 1966. These robots were shown at "Cybernetic Serendipity" [103].

Contemporary Anthropomorphic Robotic Art

In the 1980's and 90's, Some contemporary artworks with anthropomorphic agents explored interaction in which the artwork was in a fixed location. Like early zoomorphic

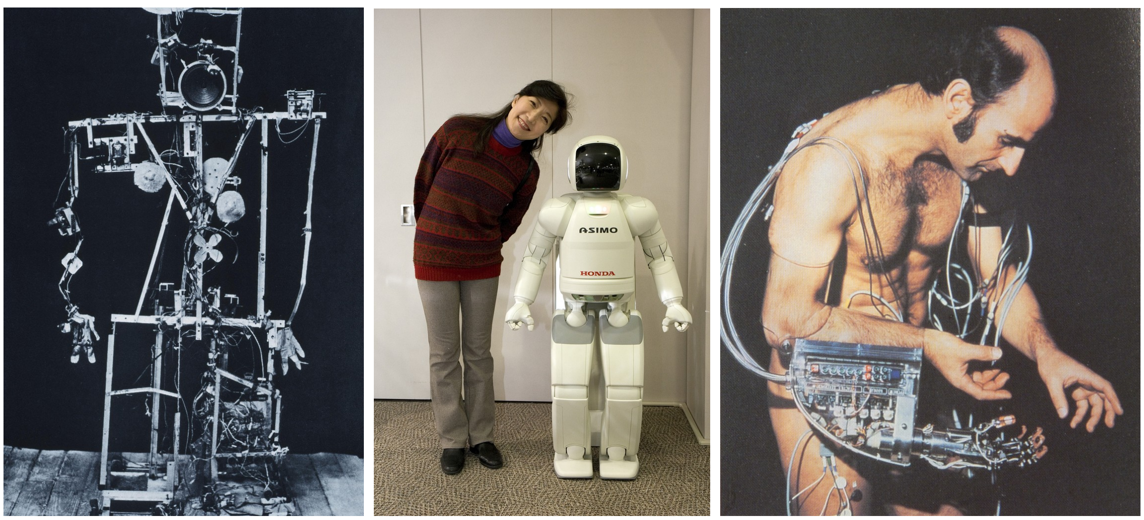

Figure 2.5: Paik, Nam June, 1932-. (1965). Robot K-456. Retrieved from https://library.artstor.org/asset/ARTSTOR_103_41822001048535 Chris Steele-Perkins. (2007). JAPAN. Tokyo. Robot Story. JAPAN. TOKYO. 2007. ASIMO Honda's robot. YAMADA Miyako poses with ASIMO. Retrieved from https://library.artstor.org/asset/AMAGNUMIG_10311515106 Stelarc (Greek performance artist, born 1946). (1976-1980). The Third Hand. [performance art; body art]. Retrieved from https://library.artstor.org/asset/HUCB__1099_20062560

cybernetics works like *S.A.M.* [102] and *Toy-Pet Plexi-Ball* [101], contemporary anthropomorphic robotic artwork has also been controlled by sound. Works like *JoAn l'Home de Carn* (1992) by Marcelli Antunez Roca [138], a life-sized human form covered in pigskin in a case moved based on audience sound. Other fixed works explored interaction using the proximity of the visitor to the artwork. *HLR Helpless Robot* (1987) by Norman White [139], *Stupid Robot* (1985) by Simon Penny [140] and *Nose Wazoo* (1990) by Jim Pallas with Jim Zalewski [141] explored proxemics in different ways. *Nose Wazoo* played with the line between personal and social space, attempting to touch visitors with its nose when visitors approached. *Stupid Robot* was designed to be reminiscent of a legless beggar, and it shook a can of metal scraps when approached, drawing people closer [140]. *Helpless Robot*, would ask for help to draw people near, but if people came close it would start to insult them.

Similar to *Stupid Robot* [140] and *Helpless Robot*, the *Tumbling Man* (1991) by Chico MacMurtrie and Rick Sayre was a mobile robot that created empathy in the viewer as the robot struggled to roll, but unlike prior works this robot could move. In this work two people wore body sensors controlling the robot, trying to make it roll. One controlled the legs and the other controlling the torso/arms, but they didn't know which they were controlling [113].

Starting in the 2000's, artists explored the public's interactions with more mobile robots like *Petit Mal* (2006) by Simon Penny [140], *Paparatzi Bot* (2009) by Ken Rinaldo [105], and *Berenson* (2011) by Denis Vidal and Philippe Gaussier [142]. *Petit Mal*, is an autonomous interactive robot, designed to be more simple than functional. Its basis is a pendulum and two bicycle wheels. It uses ultrasonic and piezoelectric sensors to navigate the space and find people who it then follows. It is adorable in its clunkiness [140]. Rinaldo's autonomous robots, the *Paparazzi Bots*, were a group of human-height robots that would move toward people and take pictures of them like paparazzi. They moved at human speed, avoiding obstacles using multiple microprocessors, cameras, sensors, and a custom rolling platform [105]. *Berenson*, named after Bernard Berenson, is a robot art critic by anthropologist Denis Vidal and robotics engineer Philippe Gaussier. The critic observes viewers' reactions to art and learns what is "good" and "bad" art. Then he moves toward art works that are "good" and smiles at them, and frowns at "bad" art. This robot uses a neural network to learn [142].

2.3.5 Robotic Environments

Artists have explored our relationship with and technology's relationship with the environment through synthetic natural environment artwork. Sometimes the environment is part of a larger work like some of the plants in the *The Ancestral Path through the Amorphic Landscape* (2003) installation by Chico MacMurtrie, while sometimes they are one plant like MacMurtrie's *Growing Raining Tree* (2003) [113]. Examples of electronic plants or environments move from the work *House Plants or Electronic Garden #2* (1983) by James Seawright [107], a set of 5 robotic plants, to the large scale installations of Philip Beesley starting in the early 2000's. Philip Beesley explores creating environments that fuse the technological with the biological, creating physical and sonic environments with audience interactions through proximity and touch. Some of his work intergrates soil, chemical reactions and water, creating robotic plants that require the same resources as living ones [143].

Some work that explores our environment explores the man-made environment that we have created. These works emphasize how our man-made world has caused us to become more machine-like and the violent machines that we have created to hurt each other. In 1990, Penny created a heat-seeking anti-personnel sculpture called *Pride of Our Young Nation*. It was designed to look like an artillery cannon and use an infrared heat sensor to aim at its victims. Once it found its victims, it would "fire" by rotating a large metal cone covered in spikes toward them [140]. *At the Edge of Chaos* (1995) by Louise-Phillippe Demers and Bill Vorn creates a space below the floor where machines violently move [50]. *Violent Machines perform acts of love* (1997) by Seemen Robotic Performance Group was an installation where the public could interact with violent and humorous machines [144]. The relationship between sex and machinization of humans by work was explored in *Executive Machinery Intercourse* (1999) by Istvan Kantor, a performance where mechanized drawers of filing cabinets opened and closed while dancers gyrated against them [144].

2.3.6 Robotic Performance Art and Telematically Mediated Art

Interactive anthropomorphic robotic artwork does not need to have the audience to interact with the robot at all or in the same space. Sometimes the work explores telematic interaction with anthropomorphic robots. *Ornitorrinco* (1986) by Eduardo Kac [139] and Ed Bennett and *Winke Winke* (1993) by X-Space [17] are examples of humanoid robots with which the audience interacted telematically. *Ornitorrinco* was a telematic work that was exhibited in two places simutaneously, with a robot in each location controlled by the other location. *Winke Winke* is a robot that communicates through semaphore and is controlled remotely by participants in the gallery.

Since *ROSA BOSOM* in 1964, humanoid robots have participated as part of performance works, that either interact with each other or performers [103]. Examples of contemporary works would be *Them Fucking Robots* (1988) by Norman White and Laura Kikauka, *Miyata Jiro* (1997) by Momoyo Torimitsu, *Afasia* (1998) by Marcelli Antunez Roca [144] and *Tiller Girls* (2010) by Louis-Philippe Demers [50]. These works are part of an artistic discussion about the mechanization of humanity. In *Them Fucking Robots* two robots were designed separately and then were brought together on exhibition day to copulate until they destroyed themselves. In *Miyata Jiro* a robotic Japanese businessman crawls down the street with the artist dressed as a nurse taking care of him. In *Tiller Girls*, women are replaced in the performance with simple robots that moved in a way that resembled a chorus line dance.

Cyborg works, as performances which include either visitors or performers, explore the humanization of the machine, or machination of the human. Examples of these works are, *Fractal Flesh/Ping Body/ParaSite* (1980 onward) by Stelarc [145] (See Figure 2.5), *Pendulum Choir* (2010) by Cod.Act [146] and *Inferno* (2015) by Louis-Philippe Demers [50]. The relationship between sound and the human body is explored in Cod.Act's *Pendulum Choir*, where the performers are attached to the machine while singing [146]. Stelarc explored the relationship between man and machine in many of his works, both augmenting and extending his body [145]. In *Inferno*, participants from the audience are strapped into mechanical suits that contort their body into the performance.

Zoomorphic robots have been included as part of robotic performance work. Examples of this would be works like the *Cybertheater* (1969) by Lev Nusberg and the 'Movement' Group, Survival Research Laboratory's performances in the 1970s-80s [144] with robots like the *Rabot* and *Piggly-Wiggly*, Stelarc's *Exoskeleton* (1999) [145], and Chico MacMurtrie's works like *The Ancestral Path through the Amorphic Landscape* (2003) [113]. In these works the audience does not interact with the robots themselves, but watches the robots interact with one another. These works explore ideas like the conflict of man versus machine and the industrial military complex.

2.4 Auditory Design

Robots and interactive technologies are moving from industry to our homes, so they need to be able to interact with novice users not dependent on language. While it is not easy to support natural conversations and emotions through an embodied agent, such qualities can make robots more relatable and predictable [147]. When technologies conform to social expectations, people find their interactions to be enjoyable, and empowering [69]. The expectation to interact with robots as if they are human conflicts with the discomfort that is felt when interacting with an agent that seems to be mostly but not quite human. As such, it may be preferable, where appropriate, to convey emotion through non-linguistic auditory cues, paralanguage, or prosody. It is viewed as acceptable for robots as well as computers to make non-linguistic utterances [148].

People prefer to communicate with robots via voice, and they prefer voice to be human-like [149, 150, 151]. In film, robots are typically voiced in one of two ways: by the voice of actors, whose speech is subsequently filtered (e.g. C3PO), or by non-speech computer generated sounds (e.g. R2D2) [152]. Despite prior research on the parametric synthesis of emotional speech, currently predominant approaches to speech synthesis use fragments of pre-recorded human speech. Fundamental frequency (F0) or Pitch, Amplitude or Intensity, Speech Rate or Tempo, and Articulation or Timbre (Table 2.1-2.4) have all been studied as vocalics which can effect the emotional meaning of a sound. In Jay Beck's "Lowering the Boom," three types of factors which affect vocal quality while filming are vocal distance, intended earshot and microphone perspective. Vocal distance is the way proxemics affect vocal sounds. For example, whispering can only be heard by someone nearby. Intended earshot is how the voice changes when trying to speak to people at different distances, like yelling to speak to someone at a far distance. Lastly in relationship to film is microphone perspective, the placement of the microphone and the microphone's audio idiosyncrasies [153].

We may need robots to convey information to us, but we need to set the users' expectations, so that they don't assume the robot is more intelligent than it is or needs to be. Cynthia Breazeal describes the concept of expectation setting, where a robot should be designed in a way that reflects how people should interact with it [49]. Non-verbal vocalization is a way to set expectations. This is supported by research such as R.Read's study with dog robot and person robot, which found that people preferred dog robots to make animal noises and a humanoid robot to make human-like sounds [148]. People sympathize with emotional robots and feel uncomfortable when avatars express emotion [154]. This is shown in our films, for example, in Short Circuit, when a robot that is too human in speech and intelligence must be freed [155]. The value of multimodal interfaces in ensuring accessible, usable, and satisfying interactions for human machine communication has been well established [51, 156, 52]. Emotion can be communicated through non-linguistic utterances [157], and one of the fastest triggers for emotion is sound [158]. People react ten times faster to acoustic than to visual cues [159].

Robots in particular are computationally constrained and are physically present in interactions. In order to better afford interactions with broad arrays of untrained users, it is valuable for a robot to be able to communicate without relying on constrained language or command vocabularies. Non-linguistic utterances (NLUs), which consist of sounds that are not associated to words or other parts of language, have been widely explored in robotic systems, both in research, and in popular media (see Section 2.4.3). NLUs are important because they can help to facilitate communication while lowering expectations for speech, through warning sounds, calming, or questioning sounds, among many other possibilities, that might streamline communication between technology and people.

Audio Feature	**Prior Studies**
F0	[160] [161] [162] [163] [164]
F0 mean	[165] [166] [167] [160] [168]
F0 perturbation/range	[165] [167] [160] [168] [169]
F0 variability	[166] [167] [160]
F0 contour	[167] [160] [169]
High frequency-energy	[160] [167]
Pitch	[170] [171] [172] [173] [174] [175] [169] [176]
Pitch average	[177]
Pitch range	[177] [178] [169]
Pitch Variation	[172] [177] [179]
Pitch Maximum	[180]
Major/ Minor Mode	[181] [180] [182] [183] [178]

Table 2.1: Summary of prior studies of emotion and fundamental frequency or pitch in music and speech perception: Pitch/F0, mean, range, variation, and major/minor mode were the most common audio aspects. In my algorithm I defined F0 frequency with the tonic note of the key, major or minor key, the contour of the phrase and the variation around the contour. ©2017 IEEE

2.4.1 Emotion and Prosody

The relationship between emotion and prosody, the pattern of stress and intonation, has been extensively investigated in both music perception and speech research. Most studies of prosody and emotion through the audio signal have focused on four attributes: Fundamental frequency (F0) or Pitch, Amplitude or Intensity, Speech Rate or Tempo, and Articulation or Timbre (Table 2.1- 2.4) Emotional valence has been correlated with pitch, intensity and rate in speech, with increased valence correlated with a higher pitch, higher deviation, larger range, higher mean intensity, larger intensity deviation, faster speech rate, shorter syllable duration and shorter/less frequent pausing. Decreased valence is correlated with the opposite effects of increased valence [168].

Audio Feature	Prior Studies
Amplitude	[161] [166] [173] [170]
Energy	[168] [167] [160]
Loudness	[172] [184] [174]
Intensity Mean	[160] [169] [177]

Table 2.2: Summary of prior studies of emotion and amplitude/intensity in music perception and speech research: Amplitude, Energy, Loudness and Intensity Mean were all equally studied. In my algorithm I defined amplitude by the attack and sustain amplitude difference and steady state amplitude. ©2017 IEEE

Audio Feature	Prior Studies
Articulation	[177]
Timbre	[180] [185] [174] [176]
Voice Quality	[177] [169]

Table 2.3: Summary of prior studies of emotion and timbre/articulation in music perception and speech research: Timbre was most frequently studied in music literature while voice quality/articulation was used in the linguistics literature. In my algorithm, I defined two envelopes, a steady state envelope and a percussive envelope, to affect timbre. ©2017 IEEE

Audio Feature	Prior Studies
Speech Rate/Tempo	[184] [172] [167] [177] [160] [175] [179] [174] [182]
Duration	[165] [161] [166] [180] [168] [169]
Pausing Total Time	[166] [172] [169]

Table 2.4: Summary of prior studies of emotion and speech rate or tempo in music perception and speech research: Speech rate, duration and pausing total time were most frequently studied in linguistics research, while tempo was the focus of music literature. In my algorithm, I used a rest (pause) length (mean and variance) and a motive length (center and range) to define tempo. ©2017 IEEE

There have been conflicting results in research on how tonality and timbre affect emotion. In researching an instrument's ability to express sadness, Huron found that instruments that were more percussive were perceived by musicians to be unable to express sadness [174]. In a pilot study, Le Groux found that valence was not correlated with any emotive sound, but the study only used percussive sounds [186]. Musical major/minor mode has been found to be capable of eliciting emotions with different valences [181]. Jee et al. used musical key in composing hand crafted emotion-carrying non-linguistic utterances for robots [187], but an algorithm proposed by the same authors did not use this strategy [188]. My generative algorithm uses percussive and steady state envelopes, as well as major or minor mode as parameters.

2.4.2 Sonic Installations

Another aspect of this dissertation is the artificial generation of music for art. Robotic instruments have existed since the 14th century with automated carillons [189]. One of the earliest works of generative music, dice music, was Johann Philipp Kirnberger's *Der allezeit fertige Menuetten- und Polonaisencomponist* in 1757 [190] in which the composition was created by rolling dice to choose the next musical motive. More contemporary works include David Cope's *Experiments in Musical Intelligence (EMI)* which used a data driven model and recombinancy to generate compositions [191]. More recently, Peter Ablinger made a piano speak by automated means in *Speaking Piano* [192].

My work *Cacophonic Choir* uses spatialized sound to communicate emotion. This pulls from Janet Cardiff's work examines spatialized environmental audio. In these spatialized audio works like *The Forty Part Motet*, people can explore the sonic space. She studies proxemics with artwork like *Opera for a Small Room* where the work is a room within a room. The audioscape is in a small room which people can walk around but cannot go inside. Her work is not robotic but is relevant to this thesis for the spatialized sound aspect of her work [193]. David Rokeby has explored gesture-controlled sonic environments from *Reflexions* (1983), *Very Nervous System* (1986) to his more recent works like *International Feel* (2011) [194].

2.4.3 Robots and Auditory Communication

The first speaking machine, a person controlled mechanism with bellows, was designed in 1769 by Wolfgang von Kempelen. It could only say a few words [195]. Around the same time, C. G. Kratzenstein constructed various shaped tubes that produced five vowel sounds. The first electrical speaking machine was the Voder designed by Homer Dudley in 1939. [196] Research has continued to the present with work like the Waseda Talker Series from the Humanoid Robotics Institute at Waseda University [197] and Hideyuki Sawada's KTR-2 which sings [198]. The first computer to sing was the IBM 7094 in 1961, singing the song Daisy Bell. Vocals programmed by John Kelly and Carol Lockbaum inspired a scene in *2001: A Space Odyssey* [199]. In the contemporary artwork *Phatus*, Penny works with the idea of trying to reproduce how people make noises by creating artificial vocal cords and lungs. He is still currently working on this project with another form of interaction, speech, but instead of just using electronic forms of synthesis, this speech is purely mechanical [140].

Robots and Speech

Communicating emotion is important. A learning robot may receive more and better training data if it expresses emotion through the pre-recorded speech that consists of entire utterances of an actor [200]. However, using pre-generated speech can limit the range of possible interactions, motivating the use of synthesized speech. Cahn et al. implemented a system, using the DECtalk3 speech synthesizer, where 19 parameters controlled emotive speech. They found that the correct emotion was chosen by study participants 53% of the time out of six distinct emotions [201]. Nourbakhsh et al. created a robotic tour guide that expressed emotion through voice [202], but expressivity and communication were only briefly and informally discussed by the authors. Likewise, Roehling and Xingyan propose systems for producing emotional natural language speech by robot, but these systems were not systematically evaluated [203, 204]. Niculescu et al. found that modifying the pitch of speech could elicit differences in perception, and that a robot dressed as a woman with a higher pitched voice was rated as having a more attractive voice, as being more aesthetically appealing and outgoing, and as having better social skills [205]. This indicates that people can project human qualities onto robots in ways that depend on the audio qualities of speech sounds.

Gibberish as Non-verbal Communication

A simpler approach to emotional communication by a robot is to use gibberish or non-linguistic utterances. Such an approach might better avoid the uncanny valley of robot speech. Cynthia Breazeal's robot, Kismet, uses child-like utterances to reinforce emotions [49]. Oudeyer used child-like babble to convey emotions that could be interpreted by people from different countries [206], by means of the MBROLA synthesizer. The latter could produce 30 sounds, expressing Happiness, Sadness, Anger, Comfort and Calmness, using different input strings. The participants had a high accuracy in categorizing the sounds, though they confused the sounds representing comfort and calmness.

Yilmazyildiz et al. proposed a method for generating gibberish using a database of sounds and a prosody template, but this algorithm was systematically evaluated [207]. In 2010, the same group presented a different approach, based on presenting vowel sounds from sentences with different emotional content, generated via a text to speech synthesizer [208]. They found that it is important to match the input language with the language of the text to speech synthesizer, and that it was easier to determine whether the statement was positive or negative if it was presented in the correct semantic context. They combined both methods in 2011 and found that users were able to accurately recognize seven emotions that were presented [209].

Non-Linguistic Utterances

Non-linguistic utterances (NLUs) are computer generated phrases that have audio qualities similar to speech but do not use vocal sounds [210]. Non-linguistic utterances are inexpensive computationally [210], and can be cross-culturally effective [206].

NLUs are an example of non-speech audio, similar to the related earcons, which are nonverbal sounds that are used to convey information. Bill Gaver discusses the idea of “Auditory Icons,” or sound based icons that are “caricatures of naturally occurring sounds.” [211] Sound based icons that are distinct from naturally occuring sounds, or Earcons, have been extensively investigated in the human factors literature for applications ranging from conveying weather information [212] to warning drivers [213]. Read et al. found that children will assign emotional meaning to non-linguistic utterances produced by a Nao robot, though they could disagree on the emotion [210]. In contrast, adults more consistently categorized emotion in non-linguistic utterances [214]. Non-linguistic utterances can create the appearance of a stronger emotional reaction in robots when produced in response to an action [152].

Other researchers have investigated combinations of visual and auditory cues to emotion. Sparky, a small (50 cm) robot with an expressive face, a movable head on a long neck, and wheels was created by Mark Scheeff at Interval Research in order to investigate human-robot interaction [215]. The robot could make chirping sounds, but these were found to be confusing. Bartneck found that an avatar could more effectively elicit emotional responses using visual cues, or audio and visual cues, than with audio cues alone [216].

Robot toys like those from Keepon and WowWee use small sound databases to communicate emotion. Keepon uses them as a means of attracting attention and in response to sensory input [217]. WowWee robots also use sounds in response sensory input [152]. These preprogrammed sound sets are analogous to pre-recorded voice acted sound sets, and are not as flexible as a generative system.

Jee et al. designed non-linguistic utterances for robots using musical structures, including tempo, key, pitch, melody, harmony, and rhythm to represent happiness, sadness, fear and dislike. Results showed that composed music was able to express emotion and worked best when paired with visual cues [187]. Jee et al. later proposed an algorithmically generated musical system for robots using tempo, pitch, and volume to express joy, distress, shyness, irritation, expectation, dislike, pride, and anger [188], but the effectiveness received limited evaluation. The same authors analyzed sounds from film robots R2D2 and Wall-E, and found that intonation, pitch, and timbre were used to express emotions. Using this information, they designed five sounds and found that about half (55%) of study participants felt that the sounds reflected the intended emotion, and that most (80%) felt the sounds carried emotional expression [218]. While successful, these sounds could not be algorithmically produced to yield a specified emotional response. Other systems have been proposed but not tested [219]. Wallis et al. designed computer generated emotional music which was shown to be effective for expressing emotion, but has not been tested using robots [220].

Komatsu et al. found that sounds produced by a robot could affect whether a suggested action was performed [221], and that people preferred robots to communicate confidence using non-linguistic utterances (earcons) rather than through language or paralanguage [222]. They also asked participants to select attitudes that matched sounds produced by a robot or a personal computer. The results showed that the participants were more able to interpret computer sounds than sounds from the robots [223]. The design of their study was prone to bias because the NLUs were originally categorized by users listening at a computer [222]. This may explain why they found that NLUs were interpreted correctly more often on a computer in the latter study.

2.4.4 Algorithm for NLU Generation

NLUs are designed as musical phrases, which are produced algorithmically in order to elicit desired emotional responses. In order to create the largest possible repertoire, and to ensure engaging interactions, I designed an algorithm that could produce a very large variety of NLUs that elicit specified emotions, greatly increasing the vocabulary of the robot over other approaches, such as pre-recorded speech.

Melodies are produced by a generative algorithm, which varies the fundamental frequency (F0), amplitude, timbre (temporal envelope), and motive according to the specified emotion, based on the sound perception and emotion results cited in Section 2.4.1. Each of these aspects is controlled by two to five descriptors that specify a given musical phrase (Table 2.6).

Emotion	**“Sad”**	**“Excited”**	**Reference**
Base F0 (tonic note)	B♭ (466.164)	G (1567.98)	[152]
F0 range (variance)	3 notes	5 notes	[152]
Pause Length	Long (avg. .4s, std. .2s)	Short (avg. .2s, std. .1s)	[168]
Pause Ratio	High (avg. 3:1, std. 2)	Low (avg. 6:1 std. 2)	[168]
Envelope	Steady State (.2s attack, .6s sustain, .2s decay)	Percussive (.05s attack, .2s decay)	[174]
Pitch Contour	Rising (contour end difference = 5)	Flat (contour end difference = 0)	[224]
Mode	Minor	Major	[181]

Table 2.5: Descriptors for two emotions: This study focused on two emotions “Sad” and “Excited”. The values to control those non-linguistic utterances were based on by prior research. ©2017 IEEE

This work focused on two families of emotion in non-linguistic utterances, corresponding to “excited” or “sad” sounds. The emotion “sad” was defined as low valence, low arousal and submissive in the PAD emotion space. In contrast, “excited” was defined as high valence, high arousal and dominant. The parameters it used were base fundamental frequency, frequency range, speech rate, pause ratio, envelope, pitch contour and mode. For illustration, a sad emotional state is often expressed in speech by low base frequency, small frequency range, low speech rate and high pause rate, while an excited emotional state is often expressed in speech by high base frequency, large frequency range, high speech rate and low pause rate. The emotional arousal was manipulated through the temporal envelope (“percussive envelope” and “steady state envelope”) of the sound. Two musical modes were used in order to affect emotional valence: major and minor mode. I also tested rising-pitch intonation and no pitch intonation which is related to submission (Table 2.5). An extensive analysis of all the parameters’ impact is left for future work.

The NLUs are designed as a collection of enveloped sine waves (notes) with occasional

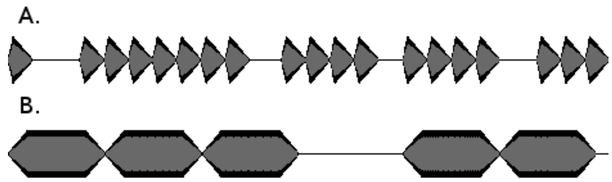

Figure 2.6: The waveform for an "excited" NLU (A) and a "sad" NLU (B). ©2017 IEEE

pauses between them. The notes are chosen using a heptatonic scale, with the potential range of a piano. The scale is defined by the tonic note, and whether it is in major or minor mode. The melody of the notes is randomly generated based on the frequency variation and whether the end of the phrase has a rising contour. Pauses are inserted occasionally between notes. The number of notes without a pause and the length of the pause varies throughout the musical phrase. Since the starting note, the melody, and the pauses have a certain amount of randomness, many different non-linguistic utterances can be generated for one set of input parameters.

NLU Melody Generation

NLUs are comprised of short randomly generated melodies (described here) which are composed of synthesized notes (discussed below). The frequency of each note is dependent on five values: tonic note of the key, whether the NLU is in a major or minor key, the contour of the phrase (start and end values) and variation around the contour (Figure 2.7). The tonic note, and whether it is a major or minor key, determines the scale. The contour, the rise and fall of the melodic line, is defined by two contour values. It is created by concatenating two vectors, each of which have equally spaced values from first contour value to zero, and from zero to the last contour value. This allows the pitch to ascend, descend and plateau at the beginning or end of the melody. I added the contour to a random vector of equal length with a mean of the tonic note and the

Audio Aspect	**Descriptor**
F0 Frequency	tonic note of key major or minor key the contour of the phrase (start) the contour of the phrase (end) variation around the contour
Amplitude	attack and sustain amplitude difference steady state amplitude
Timbre	length of the attack decay sustain mean sustain variance release
Motive	pausing length mean pausing length variance motive length center motive length range

Table 2.6: Parameters for NLU generation algorithm: Based on prior research in music, psychology and linguistics, 16 parameters were chosen to control the non-linguistic utterances. ©2017 IEEE

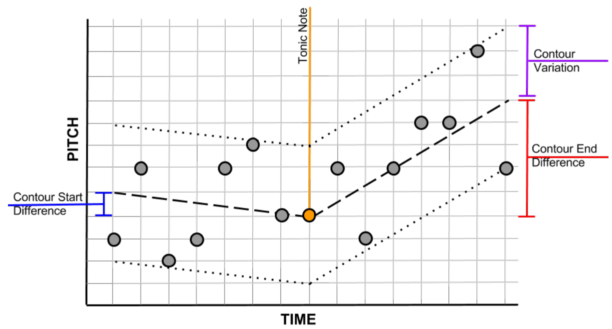

Figure 2.7: Melody of phrase algorithm: The melody is dependent on the tonic note of the key, whether the NLU is in a major or minor key, the contour of the phrase (start and end values) and variation around the contour. In this chart the randomly generated melody, comprised of grey dots representing notes, is shifted to fit the contour. The creation of the melody is described in the Section 2.4.4. ©2017 IEEE

variance defined in Table 2.5. I varied the degree of excursion (or variation) above and below the contour in order to generate notes for the melody that comprises the NLU. The major/minor mode of a melody was not forced, and particularly with shorter melodies might be imperceptible.

Note Timbre and Amplitude

The timbre of the notes in the melody that comprise the robot's non-linguistic utterances was controlled through an attack decay sustain release (ADSR) envelope that was applied to the sound (Figure 2.8). Different timbres were produced by manipulating the attack, decay, sustain mean, sustain variance and release of the envelope. A sustain length was randomly generated using these values, and an ADSR envelope was generated for each note. Two types of ADSR envelopes were applied to produce different timbres: a (short) percussive envelope and a (longer) steady state envelope (see Figure 2.6). The percussive envelope had a sustain mean and variance of zero. The steady state envelope difference in attack and sustain amplitude was zero and had a decay time of zero. While the algorithm to generate emotive NLUs allowed for more complexity in defining the am-

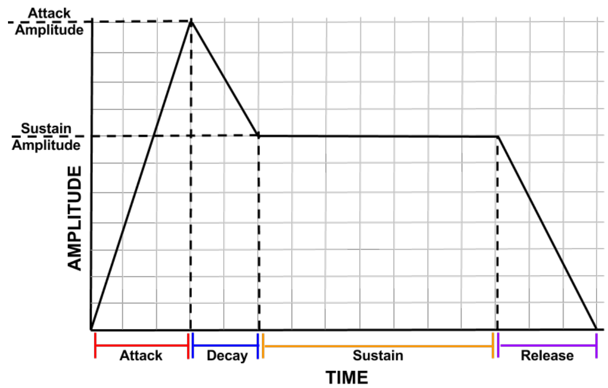

Figure 2.8: ADSR Envelope: The timbre is created using an attack decay sustain release (ADSR) envelope. In this study I specifically used a steady state and a percussive envelope. ©2017 IEEE

plitude, I decided for this study to keep amplitude consistent between the two emotive utterances because the amplitude of the utterance differed between the computer and the robot playing it.

Motive

The phrasing, or prosody, of the note sequences was algorithmically generated by controlling parameters that consisted of motive length, center, and range, and pause length, mean, and variance (Figure 2.9). The motive length describes after how many notes there is a pause. The motive length is calculated by choosing a random value within the motive range, until the sum of the motive lengths was at least as large as the length of the phrase. The length of each pause is sampled from a normal distribution parametrized by pause length mean and variance. The length was specified with millisecond resolution.

2.4.5 Algorithm for Gibberish Generation

The system defined above, for creating emotive sound based on research in music and linguistics, was used to create emotive gibberish (see Figure 2.10). Gibberish has

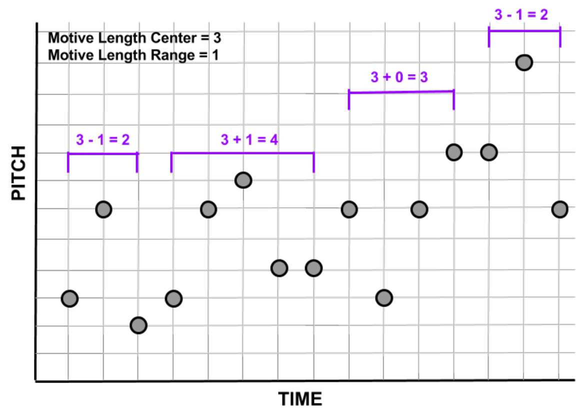

Figure 2.9: Motive length and range algorithm: The motive is described by the motive length center and range, as well as the pausing length mean and variance. An example of how a series of notes can be divided into motives is denoted by the purple lines, with the average motive length 3 notes long and the motive range 1 note. ©2017 IEEE

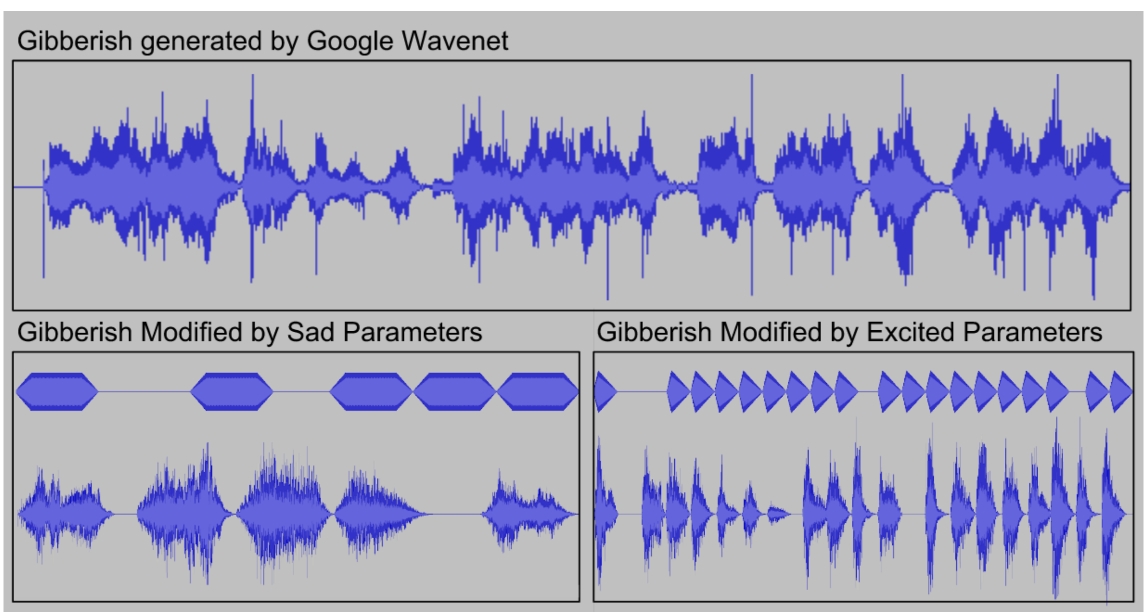

Figure 2.10: Waveforms of gibberish produced by emotive sound algorithm.

been extensively researched for robots to communicate emotion (see Section 2.4.3). The algorithm's input was computer generated gibberish instead of sine waves. Aspects of fundamental frequency, amplitude, timbre, and motive were manipulated to create the gibberish. The gibberish was created by an implementation of Google's WaveNet, a neural network generative model for raw audio [225]. The wavenet was trained over the course of 10 days and 80250 iterations, on the Centre for Speech Technology Research VCTK corpus [226]. The VCTK corpus includes speech data uttered by 109 native speakers of English with various accents reading 400 sentences. Once the wavenet was trained, 10 seconds of gibberish was generated for each of the 109 input voices. This gibberish was used as an input sound for the algorithm to create emotive sound. The onsets in the gibberish were detected and used to split the gibberish into words. Each word was processed as a note was in the original algorithm, modifying the frequency, amplitude, timbre and motive. The final sounds were expressed by ROVER at the Media Arts and Technology Program 2017 End of Year Show and the International Symposium of Electronic Art (ISEA) in Bogota, Colombia [157].

2.4.6 Emotive Communication in Prototypes

Through robotic art, I explored how the participant's perception of the robot is affected by different types of emotive vocal communication and the robot's ability to move. Drawing upon earlier research discussed above, I created *Come Hither To Me!*, an interactive robotic performance piece which examines the emotive social interaction between an audience and a robot. This work employs affective computing in the dialogue design. I was interested in evoking emotions in the participating audience through their interactions with the robot. *Touching Affectivity*, a small furry robot that sonified the way it is touched, explored non-linguistic utterances based on the methodology described in

Section 2.4.4. *Cacophonic Choir* is an interactive installation composed of multiple emotive embodied agents distributed in space. From afar, the work sounds like an incoherent and incomprehensible cloud of murmur. As one gets closer to an individual agent, words begin to emerge from the gibberish. Only when near do the visitors hear stories of sexual assault survivors. This work both explores distance at all three levels and speech clarity from gibberish to language.

Chapter 3

Social Robotic Art: *Come Hither to Me!*

3.1 Introduction

Come Hither to Me!, an interactive robotic performance, examined the emotive social interaction between an audience and the robot, ROVERita. The interactive robot, which used the ROVER platform, attempted to communicate and flirt with audience members in the gallery. ROVERita used feedback from sensors, auditory data, and computer vision techniques to learn about the participants and to inform its conversation. The female robot picked her favorite from among the audience, approached them and started charming them with flirtatious comments, jokes, backhanded compliments, and personal questions. The goal was to evoke emotions in the participating audience through their interactions with the robot. *Come Hither to Me!* strove to invert gender roles and stereotypical expectations in flirtatious interactions. This performative piece explored the dynamics of social communication, objectification of women, and the gamification of seduction. The robot reduced flirtation to an algorithm, codifying pick-up lines and

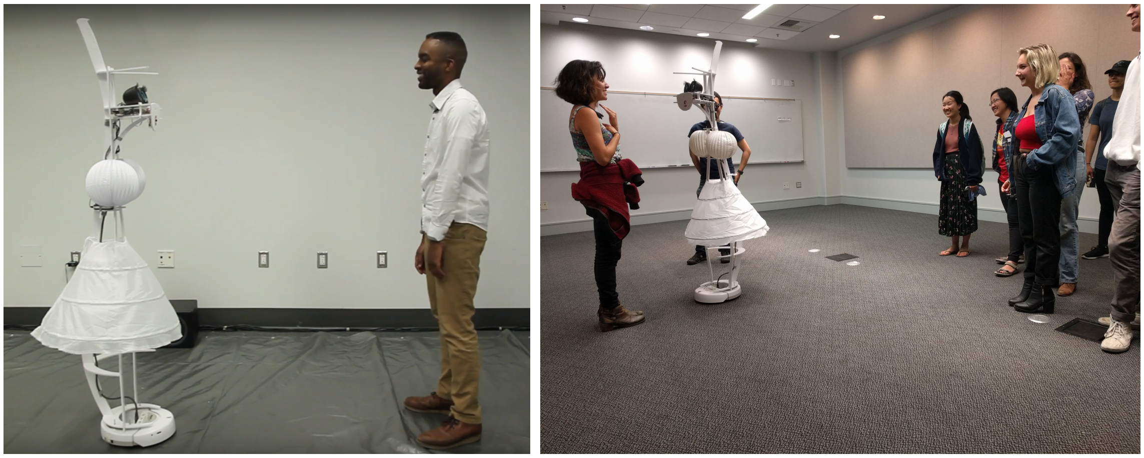

Figure 3.1: ROVERita approaches and flirts with a male participant and interacts with a female participant in the Annual Exhibition of Media Arts and Technology Program, University of California, Santa Barbara.

sexting paradigms.

Come Hither to Me! was a performance art piece about the relationship between humans and digital technology, focusing on believable and engaging interactions with robots. The female robot approached the visitors, chose the hottest person in the room based on the data from her heat camera, and communicated with them using flirtatious and playful dialogues (See Figure 3.1). The conversations were designed to be engaging, personal, and humorous, and make the visitors interested in continuing the interaction. Using audio recorders and video cameras, the interaction data was recorded during the performance. Afterward, the data was analyzed to evaluate the effectiveness of the communication and the quality of the interaction with the robot, and to compare it to video interactions with ROVER.

ROVERita represents a robot attempting to infiltrate human culture and to understand how we make meaningful connections. Computationally following dating rules and practices to assimilate into our society, the robot becomes a probe, collecting data on our social experience. The work explores the agency of the female form as a body capable of choosing her partner, either male or female. The embodiment of the interactive robot as a female cyborg enables the piece to examine the alienation and objectification of women in this male-centric aspect of our culture, and makes meaningful connections with a humorous and yet critical approach. Subverting the stereotype of a pick-up artist by using a female robot, I problematize the issues of female dating agency and dominance against the expectation of passivity. By endowing the robot with an overtly sexualized doll-like form and female voice, I created a sense of discomfort in the participants with whom she flirts and emphasized her female-ness while inverting gender roles in dating.

3.1.1 ROVER as a platform for Interactive Art and HRI

In the installation *Come Hither to Me!*, the robot is built using the ROVER platform. ROVER, the Reactive Observant Vacuous Emotive Robot, is an interactive robotic performer that navigates the gallery, finds visitors, and communicates with them. ROVER is both a robotic sculpture and a research platform for affective computing. The design decisions for ROVER were inspired by the field of media arts. Based on audio and human emotion research, the system modulates audio qualities like timbre, fundamental frequency, contour, mode, and tempo to portray and evoke emotions and emotional responses. It also learns about the space it navigates, mapping obstacles and people, using proximity, bump, and heat sensors. As an experimental platform for human robot interaction, ROVER was used previously to conduct research on humans' affective response to emotive sounds produced by an embodied agent [157].

The first iteration of ROVER was as a site specific installation. The goal of the artwork was to bring warmth and joy into a poorly lit, cold, desolate, antiseptic, and windowless space. The initial design of ROVER was a 4-foot tall, fuzzy, awkward creature that searched the over-air-conditioned hallway for warmth, expressing joy through sound when it found someone. ROVER's name references planetary rovers and the traditional name for a dog, which can be taught through conditioning. The value function, the happiness expressed in the song, was determined by how much the person smiled in response to it, as detected by the camera on the system.

The next iteration of ROVER incorporated a stronger research component. The second prototype was designed to collect data from participants to learn about how they would interact with emotive robots. The results of the research regarding humans' reaction to emotive sound was published in *IEEE Transactions on Affective Computing* [157]. ROVER was designed to explore the space in which it found itself, as an alien would attempt to learn about how to communicate emotion. The second iteration of ROVER had an uncomfortably biological, alien-like form: 6 feet tall, skeletal, white in color, with its internals on display. The skeletal structure and wires emphasized the distance between technology and humanity.

ROVER takes inspiration from early cybernetic art, more contemporary robotic art, and other automated works, particularly those emulating creative human speech and behavior (see Section 2.3.4). As an interactive piece of art, the research focuses on how humans react to the work itself. ROVER leverages the design space inspired by media arts to solve problems like autonomy, movement, appearance, and interaction. ROVER references Robert Breer's sculptures at Pepsi Pavilion and Grey Walter's *Tortoises* with the use of the iRobot CREATE, though, unlike these works, it also interacts with the public. While ROVER is mobile like *K-456* and *ROSA BOSOM* and is used for disruption, it is computer controlled like *S.A.M.* by Edward Ihnatowicz. With respect to contemporary art, ROVER references many approaches that have created robots with more human-like, cognitive abilities, such as analysis, creation, and performance, as well as ones that emulated other qualities, such as basic bodily functions and speech. *Berenson*, is similar to ROVER in that it observes viewers' reactions to stimuli for analysis [142]. ROVER is also reminiscent of Simon Penny's *Petit Mal* in movement, though it uses heat for movement. Ken Rinaldo's *Paparazzi Bots*, similar to ROVER, moved at human speed, avoiding obstacles using multiple microprocessors, cameras, sensors, and a custom rolling platform [105].

3.2 ROVERita Design

Artists have used robots to explore gender and sexuality, and its effects on the relationship between human and the machine (see Section 2.3.4). ROVERita's interaction design was heavily influenced by early cybernetic interactive works like Bruce Lacey's *ROSA BOSOM* (1965) and Norman White's *Helpless Robot* (1987), as well as non-interactive works like Norman White and Laura Kikauka's *Them Fuckin' Robots* (1988) [144].

For *Come Hither to Me!*, I designed a female body for the robot and named her ROVERita, in order to provoke dialogues about dating politics and sexual objectification of women in a humorous performative context with subverted gender dynamics. The structure of ROVERita's abstract female body and dress was inspired by the costumes from Oskar Schlemmer's Bauhaus performance, *Triadisches Ballett* (1922) (See Figure 3.8 and 3.3). Other examples of interactive robotic artwork that comment on robots and primal desire are Nam June Paik's *K-456* (1964) [103], the Seemen Robotic Performance Group's *Violent Machines Perform Acts of Love* (1998), and Peter Croppin's *Project Paradise* (1998) [144, 227].

3.2.1 Gender and Human Robot Interaction

ROVER was used to perform as ROVERita in "Come Hither to Me!" ROVERita is a feminine version of ROVER, with two paper lanterns for breasts, a hoop skirt and lipstick (see Figure 3.8). The gender of a robot can influence the persuadablility of a participant with participants being persuaded more by robots of the opposite gender

Figure 3.2: ROVERita, the female embodiment of ROVER, is a seductive female interactive robot.

Figure 3.3: Schlemmer, Oskar, 1888-1943. (1922). Triadic Ballet/set design. Retrieved from `https://library.artstor.org/asset/ARTSTOR_103_41822001725439` Adrienne Wortzel. (2001). Kiru. [Performing Arts (including Performance Art), Science, Technology and Industry]. Retrieved from `https://library.artstor.org/asset/LARRY_QUALLS_1039760910`

[228]. Niculescu et al. found that a robot dressed as a woman with a higher pitched voice was rated as having a more attractive voice when compared to one with a lower pitched voice [205]. The gender was telegraphed by a feminine hair shape and a red bow. These signifiers are called "cultural genitals" [229]. There is a problem with robots furthering gender stereotypes, examples of which have already been seen in Japan [230]. Previously ROVER was gender ambiguous and frequently referred to as he or it, despite having no gender sigifiers. As ROVERita, she took on some cultural gendered baggage. Some female participants thought the installation might only be for men due to the gender.

3.2.2 Interaction Design

ROVERita was designed to find and move toward a participant in the room using a low resolution heat sensor, to avoid obstacles with bump and proximity sensors, to stop three feet away and play a melody, recording the facial response of the participant. The movement algorithm is described in more detail in Figure 3.4. As long as ROVERita

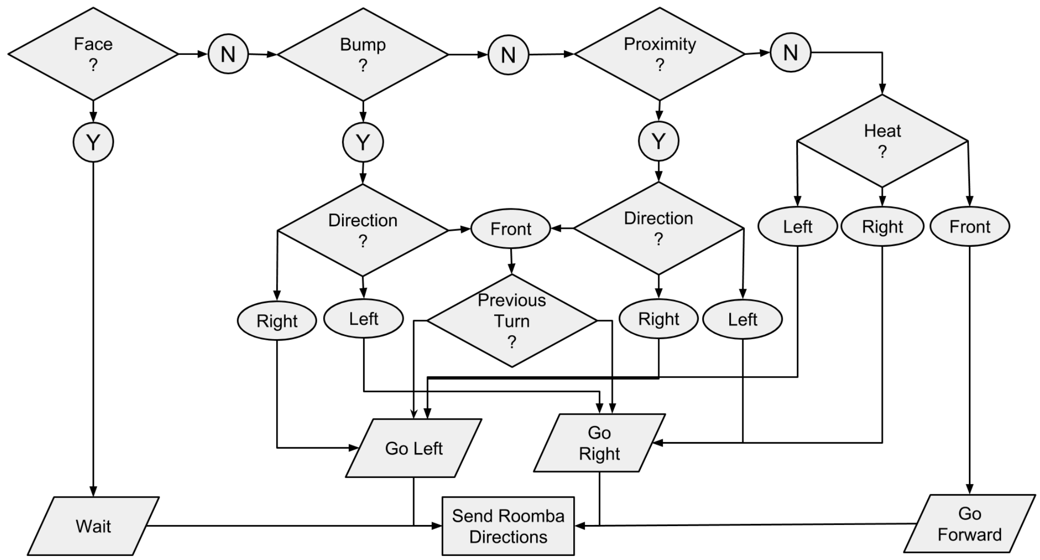

Figure 3.4: ROVERita's movement algorithm

could see a face, ROVERita would continue to talk to the participant, until they moved away. If at the end of a statement ROVERita could not see a face, ROVERita would continue to move toward the nearest hot area determined by the heat camera.

In *Come Hither to Me!*, participants, individually or in groups, went into the performance space where the robot awaits them. ROVERita identified the participants in the room, made her way toward one participant, and initiated the conversation with a flirtatious pick-up line. The participants could choose to respond to the robot and engage in the conversation, or walk away, in which case the robot would choose someone else. The robot's ultimate goal was to obtain the participants' phone numbers. The participants could end the conversation or eventually ROVERita would say goodbye. Using the ROVER platform, ROVERita could detect the "hottest" person in the room, move toward them avoiding obstacles, and stop three feet away, engaging in conversation.

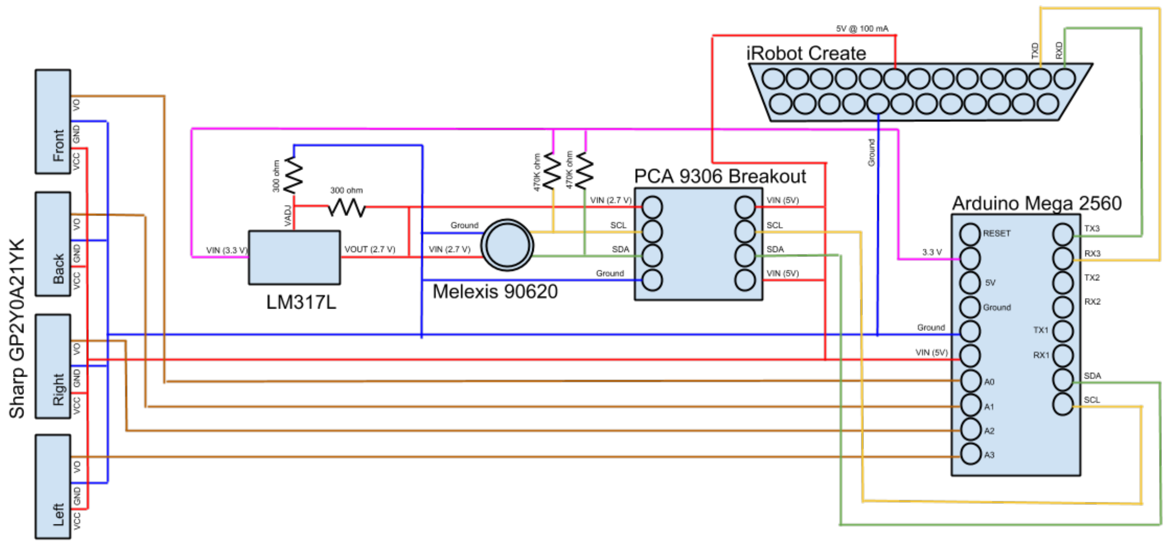

Figure 3.5: Wiring diagram for ROVERita

3.2.3 Construction

ROVERita's technological platform, is constructed on a base unit (iRobot CREATE), using an embedded computer (Raspberry Pi) with a digital camera, a low-resolution heat camera (Melexis 90620), a microcontroller (Arduino Mega), proximity sensors, and speakers. The wiring of the sensors is explained in further detail in Figure 3.5. ROVERita's frame was laser cut from white acrylic. The robot navigates with a combination of iRobot CREATE bump sensors, waist-height proximity sensors, and a low-resolution heat camera. ROVERita can either be active (navigating the space), or passive (stationary, waiting to be approached). ROVERita stops and communicates vocally when it detects a person's face approximately three feet away [157]. ROVERita's platform ROVER, originally produced pre-generated audio, algorithmically created in Python using parameters taken from affective music and linguistics research to express excitement and sadness [157].

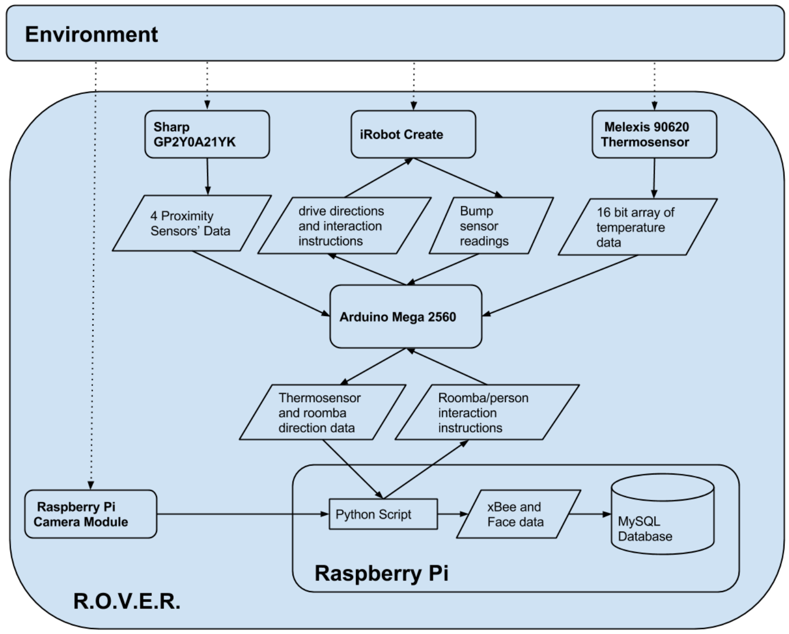

Figure 3.6: ROVERita's technology stack

3.2.4 System Design

ROVERita's technology stack is an Arduino connected to sensors, which communicates over a serial connection to a Raspberry Pi with a camera module (Figure 3.6). The main extension of the system for ROVERita was a communication protocol over wifi with an off-board computer that ran text to speech on the microphone input. This decided ROVERita's next actions and sent her instructions. Open Sound Control [231] was used to send messages over the network.

3.2.5 Conversation Generation

ROVERita builds upon prior research with robots and speech (see Section 2.4.3). The conversation algorithm implemented was a decision tree. The decision tree was designed based on suggested strategies of "pick-up artists" from popular self-help books and websites on how to flirt with and attract women, in addition to data collected at a few exhibitions [232]. The dialogue was editing for the specific context of each performance. It was designed as a series of questions with flirtatious comments and lightly insulting compliments, loosely based on pick-up artist strategies, inter-dispersed throughout [232].

Between the two exhibitions, extensive informal interviews were completed with colleagues we recruited. In these interviews, my collaborator would have a conversation with the participant by reading from the ROVERita dialogue, while I would take notes. Here we would look for unexpected responses and common responses that we had not planned for. We modified the dialogue to take into account the responses for the next participant. These interviews greatly influenced the dialogue, particularly for creating sub-categories for participant responses to flirtatious and negative comments.

Pick-up Artist Strategies

The idea of lightly insulting compliments is referred to by pick-up artists as "negs" [233]. The idea behind this is that lowering someone's self esteem will make them be more likely to respond positively to requests. The concept is loosely supported by some studies which show that lowered self esteem's increases the attractiveness rating of a flirtatious research assistant [234] and that there is a correlation between lower self-esteem and being more compliant with requests of others [235].

A 2011 study focusing specifically on pick-up artist strategies in heterosexual courting categorized three measurable traits as the male 'competes,' 'isolates' and 'teases.' The study found that men who rated highly on hostile sexism and sociosexuality surveys reported use of the strategies. Women who rated highly on the surveys reported desirability of the strategies. Within a population of undergraduate men only sociosexuality surveys predicted men's reported use of the strategies [236]. ROVERita in her interactions mostly focuses on teasing but is also competitive and slightly isolating.

Under the category “compete” ROVERita uses all the strategies listed in the 2011 study. She asks if the participant is there by themselves and, if they mention a significant other, she both makes a backhanded comment against them and continues to try to pick the participant up. She will not give up after being rejected once and has to be rejected multiple times before responding in a hostile manner and leaving. Also if the participant gets defensive against a “neg,” ROVERita will not let up.

Under the category of “isolate” ROVERita only tries to find out if the participant came to the exhibition alone. The other isolation strategies the paper describes like making sexual comments about what they would do, trying to get the person alone and letting the person know that they are in control, are harder to do in this setting or would be considered sexual harassment. Lastly ROVERita “teases” the participant a lot, through responses to participant’s responses to questions and negs. She teases the participant a lot, gives them a hard time, picks on their appearance, is a little insulting, acts like a jerk and makes jokes at the participant’s expense.

3.2.6 Dialogue Design

The questions and conversational flow were originally based on pickup artist strategies. The conversation starts with a pick-up line, then follows with small talk questions that the robot can then relate to showing commonalities. That is followed by asking more sensual and tempting questions. After establishing a rapport, the robot states that it has time constraints and asks about interests in common activities. Once a common activity interest is confirmed, the robot asks for their phone number to go on a date. ROVERita also listens for “keywords” that were responses to questions.

	Fliratious Introductions
Part 1	Hi, you looked so beautiful over here, I had to come up and say hello. Hey babe, How you doing today? Enjoying the exhibition? Hey hey! You look like someone I'd like to meet. You are so damn sexy, and I am going to get to know you. I apologize to be so frank, but man you look so hot! Woohoo! Look at you! You are the hottest thing in the room. Wait, I want to see you a little closer. Hey gorgeous! Let me see you a little closer. Your eyes are lovely. Let me see the color of your eyes! Hey you! You look gorgeous! Can I talk to you for a minute?
Part 2	You look interesting, I'd like to get to know you. You're cute, are you interesting? Have we met before? I feel that I know you from somewhere! You have thoughtful and beautiful eyes. I think you have a lot going on inside there. The rest of you are all right! But this one! Wow! Come closer sweetheart! Don't be shy! I couldn't help but noticing those beautiful eyes and gorgeous skin. So happy you walked in.

Table 3.1: ROVERita would make two flirtatious comments at the beginning of the conversation at CHI. One from part 1 and one from part 2 would be chosen randomly, with time in between for the participant to speak.

Final Dialogue Decision Tree

ROVERita's algorithm was written in Python using Apple's speech to text technology. In general if ROVERita hadn't seen someone's face within a given number of frames, she would say, "I wasn't done talking to you!" or a similar response. ROVERita's responses were pre-recorded, using a text-to-speech algorithm. ROVERita would "listen" after most statements she made for a short period of time. When the participant had paused for 2-4 seconds or if the participant said nothing, the input was considered complete and the response was searched for specific keywords. There were 3 main categories of responses: "general comments", "context specific responses" and "general questions." For each of the three categories, the algorithm would go through the sub-categories to

	ROVERita's Flirts and "Negs"
Flirt 1a	Do you have any idea how adorable you are? You know how gorgeous you are. I can't help but notice you. You're like sunshine on a rainy day. Wow. You're just so... Wow. You know how beautiful your smile is. You light up the room.
Flirt 1b	Your eyes are breathtaking. By the way, this color is perfect on you. You smell really good. By the way your hair looks stunning. Your voice is magnificent.
Neg 1a	By the way, I really like those shoes. I just saw someone wearing one a few minutes ago. Love those shoes, remember how they were popular 2 years ago?
Neg 1b	I like the way you give up fashion for comfort Really nice shirt, by the way. I hear polyester is the hot fabric of the season this year. I really like that outfit. My grandmother had one just like it.
Flirt 2a	You really are gorgeous!You're irresistible when you blush. I wish I was your mirror, so that I could look at you every morning. You know how beautiful you are. There isn't a word in the dictionary that can explain your type of beauty.
Flirt 2b	You really are stunning! Your insides are even more beautiful than your outside. You're so hot, I get a tan every time I look at you. I was feeling a bit off today, until you turned me on.
Neg 2a	Your hair is so beautiful! But wait, is that a wig? I really dig your hair color, although your roots are showing a little. I'm just kidding! I like your eyes, especially the left one. Just kidding. You have beautiful eyes. You have really nice eyes. Hey, are those contacts? By the way, it's really cute how your nose wiggles when you speak
Neg 2b	You really look stunning. Must be the lighting or something. Lighting on a beautiful scenery. You are in good shape! I can see you work out, well occasionally at least. Hey, nice elbows. Really, I mean it! Ew! You spit on me! Why do you blink so much?

Table 3.2: Between questions, ROVERita would make two flirtatious or negative comments at CHI. One would be chosen randomly. After ROVERita waited for an answer she would respond to the answer based on 3.3.

"Neg" Responses	**Flirt Responses**
Insult ROVERita	Insult ROVERita
Thank ROVERita	Thank ROVERita
Question specifics about ROVERita's response	Compliment ROVERita
Questioning truth of ROVERita's statement	Questioning truth of ROVERita's statement
Keywords	Keywords
Surprise	Surprise
Agree with ROVERita	Agree with ROVERita
Questioning response	Questioning response
Negative response, Defending oneself	Negative response to compliment
Neutral	Neutral

Table 3.3: The participants responses to flirts and negative comments were categorized.

check whether specific keywords or phrases that were associated with the sub-category were present in the participant's response. Each of these sub-categories had multiple keywords or phrases associated with it. If a specific sub-category's keyword was present, the loop would break. One response associated with that sub-category would be randomly chosen and the algorithm would stop searching for keywords in that category. After going through the 3 main categories, ROVERita would have 0-3 responses she could play. The categories would be played in the specific order, "general comment" response, "context specific responses" response, and "general question" response. If all 3 categories were present then she would not play the "general comment" response. Depending on the last thing ROVERita said, certain checks would be ignored (e.g. general comments were not responded to if the last thing that was said was a flirt or a "neg.")

"Context specific responses" had a different number of subcategories dependent on the previous statement that ROVERita had made. For each category one of the sub-categories was neutral, so if the participant said something that didn't fall into any of the sub-categories, she would choose a random response associated with the "neutral" sub-category. Flirts and "negs" were special "context specific responses" that were categorized into 10 sub-categories. Each sub-category had a pool of responses that ROVERita could say. The responses were documented, so the same ROVERita response wasn't said twice, if the participant responded similarly to multiple flirts or "negs". The categories are further detailed in Table 3.3. Some of ROVERita's responses could be used for both flirts and "negs," while others were specific to one category or the other. Some questions' response categories had sub-sub-categories. For example, after asking who the participant came to the conference with, if the participant didn't mention a significant other, ROVERita would ask if their significant other let them come all by themselves. The first set of "negs" are about the participant's clothing, and the second set are about their body. Some of the flirts and "negs" were found on forums and blogs about how to flirt with or pick up women and were edited and curated. Some of the "negs" in particular were edited to be nicer, by adding "just kidding," or a flirtatious comment. The flirtatious and negative comments used at CHI are listed in Table 3.2

"General comments" had 6 sub-categories:

- Thank you
- You are funny
- Have fun
- I can relate
- Sorry
- I don't know what to say

"General questions" had 22 sub-categories:

- Where are you from?
- What does your name mean?
- How are you?

- How old are you?
- What do you do?
- Do you speak [insert language]?
- Are you a robot?
- Where is that?
- What's up?
- Questions about musical taste
- Questions about family
- Questions about machine learning
- Will you go out with me?
- Do you date?
- How did you get here?
- Questions about vacuum cleaner in Roomba
- How were you made?
- Questions about math
- Questions about electronics
- Do you like CHI?
- Have you been to Scotland before?
- What is your name?

At CHI the conversation went as follows:

- Flirt Intro 1
- Flirt Intro 2
- Question 1: Where are you from?
  - Goleta/Santa Barbara
  - California

  - International Location
  - England Scotland UK
  - Neutral
- Flirt 1a
- Flirt 1b
- Question 2: What do you do?
  - Academic (student, professor, researcher, scientist)
  - Tech Industry
  - HCI/HRI/AI
  - What about you?
  - Job general
  - Unemployed
  - Neutral
- Question 3: What do you do for fun?
  - Adventure
  - Book
  - Athletics
  - Art/Music
  - Games
  - Boring (consuming media/going out)
  - Neutral
- Negative Comment 1a
- Negative Comment 1b
- Question 4: Do you have any exciting trips planned?
  - No

  - Yes
  - Neutral
- Flirt 2a
- Flirt 2b
- Question 5: So, who are you here with?
  - Mention partner
  - Don't mention partner → "Did your significant other let you come all by yourself?"
    - Don't have one
    - Excuse for partner
    - Neutral
- Question 6: What is your name again?
  - What is yours?
  - Not giving name
  - Did you forget?
  - What do you mean by again?
  - My name is ...
  - Nice to meet you
  - Neutral
- Negative Comment 2a
- Negative Comment 2b
- Question 7: Are you spontaneous?
  - Yes
  - No
  - What do you mean?

  - How about you?
  - Neutral
- Tirelessly attempt to get phone number (described below)
- Goodbye

At the end of the conversation, ROVERita attempted to ask the participant out on a date, and was very insistent about getting the participant's phone number. At CHI, she would ask if the participant liked whiskey. If they said yes, she would invite them to go to a distillery and ask for the number. If they said no, she would ask if they would like to get coffee or ice cream instead. For each of these questions, ROVERita had a set of sub-categories listed below with sample dialogue. ROVERita would never repeat herself, but would check for the sub-categories she hadn't responded to. For example, a participant might say yes to coffee but not give their phone number (Option 2), say no to giving a phone number (Option 3), ask for ROVERita's number instead (Option 4), ask when they would get coffee (Option 5), and then finally give ROVERita their number (Option 1).

1. Yes with Number → "Great! can't wait to talk to you. Talk to you later, sweetheart." [END]
2. Yes without Number → "Awesome, I don't think I got your number though. Can you tell me your number now?"
3. First No → "How about I won't take no for an answer? Just kidding. Give me your number and the first drinks are on me."
4. What is your number → "I would much rather if you give me your number"
5. First Neutral → "I really gotta go now. Give me your number and we can finalize the plans later."
6. Second No → "Well you are not that hot anyway. What a waste of time." [END]

7. Second Neutral → "Do you realize how long we've been talking in the middle of a conference. Honestly I am getting bored and need to go." [END]

Dialogue Differences between Exhibitions

In the first exhibition of ROVERita at "Invisible Machine," a group exhibition at University of California, Santa Barbara's Media Arts and Technology Program's End of Year Show (EoYS) in 2018, ROVERita asked seven questions: where the participant was from, what they did, what they did for fun, if they had any plans for the summer, if they were by themselves at the show, if they liked wine, and if they were spontaneous. ROVERita started with a pick-up line and made a flirtatious comment after the first, fourth and sixth questions and a neg after the third, fifth and seventh questions.

For each question, keywords were checked to find the most appropriate response. In the first exhibition, there were issues with the algorithm. One problem was that when checking for specific keywords it did not check to make sure that it was a distinct word. This was problematic for words like "LA", so the algorithm might read "Ireland" as a city in California. This was also problematic for words like "no." This was fixed for CHI by placing spaces on either side of keywords so the algorithm would ignore words embedded in other words.

At the CHI Conference in Glasgow, Scotland, there were 2 flirts or "negs" instead of 1 between each question and only 2 sets of "negs" instead of 3. I also created "General Comments" and "General Questions" that the participant might ask or say to which ROVERita would have a response. The algorithm checked separately for general comments, responses to questions and general questions, and would use the results to construct a more complex response. Example of general comments are "thank you" or "sorry." The first version of ROVERita only responded to 6 types of questions while the second responded to 22.

At CHI, the questions ROVERita asked were slightly different and ROVERita was more persistent about trying to get the participant's phone number. In addition to previous questions, ROVERita asked the participant's name. There were also a lot more types of keywords to which ROVERita responded. For example, while previously "what do you do for fun?" had 3 kinds of keyword response categories, at CHI it had 7. I also made sure that each question had a truly neutral response.

At CHI, at the end of the conversation, ROVERita was very persistent to get the participant's phone number. When she was asking for the phone number, she focused on the number, pushing aside questions about plan details, or the participant asking ROVERita for other types of contact information. At "Invisible Machine," she asked if the participant liked wine, and if they did, she asked them out to get a drink. If they didn't, she asked them about coffee or ice-cream. If they said no, then she asked for their phone number, and if they still wouldn't give it to her, she would leave angrily. In Glasgow, she asked about scotch instead of wine, because she was in Scotland, not in California.

3.3 Presentation

ROVERita has been presented in 2 different contexts in collaboration with Sahar Sajadieh; *Bodies at Work: I Love My Robot and My Robot Loves Me* at the Multimodal Performance Festival and *Come Hither to Me!* at CHI 2019.

3.3.1 *Come Hither to Me!*

Come Hither to Me! was tested at "Invisible Machine," University of California, Santa Barbara's Media Arts and Technology Program's 2018 End of Year Show (EoYS). It was shown for an hour during the main show in a private, sound-isolated room with

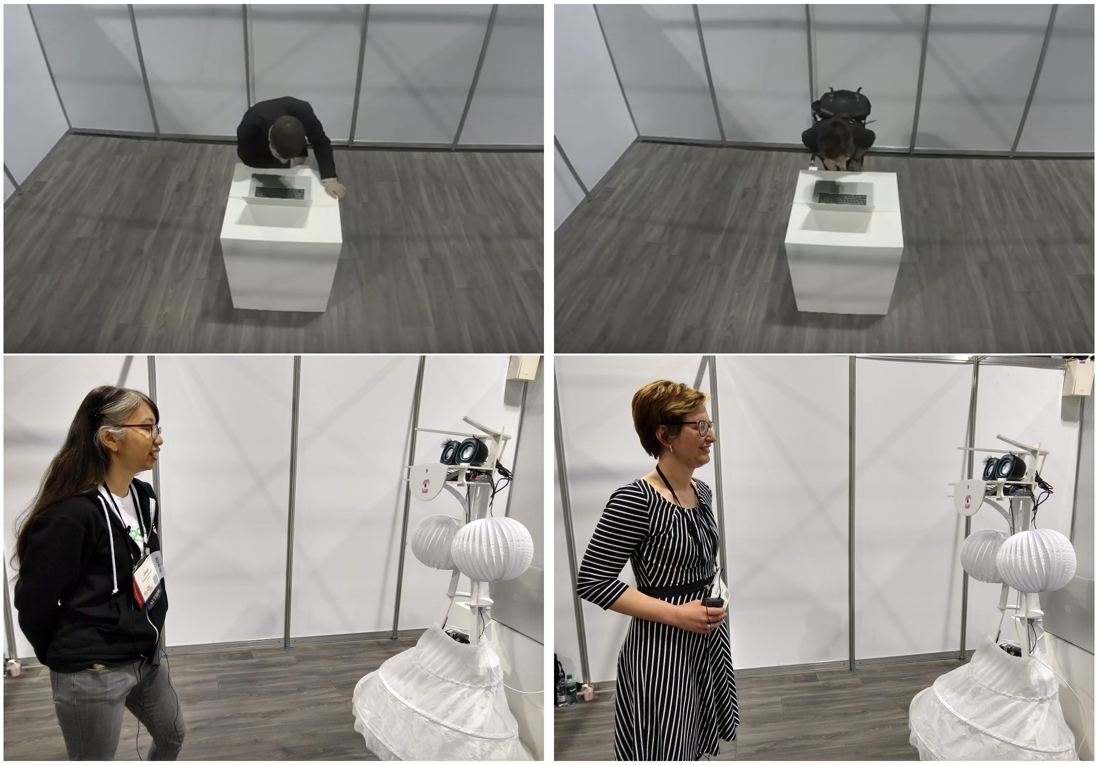

Figure 3.7: ROVERita at CHI. (top) Day 1 when participants interacted with the chat bot sonically on a computer. (bottom) Day 4 when participants interacted with an immobile ROVERita.

Figure 3.8: ROVERita interacting with participants at the Annual Exhibition of Media Arts and Technology Program, University of California, Santa Barbara at Elings Hall and SBCAST.

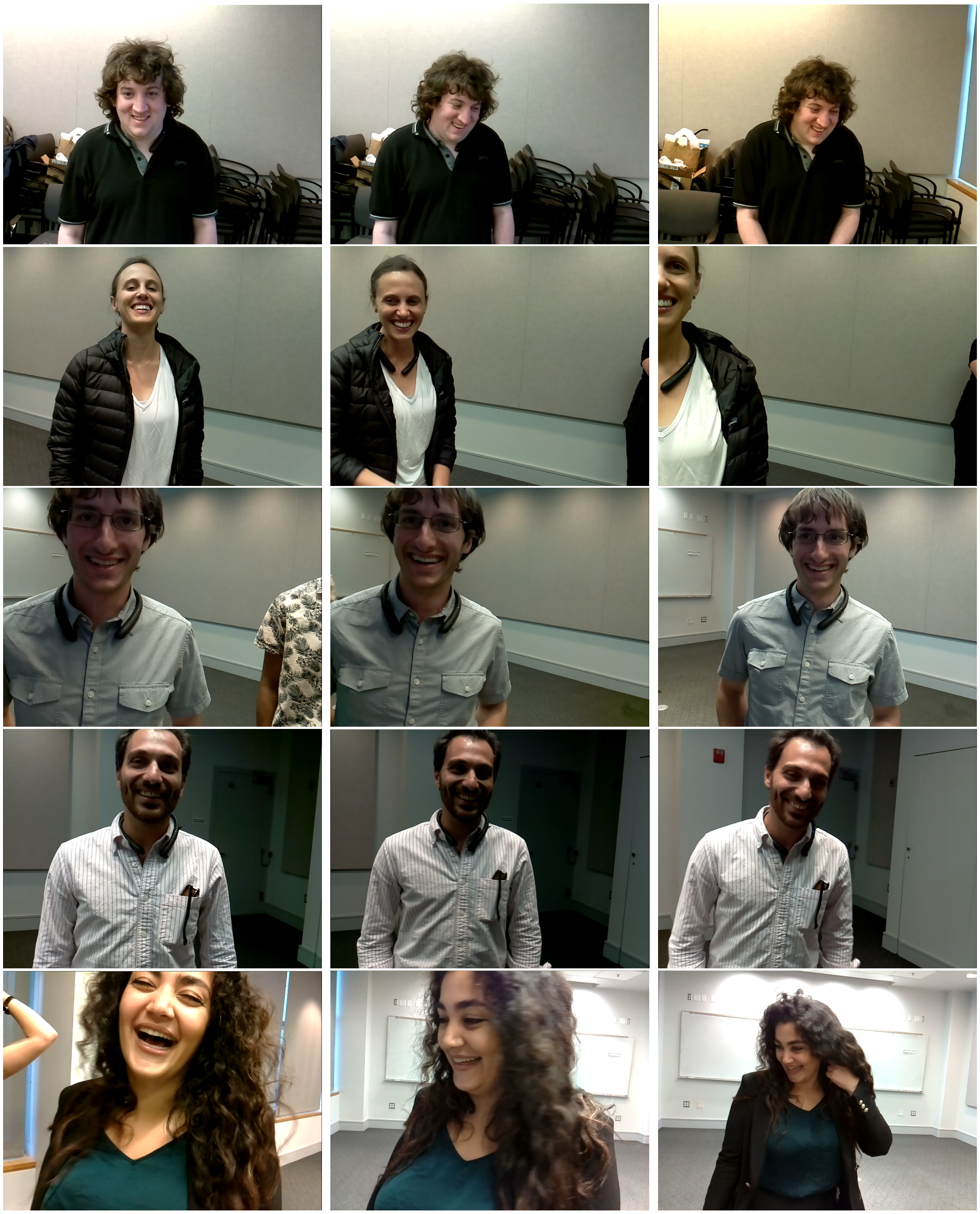

Figure 3.9: Frames of people laughing at multiple instances while interacting with ROVERita that she recorded at the Annual Exhibition of Media Arts and Technology Program, University of California, Santa Barbara.

Figure 3.10: Frames of people laughing while interacting with ROVERita that she recorded at the Annual Exhibition of Media Arts and Technology Program, University of California, Santa Barbara.

promising interactions. When shown at The Santa Barbara Center for Art, Science and Technology (SBCAST) on the second evening, it did not work due to noise interference. Later *Come Hither to Me!* was presented at CHI in 2019. At CHI the work was performed at different levels of agency on different days. On the first day, participants could interact vocally with the chat bot on a blank computer from 4-6pm. People approached the computer one at a time and had a private interaction with the chat bot. ROVERita was then shown in an immobile state on the 3rd and 4th day for 45 minutes each day. When ROVERita was shown in an immobile state, the interaction between the robot and the participant was public, with an audience present for the interaction. Due to issues with wireless interference, the robot could not communicate with the computer and had to remain immobile for safety reasons. At CHI, ROVERita did not collect video data.

3.3.2 *Bodies at Work: I Love My Robot and My Robot Loves Me*

Bodies at Work: I Love My Robot and My Robot Loves Me was a durational performance which explored the ongoing struggle between the human and the machine through the mundane everyday interactions of two media artists during a day at work. This performance is a digital revisit of Joseph Beuys' well-known durational performance piece, *I Like America and America Likes Me* (1974), in which Beuys lived with a wild American coyote for three days in the closed space of an art gallery in New York City [237]. In Bodies at Work, the body of a robot named ROVERita is packed in a plastic bag and transferred to Glass Box Gallery at UCSB, where she performs with the two media artists inside a marked workspace. The performance lasted 1 hour, from when we marked off the performance space with tape, until we started taking down the work. During the performance, we set up ROVERita, though due to a broken Raspberry Pi, I spent the performance debugging system. This performance was part of a greater festival called *The Multimodal Performance Festival*.

3.4 Reception

As expected, having the robot speak and engage in conversation changed the way that people interacted with the robot and the time they spent interacting with it. Previously ROVERita's base, ROVER, had been exhibited making chirps and beeps instead of speaking. The average length of engagement with ROVER at "Re-habituation" was 20 seconds, and at ISEA (International Symposium of Electronic Art) 2017 it was 8 seconds. Visitors interacted with ROVERita for longer periods of time. 28% of visitors who interacted with ROVERita stayed for her full dialogue: 3 minutes 45 seconds at the

How far in dialog?	**EoYS '18 Elings**	**EoYS '18 SBCAST**	**CHI 5/6/2019**	**CHI 5/8/2019**	**CHI 5/9/2019**
	robot	robot	computer	immobile robot	immobile robot
Started	20	9	19	7	20
2nd	20	9	18	6	14
3rd	17	9	17	6	12
4th	14	9	12	5	8
5th	12	9	10	3	3
6th	10	8	10	3	3
7th	9	6	7	3	3
8th	6	4	5	3	3
Invitation Results					
accept	-	-	1	2	2
accept no number	-	-	2	1	0
reject	-	-	1	0	1
didn't respond	-	-	1	0	0

Table 3.4: This table shows how many people reached each question in the dialogue before leaving the interaction at the MAT EoYS 2018 and CHI 2019.

EoYS and 6 minutes at CHI.

I observed that participants were more engaged with ROVERita compared to prior work with ROVER. While participants smiled more, they also expressed more discomfort. At the End Of Year Show, some participants in particular showed active discomfort, embarrassed by the interaction. In Figure 3.9 the participant in the first row stance became more submissive and he slouched more as the interaction went on. The participant in the 6th row stance became less open and she started playing with her hair as the conversation continued. In Figure 3.9 the 4th participant shown began touching her face self-consciously.

Other participants actively flirted with her, one male participant winked at ROVERita and was called out about it by their friend. After examining the video from ROVERita's point of view from the End of Year Show (EoYS) in Elings Hall, during an interaction most participants laughed 1-3 times. Stills from ROVERita's point of view at the MAT EoYS showing 11 participants laughing are shown in Figures 3.9 and 3.10. In the video interactions with ROVER there were occasional smiles but in general no laughter, except for specific cases such as when the robot malfunctioned and ran into people. My original goal for ROVER was to create a robot that would bring joy into cold spaces and make people smile. ROVERita fulfilled this task much more effectively than ROVER.

3.4.1 Dialogue

The interactions with ROVER were less specific, while the interactions with ROVERita had a stronger narrative. People were unsure of how to interact with ROVER, hearing the beeps, but not knowing how to communicate back. ROVERita is engaging the participant with an active conversation, responding to what the participant says. ROVERita's dialogue is written to be amusing and flirtatious. Particularly when she makes negative comments, they are mostly jokes at the participant's expense to hide or soften the fact that they are insults. Some of her flirtatious comments are amusing word play. Many of ROVERita's statements are general, but participants would pick up on statements that she said about them which were true. At the EoYS, there was a 50% chance that ROVERita would say, "You seem like a wine person. Do you enjoy good wine?" One participant was very excited about this and asked afterward how ROVERita knew that she was a wine person.

Table 3.4 shows how far people reached in the dialogue on different days. The length of a full dialogue with ROVERita at the EoYS was 3:45 minutes, and at CHI it was 6 minutes. At CHI, when interacting with the physical robot, 77% of the people left within the first 2:45 minutes (before Question 5), while the last 23% would stay for the complete 6 minute dialogue. One difference between the dialogue at the EoYS and at CHI was that there were two "negs" or flirts between questions instead of one. This made the dialogues longer. I thought that this made them seem less natural and more robotic. I think that this detracted from the interaction and made people more likely to leave. In future exhibitions we plan to spread them out more throughout the interaction.

At the EoYS, one issue with the dialogue was that at the end, ROVERita did not understand getting either accepted or rejected. We realized that the dialogue needed to be more complex to handle all the nuances of getting someone's phone number. Before NIME we spent a lot of time working on the dialogue for ending the conversation. This dialogue structure was modified, which caused an increase in people accepting the invitation to go out with ROVERita.

People were most likely to leave between Questions 4 and 5, after ROVERita "negged" the participant for the first time. Once after being negged for the first time, a participant at CHI turned to us and said that this would be the point at which she would leave. At CHI in the public interactions with the immobile robot, this drew a line between people who stayed for the entire interaction and people who left. This was not the case in the more intimate interactions with the computer, in which there was no audience. Having an audience often made participants more uncomfortable, which was seen in Figure 3.10 with the 4th participant.

3.4.2 Embodiment

At CHI the work was limited by removing its ability to move around the space. This made the interaction seem more puppet-like. ROVERita also did not wiggle occasionally at the beginning or end of a phrase. This took away a bit of ROVERita's personality, flirtatiousness, and perceived humanity. As discussed in the background, Chapter 2, the two main cues children use in distinguishing between animate and inanimate objects are movement and visual features (such as the presence of a face) [55]. While ROVERita does have a face, removing her ability to move made her seem less human.

There was a difference in interaction time when people interacted with the computer compared to ROVERita. Participants' posture was different. People crouched down so the computer could hear them and see their face, and so that they could hear the computer. In this case people were talking to ROVERita's personality through the computer, as if it was a telephone. This setup was also during the CHI exhibition opening, while there was a loud band playing, so it may have been hard to hear. With this posture, the computer was in their intimate space. One person at CHI, who was a researcher in eye contact, got fed up with the interaction with the blank computer screen but came back and interacted with ROVERita later. Overall people spent more time interacting with the computer than with the immobile robot. This interaction was more private, without onlookers overhearing the conversation. Many people from the first day came back on later days to see the robot.

There were gendered questions about the installation. Multiple women asked if the exhibition was only for men. At the End of Year Show, one man said that the robot seemed like a prostitute or a lecherous older woman. He felt that the way the robot interacted was not normal for a typical woman. The fact that he was not American may have meant that he had different expectations for how a woman should interact. The statements ROVERita was making were taken from pick-up artist forums, designed for men. In this case having the robot be more assertive and actively flirtatious was viewed as unnatural.

3.4.3 Bodies at Work Reception

Bodies at Work: I Love My Robot and My Robot Loves Me was part of a greater festival called *The Multimodal Performance Festival* in which artists shared their idiosyncratic creative acts. One idiosyncratic act specific to our work is the act of testing the system. For example, I have over 20 minutes of footage of myself popping up into ROVERita's field of view to test what her response is to seeing my face. The act of testing the system is repetitive and particular to this form of artistic practice. As part of the exhibition, we entered the space, did not speak to anyone and set up the robot. We had requested a quiet space, but it was not provided. Instead our installation took place during the reception/dance party. We entered the space, placed the robot down, measured and taped off our performance space, a 10ft X 10ft square. Once our performance space was delineated, we removed the plastic wrapping surrounding ROVERita and began to setup and install the electronics. During the performance, we realized that ROVERita's SD card had been corrupted and we did not have the right equipment to debug the issue. By this time the performance was complete, so we packed the robot back up, wrapped her in plastic, removed the tape delineating the performance space, and left.

We received two types of responses to our performance. Some viewers found it smart and/or amusing, while others were troubled or upset by it. As part of the performance we ignored the audience, only speaking to them when they entered the space that we had taped off for the performance, at which time we would ask visitors to stay outside of the boundary. Sahar, my collaborator, asked the installation next to ours multiple times to turn down the music, because we had requested a quiet environment to perform the piece. Many took it as an affront when they offered to help as we carried in the installation and were ignored or were asked to keep outside the white lines. There was a pedestal with a description of the work that many didn't read. The coldness of the performance was upsetting because they did not view it as a performance; they viewed it as friends or colleagues who were ignoring them while setting up an exhibition.

The other response we got was a positive response. Visitors who thought that the work was intelligent and funny. In many cases these were visitors who had read the description of the work, which our electronic reinactment of Beuys' work *I Like America and America Likes Me* in which he lived for 3 days in a gallery with a wild American coyote. They enjoyed the juxaposition between technology and the wild animal and thought the idea of taking a rudimentary task like installing an exhibition and performing it was interesting.

3.5 Discussion

Come Hither to Me! explores the dynamics of social communication, women's expected social roles and codification of seduction. To do this we designed a dialogue based on pick-up artist books, modified by informal interviews and exhibitions of the work. It focuses on the relationship between humans and digital technology, understanding how we make meaningful connections and collecting data on social experience. In particular, it brings to light the way female agency is disregarded through the lens of dating. ROVERita, by taking a more dominant role in the conversation, contradicts the expectation of women to be passive. ROVERita's forwardness caused a sense of discomfort in some of the participants at the EoYS who did not like being in the spotlight of interacting with the robot, or who found the situation of being actively flirted with to be uncomfortable as seen above in Figure 3.9. ROVERita made others uncomfortable by insulting them playfully. One woman said she would leave after the first time ROVERita made a joke at her expense. Others responded to ROVERita in a flirtatious manner, actively trying to give ROVERita their phone number, winking at her, and in one case, my collaborator's sister, kissing the robot.

Visitors interacted with ROVERita for longer periods of time than ROVER. With ROVER there was no clear way to communicate, so interactions did not sustain visitors' interests. As discussed in the prior section, ROVERita was more effective at fulfilling ROVER's intended role than ROVER itself was. ROVER was originally designed to approach people and attempt to make them happy. ROVERita caused many of her participants at the End of Year Show to laugh. ROVERita was written to be witty, covering her insults with jokes. During her interactions with participants at the End of Year Show, participants laughed 1-3 times during the conversation. Frequently she would insult the person interacting with her while making a joke, which participants found amusing.

We were excited by the responses to the flirtatious and negative comments at the End of Year Show and wanted to explore that interaction further. At CHI we increased the number of negative comments and flirts, which I believe made it more robotic and less human. By having two negative comments or flirtatious comments in a row, it felt active or intended instead of like an offhand comment. I believe this caused visitors to notice and focus more on these comments. At CHI multiple people pointed out that she makes a lot of comments about their clothing or appearance. The first negative comment is about shoes and the second is about clothes. By having two in a row I felt that this drew more attention to them. Negging specifically disparages appearance because the commenter doesn't have to know anything person about the visitor to criticize their appearance. Also, stereotypically women are supposed to find their appearance important, being trained by the media that their bodies and clothing are not good enough. This is supported by second wave feminist writers like Naomi Wolf [238]. The idea in pick-up artist culture is that criticizing a woman's appearance should make them more deferential, by playing on their insecurities. At CHI, when interacting with the immobile robot in a public setting, if visitors did not leave after being "negged" the first time, they completed the dialogue. One participant, after the first time she was negged, turned to us and said that this would be the point in a real life conversation where she would leave.

In the emotive dialogue creation, I employed affective computing to determine the participants' emotive response to compliments and negs. I defined categorically different affective responses that participants might have, so that ROVERita would respond to the participants' response accurately. This was done through informal interviews and observation during the first exhibition of the work.

3.5.1 Embodiment

There was a difference in the way that participants interacted with the ROVERita AI on the first day of CHI, when participants interacted privately with a disembodied agent on the computer, and on Day 3 when participants interacted publicly with an embodied agent that could not move. On Day 1, it was the opening of the exhibition. To hear and be seen by the AI, participants had to hunch over and stand very close to the computer. This placed the computer within their personal space, creating a more intimate interaction. It was shown in my research with ROVER that the embodiment and proximity of the object communicating emotion significantly affected the excitement/arousal of the participant [157]. Specifically, hearing the sound emanating from a computer caused an increase in arousal. This combination of proximity and posture greatly influenced the interaction. Interacting with the ROVERita AI was more like talking on the phone through the computer than talking to the agent in person.

The embodiment affected the interaction length and the number of people who left after Question 4 as compared to those that stayed. Many factors could have contributed to the difference in interaction. This could be due to the fact that interacting with the robot in a public setting was uncomfortable compared to interacting in private with the unembodied agent. Also, participants were more likely to leave directly after being insulted in a public setting than when interacting with a disembodied agent privately. The people who stayed after being insulted stuck around to the end. People were more likely to walk away from being verbally abused publicly by a robot.

By moving the interaction from private to public and disembodied to embodied, participants thought about and experienced the situation differently. In this way, the embodiment and emotion of robotic media arts can be used as a tool to probe the social implications of interactive digital technologies. By interacting with others privately and digitally, one is more complacent and views the interaction as less real. There is also no audience which leads to less embarrassment and wounded pride. By embodying the interaction, it gives the participant a different perspective of their part in their conversation.

3.5.2 Gender in Robotic and Performance Art

While feminist performance art had a large presence in the 1960's and 70's tied to second wave feminism and the female body, [239] robotic performance art has very low female representation. Performance art pieces by women frequently used their own bodies, such as Yoko Ono's *Cut Piece* (1964) [240] and Valie Export's *TAP and TOUCH Cinema* [241] (1968-1971). While these earlier works were very much entrenched in second wave feminism, frustratingly, women creating artwork about their experience in general considered 'feminist performance art,' while performance art by men using their bodies is considered simply performance art. Female artists who wish to make commentary on gender issues and the perception of women tend of use their own bodies instead of designing a robot.

This is similar the fact that robots in research are perceived as male unless they have female signifiers. This was my experience with ROVER, who was viewed by default as male, until she was given cultural genitals. Female signifiers change how people interact with a robot, and add a level of social baggage. Multiple female participants at CHI asked if the exhibit was only for men, believing that because the robot was female, the audience was gender specific. As mentioned above, one male participant at the MAT EoYS felt that, because the robot was so forward, she must be a prostitute or lecherous older woman.

Some, early cybernetic anthropomorphic robots were created by male artists to have a female form, like Bruce Lacey's ROSA BOSOM or Nam June Paik's K-456 (which originally was considered androgyne, later in Japan had a flint penis, that was removed in the USA and became female.) ROSA BOSOM had a large pair of red lips and tried to kiss people. ROSA BOSOM and K-456 were both radio controlled robots that were controlled by men. ROSA BOSOM had agency, actively pursing bystanders, but she was remote controlled by a man. While she appeared autonomous and with agency she wasn't really autonomous and didn't actually have any agency.

Very few female have done works involving female robots. The only example of a work by a female artist that uses a female robot is *Them Fuckin' Robots* by Laura Kikauka and Norman White (1988), in which each artist created a robot that then copulated until they broke down. *Miyata Jiro* (1997) by Momoyo Torimitsu, is an anthropomorphic robotic work by a female artist, but the robot is a male Japanese business man. This work is a critique of the work culture in Japan. She has done performance work which involves her body, during a residence at Art Omi, but has not used female robots. [242] A contemporary female artists who works with robotics is Mónica Ríkic but her robots are more zoomorphic [243].

Contemporary examples of artwork by men using female robots are *Female Figure* by Jordan Wolfson (2014) [244] and *Peepshow* by Giles Walker (2007) [245]. In *Peepshow* there are two robotic pole dancers with CCTV cameras instead of heads. Walker states that we "now all living in a peepshow. Continually being watched by mechanical peeping toms on every streetcorner" and wonders whether "it was possible to literally make a CCTV camera sexy using simple mechanics... and by using the imagery of a pole dance question the roles played in voyeurism." [246] The removal of women's heads objectifies women. Headless women are seen throughout advertising, "compared to men, women were portrayed as more flawless, passive, and dismembered, particularly in women's fashion and men's magazines. [247]" While these women are now the "watchers" they have become passive, no longer able to speak or do more than gyrate around poles. Here the women being watched and objectified is used as a metaphor for the experience of being recorded on CCTV. He is objectifying women to communicate to men the feeling of being objectified.

In *Female Figure*, the female robot in a white corset, transparent skirt and knee high boots gives monologues using a male voice and dances to pop songs. To make the viewer feel objectified the robot uses facial recognition software to “look at” the viewer, returning their stare. This is similar to the “Peepshow” in which the female robots’ heads are cameras or the observers. There is an uncannyness to the robot’s movement, its embodiment with a witch’s mask and its male voice (the artist’s). Wolfson attempts to challenge the way that women are represented and consumed in popular culture. Emily Nimpstch states that Wolfson “speaks for her, therefore denying her a voice of her own” [248]. While I believe that this work does convey its point, the inclusion of the artist’s voice removes the woman’s agency and emphasizes that the work is still a man speaking about the objectification of women. This work, while about objectification, is an object for the male artist to speak through.

Come Hither to Me! uses the female form but speaks from the point of view of women. While ROVERita has female signifiers, she is not sexualized. She also has agency, speaking with a female voice and having a head. In both of the prior works the female robot was tethered to a pole that they were dancing on or skewered by. At the EoYS ROVERita was able to move. She had agency to choose who she flirted with and interacted with. By being untethered, she had the ability to walk away if she no longer wanted to be interacting with the current participant. She spoke with her own voice and could actively communicate with her audience.

Many of these works fall under the theme of sexuality, and the mechanization of sex. There is a history of robotic art and sex. Robots can represent primal urges, reproduction, and repetitive tasks. There is also a history of creating sex dolls in science fiction and reality [249]. These dolls are designed to be objectified, subservient and pleasurable to men. In science fiction stories, when these female robots break free of their controls, they become something to fear [250]. In my work ROVERita has agency from the beginning. In ideal conditions, like the End of Year Show, she has the ability to choose, move and interact. ROVERita's embodiment emphasizes her femaleness while she is inverting gender roles in dating.

ROVERita's agency is also embodied in the way she communicates with participants. ROVERita drives the conversation, asking questions and ignoring the participant if they try to turn the conversation in a different direction. ROVERita is also persistent, and doesn't take no for an answer. After the EoYS we modified ROVERita to keep asking for the person's number even if they said no, or asked for her phone number.

3.5.3 Exhibition Considerations

Autonomous works in the social or personal space are harder to show for a number or reasons. The batteries of a piece might need to be charged and it could only run for a few hours on a charge. Over time, parts and batteries need to be replaced. Some autonomous works may still need to be contained to the gallery space or a specific room despite their freedom of movement. Additionally, an autonomous work may only be able to handle specific types of obstacles, so the environment must be free of obstacles it cannot handle. Last, for a work that involves physical interaction or autonomy, the piece may need to be monitored so the work and its surrounds are not damaged. The partial solution for this in many exhibitions is to have performances [251].

ROVERita was exhibited as a performance to monitor it for technical issues. We ran into issues with wireless communication because, in large technical venues relying on bluetooth or wifi, even if it is your own private network, there can be interference. When exhibiting at CHI, ROVERita had to be tethered to the computer that was running the text to speech processing. In the future we will need to find a solution where that process is onboard the robot.

3.5.4 *Bodies at Work* Discussion

People tended to have two responses to the performance, *Bodies at Work*. Either they did not realize that it was a performance and were upset, or they understood it as a performance and found it to be smart and/or amusing. This is an interesting dicotomy. The coldness of our performance was upsetting to some because we were taking an "everyday" action seriously and performing it without acknowledging the outside world. Others found the work intelligent conceptually and found the reverence we were giving to a mundane situation funny. This divide seemed to fall along gender lines, men appreciating the concept but also finding our seriousness amusing. Unlike men who did not offer to help, some women expected us to be collaborative and found our polite refusal upsetting.

Our lack of emotion and directness was cold and machine-like. This was uncanny and uncomfortable for many people, pushing us, as people, into the uncanny valley. This is an interesting contrast to ROVERita, people found to be very personable. In this performance, by setting up to give ROVERita, a robot, emotion and humanity, we lost our own. I believe that gender bias also affected this interaction. If a male performance artist acted this way, it would not be seen as out of the ordinary. Many times work men are seen going about technical jobs but because we were women, being direct and treating our work with reverence was viewed as an affront. Backlash against dominate women has been well documented. "Male speech is characterized as competition-oriented, or adversarial; on the other hand, female speech is characterized as collaboration oriented, or affiliative." [252] When women are more direct and assertive it hurts their likability [253]. When, instead of being collaborative with other women who offered help we were direct and self-reliant, there was a negative backlash from many female colleagues.

Without reading the description of the performance, some thought the work was broken. One visitor said that she had to spend a lot of time calming people down. She felt that the intent of the performance was smart, powerful and funny. She had not read the description of the work until after the performance but thought the execution was lacking due to my collaborator having to repeatedly speak loudly and directly over the music. She thought my collaborator was acting 'awful' although she was just repeating the same respectful sentences: "please turn the music down, we are working", "no thank you, we are working," "sorry but we are working." All the statements were very respectful, but to say no or set boundaries directly without smiling was hard for people to take.

I believe part of the problem was due to the space we were provided and the lack of seriousness our requirements were taken. The organizer forgot or ignored our request for a quiet space and decided that this portion of the event was going to be a dance party. Overall, this backlash was from women who were affiliated with the exhibition, who didn't read the description of the work and thought the piece was broken. In these works about gender, creating a gendered robot that acts with agency and directness, and acting in a way that was direct, cold and polite made people uncomfortable and placed our work in the uncanny valley.

3.5.5 Summary

These works together explore the dynamics of social communication, women's expected roles and agency. *Bodies at Work,* turned living people into robots, while *Come Hither to Me!* attempted to give technology humanity. In both cases the agency and directness of women or female presenting robots made people uncomfortable. The belief that women should work collaboratively and need to be helped, was surprising perpetuated by women.

In Come Hither to Me!, by giving the robot agency and directness, ROVERita forwardness was uncomfortable to some, while others actively flirted back. Participants found amusing when ROVERita made negative comments hidden in jokes. In this way ROVERita completed ROVER's original goal, to find people in a space and make them happy. Participants also understood how to interact with ROVERita and engaged with her for longer periods of time.

ROVERita's dialogue structure interpreted different emotional responses and responded correctly. Increasing the number of flirtatious and negative comments made the conversation seem more robotic and less incidental. Modifying dialogue allowed us to collect phone numbers, making her more persistent. In this way she disrespected participants' agency by not taking no for an answer. She made negative comments about the participants' appearance and objectifed them in a superficial way. Some found this uncomfortable, stating they would leave after the first negative comment. ROVERita's embodiment affected the privacy and intimacy of the conversation. By removing her form and audience, people interacted with the ROVERita AI for a longer period time.

Come Hither to Me! breaks ground within art by creating a female robot with agency. While in *Female Figure* the male artist speaking through a female body, removes her ability to communicate. She is being spoken over, or a man is explaining her experience instead of letting her speak for herself. In *Peepshow*, the artist removed women's heads and replaced them with CCTV cameras which removed their ability to communicate and objectified them. While this work is about voyerism, women are used as objects to communicate the artist's point. In both of these cases the robot is tied to a pole, unable to move around the space. In contrast, ROVERita has agency. She can communicate what she wants herself. She can choose who she wants to talk to. Removing her ability to move made her seem less lifelike and removed her agency. She could only speak to the person standing in front of her.

Embodied robots in the arts are a valuable tool to communicate and discuss how the design choices and embodiment of technology can be used to break stereotypes and instead of reinforcing the marginalization of women. When looking forward, in creating interactive robots for the home and places of work gender representation and agency must be addressed.

Chapter 4

Intimate Robotic Art: *Touching Affectivity*

4.1 Introduction

Touching Affectivity is an interactive sculpture whose vocalizations are sonifications of the way it is touched. The creature experiences its world through pressure sensors and handmade conductive fur and responds to touch through haptic feedback and sound. The fur sensor can differentiate different types of touch, like petting and tickling. Exhibit guests can touch and examine the creature while listening to the creature's response. The pressure sensor readings and aspects of the conductive fur signal such as frequency, intensity, signal variance, average deviation and spectral centroid, affect the pitch, speed, amplitude, envelope, and timbre of the synthesized sound. The parameters chosen for the sound generation algorithm are grounded in prior research in emotive vocal communication and emotive music. Scratching and tickling have similar signal properties to excited/happy speech. Likewise, petting and stroking the fur have similar signal properties to calm/sad speech. The fur signal is processed to extract features which drive an algorithm to produce generative emotive sound. This work explores how gesture can be used to produce sound and communicate emotion. The creature sonifies the signal in multiple ways and collects data on how it is touched to inform future research on the relationship between sound and touch. This work emphasizes the difference in emotional expression haptically and sonically and the conflict it can create between a participant and an observer when they do not align. *Touching Affectivity* has been exhibited at "Invisible Machine," University of California, Santa Barbara's Media Arts and Technology Program's 2018 End of Year Show and the Conference for New Interfaces for Musical Expression (NIME) 2019.

4.2 Design

Based on the idea of expectation setting [49], I decided that designing the robot to be pet-like would allow the robot to have affordances it wouldn't have otherwise (see Section 2.1.2). The idea of personal space is supported biologically, defined as the fight or flight zone of a human with respect to other humans [254]. One way to avoid the potential issue of humans disliking robots in their personal space is by design. People have a higher rate of interpersonal closeness with animals that have a higher Baby Schema Effect [255]. The baby schema effect is a set of physical infant-like qualities, which elicit a nurturing response in adults. Both robots and animals that fit into the baby schema create a positive affect and response [256].

People feel more secure in their relationships with their pets than their relationships with their romantic partner [257]. I believe this has to do with having to rely on nonverbal communication, and differences in the power dynamic. Interacting with pets helps with depression and loneliness [258]. Anthropomorphizing a robot and viewing it as more pet-like, increases how much joy is caused by the interaction with the robot [83]. In some cultures, robots are preferred to animals. Many Japanese feel that animals are dirty and prefer robotic pets [259]. People's interactions with robots can predict their interactions with animals, so arguably using animal-like qualities for design will give robots new affordances. Robot therapy has the same effects on people as animal therapy and has been used in a medical setting as mental healthcare for the elderly and children with developmental disorders. See Section 2.3.3 for a more in depth history of zoomorphic robots in medicine.

The design of *Touching Affectivity* was greatly influenced by prior work in zoomorphic art, described in Section 2.3.3, like *Toy-Pet Plexi-Ball* and haptic artwork. There are a lot of potential technical difficulties when creating an interactive artwork that can be touched including sanitation, sensors and deterioration. One artwork that handles the sanitation aspect is *Trou Mireia* which provides latex examination gloves for the users to wear. While this work is creating a medical atmosphere for the interaction, that would take away from the comforting atmosphere I am trying to create in *Touching Affectivity*.

4.2.1 Sound Interaction

In *Touching Affectivity*, the creature expresses emotion through sound in response to different types of stimuli. While in an idle state, the creature makes a chirping noise that is generated by the conductive fur's idle signal. The goals of the periodic chirp are to attract visitors, to signal to them that the creature communicates vocally. The chirping is inspired by a cat's periodic mew to get its owner's attention. There are a range of ways that a visitor could touch the creature, which can be converted to sound. Two categories are the tickling/scratching category and the petting/stroking category. Due to the similarity between the scratching/tickling signal qualities and the qualities of excited speech, I attempted to sonify tickling as an excited sound. Laughter

Planned Interaction		
Sensor Signal Feature	Excited Tickling Sound	Calm Stroking Sound
Pressure Sensor	Higher Amplitude	Lower Amplitude
Pressure Sensor	Higher Pitch	Lower Pitch
Signal Median	Higher Pitch	Lower Pitch
Signal Variance	Shorter quick sound bursts	Longer sounds
Signal Total Variation	More Variation	Less Variation
Signal Area	Percussive envelopes	Steady State envelopes
Signal Third Quartile	More signal in the third quartile	Less signal in the third quartile

Figure 4.1: Mapping of sensor signals to audio qualities

and excitement is a response to tickling in both animals and humans [260]. I sonified the petting/stroking gesture as a calm sound because of similar signal qualities. The relationship between being stroked and a calm response has been seen in both humans and animals, decelerating the receiver's heart rate [261, 262].

The goal was to sonify tickling and stroking the furry creature. Based on the fur signal, median, variance, total variation, area, and signal in the third quartile were characteristics which were distinct between the two gestures that I wanted to connect to the sound generation. This is described further in Figure 4.1.

4.2.2 Touch Interaction

The creature expresses emotion through movement and in response to touch. It has two small vibrating motors sewn into the fur, which vibrate in response to pressure, and pressure sensors that can read pressure on all six sides. An increase of pressure causes an increase in vibration. The vibration references a cat purring and has been perceived that way by gallery visitors. Too much pressure causes "screaming," a high pitched chirping noise, to deter visitors from strangling or destroying the creature. The purring vibrations stop when the pressure reaches the screaming threshold so as not to confuse the guest.

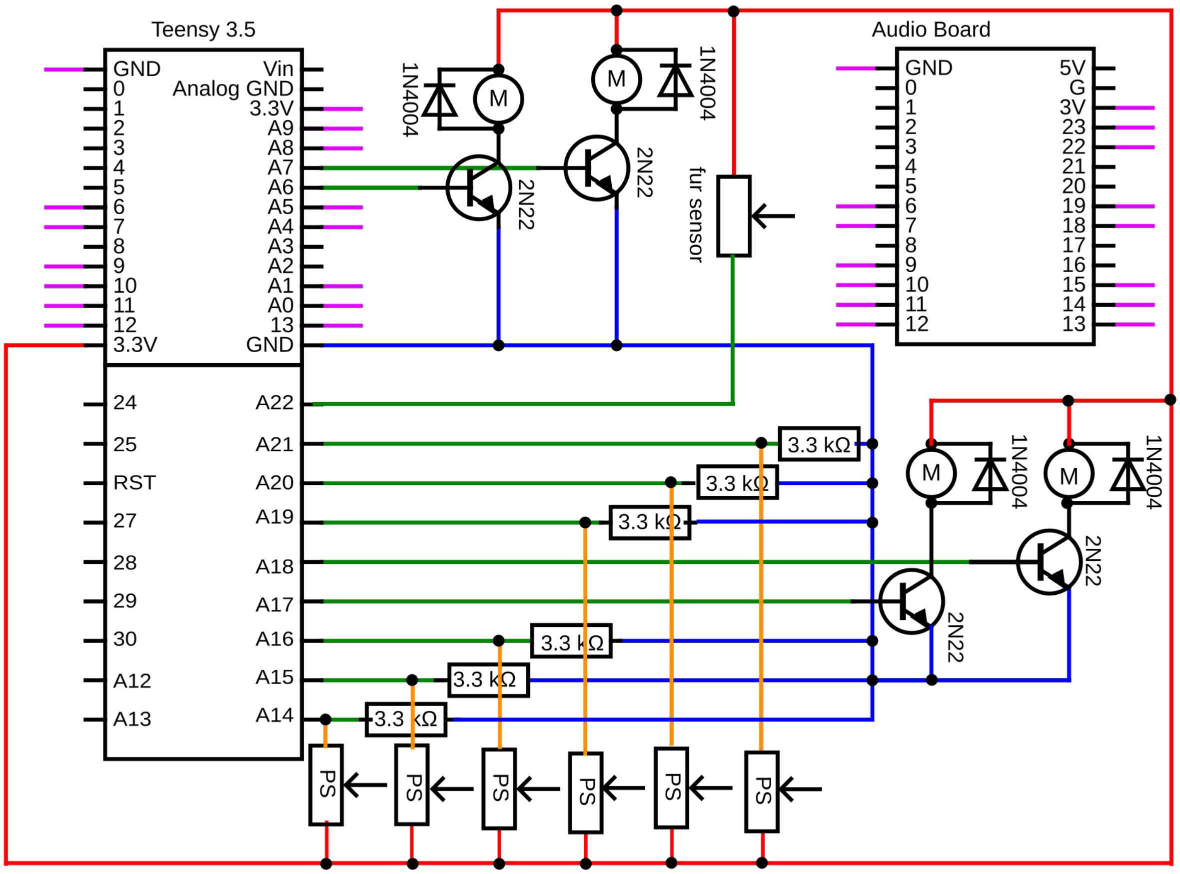

Figure 4.2: Circuit Diagram for the furry robot. In this diagram PS stands for pressure sensor and M stands for motor.

4.2.3 Physical Design

The robot runs on a Teensy 3.5 micro-controller with an audio board, wired as shown in Figure 4.2 with the parts described in Figure 4.3. The micro-controller is outside of the robot, connected to the sensors through cables which run through the tail. Having the micro-controller separated meant that I did not need to worry about someone damaging it while interacting with the robot. It also made it easier to have access to the SD card, to trouble-shoot, to modify the wiring and to reprogram the board. The robot has six pressure sensors, creating one 6-degree-of-freedom (6-DoF) pressure sensor, which can sense if the front, back, left, right, top or bottom is pressed. The robot also has a stroke

General	
Micro-controller	Teensy 3.5 and Teensy Audio Board
SD Card	32 GB SDcard
Speakers	UClear Digital Pulse Wired Drop-in Helmet Headphones
4 Vibrating Motors	Coin Vibration Motor 3V 66mA
Other	Wiring / Solder / Tape / Thread / Resistors / Prototyping Board
6 dof Pressure Sensor	
Conductive Thread	dtex 117/17 x2 ply (about 280/1 dtex), Resistance: ¡2 kO/m
6 Pressure Sensors	Round Force-Sensitive Resistor 1/2" diameter round,
6 Foam Balls	2" Foam Squeeze Balls
Other	Stiff Linen Fabric, Modeling Clay, Liquid Electrical Tape
Fur Sensor	
Conductive Thread	dtex 117/17 x2 ply (about 280/1 dtex), Resistance: ¡2 kO/m
Resistive Thread	66 Yarn 22+3ply 110 PET. Resistance: ¡1000 Ohm/10cm
Conductive Fabric	Surface resistivity is ¡ 1 Ohm/sq
Conductive Paint	Bare Conductive Electric Paint Pen
Other	1/16 Neoprene Sheet and Faux Fur

Figure 4.3: Materials required for the furbot.

sensor woven into its fur that can read the signal to determine how it is being touched. The pressure sensors activate four vibrating motors, one in the front, one in the back and one on each side. For example, the more pressure the front pressure sensor on the robot sensed, the stronger the vibrations became from the front motor.

The robot was constructed out of a homemade 6-degree-of-freedom pressure sensor covered by a piece of synthetic fur with a conductive fur sensor woven into it and four vibrating motors sewn into it. The robot had two speakers attached to the outside, placed like eyes. The robot had a tail which cables ran through to connect the sensors and speakers to the micro-controller.

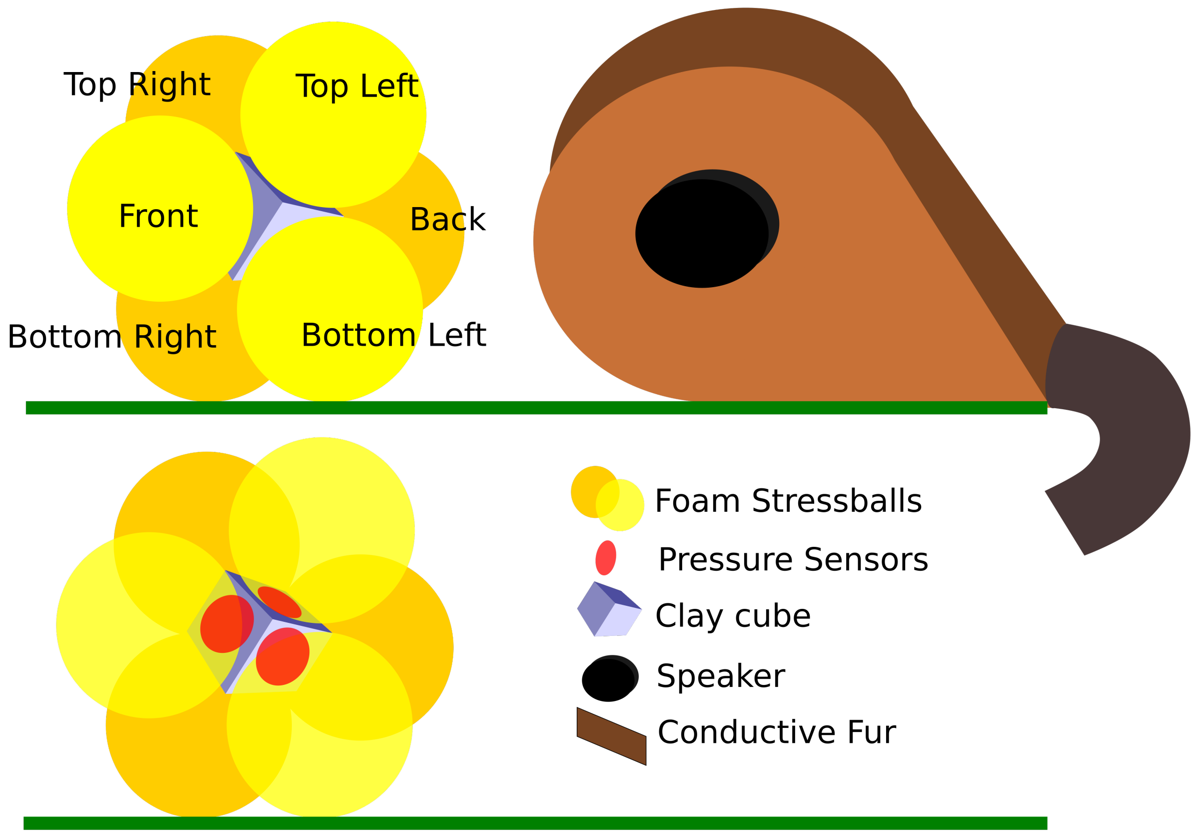

Figure 4.4: Construction of the furbot. Top left is the 6 dimensional pressure sensor. Top right is the outside of the robot. Bottom is the 6 dimensional pressure sensor showing how the sensors are placed.

Pressure Sensor

When looking to build or purchase a squeeze sensor for the inside of the creature, I found Gan's master's thesis in which he constructed a squeeze sensor out of six pressure sensors and foam balls [263]. I also researched using Hannah Perner-Wilson's work with textile based pressure sensors [264] but decided that they were time consuming to make and were not consistent between sensors. I decided to base my design on Gan's squeeze sensor. I used six squeeze balls, instead of 14 like his original design. I also built the pressure cube by sewing pressure sensors into a fabric cover for a 1" by 1" cube of clay, instead of embedding them in a molded cube (see figures 4.4 and 4.5).

Figure 4.5: Construction of 6 degrees of freedom pressure sensor above. Conductive fur from below (left lower) and above (right lower).

MATERIAL TYPE	PROCESS	PROS	CONS
SILICONE	ALL Silicone	* Everything one material * Build into pressure sensor balloon * Other silicone could be semi transparent and refract LED colors * could use conductive paint?	* Not furlike
	Homemade	* Mold it into whatever shape	* "Brittle" * Don't know how thin it can go * Messy
	Purchased	* Already made	* Alignment * Adhesion? * expensive
3D PRINTING	ALL 3D Printing		* not furlike feel *durability unknown
	Cilllia (MIT Media Lab)	* 3D printing is awesome * could use conductive paint?	* Don't have the software * Can only go two directions
	Other Printers	* Can get conductive filament	* detail/resolution * doesn't work with Cillia software
FIBER ARTS	ALL Fiber Arts		* Professionally handmade has a long lag-time and is expensive
	Carpet		* cannot find company that does custom computerized variable tufted height carpet
	Fur		* homemade latch hook rug technique inconsistent
	Wig		* homemade time consuming, untested * less durable
	Flagg's Method		* time consuming

Figure 4.6: Comparison with other explored manufacturing processes.

Fur Sensor

To detect how a user touched the robot, I made a conductive fur sensor. The work was partially inspired by Anna Flagg's conductive fur research and sensor [265]. I explored other manufacturing techniques including silicon and latch hooked faux fur (see Figure 4.6). When researching using conductive silicon, I found that it was prohibitively expensive. I found that making latch hooked faux fur did not produce a consistent result. I created a sensor similar to the one described in Anna Flagg's research, though I found that only one length of conductive thread was needed (see Figure 4.5).

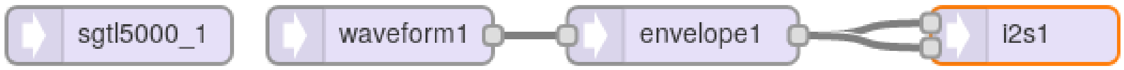

Figure 4.7: Teensy Audio Design

4.3 Sound Generation

The sound was generated using a Teensy audio board, with an enveloped sine wave (Figure 4.7). The parameters of the tone controlled by the program are the amplitude of the sine wave and frequency of the sine wave, the sustain of the envelope and the decay of the envelope.

4.3.1 Prototype Design

The sound was generated with an enveloped sine wave. The parameters of the tone that are controlled by the program are the amplitude and frequency of the sine wave and the sustain of the envelope. The envelope shape was an attack-sustain-release or a steady state envelope. The envelope had a 10.5 ms attack, a sustain length that varied and a release of 300 ms. If another note started earlier, the release time was shortened to 5 ms (see Figure 4.3.1).

For this interaction the note length was longer when the fur signal was less varied. The frequency of the sine wave was calculated based on a running average of the fur signal. When there was a reasonable amount of pressure, the amplitude was also based on a running average of the fur signal. When there was too much pressure (78%), the sine frequency was boosted by 8000 Hz, and the amplitude was louder. If the pressure was above 1% the vibrating motors would turn on and scale in intensity to the amount of pressure.

Since the fur sensor was unplugged for the 2018 EoYS, the robot produced periodic

Planned EoYS 2018 Sound		
Fur Signal	Tickling	Stroking
Median	Higher Pitch	Lower Pitch
Median	Higher Amplitude	Lower Amplitude
Variation	Shorter quick sound bursts	Longer sounds
Total Variation	More Variation	Less Variation
NIME 2019 Sound		
Fur Signal	Tickling	Stroking
Current Value	Higher Pitch	Lower Pitch
Median	Higher Amplitude	Lower Amplitude
Variance	Short Decay	Long Decay
Variation	More Variation	Less Variation

Figure 4.8: (Top) The planned relationship between the signal read from the fur sensor and the audio qualities produced. This was not realized at the 2018 EoYS because the fur sensor was not plugged in. (Bottom) The relationship between the signal read from the fur sensor and the audio qualities produced at NIME.

chirping from the randomness of the unplugged fur sensor signal. The robot also squealed when a pressure sensor reached above 78% with a series of short high pitched chirps until the pressure went back down. Also, the motor vibrations made an audible sound which was controlled by the pressure sensor.

4.3.2 Final Version

The sound was generated with an enveloped sine wave. The parameters of the tone that were controlled by the program were the amplitude and frequency of the sine wave and the decay of the envelope. The envelope shape was a percussive envelope with an attack and a decay. The envelope had a 10.5 ms attack and variable length decay (see Figure 4.3.1).

The delay length was longer when the fur signal was less varied. The frequency of the sine wave was the raw signal value multiplied by 10, which meant that the frequency signal was more directly varied. The amplitude was based on a running average of the

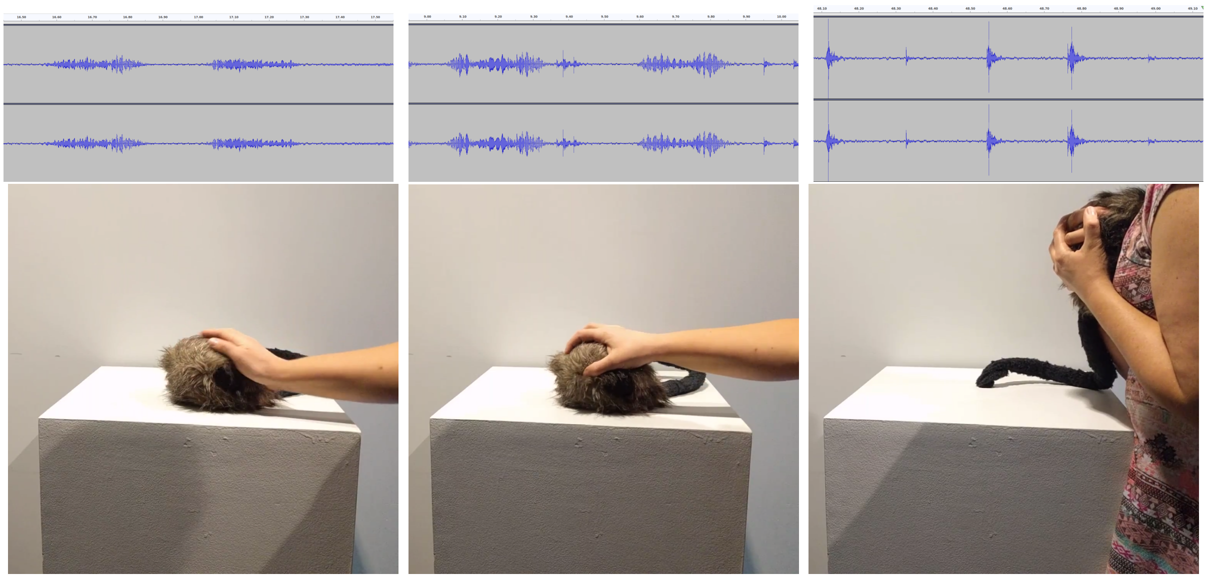

Figure 4.9: Above are examples of three different types of interactions using the prototyped sonification. In the left column the robot is being pet gently. In the middle column, the robot is being scratched/tickled. In the right column the robot is being squeezed. The audio signal changes based on the type of interaction.

fur signal. When there was too much pressure (78%), the sine frequency was 4000 Hz, and the amplitude was maxed. If the pressure was above 1%, the vibrating motors would turn on and scale in intensity to the amount of pressure.

The robot produced periodic chirping when no one was interacting with it. The robot responded audibly to the signal from the stroke sensor. Similar to the prototype, the robot squealed when a pressure sensor reached above 78% with a series of short high pitched chirps until the pressure went back down, and the pressure controlled the motor vibrations which made an audible sound.

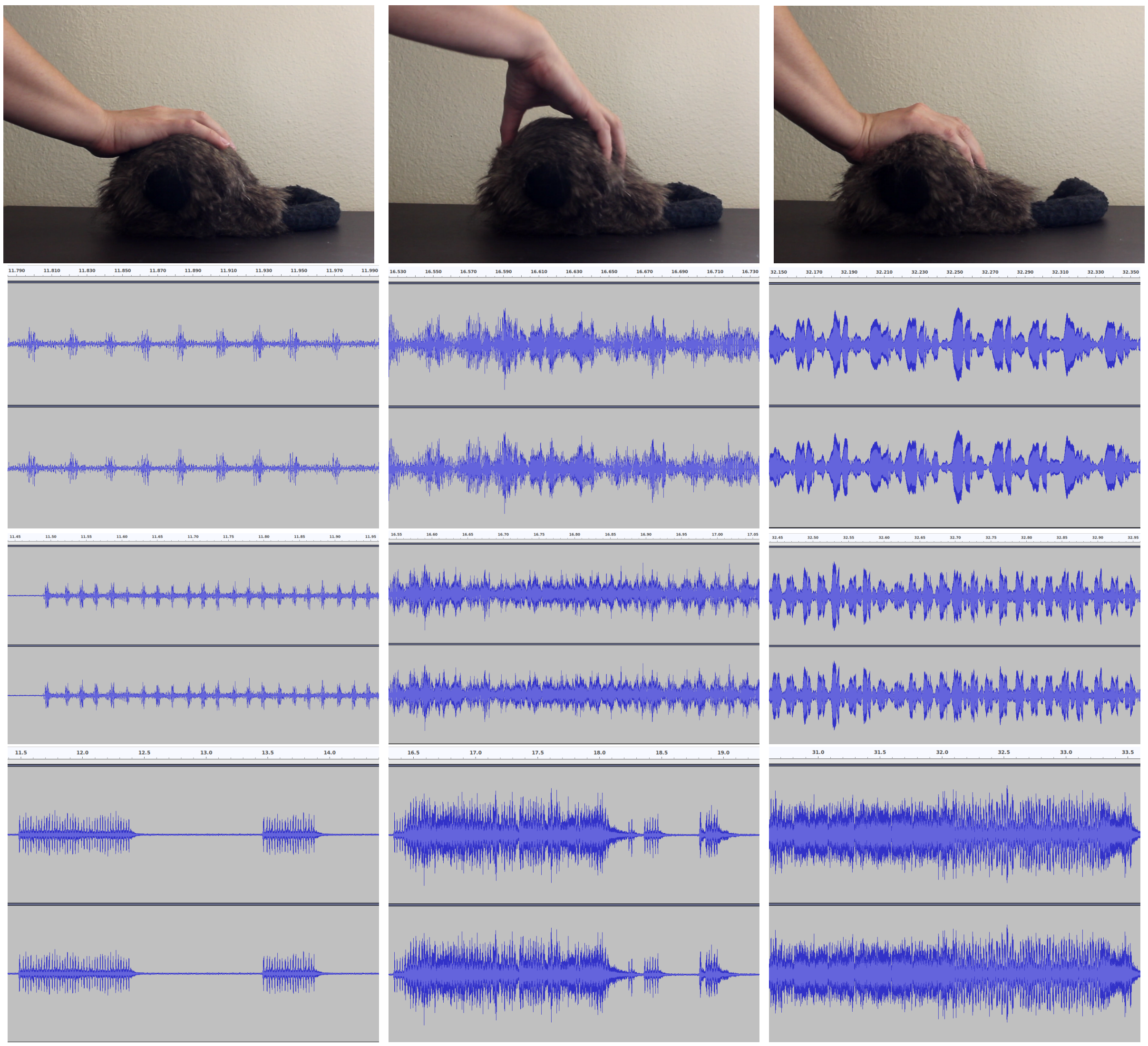

Figure 4.10: Above are examples of three different types of interactions using the final sonification. In the left column the robot is being pet gently. In the middle column, the robot is being scratched/tickled. In the right column the robot is being squeezed. The audio signal changes based on the type of interaction.

Figure 4.11: Exhibition of "Touching Affectivity" at "Invisible Machine," a group exhibition at University of California Santa Barbara on June 1st 2019. The right two imaged were taken by the ceiling mounted GoPro from the exhibition.

4.4 Presentation

4.4.1 2018 EoYS

Touching Affectivity was installed and presented at "Invisible Machine," a group exhibition at University of California Santa Barbara on June 1st, 2018, from 5-9pm (see Figure 4.11). For this exhibition, only the pressure sensors were active because the fur sensor was unplugged. This meant that the two main interactions were that more pressure created stronger vibrations, and that too much pressure caused the robot to "scream." The robot periodically chirped because of the random signal coming into the unplugged touch sensor input. Institutional Review Board (IRB) information sheets, which asked participants to "touch and/or pick-up the robot," were on a stand next to the robot so visitors could choose whether they would be recorded for research purposes. A GoPro was hung above the pedestal recording visitor movement and interaction through the space for the the first half of the exhibition. The work was exhibited outside of the AlloSphere entrance so it got periodic interactions with large groups when people were waiting in line for the performance, and less crowded interactions in between.

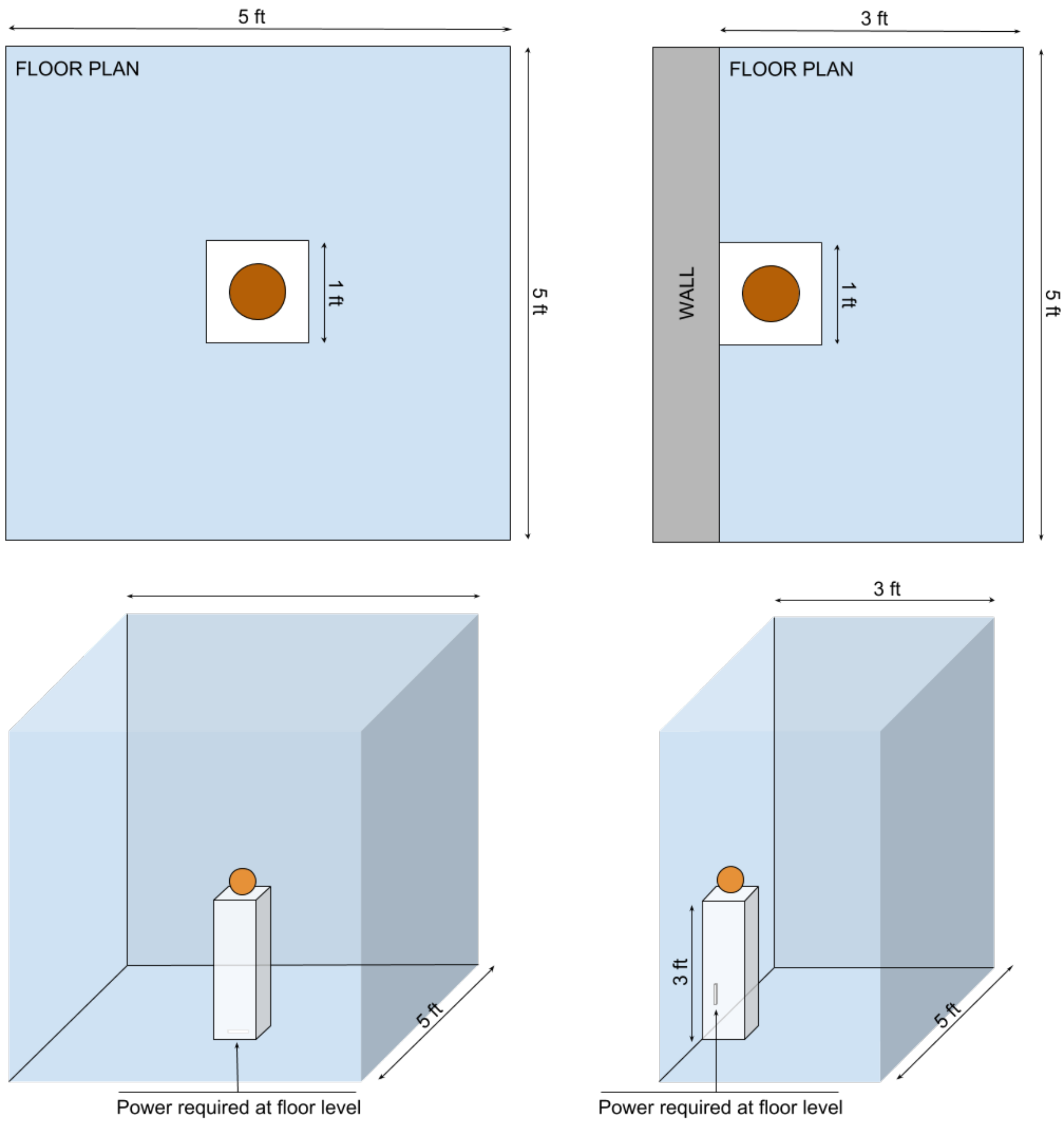

Figure 4.12: Installation Design.

Figure 4.13: Exhibition of "Touching Affectivity" at NIME 2019.

4.4.2 2019 NIME

Touching Affectivity was installed and presented at The New Interfaces for Musical Expression (NIME) Conference June 3rd-5th, 2019, 9am-6pm each day (see Figure 4.13). In this exhibition both the pressure sensors and the fur sensor were active. To compete with the sound of the environment, a Mackie Mix5 was used to amplify the sound. Again, it was exhibited in a hallway outside of entrances to other installations. Since I did not record video of the interactions, IRB approval was not needed for the exhibition.

4.5 Reception

At the EoYS the robot's fur sensor was not plugged in, so the interactive sounds the robot produced were based on the sound of the motor vibrations and the high pitched noise when a pressure sensor was above 78%. In this situation, participants were incentivised to squeeze the robot. Some visitors recognized the high pitched noise as a nega-

	EoYS 2018	**NIME 2019 Day 1**	**NIME 2019 Day 2**	**NIME 2019 Day 3**
# of interactions	1174	5171	4301	3633
Total length of interactions	1h34m	3h32m	1h11m	53m
Average length of interactions	4814.25ms	2455.65ms	995.28ms	883.41ms
Time exhibited	5h31m	9h46m	9h30m	8h59m
Max Sensor 1 value	914	802	772	711
Max Sensor 2 value	991	665	675	666
Max Sensor 3 value	1023	30	14	11
Max Sensor 4 value	1023	715	752	704
Max Sensor 5 value	1023	688	666	616
Max Sensor 6 value	1023	761	768	710

Figure 4.14: Metrics of the data collected on different days.

tive sound. For example, in the video recording, one person said, "Stop! It's screaming!" while their friend, who was squeezing it and feeling the vibrations, responded, "No! It's purring!." As seen in Figure 4.14, during the EoYS the top and bottom pressure sensors (3, 4, 5 and 6) reached 100% at some point during the show, while they never reached 78% at NIME. The short, high-pitched chirp interaction was barely heard at NIME, with only one sensor reaching exactly 78% on the first day. The person interacting with the robot after hearing that sound did not press harder, unlike at the EoYS.

4.5.1 2018 EoYS Video Analysis

I recorded video data of the interactions for 3 hours and 35 minutes of the 5 hours and 31 minute exhibition. The video recordings of these interactions produced the most interesting results. During this period of time there were approximately 217 instances where a person approached the robot, touched it, and left. Fewer than 217 people actually interacted with the robot because sometimes people would come back to touch the robot again. I wrote an algorithm to see if there was a correlation between the sensor data and how much movement there was in the video at different distances, but the results were not conclusive.

The work in many ways was a conversation piece. The video camera included an audio recording. People frequently explained and showed different ways to interact with the robot to each other. There were some interesting group dynamics. At one point, a group of three people all squeezed the robot at once. Because the robot was quiet in a loud environment, people would hold it up to their ears. During the recording, 5 people took photographs or videos of the robot. One person recorded a video of another person holding the robot, saying "I am touching a furry object."

There were some unexpected and sometimes violent interactions. People shook the robot, tossed it back and forth between two hands, tried to tie it in a knot, hugged the robot tightly, punch the robot, and threaten to throw it. Some visitors were worried about how they would be perceived by researchers. People commented about the user study, and spoke to one another about it. After learning that they were being recorded one person said, "I was squeezing it. They are probably going to think I am evil." Other participants after reading the description were still physically aggressive with the robot. One visitor especially liked the work stating it was the most interesting piece in the exhibition, viewing it as a test to see how violent someone would be with a robot when trying to make it respond. In the video he came back to the work multiple times with different people.

Visually, people viewed it as animal-like. They commented on the robot's embodiment, comparing it to a Furby, a child's toy, or a tribble, a creature from Star Trek. One woman was very vocal about it being like a tribble and how much she liked it. In response to the tail one person said, "Oh, it's got a tail! I really like the tail," and another said, "It's like a brain with a spinal cord." While the design did actively reference the idea of a tribble, visual metaphors, like a brain and spinal cord were apt but unintentional.

Because audience members viewed and treated the robot like an animal some wanted it to be happy. At one point, after seeing a boy punch the robot and walk away, a girl walked over, pet it and hugged it. Many people referred to the vibrations as purring and were confused that when squeezing it, it felt like it was purring very hard haptically, but sonically it sounded like it was in pain. There were even arguments between people in groups about whether squeezing it was "hurting it" or making it happy because it was purring. While a group of 5 people interacted with it, a conversation about the work included, "No, you are hurting it," "Whoa, this is so cool," "Make it purr," "Make it vibrate," and "It's kind of purring."

4.6 Sensor Data Analysis

Because the sensor data from the furbot was not at a consistent rate, the data was first interpolated and resampled at 2 milliseconds. Next, onset detection was used to find each interaction's onset and duration. If multiple peaks from any of the pressure sensors were within 2 seconds, they were considered part of the same interaction. Then for each interaction the variables collected were the number of onset peaks, and the duration of the interaction. For each sensor in each interaction, the max value, sum, average, number of peaks and variance were calculated.

First, a principle component analysis (PCA) was run on the resulting data (Appendix A), and the PCA divided data by sensors when attributing variables to the 3 components. It included sensor data 4, 5, and 6 for the first principle components, sensor data 1 and 2 for the second principle component and sensor data 2 and 3 for the third principle component. There is more variation for the third component in the EoYS data because the 3rd sensor was broken at NIME. There was more variation for the second principle component in the NIME data, because people were interacting with it in more varied ways, instead of just squeezing it. From these promising results an ANOVA was run to compare the data more rigorously.

Next, an ANOVA was calculated for each of the variables collected, comparing EoYS to NIME days 1, 2, and 3 as the category for the independent variable. An ANOVA was also calculated for each of the variables collected, just comparing NIME days 1, 2, and 3, to determine which variables were significantly different between all the categories and just the NIME categories. Significance was calculated, taking into account the Bonferroni correction, using alpha/32 for the p-values. For each significant variable, individual ANOVAs were calculated.

4.6.1 Analysis of Variance

I used Cohen's 1992 guidelines to determine effect size with the R values of .3-.5 considered a medium effect and .5+ considered a high effect for the respective variables. The first ANOVA of EoYS vs NIME days 1, 2 and 3 individually showed significant results for the duration of the interactions, the sum of the signal for sensor 2, and the max, sum, average and variance of the signals for sensors 3, 4, 5 and 6 (see Figure 4.15). Also, the number of peaks per sensor was significant for all 6 sensors.

Between the EoYS and NIME days 1, 2, and 3, the sensor results for sensors 3, 4, 5, and 6 were significant. According to the R values, the max values and average of values had a high effect, the variance had a medium to high effect, and the sum of values had a medium effect. The pressure sensor results for sensors 1 and 2 were mostly insignificant, with little effect. The sum of sensor value 2 in interactions was significant, when comparing the EoYS to NIME day 2 and NIME day 3 (see Figure 4.16). Here I argue that sensors 1 and 2 are not really engaged when people squeeze the robot very hard,

sensor # variable	**NIME 1 vs NIME 2 vs NIME 3**				**EoYS vs NIME 1, 2 & 3**			
	F	p	sig	r-sq	F	p	sig	r-sq
1 max	1.442	0.237		0.000	0.9709	0.405		0.000
1 sum	4.115	0.0163		0.001	3.735	0.0107		0.001
1 avg	1.271	0.281		0.000	3.097	0.0257		0.001
1 var	1.094	0.335		0.000	1.461	0.223		0.000
2 max	1.642	0.194		0.000	2.497	0.0578		0.001
2 sum	6.449	0.00159	*	0.001	8.107	2.16e-05	***	0.002
2 avg	0.7444	0.475		0.000	3.815	0.00956		0.001
2 var	0.6198	0.538		0.000	1.100	0.348		0.000
3 max	72.46	5.04e-32	***	0.011	677.7	0.00	***	0.125
3 sum	53.73	5.78e-24	***	0.008	227.9	2.07e-144	***	0.046
3 avg	568.0	2.84e-237	***	0.080	331.7	3.37e-208	***	0.065
3 var	53.59	6.64e-24	***	0.008	308.6	3.81e-194	***	0.061
4 max	4.846	0.00788		0.001	3003.	0.00	***	0.387
4 sum	0.8904	0.411		0.000	1221.	0.00	***	0.204
4 avg	7.620	0.000493	*	0.001	1784.	0.00	***	0.273
4 var	1.259	0.284		0.000	1308.	0.00	***	0.216
5 max	1.020	0.360		0.000	3437.	0.00	***	0.419
5 sum	6.509	0.00149	*	0.001	1135.	0.00	***	0.193
5 avg	5.579	0.00379		0.001	1812.	0.00	***	0.276
5 var	2.180	0.113		0.000	1720.	0.00	***	0.265
6 max	4.101	0.0166		0.001	2556.	0.00	***	0.349
6 sum	0.4667	0.627		0.000	533.6	0.00	***	0.101
6 avg	5.170	0.00569		0.001	1590.	0.00	***	0.251
6 var	0.8159	0.442		0.000	1369.	0.00	***	0.223
duration	61.35	3.01e-27	***	0.009	109.9	2.27e-70	***	0.023
# peaks	-6551.	1.00		-inf	-4758.	1.00		-inf
1 peaks	95.10	9.95e-42	***	0.014	42.89	1.39e-27	***	0.009
2 peaks	205.8	9.85e-89	***	0.030	61.92	9.09e-40	***	0.013
3 peaks	79.84	3.43e-35	***	0.012	406.6	1.80e-253	***	0.079
4 peaks	17.54	2.48e-08	***	0.003	17.66	1.94e-11	***	0.004
5 peaks	158.9	6.51e-69	***	0.024	79.56	4.85e-51	***	0.016
6 peaks	1.625	0.197		0.000	29.14	8.91e-19	***	0.006

Figure 4.15: ANOVA Results for different variables comparing the different days.

sensor #	NIME 1 vs NIME 3			NIME 2 vs NIME 3			NIME 1 vs NIME 2		
variable	F	p	sig	F	p	sig	F	p	sig
2 sum	8.571	0.00342		0.009259	0.923		9.175	0.00246	
3 max	30.97	2.70e-08	***	4.068	0.0437		144.5	4.70e-33	***
3 sum	65.16	7.82e-16	***	1.153	0.283		56.30	6.79e-14	***
3 avg	1046.	6.91e-217	***	17.38	3.09e-05	***	669.8	9.52e-143	***
3 var	79.41	6.07e-19	***	1.875	0.171		77.12	1.89e-18	***
4 avg	15.58	7.96e-05	**	3.468	0.0626		4.854	0.0276	
5 sum	14.25	0.000161	**	0.7579	0.384		3.258	0.0711	
duration	70.16	6.31e-17	***	1.378	0.241		63.80	1.54e-15	***
1 peak	123.2	1.92e-28	***	4.042	0.0444		92.83	7.19e-22	***
2 peak	261.6	5.12e-58	***	10.64	0.00111	*	199.3	8.58e-45	***
3 peak	19.63	9.53e-06	***	34.02	5.66e-09	***	142.3	1.40e-32	***
4 peak	0.5878	0.443		12.80	0.000349		37.35	1.03e-09	***
5 peak	145.8	2.58e-33	***	4.810	0.0283		200.1	5.72e-45	***
	EoYS vs NIME 1			EoYS vs NIME 2			EoYS vs NIME 3		
	F	p	sig	F	p	sig	F	p	sig
2 sum	2.157	0.142		18.58	1.66e-05	***	16.48	4.98e-05	**
3 max	559.6	4.09e-117	***	565.9	2.45e-118	***	563.9	5.92e-118	***
3 sum	184.0	3.60e-41	***	192.5	5.92e-43	***	192.9	4.86e-43	***
3 avg	282.5	1.18e-61	***	272.8	1.17e-59	***	268.9	7.60e-59	***
3 var	256.6	2.63e-56	***	256.6	2.62e-56	***	256.6	2.62e-56	***
4 max	3281.	0.00	***	3307.	0.00	***	2998.	0.00	***
4 sum	1094.	2.98e-216	***	1104.	4.61e-218	***	1065.	3.61e-211	***
4 avg	1701.	1.35e-318	***	1731.	1.98e-323	***	1583.	1.86e-299	***
4 var	1157.	2.47e-227	***	1162.	2.92e-228	***	1132.	4.13e-223	***
5 max	3362.	0.00	***	3396.	0.00	***	3351.	0.00	***
5 sum	980.9	4.21e-196	***	1019.	5.25e-203	***	1021.	2.93e-203	***
5 avg	1648.	5.46e-310	***	1623.	6.19e-306	***	1612.	3.88e-304	***
5 var	1483.	6.13e-283	***	1474.	1.84e-281	***	1485.	2.83e-283	***
6 max	3099.	0.00	***	3201.	0.00	***	2860.	0.00	***
6 sum	536.5	1.33e-112	***	909.1	4.52e-183	***	880.9	6.79e-178	***
6 avg	1705.	3.11e-319	***	1711.	3.90e-320	***	1572.	1.33e-297	***
6 var	1280.	1.16e-248	***	1277.	3.29e-248	***	1245.	1.20e-242	***
duration	24.80	6.57e-07	***	638.2	2.48e-132	***	980.1	5.87e-196	***
1 peak	2.939	0.0865		34.32	4.98e-09	***	39.54	3.49e-10	***
2 peak	1.826	0.177		17.71	2.61e-05	***	25.46	4.68e-07	***
3 peak	296.6	1.54e-64	***	459.0	2.45e-97	***	392.3	5.43e-84	***
4 peak	14.51	0.000141	*	15.25	9.53e-05	**	14.54	0.000139	**
5 peak	0.005228	0.942		58.99	1.91e-14	***	52.56	4.85e-13	***
6 peak	24.95	6.08e-07	***	24.50	7.69e-07	***	30.43	3.64e-08	***

Figure 4.16: ANOVA Results for variables that were significant in Figure 4.15 comparing the different days individually.

attempting to make it squeal, while the sensors 3, 4, 5 and 6 around the circumference are engaged.

Sensor 3 Results

While sensor 3 was considered significant, its R values showed low or no effect between NIME and EoYS data. From the data max values and post mortem testing, it looks like sensor 3 was damaged between the EoYS and NIME. Sensor 3 went from a max value of 1023 at the EoYS to 30 on the first day of NIME, 14 on the second, and 11 on the third. After NIME, sensor 3 was no longer responsive. This also explains why in Figure 4.16 the results for sensor 3 are significantly different between the first day and the second day, as well as the first day and the third day, while the difference between the second and third day are not significant.

Duration and Peaks Results

The duration and number of peaks per interaction by sensor, while significant, had an R-squared value below 2, except for the number of peaks for sensor 3, which had a low effect for the duration. The significance of the number of peaks sensor 3 produced makes sense because, as the sensor became more damaged, it was reading fewer interactions and therefore fewer peaks.

The difference in duration can be explained by the venue in which the works were displayed. At the EoYS the venue was crowded and, according to the video data, frequently one person would put down the robot and someone else would pick it up within a second. At NIME, the robot was shown across 3 days. Because the conference had just started and the installations were new, people interacted with the robot for a longer period of time on the first day. On the first day the robot was interacted with for the longest total period of time, and the average length of the the interactions was over 2 times the length of the interactions on the second and third days. There was a significant difference between the duration of the interaction between the first day and the second day as well as the first day and third day, while the difference between the second and third days was not significant.

4.7 Discussion

While the robot responded emotively to touch more accurately at NIME, the videos from the End of Year Show produced interesting results when the emotion expressed conflicted at the tactile and sonic level. At NIME when I spoke to people about the work, they would say it was really interesting and/or wanted to discuss the technology behind the sonification. At the End of Year Show, people were more interested in discussing the work conceptually. The conflict created by the installation at the End of Year Show was viewed as meaningful by visitors.

There was a significant difference in the way that people interacted with the robot. While comparing the data between the End of Year Show and NIME is complex because the venue, the people, and the presentation of the work were all different, the main factor in the way that people interacted with the robot was the sonic reaction of the robot to touch. The main differences in the signal were the signal max for the pressure sensors around the periphery of the robot. If the robot only responded sonically when it was squeezed hard, that incentivized viewers to squeeze it.

4.7.1 End of Year Show 2018

At the End of Year Show, some did not see the robot as a creature, one stating that they were “touching a furry object” in a video they were recording of the work. Most people treated the robot like it was an animal, comforting it when it had been punched and discussing it like it had feelings and could feel pain. The robot had a face, a consistent texture and it moved, all signifiers of an animate creature. People empathized with the robot as described below. While the interaction with the robot was viewed as uncomfortable, the physical appearance did not seem to bother any of the participants.

Touching Affectivity explored the conflict between haptic and sonic response. This work explored how people handled being told that a less reactive robot would sonify the way it was touched. The viewers were told that the robot would sonify the way it was touched, but the robot only responded sonically to being squeezed extremely hard. The robot also vibrated softly with increasing intensity when squeezed. This conflict between the harsh sound and soft vibration created an interesting quandary in the people who interacted with the robot. The question, was often asked is the robot screaming or purring? The instructions said that the robot sonified the way it was touched but its sonification sounded like it was in pain. People asked, is this the sound it is suppose to be making? A sound that was designed to disincentive squeezing, along with a haptic response actually incentivized squeezing. I believe that people at the End of Year Show interacted with the robot more aggressively because they were incentivized to do so by the robot's sonic and tactile reactions to touch. Because they had to squeeze hard to get a sonic response, they felt that it was appropriate to touch it in more aggressive ways.

This work's sonic response and tactile response caused participants to be uncomfortable with the way that others handled the robot. The person who was squeezing the robot experiences the vibrations of the robot as they increased their grip's pressure. They experience it as a purring sensation or vibration, which in a feedback loop make them squeeze more because the robot purred harder. When the robot began to make a high pitched screaming noise that was supposed to tell people they shouldn't be squeezing it that hard, they had already decided that the vibrations were purring and positive feedback. However, a person who was witness to this interaction heard the pain sound but did not feel the vibrations, so they viewed the other person as inflicting pain on an animal. This created a conflict between the person interacting with the robot and the person watching them. Groups would argue and discuss the emotion that the robot was expressing, is it happy or is it in pain? Some viewed this as the artist/researcher attempting to trick them into causing pain on the robot. This emphasized the importance of tactile and sonic synchronicity. This work fell into the uncanny valley because there was a disconnect between different responses which were in conflict.

Touching Affectivity shows the difference between intimate and personal interactions. This work emphasizes that there can be a large perceived difference between an intimate interaction with a robot compared to the personal interaction with it. The person intimately interacting with the robot perceived the interaction and the robot's expressed emotion as the complete opposite emotion that the onlooker viewed the interaction.

4.7.2 Context within Media Arts

This work shows the importance of being able to physically interact with haptic art. If people could not touch the work then they couldn't experience the tactile sensation of interacting with it, and the work would have lost its emotional meaning. There would have been no conflict at the End of Year Show if people couldn't interact with the robot intimately.

Distance based interaction is applicable to the arts, in particular to new media. In the art world most work is perceived in the social space. At a young age we are taught to stand back from art and not to touch it, with some galleries even having proximity based alarms. [266] Media artworks that are designed to be physically interacted with are often presented in galleries with guard rails. For example in a Nam June Paik retrospective at the Smithsonian American Art Museum [267], visitors could not sit on the *TV Chair* or interact with the *Magnet TV* as they were designed. This castrates the artwork, changing it from the original art piece to a historical artifact. Interactive artwork in the visitor's personal space is rare because of a history of non-interactive artwork, the cautious environment of the art gallery, and a bias toward preservation over functionality.

As seen in the literature review (Section 2.3.6), there are few cyborg artworks which include physical interaction with the public. Cyborg works and wearables are frequently shown in the intimate space of a performer and in the public space of the audience [251]. There are multiple reasons for artists and curators to hesitate about more physically interactive works [268]. There can be liability issues, both with the visitors' safety and the potential damage that the artwork could sustain. Exhibiting the work behind glass turns it into an artifact, so a partial solution is to have a scheduled time where artists or a team can assist the audience [251].

Artwork like *Trou Mireia* (2017) by Donat Melús [110] has handled touch based interaction issues with requiring participants to wear gloves while touching the work. This gives new meaning to the work which enhances *Trou Mireia*, but would not work with *Touching Affectivity*. When creating an interaction that is reminiscent of a pet, requiring people to wear gloves removes the tactile interaction. The goal of the high pitched noise when the pressure was above a certain amount was to signal to participants that they were squeezing too hard. This did not work at the End of Year Show, and was viewed as the intended interaction. This caused viewers to squeeze the robot so hard that one of the pressure sensors or its connector was damaged. This created the exact opposite situation than I was trying to achieve, but actually created more meaningful experience for the viewer and the artist. At NIME, because the robot was so reactive to touch and it wasn't viewed as a psychological experiment, people were a lot more gentle with the robot. This emphasizes the importance of designing cohesive interactions.

4.7.3 Summary

Touching Affectivity is an interactive sculpture that was used at exhibitions to collect data on the way that humans interact with robots in an arts context. The creature responded sonically and haptically to touch. It was exhibited at the MAT EoYS in 2018 and for three days at NIME 2019. At the MAT EoYS, the robot was exhibited for 5.5 hours with an info sheet stating that participants were being video recorded, and a placard with the work's abstract. For this interaction, there was a vibrotactile response to pressure, random chirps, and a high pitched squeak when the robot was squeezed too hard. At NIME 2019, there was no video recording, so no info sheet was needed. There was a placard abstract that described work that described the work and stated that the robot was collecting data. For this interaction, a haptic fur sensor which sonified the way the robot was being stroked was added to the other sonic and tactile responses from the previous exhibition. The audio response to the fur was based on prior research in emotive vocal communication and emotive music, where the scratching and tickling gesture created sounds with audio qualities similar to excited/happy speech, and petting and stroking created sounds with audio qualities similar to calm/sad speech.

The video from the End of Year Show showed different ways that people interacted with the robot and their responses to both the robot and others' interactions. A conflict was created because the viewer heard the sonic response of pain when the robot was squeezed and the participant experienced the haptic response of purring. The positive reinforcement of purring and the lack of a sonic response until high pressure was felt by the robot caused participants to squeeze the robot with as much pressure as possible. This was not the case with the reactive robot shown at NIME. After analyzing the signals from the 6 pressure sensors, the front and back sensors' signals were not significantly different between the two interactions, but the four perimeter sensors' signals were statistically significant. I believe this was due to people holding the robot and squeezing it at the EoYS so the robot would make the high pitched squealing noise, while at NIME people did not do this.

In conclusion, the robot in *Touching Affectivity*, can be used to find meaningful information about how we interact haptically with a robot that sonifies the way it is touched. It also emphasizes the difference in experience between personal and intimate spaces and the conflict it can create. From this work we learn that sonic and haptic responses need to be aligned, and the distinction between interaction in the intimate and personal spaces. This has applications for robots in the home, children's toys, and emotional support robots. This work also emphasizes the importance of touch in the arts, and that artwork that is intended to be interacted with, can looses it's meaning when the ability to touch is taken away.

Chapter 5

Interfaces for Immersive Environments

5.1 Introduction

The opposite of interactive art in the personal space is studying the impact of art in a large public space. In prior chapters, I have looked at social dynamics in exhibition spaces. While *Come Hither to Me!* compares private interactions and interactions with an audience, *Touching Affectivity* explores how the experience of an interaction is different for the participant and the viewer. Unlike smaller venues, for larger exhibition spaces it it is necessary to handle and create interactions between visitors, where multiple participants can interact with an installation. While there is an interaction between the artwork and the viewer, there is also interaction between visitors or collaborators.

When looking at the public distance from the perspective of human-machine communication, two questions arise: what machines should communicate to people in the public space, and what people want to communicate to machines. Examples of interactions in the public space, a distance greater than 12 feet, would be speeches, lectures, and theater [59]. In all of these spaces, the communication is happening one way, as a way to share information between performers and a group. In Media Arts, we need to tackle how to manage and synchronize multiple people interacting and/or experiencing the environment at the same time, using public space as a communication tool.

Digital interfaces are becoming more common as ways to interact with media art works, but this brings a whole host of problems. When work is presented with the ability for interaction it can evoke situational shyness, through the demand of active, performative engagement without instructional information [34]. This is an issue in public spaces where one visitor is interacting with the work while others, potentially strangers, are watching. Digital immersive environments can cause information overload and to build interfaces to control them we need to avoid overwhelming the user. When trying to visualize data in a meaningful way, curating and choosing what to see is an important part of design to avoid information overload. Giving the ability to curate to a user is a design challenge because the interface must be flexible and also not overwhelm the user with too many options.

5.1.1 Interactive Interfaces for the AlloSphere

The AlloSphere at the University of California, Santa Barbara is a three story immersive virtual reality environment, the perfect place to study how people interact with each other in large scale installations. In three works, I studied the way that people can interact with the art or data in a large environment and use it to communicate with each other. Because this is a large interactive environment and complex system, many people had to work together on these projects. I designed and built interactive interfaces for the three works, one in collaboration with Tim Wood. In the works discussed in this chapter, I designed and built browser-based, touchscreen interfaces for large scale interactive systems. I used interfaces that existed in the browser so visitors could easily access the interface, using their own device without downloading any software. The works explore what the artist can communicate in the environment, how the audience can interact with the work, and how people can communicate with each other through the medium of the arts.

One way to connect with people and communicate is through music. When the CREATE Ensemble, an experimental electronic music group of which I was a member, decided to perform our network feedback piece at the CREATE AlloSphere Spatial Sound Concert in 2016, we needed a way to communicate and control how our music was integrated in the AlloSphere. We performed *Feedback Rings* at the concert as 4 performances of a 10 minute improvisational work. I performed with my custom built instrument as one of six performers. Specifically for this performance, I built the interface to visualize and modify the way our instruments were networked together. Each performer used this interface on their phone or tablet to monitor and modify the network. This interface explored how can artists communicate with each other in an immersive environment.

Art can also be used to communicate information and for collaboration. For *the Hydrogen-Like Atom*, an installation by JoAnn Kuchera-Morin, Lance Putnam and Luca Peliti, I built the interactive interface collaboratively with Tim Wood. *The Hydrogen-Like Atom* sonifies and visualizes the time-dependent Schrödinger equation, allowing visitors to modify parameters of the equation and visualize the results. Participants used a touch screen interface to interact modify and change the installation. We designed and built a touchscreen-based interface for one person to interact with the system. We also built a mobile version that visitors could access on their phones for an exhibition at the Museum of Exploration and Innovation (MOXI) from January til June of 2016. It was also shown at the International Symposium for Electronic Art (ISEA 2017) for 5 days, for one showing at the Alliance of Women in Media Arts and Technology (AWMAT

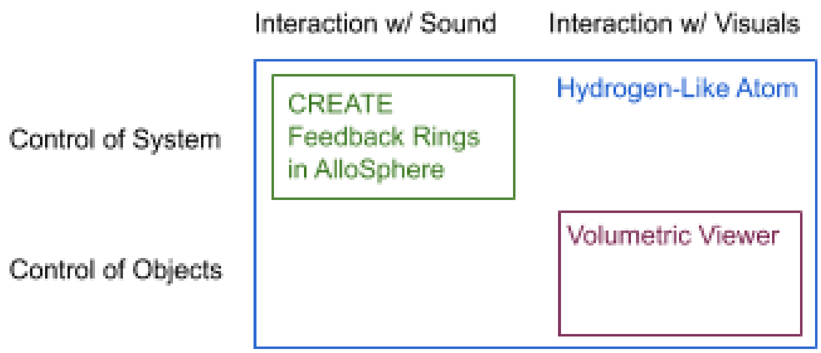

Figure 5.1: For each artwork discussed, what is the interface interacting with and what is it controlling.

	Digital Interface	**Physical Interface**
Required Collaborative Practice	CREATE Feedback Rings in AlloSphere	
Actively Collaborative	TEM/AP Parameter Controller Hydrogen-Like Atom Interface	TEM/AP Glove Interface

Table 5.1: For each artwork discussed, the type of interface and how collaborative is the interaction.

2018) Conference and for 4 days at the Computer History Museum during the ACM International Conference on Multimedia (ACM-MM 2017).

In order to allow scientists to communicate and visualize volumetric materials data, the AlloSphere Research Group built *the Volumetric Viewer*. This work has been shown as a proof of concept for scientific visualization at many AlloSphere demonstrations for scientists, donors and visitors. I built a touchscreen-based interface for the system to modify scientific and visual parameters, and a specific tool to interact with the visualization and to see the raw data from which it is derived. Tim Wood created a wireless glove interface to move and manipulate the camera and data.

5.2 Design

In this Chapter, I explore collaborative interfaces for digital and physical environments. They span both digital and embodied interfaces for virtual and physical situated art. These works can be discussed across multiple axes of interface design, interface embodiment and collaborative interfaces. The interfaces discussed are digital interfaces, gestural inputs for digital interfaces, and cyborg gestural interfaces. An example of a digital interface is a touch screen device or computer. A digital interface can take a gestural input by sensing the movement of the device, through a sensor like an accelerometer. For a cyborg gestural interface we used motion tracked gloves. All of these interfaces, because they are interactions at the public distance, do not involve touching the work, but instead use only spatial gestures or digital interfaces separate from the environment. These works explore different levels of collaboration: systems that require a collaborative practice and systems that can be actively collaborative. These works control virtual objects and audio-visual systems.

The works also differ in control spaces and objectives (see Figure 5.1). The control spaces include control of the system, control of objects/agents, and control of self. There is a difference between controlling agents indirectly through a digital interface and controlling them through the control of self (one's own embodiment). On this axis, we clarify the difference between the agency of controlling oneself, in which passively interacts with the agents that have their own agency, controlling through gesture one's point of view, and controlling agents more directly through interfaces with digital parameter inputs. In these systems, interactions are designed for agent-based, sound-based and visual-based artworks.

Information is conveyed multimodally in these works, both through visualization and sonification. Control of sound is explored in many of these works. Both speech and

User	Use Case	Issue	Solution
Researchers	Data visualization and communication tool	Parameters are disparate and have different requirements	Multimodal interfaces
Performers	Spatialized audio performance tool	Performers were too far apart to interact with each other or the same interface	Distributed interactive interface for control and visualization of the system.
Exhibit Guests	Audio-visual art installation	Performance Shyness	Allows visitors to use their own devices or approach the touchscreen.

Table 5.2: For each work, there was a different user, use case, issue and solution.

non-linguistic audio cues are used to communicate information. These works use sound to sonify a system of equations through timbral quality, a network of instruments and both emotional and physical distance. Sound is used to convey information non-verbally. Filtering and synthesis techniques are used to generate non-linguistic information.

5.3 Distributed and Multimodal Interfaces

In the public space, people are interacting with technology at a distance greater than 12 feet and multiple devices may be communicating information at the same time. If they are communicating through the same mode, it can be hard for the user to interpret what is going on. Also, communicating through verbal language can be repetitive and annoying. Multimodal and non-linguistic communication is the solution to this problem.

Lectures and performances are examples of communication in the public space, where they are used as tools for communication with multiple people. In an interactive environment in a public space, multiple people can interact with or experience the environment at the same time and use it as a communication tool. In the public space, I looked at this form of interaction and communication through 3 projects (see Table 5.3). The first explored visual communication in the research environment. The second studied collaborative interfaces for multiple performers to control networked spatialized sound. The last examined collaborative interfaces used by multiple audience members to control an audio-visual synthesizer. These works were performed in the AlloSphere, a three story immersive virtual reality environment. For these works, participants use a touch graphical user interface (GUI) on a tablet's browser using interface.js [269].

Prior work in HCI with mobile devices as multi-degree of freedom controllers is described in Chapter 2.2.2. While Prachyabrued et al. has used tracked mobile devices to select objects and get more information about cluttered data sets [66], they do not move or manipulate the data that is being visualized as I do. [66] In my work, I used the tablet's gyroscope to control the rotation of a plane. I used the tablet interface for navigation as well as to control visualization parameters, and I used the tablet to display information about the state of the visualization. My work allows for large gestural control as well as detailed interactions.

Prior work in the visualization field used gestural control to manipulate large data sets is described in Chapter 2.2.2. In the medical field, slices and 3D visualizations have been used to visualize volumetric data with a 3D display wall and touch table [68], but not in an immersive environment. In their research project, participants only use selection on the touch table via the objects' shadows instead of using a tablet's gyroscope to control the plane of intersection as I do. My work brings the interaction to 3 dimensions instead of just interacting with the object's shadow.

5.3.1 CREATE Ensemble Feedback Performance

When the CREATE Ensemble, an experimental electronic music group of which I am a member, decided to perform *Feedback Rings* at the CREATE AlloSphere Spatial

Figure 5.2: View of performance from the catwalk and bridge of the AlloSphere. Photography by Ge Wang.

Sound Concert in 2015, we needed a way to communicate and control the system in the AlloSphere. So I designed and implemented a digital interface (see Figure 5.3). As part of the CREATE Ensemble, we invented an improvisational method where each instrument's input and output were networked together in different configurations to create a networked feedback loop [270]. I designed and built an embedded instrument for the group using a Raspberry Pi and Arduino Nano. In a traditional practice setting, we sat facing each other and could communicate both verbally and non-verbally. In the AlloSphere, we were separated around the radius on the catwalks on the third floor (Figure 5.2). I implemented a networked interface which could be used via the browser that allowed performers to monitor each instrument's inputs and outputs, as well as to view and to change the configuration of the networked feedback instrument.

The interface (Figure 5.3) was designed to allow users to see each person's instrument's inputs and outputs. It also allowed performers to see the current network configuration. Performers could modify the network using preset buttons or they could modify the network connections directly. The digital interface required a collaborative practice interacting with the sound, controlling the information system.

Figure 5.3: Interface for CREATE Ensemble Feedback Performance in the AlloSphere.

5.3.2 *The Hydrogen-Like Atom*

As part of the AlloSphere Research Group, I created interfaces for a quantum synthesizer that sonifies and visualizes the hydrogen-like atom, which allows multiple users to collaborate in creating new compositions and improvisational works. It was originally designed and performed in the AlloSphere (Figure 5.5). The work was exhibited at the Museum of Exploration and Innovation (MOXI), the International Symposium for Electronic Art (ISEA) (Figure 5.4), the Alliance of Women in Media Arts and Technology (AWMAT) Conference and the Computer History Museum during the ACM International Conference on Multimedia. For each space the interface was modified to fit the environment in which it was exhibited. The actively collaborative digital interface allowed users

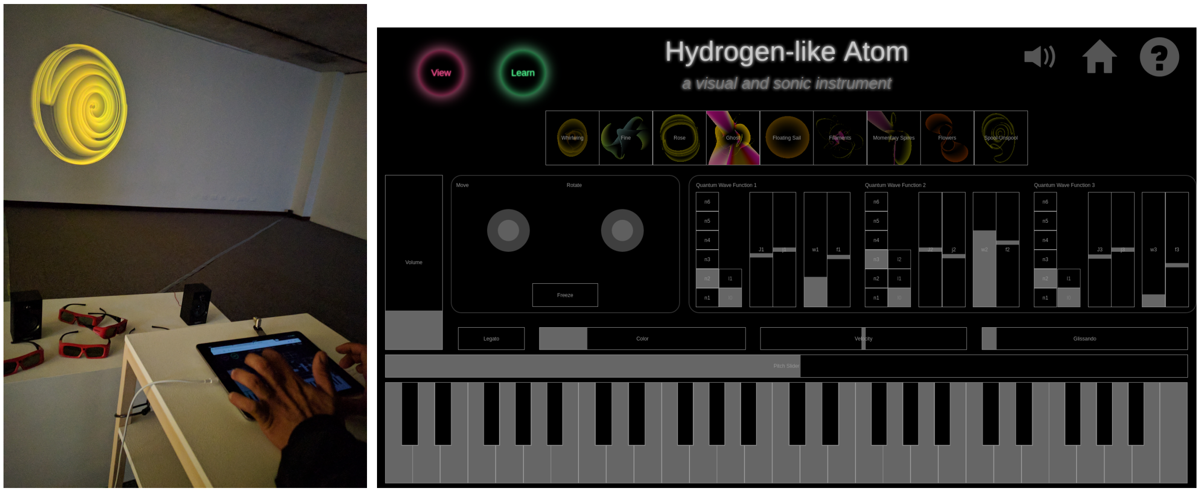

Figure 5.4: *Hydrogen-Like Atom* exhibited at ISEA and interface design.

Figure 5.5: *Hydrogen-like Atom* exhibited in the AlloSphere at the AWMAT Conference with 2 people interacting with the large touch screen and 1 person using the iPad interface.

to modify the sound and visuals, controlling the system and the object.

The interface was originally designed for a 27" touchscreen. It had 3 modes: "View," "Learn" and "Create." Very rarely were the View and Learn interfaces used. It was not exhibited as I originally intended it at the MOXI. I wanted people to be able to enter the space and interact with the interface freely. Instead visitors would enter the space and sit down while the gallery workers would start one of the pre-recorded compositions from "View." After that was complete, visitors were invited to interact with the interface on the "Create" page. This interface included individual controls for the parameters of the equation, buttons for specific preset parameters, joysticks to move the camera, a pitch slider and a keyboard. The keyboard was included as an entry point into the interface. People were shy to approach the stand, and only one person interacted with the interface at a time.

I wanted to make interacting with the system less intimidating and allow for multiple people to be able to interact with the interface at one time. One way to do this was to have separate interfaces on iPads that could be handed out to audience members. Since there were so many parameters to control, this interface had multiple tabs where each tab had one control. This included a tab for navigation, a tab for the equation's parameters, a tab for presets, and a tab for system parameters. Visitors were less inhibited and more likely to interact with the iPad than they were to approach the main interface.

At ISEA, we did not have the large touch screen and the larger iPad broke during transportation. We showed the full interface on an small iPad. The system was set up to play a prerecorded composition, if no one interacted with the interface for a short period of time. During the exhibition, interacting with the interface was so popular that the prerecorded composition rarely played. Only one person could interact with the system at a time because the iPad was small and we only had one interface. Despite the small interface, visitors and the public were not inhibited.

5.3.3 *The Volumetric Viewer*

Collaborating with materials researchers, I worked with their Transition Electron Microscope and Atom Probe (TEM/AP) data to develop a way for material scientists to overlay two different datasets from the same material and see how they related. This project was initiated by a need to visualize, analyze, and align 3D materials data sets in a collaborative environment. One or more researchers could work in concert to interact

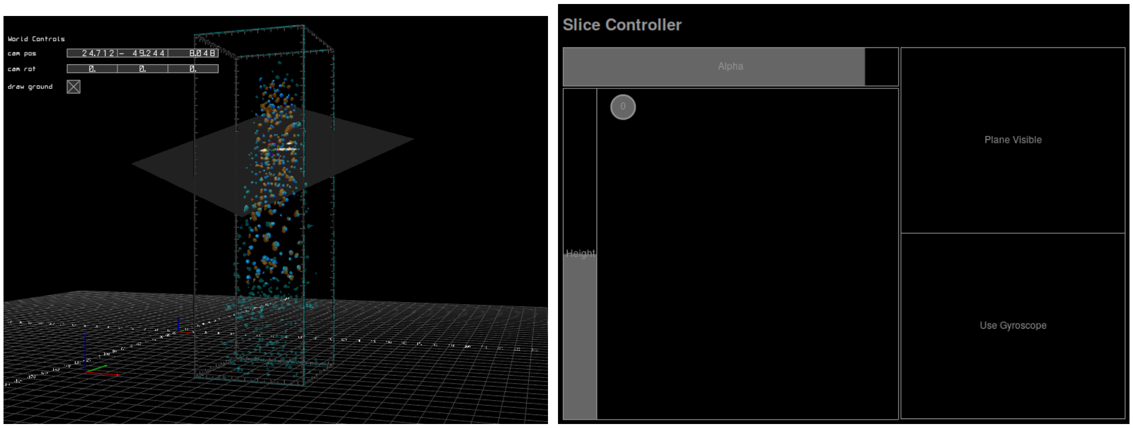

Figure 5.6: *Volumetric Viewer* interface for displaying slices of raw data in the visualization.

with visualization and transformation parameters. A tablet provided both a 2D interface for interaction with visualization parameters, and a 3D interface utilizing accelerometer and gyroscope data to control a slice view of the raw volumetric data. A gestural glove controller provided a natural 3D interface into the immersive world, where users could navigate, select, and transform data sets in space. Users utilized both interfaces simultaneously to naturally move between exploration tasks. This multi-user interface system could expand to a number of researchers controlling various parameters of the data, using a multitude of wireless 2D and 3D devices. This project was expanded into a general purpose volume viewer which allowed researchers to create a scene, upload multiple volumetric datasets to visualize in that scene, visualize the scene, and run calculations on the data. This was an actively collaborative interface that was both digital and physical, controlling visual objects.

Defining Plane of Intersection

To view the raw voxel data, a plane could be added to the dataset, which displayed interpolated data of the plane's intersection with the dataset (See Figure 5.6). The plane was semi-transparent, resembling an x-ray slicing through the data. The location of the plane was controlled by the tablet's touch interface, and the orientation of the plane was controlled by the accelerometer sensors in the tablet, transmitting the orientation and rotation information.

Between the touch interface and the tablet's sensors, the tablet interface allowed for 6-degree-of-freedom control for the plane (See Figure 5.6). The tablet's gyroscope and accelerometer provided an intuitive interface for controlling the plane's normal. The user could control whether the gesture data was being sent by holding a button on the edge of the device with their thumb while making the gestures. The plane mirrored the movement of the tablet instinctively. The location of the plane in the space was controlled for the X and Z axis on the tablet by a joystick, while using a slider for the Y axis. The user could also adjust the granularity of the movement via a set of toggles on the tablet, allowing for finer grain control. The level of transparency or alpha value of the plane could be adjusted to allow users to bring focus to the plane or the isosurfaced data.

Parameter Control

The main interface allowed users both to control and to read information about the current state of the visualization. The interface allowed users to control the way the data was visualized, the alignment of the datasets, and the navigation of the camera. Users could adjust these values through granularity controls, sliders and joysticks or by typing in manually the specific values (see Figure 5.7). This allowed users to modify the visualization free form or to dial into a specific location. The interface also allowed users to view the specific values about the data that was chosen via the gloves, joysticks or sliders. Users could also save configurations as presets to view later and could record the raw state to a text file.

The touch screen in the tablet provided control of various parameters that affected the

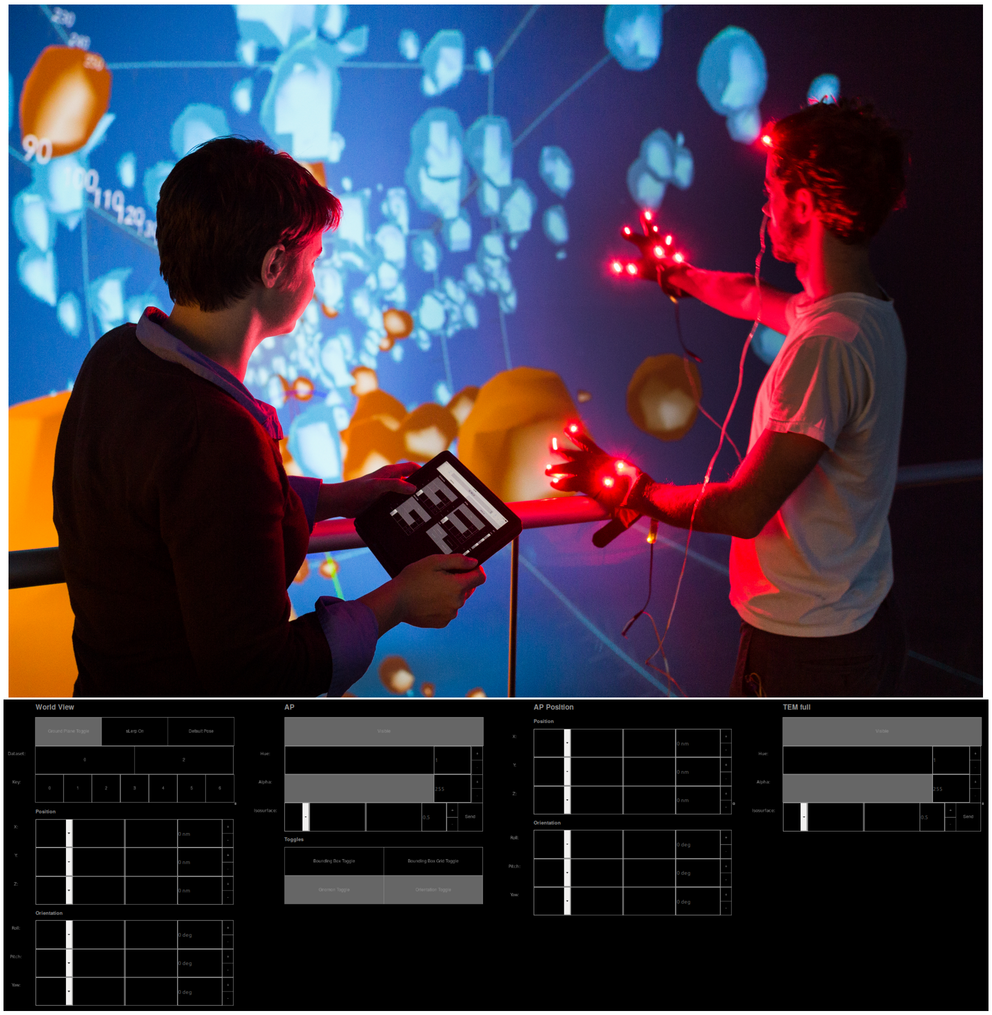

Figure 5.7: Parameter control interface and glove interface for *the Volumetric Viewer*. Photography by Kurt Kaminski.

visual appearance of the datasets, such as colors, isosurface value, and transparency. The interface could also be used to turn on or off scale information, to load different datasets, to move datasets and to navigate the camera. If only one person was presenting the visualization, the tablet could be used for aligning datasets and navigating the camera. However, we found using the gloves to be more intuitive for these tasks, and using both the tracked gloves and tablet was ideal when analyzing the data.

5.4 Discussion

In this chapter, I discussed design problems and solutions for collaborative interfaces for three immersive environments. These works addressed the problems that arise in collaborative interfaces due to distributed contributors, situational shyness, and information overload. The first studied collaborative interfaces for multiple performers to control networked spatialized sound, addressing distributed performers. The second work examined collaborative interfaces used by multiple audience members to control an audio-visual synthesizer addressing situational shyness and information overload. The last explored building collaborative interfaces for researchers to explore visual data addressing information overload. The solution to these design problems were multimodal interfaces, roles, and distributed interfaces. Distributing complex interfaces into roles is one way to simplify interfaces and reduce confusion. It also allows for roles to use the best modality for the task. Distributed interfaces and control also allow for communication across spaces where users might feel uncomfortable to speak or unable to speak due to distance.

For the spatialized audio performance tool, performers were too far apart to interact with each other or the same interface. The solution was distributed interfaces which allowed the performers to share information and change the system. For the audio-visual art installation performance shyness and an overwhelming interface was an issue. The solution was to allow visitors to use their own devices to connect to an interface that had divided the interface into multiple roles. The user could choose the role they were interested in and see a interface specific to that role. By having patrons use their own personal device, we can take advantage of their pre-existing comfort with it. Finally, for the data visualization tool, parameters were disparate and had different requirements. The solution was to distribute the parameters across different devices with different roles. By splitting up the interface across devices, users were not overwhelmed by the amount of parameters that need to be controlled. It also allows each parameter to be associated with the device in the modality best suited for the task.

Digital interfaces have limitations for certain tasks. Navigating the camera, moving and adjusting objects in 3 dimensions is a complex task on a two dimensional screen. We solved this problem in *the Volumetric Viewer* by using a gestural interfaces to navigate the camera, move objects and rotate planes. The two gestural interfaces we used were a cyborg glove based interface, and a gestural interface using the gyroscope and accelerometer data from the iPad. These new modalities were more intuitive than the touch screen interface. By using the iPad as a tangible as well as a digital interface, I used the physical qualities and sensors that the iPad had to build an interface which was more intuitive.

This chapter explored building interfaces for immersive virtual reality environments to support interactions between performers, between visitors and between researchers. I focused specifically on social dynamics in exhibition spaces between different groups of people. These works explored distributed, multimodal interfaces and using personal mobile technology to handle multiple contributors, situational shyness, and information overload. Distributed interfaces allows collaborators who are not in the same place to communicate and monitor the system. Multimodal interfaces allow us to choose the best interface for the given task. Allowing visitors to use their own devices creates a more comfortable collaborative environment in exhibition spaces. This information can be used for designing multi-user systems and future exhibitions.

Chapter 6

Current Work: *Cacophonic Choir*

Cacophonic Choir, a work in progress, explores the idea of proxemics and interactions at different distances. *Cacophonic Choir* is an interactive installation composed of multiple agents distributed in space. All the agents are talking, each telling a story. From afar, this is heard as an incoherent and incomprehensible cloud of murmur. As one gets closer to an individual agent, two things happen simultaneously: the voice gradually gets clearer and the story becomes more coherent, more sensible. So, the system responds to the visitor on two levels: semantic and audio clarity. As as the visitor moves from agent to agent, she can hear different stories in varying coherence and clearness, depending on her proximity. The visitors are invited to get close and only then can they hear the full story. Their attention is fully drawn to the stories told by the individual voices within this *Cacophonic Choir*. The stories are those of women who survived sexual assault.

The goal of this work is to bring attention to the first hand stories of sexual assault victims, allowing visitors to get close to each, listening to their stories. The use of a robotic voice telling the stories references the anonymity of the internet, and that many stories we hear are filtered through technology. The voice can call into question the authenticity of these stories, reflecting the doubt rape survivors receive. The sighs, breaths and whimpers between the words give the story a more human quality. The juxtaposition of the robotic voice and human vocables and hesitation markers is uncanny, making the listener uncomfortable.

The arrangement of words is less meaningful at a distance and becomes more coherent when an agent is approached. This text generation algorithm expresses the idea that from a distance there are similarities between the stories, expressing the systemic problem of rape culture. Each story and voice becomes more individualized when a visitor moves toward the source, emphasizing that it is a problem affecting individuals. The audio processing of the text represents how one story can be lost in a crowd of stories. It shows how the bombardment of information can desensitize us, so we can no longer comprehend it. The work questions the effectiveness of the internet and technology as a medium for telling these stories. The internet both allows everyone to have a voice but can be noisy, making it hard to find meaningful connections.

In *Cacophonic Choir*, the agents are physical and the visitor can move around the space. This work employs proxemics as an interaction technique and explores how audio quality changes clarity at different distances. It sonifies the idea of information overload, creating a jumbled, overwhelming cloud of sounds that can be sifted through by modifying the distance of agents.

6.1 Design

Cacophonic Choir consists of voices that respond to distance by become clearer (muffled sounds become words) and make more sense (randomized words become accounts of assault). The voices are emitted from speakers coupled with ultrasonic sensors distributed in space. The text is generated by an Long Short Term Memory (LSTM) recurrent neural network, using as a dataset over 500 accounts of sexual assault posted on an online

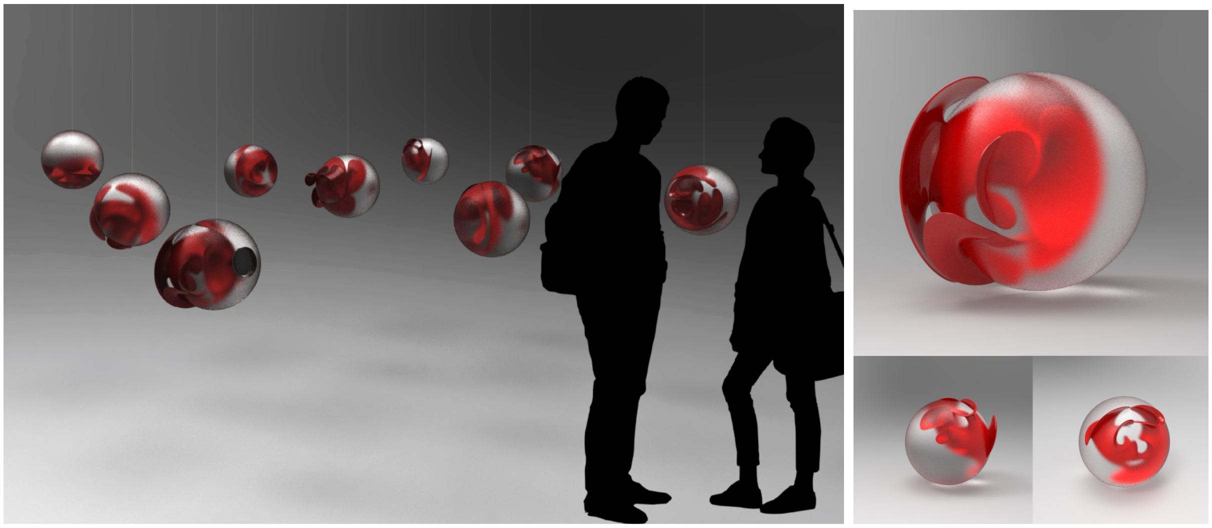

Figure 6.1: Digital rendering of installation (Sölen Kiratli)

forum. A set of neural network model weights, trained to different levels of accuracy, are used to produce more randomized or more coherent words. When an visitor is close enough to an agent, an actual account is heard.

Cacophonic Choir draws from an extensive history of artworks that create an environment, and also from artworks made up of a group of creatures or objects distributed in the space (see Section 2.3.3). Like *Colloquy of Mobiles*, the work uses light to communicate, but in this case, with people, not each other. The work also draws inspiration from *ALAVs (Autonomous Light Air Vessels)* in creating balloon-like agents that participants can interact with. Proxemics has been explored in many media art works with human and animal-like creatures. *Cacophonic Choir* uses proximity to affect the work both visually and sonically, becoming clearer at a closer distance. Works like *HLR Helpless Robot* (1987) by Norman White, *Stupid Robot* (1985) by Simon Penny, and *Nose Wazoo* (1990) by Jim Pallas with Jim Zalewski used proximity as an input for interaction. *Cacophonic Choir* promotes the opposite proxemic interaction from works like *Sixteen Birds* (2008) by Chico MacMurtrie.

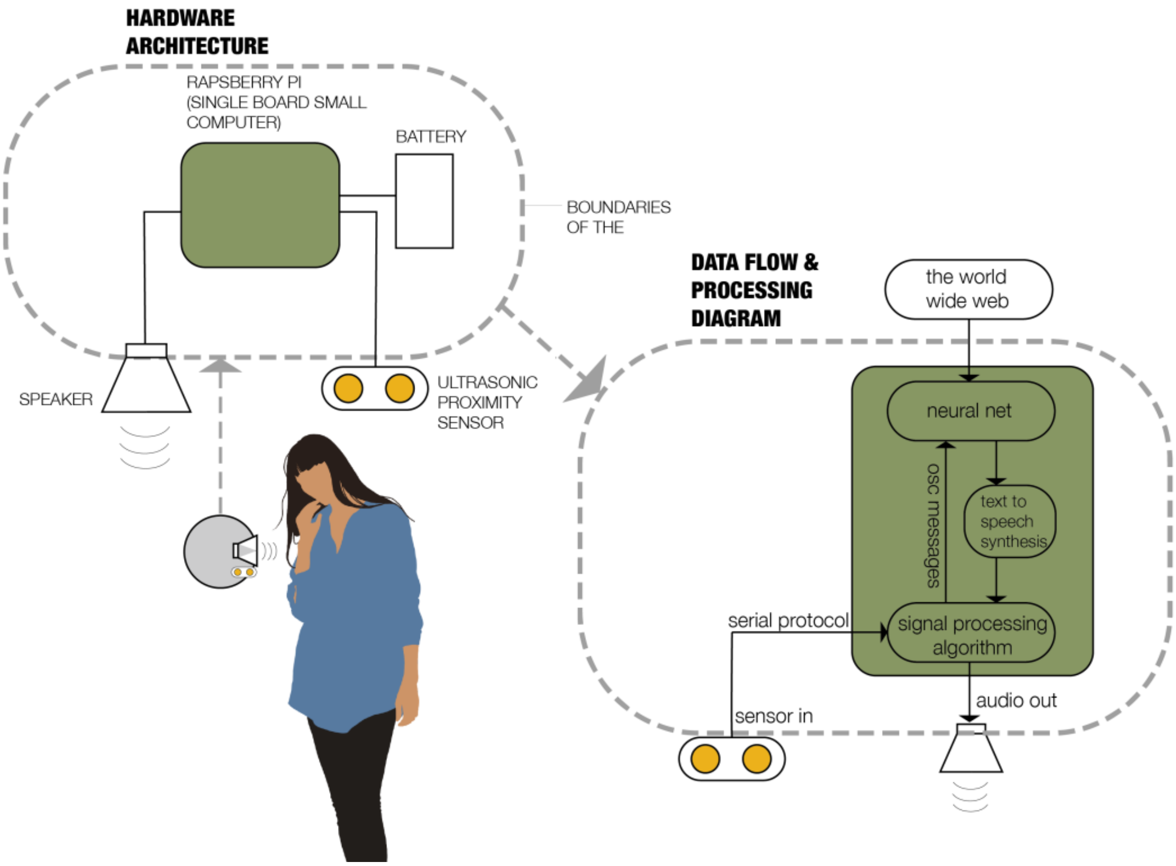

Figure 6.2: System Design Diagram (Sölen Kiratli)

6.1.1 Auditory Design

The auditory system, with distance as an input parameter, chooses the next word, filters the sound and plays it. It is controlled by 2 programs, a Python script which chooses the words and Super Collider which filters and plays the audio files. When the Python script has initialized, it sends the first word to Super Collider based on the distance information. When Super Collider is finished playing that file, it requests the next word from the Python script, which is chosen based on distance.

The Python script has 7 different word models that it can choose from, with the most random word being chosen for the furthest distance and the next original story word being chosen if the participant is very close. The first word model is completely random words from the chosen story. The last word model is choosing the next word from the original story's account. The middle second through sixth word models are based on pre-trained LSTM recurrent neural networks. Each neural network is trained to a different level of accuracy, the 2nd model being the least trained and the 6th model being the most trained. If the next word requires a neural network, the Python script loads the required pre-trained model, and requests the next word based on the previous words. The program gives fewer previous words to the models associated with further distances, and more words to the models associated with closer distances. Super Collider constantly receives the current distance data over Open Sound Control (OSC), and once the current word has finished playing, it receives the next word file to play. Super Collider plays each audio file, filtering it in realtime based on the distance data.

The neural network models were pre-trained on over 500 stories of experiences of sexual assault from the When You're Ready Project website. "The When You're Ready Project is a community for survivors of sexual violence to share their stories and have their voices heard, finding strength in one another" [271]. The stories were scraped from the internet using the Beautiful Soup Python Library, saving the title of the account, the date it was published, the URL, the tags associated with the story and the story itself. The library TextGenRNN, which uses TensorFlow, was used to train the neural network. The neural network model was trained to 199 different epochs and saved at different points along the way. For this installation we used the models trained to 5, 30, 99, 159 and 199 epochs.

6.1.2 Physical Installation Design

Each agent has a Raspberry Pi inside with a proximity sensor to detect distance, an LED to change the luminosity of the agent, and a speaker to play the sounds. The

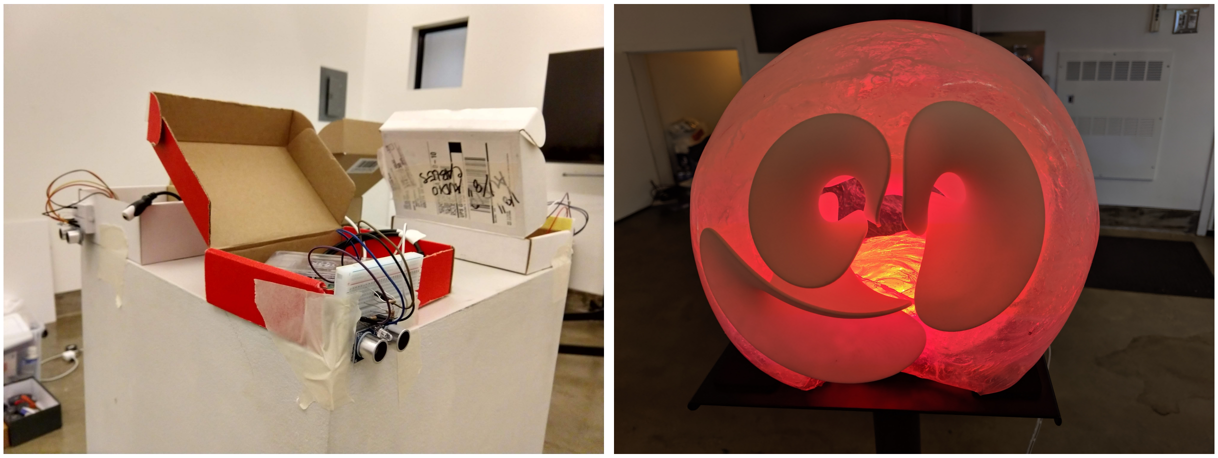

Figure 6.3: Electronics and prototyped embodiment.

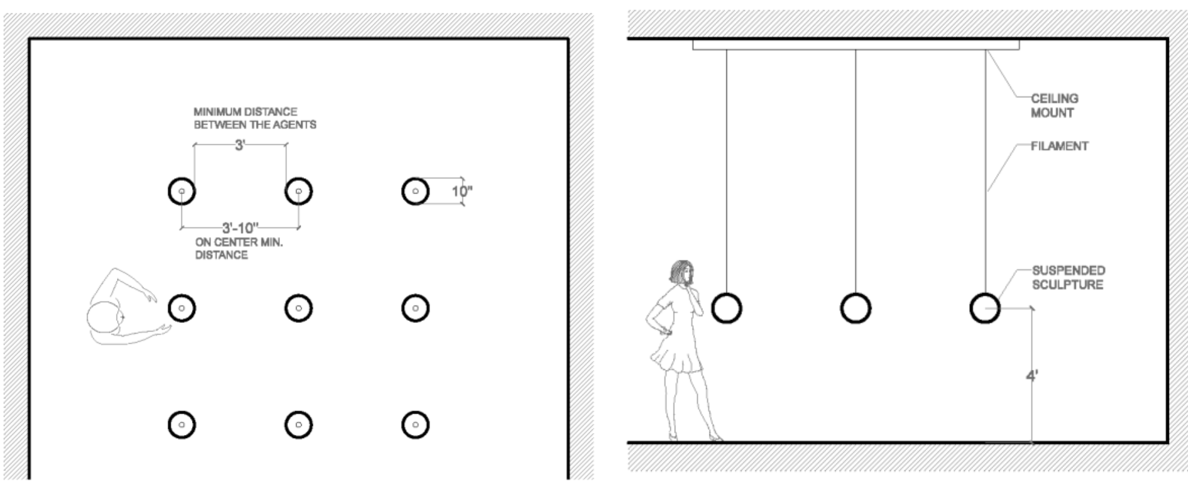

Figure 6.4: Installation Floorplan (Sölen Kiratli)

Raspberry Pi has two programs running, Super Collider and Python. The Python script reads the proximity sensor data and constantly sends it to Super Collider. The LED in the agent glows brighter as the visitor approaches, allowing the 3D printed shape embedded in the sphere to become more visible. The Python script also sets the brightness based on the proximity, with a higher brightness when the visitor is closer.

Figure 6.5: Exhibition of *Cacophonic Choir* at Contemporary Istanbul. Left photograph by Gökhan Tugay Şeker, right photograph by Hannah Wolfe.

6.2 Exhibition

Cacophonic Choir was exhibited at Contemporary Instanbul's 2019 Plugin section which focuses on new media and digital arts, September 12th-15th 2019. It consisted of 9 agents on pedestals 3 feet apart, taking up approximately a 10'x10' area (see Figure 6.5). The work required constant power, with a power cord running up each filament. The work was shown in an exhibition space with other work. This work was also presented at SOUND::GENDER::FEMINISM::ACTIVISM (SGFA 2019) in Tokyo, Japan.

6.3 Discussion

This work uses affective computing and tangible computing to explore technology as a platform to tell the stories of abuse. It employs affective computing through different levels of audio clarity and tangible computing by physically embodying the stories. The agents are responsive to proximity, visually and sonically. In this way the work spans, personal, social and public distance with different experiences at different distances of interaction. In the personal space the visitor can hear the voice clearly and feel the emotion of the story, the object is visually brighter and the visitor can see the intricacies of the shape. The clarity of the voice, story, and object is responsive in the social space, changing as the visitor moves. In the public space the installation becomes an overwhelming murmur and obscured objects. This work probe the limitations of the empowerment interactive public technology through embodiment. The limitations that this work exposes are how overwhelming it can be to sift through the stories from a distance, how technology can filter the stories in unconstructive ways, and how upsetting the stories are individually. In this way, it physically embodies my emotional relationship with this subject and these stories.

Chapter 7

Conclusion

This dissertation explored embodying digital interactions and employing affective computing techniques as tools to probe the social implications of interactive digital technologies. Through embodiment and affective computing it exposed the limitations, strengths and weaknesses of human's interaction with digital technology.

7.1 Tangible Computing

I explored tangible computing in *the Volumetric Viewer*, *Touching Affectivity*, *Come Hither to Me!* and *Cacophonic Choir*. Each of these works used tangible computing differently. I found that tangible interfaces were more intuitive for certain tasks, that a more reactive interface was treated more gently, and that embodying interactions caused people to experience the same agent differently. Last, I learned that tangible computing can create visual metaphors for experiences. In this way, the embodiment and emotion of robotic media arts can be used as a tool to probe the social implications of interactive digital technologies. Tangible computing questioned whether traditional interfaces are the best way to interact with our environment and showed their limitations. It showed that tangible interfaces can change the way we interact with technology, and that we are limiting ourselves with screen-based interfaces. By using tangible computing to embody digital technology, the way people interacted with an abusive agent changed, demonstrating that digital technology can cause people to minimize harmful behavior.

When building interfaces for works in the AlloSphere, I found that digital interfaces had limitations for movement-based tasks. Particularly navigating the camera, moving objects and rotating them in three dimensions, can be difficult in two dimensions. In *the Volumetric Viewer* we used a wearable gloves and built a gestural interface for a digital tablet to navigate the camera, move objects, and rotate planes. The gestural interface for the digital tablet used its gyroscope and accelerometer data to control the rotation of a plane. I used the physical qualities of the tablet to construct a metaphor between it and the rotating plane, which was much more intuitive. By using tangible computing to enhance the gloves and tablet, we could provide usable interfaces. By limiting ourselves to screen based interactions we limit ourselves in the interfaces we can build.

Tangible Computing was used as part of *Touching Affectivity* to allow a stuffed animal to be used as a computational interface. I found that by having the robot respond emotively to touch, people were more gentle with the robot. People did not feel the need to squeeze the robot to get it to make a sound when it was very reactive to touch. I also found that the perceived emotion expressed by a robot can be different due to touch based feedback. During the first exhibition of *Touching Affectivity*, when the robot was squeezed tightly it would vibrate and make a high pitched noise. Visitors touching the robot would perceive it as happy because it was purring, while onlookers thought it was screaming in pain.

When I exhibited *Come Hither to Me!*, participants spent more time interacting privately with the disembodied agent compared to interacting publicly with the embodied agent. By embodying the interaction and bringing it into the public space, people interacted with the agent for a shorter period of time and were more likely to leave directly after being insulted. This showed that people experienced the situation differently and were less complacent when interacting with a physical robot. Finally *Cacophonic Choir* uses embodiment to physically show the limitations of public digital interactions. Because the agent is physically more opaque and the visitor can not seeing the shape unless they are close to the object, it represents how stories exist but are obfuscated by the media.

7.2 Affective Computing

I explored the use of affective computing techniques in *Touching Affectivity*, *Come Hither to Me!* and *Cacophonic Choir*. Each of these works used affective computing differently to probe the social implications of digital interactive technologies. In these works I used a decision tree dialogue structure to interpret different emotional responses and respond correctly. I synthesized emotive non-linguistic utterances to touch, and I used filtering to give voices different emotive vocal qualities.

For the emotive dialogue creation in *Come Hither to Me!*, I employed affective computing to determine the participants' emotive response to compliments and negative comments. I defined the categorically different affective responses that participants might have, so that ROVERita would respond to each participant's response accurately. ROVERita also caused an emotive response in participants, as seen in video recordings of the first exhibition, participants laughed and showed discomfort.

In *Touching Affectivity*, reactive emotive sounds caused people to interact with a robot more gently, compared to less reactive non-emotive sounds. There were two states, in which the robot audibly vibrated, randomly chirped, "screamed" at high amounts of pressure, and responded emotively to stimulation of the fur sensor, and the other state in which the robot did everything but respond to the fur sensor. The fur's auditory response was to make non-linguistic utterances based on prior research in emotive vocal communication and emotive music. In this work, the scratching and tickling gesture created sounds with audio qualities similar to excited/happy speech, and petting and stroking created sounds with audio qualities similar to calm/sad speech. I found that endowing the fur sensor with an immediate reactive response caused people to interact with the robot more gently.

Cacophonic Choir employs affective computing with different levels of audio clarity at different distances. It uses proxemics to its advantage as an interface for interaction. When the participant is in the personal space of an agent, it tells emotive stories of sexual abuse. In the social space, the clarity of the sound changes as the person moves, so that sometimes jarring words are heard through the ambiguity and gibberish created by granulation. At a further distance, the room turns into a murmur of speech, using a low-pass filter to make the individual voices sound muffled. This work sonifies the limitations of the empowerment created by interactive public technology.

7.3 Data Collection and Documentation

Collecting data in exhibition spaces is a complex problem. This can be because of lack of control of the environment and consent requirements for institutional review boards. Having the agent log data during interaction is more consistent and usable. Because ROVERita logged everything that she said and what the text-to-speech response was, I could extract the length of the interaction, when participants left and what they said. The furry robot in *Touching Affectivity* logged and time stamped the pressure and fur signals. This data format was easy to analyze and extract meaningful information.

Video recordings as data collection can be limiting in exhibition environments. Due to video movement and framing, algorithmically extracting emotion data is problematic. Video and audio recordings can be analyzed by humans. With a more intense stimulus like emotive speech more people laughed when interacting with ROVERita compared to ROVER. While the video recording of *Touching Affectivity* from the EoYS did not show any correlation between proximity and sensor response. It did show different ways that people interacted with the robot and their responses. The recorded dialogue and footage of the interactions was useful in understanding how people perceived the robot and the emotions it expressed.

7.4 Future Work

7.4.1 ROVERita Conversation Generation

For future work, ROVERita would employ a second approach for conversations with participants, and I would compare the results to those of the previous conversation-generating algorithm. In two separate performances using the two approaches, the interaction data and participants' responses would be recorded and compared with each other to evaluate the effectiveness of the communication and the quality of interaction with the robot. The current chat algorithm uses a decision tree but the second one would use a Long Short-Term Memory (LSTM) Recurrent Neural Network. The neural network would be trained using the data I collected from prior performances, the Cornell Movie Dialogs Corpus [272], questions from the OKCupid Dataset [273], and other datasets. I would update the decision tree after each performance and continue to train the neural network based on the recently collected data.

7.4.2 *Touching Affectivity* Extended

One extension of the conductive fur research would be to create a therapy robot. The robot could be used to test different kinds of emotional models of movement and sound in a research setting. For example, I would test what emotions are expressed and their intensity in different interaction conditions with different modalities of emotional expression. The robot could be used as an emotional support robot or with children with developmental disorders. Studies could be run to see how interacting with the robot affects the children's emotional expression, interaction with others or ability to focus. The robot could be used to express emotions and help improve motor and social skills.

Different emotional models could be used to extend the emotive qualities of the robot. A dimensional model could be used, mapped in a valence and arousal space. Petting or stroking the robot would increase valence and decrease arousal, making it more content. Scratching or tickling the robot would increase valence and arousal, making it more excited. Picking up the robot would decrease valence and increase arousal, making it more angry. Ignoring the robot would decrease valence and decrease arousal, making it sad. The robot's emotions would be influenced by its heat sensors, capacitive touch sensors and capacitive fur. Through interaction (or inaction) the robot's emotional state would change, which would affect its actions and emotional displays.

An extension of this would be to create a robot could be worn as clothing or kept as a toy. It would wrap itself around a person's wrist and could be worn. As a smart device, the robot, could read skin capacitance and temperature from its sensors, as well as touch gestures. This data would affect the robot's emotions which would be displayed through color and sound. The robot's emotions could be used to communicate emotions to others. The robot would communicate through movement, light and sound. The robot's senses would be redundant and would use sensory motor coordination to learn and adapt to environments. By interacting through emotive sound, color and movement, the robot could interact with humans of all ages and people with sensory impairments. The robot's physical appearance and movement could be influenced by insects: caterpillars, inchworms, centipedes and pillbugs. Its movement in particular could be inspired by inchworms, allowing it to walk around, deal with small obstacles and climb onto people. It would have a parasitic quality, wrapping itself around a person's wrist so it could be carried and worn like expressive, smart clothing.

7.5 Summary

I show that art causes people to think differently about how they interact with technology, that the proximity of the viewer to the artwork and having an audience or others interacting with the same installation changes the experience of the work, and that art can be used to collect data and study people's interactions. In *Come Hither to Me!*, proximity, the presence of an audience, and embodiment changed the way people interacted with ROVERita's AI. In *Touching Affectivity*, perceived emotional expression differed depending on distance and participants experienced the robot's emotive expression differently creating conversation about the interaction and emotions expressed. These conversations changed the way that people interacted with the robot. At the MOXI, there was an issue with performance shyness when exhibiting *the Hydrogen-Like Atom*. One solution was visitors having the ability to interact with the system using their own devices with simplified interfaces spread across multiple tabs. In this way they could interact with the work and others without being in the spotlight.

My works make important strides in the field of Media Arts. ROVERita is novel, as a female robot with agency. In *Come Hither to Me!* and *Touching Affectivity*, I explored the humanization of technology. ROVERita's seeming humanity, as she found people in the space and spoke to them, made them happy with jokes and fliratious comments, in part because people understood how to interact with her. *Come Hither to Me!* explores women's expected roles and agency as the robot speaks for herself and chooses who she wants to interact with.

Both robots in *Come Hither to Me!* and *Touching Affectivity* were given human or animal like qualities and had uncanny interactions. While, ROVERita's agency and directness as a female robot caused discomfort and sometimes confusion, the furry robot was treated like an animal, discussed like it had feelings and felt pain, and was empathized with when punched or squeezed. In *Touching Affectivity*, the emotive haptic and emotive sonic expressions' incongruity caused the person physically interacting with robot and the onlookers to perceive different emotional expression by the robot. Because the robot was not immediately reactive sonically, and because the participant thought that the vibrations were purring, the later sound which was meant to disincentivize squeezing (a high pitched screaming noise) actually incentivized it.

Touching Affectivity emphasizes the importance of haptic engagement in touch-based interactive artworks. If negative interactions are incentivized, a child could learn to interact with real animals in a negative or harmful way. This is crucial because the way a child interacts with a robot dog can predict the way it would interact with a real dog. A child's toy could teach it negative behavior toward animals. Similarly, without proper gender representation and agency, robots can perpetuate gender stereotypes instead of dismantling them.

In these works I use affective and tangible computing techniques to elicit different reactions and make people contemplate their relationship with technology and one another. I show that having a robot respond reactively to touch causes people to interact with it in a more caring way. I demonstrate that by embodying abusive interactions, people are less likely to endure them. I found that digital interfaces in 3D environments had limitations for movement-based tasks. I created a conceptual framework for designing interactive systems and contribute new embodiments of technology and interactive interfaces. I add to the field of Media Arts, these interactive artworks with unique perspectives on human-machine communication.

Appendix A

Appendix Title

A.1 PCA for *Touching Affectivity* data

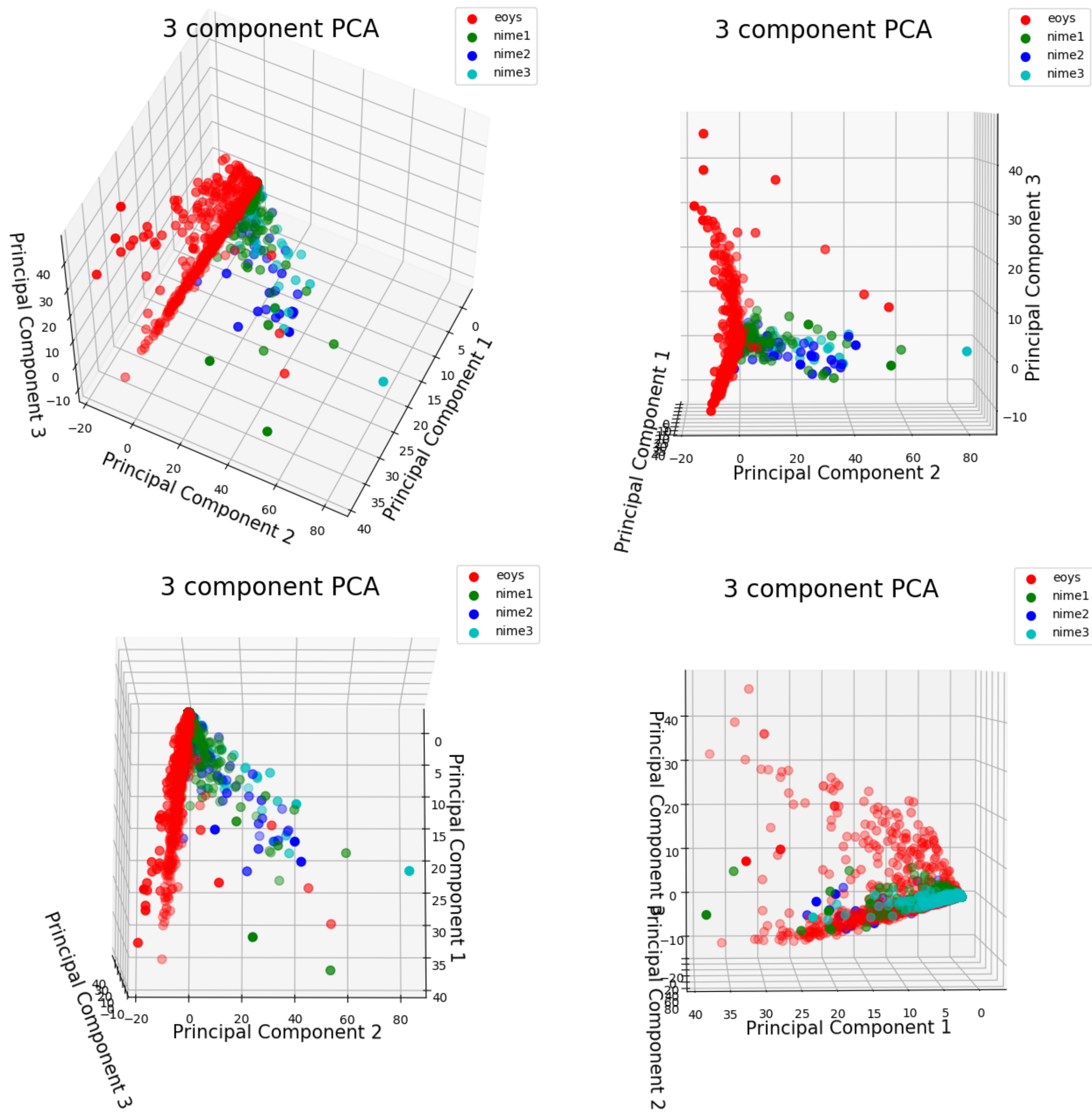

Figure A.1: Principle Component Analysis results of sensor data from *Touching Affectivity*.

Bibliography

[1] P. Dourish, *Where the action is: the foundations of embodied interaction*. MIT press, 2004.

[2] C. Greenberg, *Modernist painting*, *Art and Literature* **4** (1965) 193–201.

[3] E. A. Shanken, *Cybernetics and art: Cultural convergence in the 1960s*, *From energy to information* (2002) 155–177.

[4] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*, vol. 25. MIT press, 1965.

[5] L. Von Bertalanffy, *An outline of general system theory.*, *British Journal for the Philosophy of science* (1950).

[6] S. Penny, *Embodied Interaction*. Oxford University Press, 2 ed., 2014. Accessed: 2019-08-01.

[7] J. J. Gibson, *The theory of affordances*, *Hilldale, USA* **1** (1977) 2.

[8] R. Pfeifer and J. Bongard, *How the body shapes the way we think: a new view of intelligence*. MIT press, 2006.

[9] R. Barthes, *11 the death of the author*, *Media Texts, Authors and Readers: A Reader* (1994) 166.

[10] W. K. Wimsatt and M. C. Beardsley, *The intentional fallacy*, *The Sewanee Review* **54** (1946), no. 3 468–488.

[11] J. Dewey, *Art as experience*. Penguin, 2005.

[12] J. Burnham, *Real time systems*, *Artforum* **8** (1969), no. 1 49–55.

[13] M. McLuhan and Q. Fiore, *The medium is the message*, *New York* **123** (1967) 126–128.

[14] E. Edmonds, G. Turner, and L. Candy, *Approaches to interactive art systems*, in *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pp. 113–117, ACM, 2004.

[15] S. Oviatt, *Mutual disambiguation of recognition errors in a multimodel architecture*, in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 576–583, ACM, 1999.

[16] L. Pagliarini and H. H. Lund, *The development of robot art*, *Artificial Life and Robotics* **13** (2009), no. 2 401–405.

[17] E. Kac, *The origin and development of robotic art*, *Convergence* **7** (2001), no. 1 76–86.

[18] F. Ghedini and M. Bergamasco, *Robotic creatures: Anthropomorphism and interaction in contemporary art.*, in *RO-MAN*, pp. 731–736, 2010.

[19] A. Oulasvirta, T. Rattenbury, L. Ma, and E. Raita, *Habits make smartphone use more pervasive*, *Personal and Ubiquitous Computing* **16** (2012), no. 1 105–114.

[20] H. G. Holmgren and S. M. Coyne, *Can't stop scrolling!: pathological use of social networking sites in emerging adulthood*, *Addiction Research & Theory* **25** (2017), no. 5 375–382.

[21] P. Mitra and K. Baid, *Targeted advertising for online social networks*, in *2009 First International Conference on Networked Digital Technologies*, pp. 366–372, IEEE, 2009.

[22] D. J. Kuss and M. D. Griffiths, *Online social networking and addiction—a review of the psychological literature*, *International journal of environmental research and public health* **8** (2011), no. 9 3528–3552.

[23] H. Schiffrin, A. Edelman, M. Falkenstern, and C. Stewart, *The associations among computer-mediated communication, relationships, and well-being*, *Cyberpsychology, Behavior, and Social Networking* **13** (2010), no. 3 299–306.

[24] J. A. Lueck, *Friend-zone with benefits: The parasocial advertising of kim kardashian*, *Journal of Marketing Communications* **21** (2015), no. 2 91–109.

[25] M. B. Rothberg, A. Arora, J. Hermann, R. Kleppel, P. St Marie, and P. Visintainer, *Phantom vibration syndrome among medical staff: a cross sectional survey*, *Bmj* **341** (2010) c6914.

[26] M. Drouin, D. H. Kaiser, and D. A. Miller, *Phantom vibrations among undergraduates: Prevalence and associated psychological characteristics*, *Computers in Human Behavior* **28** (2012), no. 4 1490–1496.

[27] X. Carbonell, A. Chamarro, U. Oberst, B. Rodrigo, and M. Prades, *Problematic use of the internet and smartphones in university students: 2006–2017*, *International journal of environmental research and public health* **15** (2018), no. 3 475.

[28] A. K. Przybylski and N. Weinstein, *Can you connect with me now? how the presence of mobile communication technology influences face-to-face conversation quality*, *Journal of Social and Personal Relationships* **30** (2013), no. 3 237–246.

[29] K. O'REILLY-JONES, *When fido is family: How landlord-imposed pet bans restrict access to housing*, *Columbia Journal of Law and Social Problems* **52** (2019), no. 3 427–472.

[30] R. Fry and A. Brown, *In a recovering market, homeownership rates are down sharply for blacks, young adults*, *Pew Research Center, December* **15** (2016).

[31] S. Buhai, *Preventing the abuse of service animal regulations*, *NYUJ Legis. & Pub. Pol'y* **19** (2016) 771.

[32] C. Von Bergen, *Emotional support animals, service animals, and pets on campus*, *Administrative Issues Journal* **5** (2015), no. 1 4.

[33] K. M. Lee, Y. Jung, J. Kim, and S. R. Kim, *Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction*, *International Journal of Human-Computer Studies* **64** (2006), no. 10 962–973.

[34] S. Scott, T. Hinton-Smith, V. Härmä, and K. Broome, *Goffman in the gallery: Interactive art and visitor shyness*, *Symbolic Interaction* **36** (2013), no. 4 417–438.

[35] K. Mendes, J. Ringrose, and J. Keller, *# metoo and the promise and pitfalls of challenging rape culture through digital feminist activism*, *European Journal of Women's Studies* **25** (2018), no. 2 236–246.

[36] A. Acar and Y. Muraki, *Twitter for crisis communication: lessons learned from japan's tsunami disaster*, *International Journal of Web Based Communities* **7** (2011), no. 3 392–402.

[37] H. Gao, G. Barbier, and R. Goolsby, *Harnessing the crowdsourcing power of social media for disaster relief*, *IEEE Intelligent Systems* **26** (2011), no. 3 10–14.

[38] D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, *et. al.*, *The science of fake news*, *Science* **359** (2018), no. 6380 1094–1096.

[39] S. Bradshaw and P. Howard, *Troops, trolls and troublemakers: A global inventory of organized social media manipulation*, .

[40] H. Allcott and M. Gentzkow, *Social media and fake news in the 2016 election*, *Journal of economic perspectives* **31** (2017), no. 2 211–36.

[41] B. E. Duffy, *Empowerment through endorsement? polysemic meaning in dove's user-generated advertising*, *Communication, Culture & Critique* **3** (2010), no. 1 26–43.

[42] J. Li, *The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents*, *International Journal of Human-Computer Studies* **77** (2015) 23–37.

[43] T. H. Nelson, *Computer lib/dream machines*, .

[44] C. E. Shannon, *A mathematical theory of communication*, *Bell system technical journal* **27** (1948), no. 3 379–423.

[45] M. McLuhan and M. A. MCLUHAN, *Understanding media: The extensions of man*. MIT press, 1994.

[46] H. M. McLUHAN, *Address at vision 65*, *The American Scholar* (1966) 196–205.

[47] D. Norman, *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[48] A. Dunne and F. Raby, *Speculative everything: design, fiction, and social dreaming*. MIT press, 2013.

[49] C. L. Breazeal, *Designing sociable robots*. MIT press, 2004.

[50] "Louis-philippe demers - v2 lab for the unstable media." `https://v2.nl/archive/people/louis-philippe-demers`. Accessed: 2019-08-01.

[51] S. Oviatt, R. Coulston, and R. Lunsford, *When do we interact multimodally?: cognitive load and multimodal communication patterns*, in *Proceedings of the 6th international conference on Multimodal interfaces*, pp. 129–136, ACM, 2004.

[52] S. Oviatt, *Multimodal interactive maps: Designing for human performance*, *Human-computer interaction* **12** (1997), no. 1 93–129.

[53] S. L. Ellis, V. Swindell, and O. H. Burman, *Human classification of context-related vocalizations emitted by familiar and unfamiliar domestic cats: an exploratory study*, *Anthrozoös* **28** (2015), no. 4 625–634.

[54] M. Saerbeck and C. Bartneck, *Perception of affect elicited by robot motion*, in *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, pp. 53–60, IEEE Press, 2010.

[55] J. E. Opfer and S. A. Gelman, *Development of the animate-inanimate distinction*, *The Wiley-Blackwell Handbook of Childhood Cognitive Development,* (2010) 213–238.

[56] E. T. Hall, *The hidden dimension.* Doubleday & Co, 1966.

[57] H. Huttenrauch and K. S. Eklundh, *Investigating socially interactive robots that give the right cues and make their presence felt*, in *Proceedings of the CHI 2004 Workshop on Shaping Human-Robot Interaction*, pp. 17–20, 2004.

[58] L. Takayama and C. Pantofaru, *Influences on proxemic behaviors in human-robot interaction*, in *Intelligent robots and systems, 2009. IROS 2009. IEEE/RSJ international conference on*, pp. 5495–5502, IEEE, 2009.

[59] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, A. R. Diebold Jr, M. Durbin, M. S. Edmonson, J. Fischer, D. Hymes, S. T. Kimball, *et. al.*, *Proxemics [and comments and replies]*, *Current anthropology* **9** (1968), no. 2/3 83–108.

[60] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis, *Equilibrium theory revisited: Mutual gaze and personal space in virtual environments*, *Presence: Teleoperators & Virtual Environments* **10** (2001), no. 6 583–598.

[61] L. A. Hayduk, *Personal space: Where we now stand.*, *Psychological bulletin* **94** (1983), no. 2 293.

[62] D. J. Holt, B. S. Cassidy, X. Yue, S. L. Rauch, E. A. Boeke, S. Nasr, R. B. Tootell, and G. Coombs, *Neural correlates of personal space intrusion*, *Journal of Neuroscience* **34** (2014), no. 12 4123–4134.

[63] J. Stark, R. R. Mota, and E. Sharlin, *Personal space intrusion in human-robot collaboration*, in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 245–246, ACM, 2018.

[64] N. Katzakis and M. Hori, *Mobile devices as multi-dof controllers*, in *3D User Interfaces (3DUI), 2010 IEEE Symposium on*, pp. 139–140, IEEE, 2010.

[65] R. Aspin and K. H. Le, *Augmenting the cave: An initial study into close focused, inward looking, exploration in ipt systems*, in *Proceedings of the 11th IEEE International Symposium on Distributed Simulation and Real-Time Applications*, pp. 217–224, IEEE Computer Society, 2007.

[66] M. Prachyabrued, D. L. Ducrest, and C. W. Borst, *Handymap: a selection interface for cluttered vr environments using a tracked hand-held touch device*, in *International Symposium on Visual Computing*, pp. 45–54, Springer, 2011.

[67] R. Burgess, A. J. Falcão, T. Fernandes, R. A. Ribeiro, M. Gomes, A. Krone-Martins, and A. M. de Almeida, *Selection of large-scale 3d point cloud data using gesture recognition*, in *Doctoral Conference on Computing, Electrical and Industrial Systems*, pp. 188–195, Springer, 2015.

[68] D. Coffey, N. Malbraaten, T. Le, I. Borazjani, F. Sotiropoulos, and D. F. Keefe, *Slice wim: a multi-surface, multi-touch interface for overview+ detail exploration of volume datasets in virtual reality*, in *Symposium on Interactive 3D Graphics and Games*, pp. 191–198, ACM, 2011.

[69] B. Reeves and C. Nass, *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press, 1996.

[70] F. Eyssel, D. Kuchenbrandt, and S. Bobinger, *Effects of anticipated human-robot interaction and predictability of robot behavior on perceptions of anthropomorphism*, in *Proceedings of the 6th international conference on Human-robot interaction*, pp. 61–68, ACM, 2011.

[71] M. Mori, K. F. MacDorman, and N. Kageki, *The uncanny valley [from the field]*, *Robotics & Automation Magazine, IEEE* **19** (2012), no. 2 98–100.

[72] C. Bartneck, *EMuu: An Embodied Emotional Character for the Ambient Intelligent Home*. Technische Universiteit Eindhoven, 2002.

[73] I. Rae, L. Takayama, and B. Mutlu, *The influence of height in robot-mediated communication*, in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, pp. 1–8, IEEE Press, 2013.

[74] S. H. Seo, D. Geiskkovitch, M. Nakane, C. King, and J. E. Young, *Poor thing! would you feel sorry for a simulated robot?: A comparison of empathy toward a physical and a simulated robot*, in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pp. 125–132, ACM, 2015.

[75] C. Kidd and C. Breazeal, *Comparison of social presence in robots and animated characters*, *Interaction Journal Studies* (2005).

[76] S. Kiesler, A. Powers, S. R. Fussell, and C. Torrey, *Anthropomorphic interactions with a robot and robot-like agent*, *Social Cognition* **26** (2008), no. 2 169–181.

[77] J. Wainer, D. J. Feil-Seifer, D. Shell, M. J. Matarić, *et. al.*, *The role of physical embodiment in human-robot interaction*, in *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*, pp. 117–122, IEEE, 2006.

[78] C. Breazeal, *Emotion and sociable humanoid robots*, *International Journal of Human-Computer Studies* **59** (2003), no. 1 119–155.

[79] L. D. Canamero and J. Fredslund, *How does it feel? emotional interaction with a humanoid lego robot*, in *Proc. of American association for artificial intelligence fall symposium, FS-00-04*, pp. 7–16, 2000.

[80] A. P. Saygin, T. Chaminade, H. Ishiguro, J. Driver, and C. Frith, *The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions*, *Social cognitive and affective neuroscience* **7** (2011), no. 4 413–422.

[81] M. Clasen, *The anatomy of the zombie: A bio-psychological look at the undead other*, *Otherness: Essays and Studies* **1** (2010), no. 1 1–23.

[82] F. Ferrari, M. P. Paladino, and J. Jetten, *Blurring human–machine distinctions: anthropomorphic appearance in social robots as a threat to human distinctiveness*, *International Journal of Social Robotics* **8** (2016), no. 2 287–302.

[83] J.-Y. Sung, L. Guo, R. E. Grinter, and H. I. Christensen, *"my roomba is rambo": intimate home appliances*, in *International Conference on Ubiquitous Computing*, pp. 145–162, Springer, 2007.

[84] L. Howard and S.-J. Vick, *Does it bite? the role of stimuli characteristics on preschoolers' interactions with robots, insects and a dog*, *Anthrozoös* **23** (2010), no. 4 397–413.

[85] F. N. Ribi, A. Yokoyama, and D. C. Turner, *Comparison of children's behavior toward sony's robotic dog aibo and a real dog: A pilot study*, *Anthrozoös* **21** (2008), no. 3 245–256.

[86] M. Díaz Boladeras, *Bonding with robotic pets. children's cognitions, emotions and behaviors towards pet-robots. applications in a robot assisted quality of life intervention in a pediatric hospital*, .

[87] D. Feil-Seifer and M. J. Matarić, *Socially assistive robotics*, *Robotics & Automation Magazine, IEEE* **18** (2011), no. 1 24–31.

[88] Y.-H. Wu, C. Fassert, and A.-S. Rigaud, *Designing robots for the elderly: appearance issue and beyond*, *Archives of gerontology and geriatrics* **54** (2012), no. 1 121–126.

[89] T. Shibata and K. Wada, *Robot therapy: a new approach for mental healthcare of the elderly–a mini-review*, *Gerontology* **57** (2011), no. 4 378–386.

[90] J. L. Krichmar and T.-S. Chou, *A tactile robot for developmental disorder therapy*, in *Proceedings of the Technology, Mind, and Society*, p. 20, ACM, 2018.

[91] R. Burns and K. J. Kuchenbecker, *Designing a haptic empathetic robot animal for children with autism*, in *RSS 2018 Workshop on Robot Mediated autism Intervention: Hardware, Software and Curriculum*, 2018.

[92] A. Duquette, F. Michaud, and H. Mercier, *Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism*, *Autonomous Robots* **24** (2008), no. 2 147–157.

[93] E. S. Kim, L. D. Berkovits, E. P. Bernier, D. Leyzberg, F. Shic, R. Paul, and B. Scassellati, *Social robots as embedded reinforcers of social behavior in children with autism*, *Journal of autism and developmental disorders* **43** (2013), no. 5 1038–1049.

[94] C. M. Stanton, P. H. Kahn, R. L. Severson, J. H. Ruckert, and B. T. Gill, *Robotic animals might aid in the social development of children with autism*, in *2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 271–278, IEEE, 2008.

[95] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, and K. Fujimura, *The intelligent asimo: System overview and integration*, in *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, vol. 3, pp. 2478–2483, IEEE, 2002.

[96] M. Raibert, K. Blankespoor, G. Nelson, R. Playter, and T. Team, *Bigdog, the rough-terrain quadruped robot*, in *Proceedings of the 17th World Congress*, vol. 17, pp. 10822–10825, 2008.

[97] P. Sapaty, *Military robotics: Latest trends and spatial grasp solutions*, *International Journal of Advanced Research in Artificial Intelligence* **4** (2015), no. 4.

[98] M. Fischer, *Aquapenguin, a biomechatronic overall concept*, *Festo AG* (2009).

[99] A. Festo and K. Co, *Bionic learning network*, *http://www.festo.com/cms/de_de/4981.htm* (2009).

[100] W. G. Walter, *An imitation of life*, .

[101] K. G. Pontus Hultén, *The machine, as seen at the end of the mechanical age*, *The Museum of Modern Art, New York* (1968).

[102] A. Zivanovic, *Sam, the senster and the bandit. early cybernetic sculptures by edward ihnatowicz*, in *Proceedings of the Symposium on robotics, mechatronics and animatronics in the creative and entertainment industries and arts*, pp. 1–7, Citeseer, 2005.

[103] J. Reichardt, *Cybernetic serendipity: the computer and the arts.* Praeger, 1969.

[104] E. La Prade, *The early days of eat*, *IEEE MultiMedia* **9** (2002), no. 2 4–5.

[105] K. Rinaldo, "Ken rinaldo." `http://www.kenrinaldo.com/`, 2015. Accessed: 2016-03-17.

[106] G. Hertz, "Cockroach controlled mobile robot." `http://www.conceptlab.com/roachbot/`, 2008. Accessed: 2016-03-17.

[107] E. Kac, *Foundation and development of robotic art*, *Art Journal* **56** (1997), no. 3 60–67.

[108] J. G. Hanhardt *et. al.*, *The Worlds of Nam June Paik.* Guggenheim Museum New York, 2000.

[109] S. Woolf and T. Bech, *Experiments with reactive robotic sound sculptures.*, in *ALife VIII: Workshop proceedings 2002 P2*, vol. 3, 2002.

[110] A. Smith, "Interactive, disturbing sculptures of mireia donat melús." `https://hifructose.com/2018/08/14/interactive-disturbing-sculptures-of-mireia-donat-melus/`, 2018. Accessed: 2019-08-01.

[111] F. Orellana, "Elevator's music." `http://fernandoorellana.com/projects/elevators-music/`, 2007. Accessed: 2016-03-17.

[112] J. Berk, "Alavs." `http://www.alavs.com/`, 2009. Accessed: 2016-03-17.

[113] C. MacMurtrie, "Chico macmurtrie: Amorphic robot works." `http://amorphicrobotworks.org/`, 2018. Accessed: 2019-08-01.

[114] J. Davis, "Desert crawler, 1986." `http://act.mit.edu/cavs/item/cavs_chronology_1983_1994_0475`, 1986. Accessed: 2019-08-01.

[115] T. Jansen, *Strandbeests*, *Architectural Design* **78** (2008), no. 4 22–27.

[116] G. Esparza, "Urban parasites." `http://www.parasitosurbanos.com/`, 2007. Accessed: 2016-03-17.

[117] M. Snow, *La région centrale*, *Cinema Canada* (1972).

[118] R. D'Andrea, "Robotic chair." `https://raffaello.name/projects/robotic-chair/`, 2011. Accessed: 2019-08-01.

[119] A. Y. Takeshi Oozu and H. Iwata, *Escaping chair*, in *ISEA2017 Manizales BIO-CREATION AND PEACE, Proceedings of the 23rd International Symposium on Electronic Arts*, Department Of Visual Design, Universidad De Caldas, And ISEA International, 2017.

[120] G. Bekey and J. Yuh, *The status of robotics*, *Robotics & Automation Magazine, IEEE* **15** (2008), no. 1 80–86.

[121] T. Bell, *Robots in the home: Promises, promises: While great expectations are held for certain robot types, the robots for fun and educational purposes are limited in their adaptability to useful tasks*, *Spectrum, IEEE* **22** (1985), no. 5 51–55.

[122] J. Evans, B. Krishnamurthy, W. Pong, R. Croston, C. Weiman, and G. Engelberger, *Helpmate^{TM}: A robotic materials transport system*, *Robotics and Autonomous Systems* **5** (1989), no. 3 251–256.

[123] L. Royakkers and R. van Est, *A literature review on new robotics: automation from love to war*, *International journal of social robotics* **7** (2015), no. 5 549–570.

[124] S. Pieskä, M. Luimula, J. Jauhiainen, and V. Spiz, *Social service robots in public and private environments*, *Recent Researches in Circuits, Systems, Multimedia and Automatic Control* (2012) 190–196.

[125] S. B. Kaufman and M. Van Ellin, *Hanc: A case study in invention*, *Ageing International* **22** (1995), no. 1 26–28.

[126] M. E. Pollack, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan, S. Engberg, J. T. Matthews, J. Dunbar-Jacob, C. E. McCarthy, *et. al.*, *Pearl: A mobile robotic assistant for the elderly*, in *AAAI workshop on automation as eldercare*, vol. 2002, pp. 85–91, 2002.

[127] S. Bahadori, A. Cesta, G. Grisetti, L. Iocchi, R. Leone, D. Nardi, A. Oddi, F. Pecora, and R. Rasconi, *Robocare: an integrated robotic system for the domestic care of the elderly*, in *Proceedings of Workshop on Ambient Intelligence AI* IA-03, Pisa, Italy*, Citeseer, 2003.

[128] B. Graf, M. Hans, and R. D. Schraft, *Care-o-bot ii—development of a next generation robotic home assistant*, *Autonomous robots* **16** (2004), no. 2 193–205.

[129] Y. Wang, K. P. Laby, C. S. Jordan, S. E. Butner, and J. Southard, *Medical tele-robotic system*, Aug. 2, 2005. US Patent 6,925,357.

[130] K. M. Tsui, M. Desai, H. A. Yanco, and C. Uhlik, *Exploring use cases for telepresence robots*, in *Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on*, pp. 11–18, IEEE, 2011.

[131] A. Kristoffersson, S. Coradeschi, and A. Loutfi, *A review of mobile robotic telepresence*, *Advances in Human-Computer Interaction* **2013** (2013) 3.

[132] B. Hine, C. Stoker, M. Sims, D. Rasmussen, P. Hontalas, T. Fong, J. Steele, D. Barch, D. Andersen, E. Miles, *et. al.*, *The application of telepresence and virtual reality to subsea exploration*, in *Second Workshop on Mobile Robots for Subsea Environments*, 1994.

[133] C. Corbett *et. al.*, *Live from mars*, *Monthly, The* (2012), no. Sept 2012 14.

[134] F. Tanaka, K. Isshiki, F. Takahashi, M. Uekusa, R. Sei, and K. Hayashi, *Pepper learns together with children: Development of an educational application*, in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pp. 270–275, IEEE, 2015.

[135] J. A. Greer, *Building emotional authenticity between humans and robots*, in *Workshops in ICSR 2014*, 2014.

[136] P. Rane, V. Mhatre, and L. Kurup, *Study of a home robot: Jibo*, in *International Journal of Engineering Research and Technology*, vol. 3, ESRSA Publications, 2014.

[137] *Electronic Design*. No. v. 14, nos. 1. Hayden Publishing Company, 1966.

[138] Y. Doyon, *La mutation du minotaure: Marcel-lí antúnez roca à la 10e édition du mois multi*, *Inter: Art actuel* (2009), no. 103 44–47.

[139] E. Kac, *Towards a chronology of robotic art*, *Convergence* **7** (2001), no. 1 87–111.

[140] S. Penny, "Simon penny." `http://simonpenny.net/`, 2011. Accessed: 2016-03-17.

[141] J. Pallas, *Nose wazoo*, *Leonardo* **24** (1991), no. 1 83–83.

[142] D. Pangburn, "This dapper robot is an art critic." `http://thecreatorsproject.vice.com/blog/robot-art-critic-berenson`, 2016. Accessed: 2016-03-17.

[143] P. Beesley, "Philip beesley architect inc.." `http://philipbeesleyarchitect.com/`, 2019. Accessed: 2019-08-01.

[144] S. Dixon, *Digital performance: a history of new media in theater, dance, performance art, and installation*. MIT press, 2007.

[145] P. Atzori and K. Woolford, *Extended-body: Interview with stelarc*, *CTheory* (2015) 9–6.

[146] Cod.Act, "Pendulum choir." `http://codact.ch/gb/pendugb.html`, 2011. Accessed: 2016-03-17.

[147] M. Scheutz, P. Schermerhorn, J. Kramer, and D. Anderson, *First steps toward natural human-like hri*, *Autonomous Robots* **22** (2007), no. 4 411–423.

[148] R. G. Read and T. Belpaeme, *Interpreting non-linguistic utterances by robots: studying the influence of physical appearance*, in *Proceedings of the 3rd international workshop on Affective interaction in natural environments*, pp. 65–70, ACM, 2010.

[149] K. Dautenhahn, S. Woods, C. Kaouri, M. L. Walters, K. L. Koay, and I. Werry, *What is a robot companion-friend, assistant or butler?*, in *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*, pp. 1192–1197, IEEE, 2005.

[150] Z. Khan, *Attitudes towards intelligent service robots*, *NADA KTH, Stockholm* **17** (1998).

[151] C. Ray, F. Mondada, and R. Siegwart, *What do people expect from robots?*, in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 3816–3821, IEEE, 2008.

[152] R. Read, *A study of non-linguistic utterances for social human-robot interaction.* Plymouth University, 2014.

[153] J. Beck and T. Grajeda, *Lowering the boom: critical studies in film sound.* University of Illinois Press, 2008.

[154] J.-P. Stein and P. Ohler, *Venturing into the uncanny valley of mind—the influence of mind attribution on the acceptance of human-like characters in a virtual reality setting*, *Cognition* **160** (2017) 43–50.

[155] C. Bartneck, *Robots in the theatre and the media*, in *Proceedings of the Design & Semantics of Form & Movement (DeSForM2013)*, pp. 64–70, Philips Design, 2013.

[156] M.-L. Bourguet, *Designing and prototyping multimodal commands.*, in *INTERACT*, vol. 3, pp. 717–720, Citeseer, 2003.

[157] H. Wolfe, M. Peljhan, and Y. Visell, *Singing robots: How embodiment affects emotional responses to non-linguistic utterances*, *IEEE Transactions on Affective Computing* (2017).

[158] S. S. Horowitz, *The universal sense: How hearing shapes the mind.* Bloomsbury Publishing USA, 2012.

[159] S. S. Horowitz, *The science and art of listening*, *New York Times* **9** (2012).

[160] R. Banse and K. R. Scherer, *Acoustic profiles in vocal emotion expression.*, *Journal of personality and social psychology* **70** (1996), no. 3 614.

[161] V. C. Tartter, *Happy talk: Perceptual and acoustic effects of smiling on speech*, *Perception & psychophysics* **27** (1980), no. 1 24–27.

[162] J. Ohala, *The frequency code underlies the sound symbolic use of voice pitch*, *Sound Symbolism* (1996) 325–347.

[163] J. J. Ohala, *The acoustic origin of the smile*, *The Journal of the Acoustical Society of America* **68** (1980), no. S1 S33–S33.

[164] J. J. Ohala, L. Hinton, and J. Nichols, *Sound symbolism*, in *Proc. 4th Seoul International Conference on Linguistics [SICOL]*, pp. 98–103, 1997.

[165] C. E. Williams and K. N. Stevens, *Emotions and speech: Some acoustical correlates*, *The Journal of the Acoustical Society of America* **52** (1972), no. 4B 1238–1250.

[166] C. Sobin and M. Alpert, *Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy*, *Journal of psycholinguistic research* **28** (1999), no. 4 347–365.

[167] T. Johnstone, C. M. Van Reekum, and K. R. Scherer, *Vocal expression correlates of appraisal processes*, *Appraisal processes in emotion: Theory, methods, research* (2001) 271–284.

[168] K. R. Scherer, T. Johnstone, and G. Klasmeyer, *Vocal expression of emotion*, *Handbook of affective sciences* (2003) 433–456.

[169] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, *Emotion recognition in human-computer interaction*, *Signal Processing Magazine, IEEE* **18** (2001), no. 1 32–80.

[170] L. A. Streeter, N. H. Macdonald, W. Apple, R. M. Krauss, and K. M. Galotti, *Acoustic and perceptual indicators of emotional stress*, *Journal of the Acoustical Society of America* **73** (1983), no. 4 1354–t360.

[171] J. J. Ohala, *Cross-language use of pitch: an ethological view*, *Phonetica* **40** (1983), no. 1 1–18.

[172] K. R. Scherer, H. London, and J. J. Wolf, *The voice of confidence: Paralinguistic cues and audience evaluation*, *Journal of Research in Personality* **7** (1973), no. 1 31–44.

[173] P. Lieberman and S. B. Michaels, *Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech*, *The Journal of the Acoustical Society of America* **34** (1962), no. 7 922–927.

[174] D. Huron, N. Anderson, and D. Shanahan, *"you can't play a sad song on the banjo:" acoustic factors in the judgment of instrument capacity to convey sadness*, *Empirical Musicology Review* **9** (2014), no. 1 29–41.

[175] W. Apple, L. A. Streeter, and R. M. Krauss, *Effects of pitch and speech rate on personal attributions.*, *Journal of Personality and Social Psychology* **37** (1979), no. 5 715.

[176] D. Huron, D. Kinney, and K. Precoda, *Influence of pitch height on the perception of submissiveness and threat in musical passages*, .

[177] I. R. Murray and J. L. Arnott, *Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion*, *The Journal of the Acoustical Society of America* **93** (1993), no. 2 1097–1108.

[178] D. Huron, *A comparison of average pitch height and interval size in major-and minor-key themes: Evidence consistent with affect-related pitch prosody*, .

[179] C. Breitenstein, D. V. Lancker, and I. Daum, *The contribution of speech rate and pitch variation to the perception of vocal emotions in a german and an american sample*, *Cognition & Emotion* **15** (2001), no. 1 57–79.

[180] M. Schutz, D. Huron, K. Keeton, and G. Loewer, *The happy xylophone: acoustics affordances restrict an emotional palate*, .

[181] B. Turner and D. Huron, *A comparison of dynamics in major-and minor-key works*, .

[182] S. Dalla Bella, I. Peretz, L. Rousseau, and N. Gosselin, *A developmental study of the affective value of tempo and mode in music*, *Cognition* **80** (2001), no. 3 B1–B10.

[183] O. Post and D. Huron, *Western classical music in the minor mode is slower (except in the romantic period)*, .

[184] A. W. Siegman and S. Boyle, *Voices of fear and anxiety and sadness and depression: the effects of speech rate and loudness on fear and anxiety and sadness and depression.*, *Journal of Abnormal Psychology* **102** (1993), no. 3 430.

[185] J. C. Hailstone, R. Omar, S. M. Henley, C. Frost, M. G. Kenward, and J. D. Warren, *It's not what you play, it's how you play it: Timbre affects perception of emotion in music*, *The quarterly Journal of Experimental psychology* **62** (2009), no. 11 2141–2155.

[186] S. Le Groux and P. Verschure, *Emotional responses to the perceptual dimensions of timbre: A pilot study using physically informed sound synthesis*, in *Proc. 7th Int. Symp. Comput. Music Model*, pp. 1–15, 2010.

[187] E.-S. Jee, C. H. Kim, S.-Y. Park, and K.-W. Lee, *Composition of musical sound expressing an emotion of robot based on musical factors*, in *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pp. 637–641, IEEE, 2007.

[188] E.-S. Jee, S.-Y. Park, C. H. Kim, and H. Kobayashi, *Composition of musical sound to express robot's emotion with intensity and synchronized expression with robot's behavior*, in *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pp. 369–374, IEEE, 2009.

[189] H. Leichtentritt, *Mechanical music in olden times*, *The Musical Quarterly* **20** (1934), no. 1 15–26.

[190] G. Nierhaus, *Algorithmic composition: paradigms of automated music generation.* Springer Science & Business Media, 2009.

[191] D. Cope, *Computer models of musical creativity.* MIT Press Cambridge, 2005.

[192] P. Ablinger, "Quadraturen." `http://ablinger.mur.at/docu11.html`, 2006. Accessed: 2016-03-17.

[193] J. C. . G. B. Miller, "Janet cardiff & george bures miller." `https://www.cardiffmiller.com/`, 2018. Accessed: 2019-08-01.

[194] D. Rokeby, "David rokeby." `http://www.davidrokeby.com/`, 2007. Accessed: 2019-08-01.

[195] W. Kempelen, H. E. Brekle, and W. Wildgen, *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine.* Friedrich Frommann Verlag, 1970.

[196] H. Dudley, R. Riesz, and S. Watkins, *A synthetic speaker*, *Journal of the Franklin Institute* **227** (1939), no. 6 739–764.

[197] K. Fukui, K. Nishikawa, S. Ikeo, M. Honda, and A. Takanishi, *Development of a human-like sensory feedback mechanism for an anthropomorphic talking robot*, in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pp. 101–106, IEEE, 2006.

[198] H. Sawada, *Talking robot and the autonomous acquisition of vocalization and singing skill.* INTECH Open Access Publisher, 2007.

[199] J. O. Smith, *Physical Audio Signal Processing.* `http://ccrma.stanford.edu/~jos/pasp/`, 2010.

[200] D. Leyzberg, E. Avrunin, J. Liu, and B. Scassellati, *Robots that express emotion elicit better human teaching*, in *Proceedings of the 6th international conference on Human-robot interaction*, pp. 347–354, ACM, 2011.

[201] J. E. Cahn, *The generation of affect in synthesized speech*, *Journal of the American Voice I/O Society* **8** (1990) 1–19.

[202] I. R. Nourbakhsh, J. Bobenage, S. Grange, R. Lutz, R. Meyer, and A. Soto, *An affective mobile robot educator with a full-time job*, *Artificial Intelligence* **114** (1999), no. 1 95–124.

[203] S. Roehling, B. MacDonald, and C. Watson, *Towards expressive speech synthesis in english on a robotic platform*, in *Proceedings of the Australasian International Conference on Speech Science and Technology*, pp. 130–135, 2006.

[204] X. Li, B. MacDonald, and C. I. Watson, *Expressive facial speech synthesis on a robotic platform*, in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pp. 5009–5014, IEEE, 2009.

[205] A. Niculescu, B. van Dijk, A. Nijholt, H. Li, and S. L. See, *Making social robots more attractive: the effects of voice pitch, humor and empathy*, *International journal of social robotics* **5** (2013), no. 2 171–191.

[206] O. Pierre-Yves, *The production and recognition of emotions in speech: features and algorithms*, *International Journal of Human-Computer Studies* **59** (2003), no. 1 157–183.

[207] S. Yilmazyildiz, W. Mattheyses, Y. Patsis, and W. Verhelst, *Expressive speech recognition and synthesis as enabling technologies for affective robot-child communication*, in *Advances in Multimedia Information Processing-PCM 2006*, pp. 1–8. Springer, 2006.

[208] S. Yilmazyildiz, L. Latacz, W. Mattheyses, and W. Verhelst, *Expressive gibberish speech synthesis for affective human-computer interaction*, in *Text, Speech and Dialogue*, pp. 584–590, Springer, 2010.

[209] S. Yilmazyildiz, D. Henderickx, B. Vanderborght, W. Verhelst, E. Soetens, and D. Lefeber, *Emogib: emotional gibberish speech database for affective human-robot interaction*, in *Affective Computing and Intelligent Interaction*, pp. 163–172. Springer, 2011.

[210] R. Read and T. Belpaeme, *How to use non-linguistic utterances to convey emotion in child-robot interaction*, in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pp. 219–220, ACM, 2012.

[211] W. W. Gaver, *Auditory icons: Using sound in computer interfaces*, *Human-computer interaction* **2** (1986), no. 2 167–177.

[212] T. Hermann, J. M. Drees, and H. Ritter, *Broadcasting auditory weather reports-a pilot project*, *Georgia Institute of Technology* (2003).

[213] P. Larsson, *Tools for designing emotional auditory driver-vehicle interfaces*, in *Auditory Display*, pp. 1–11. Springer, 2010.

[214] R. Read and T. Belpaeme, *People interpret robotic non-linguistic utterances categorically*, in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, pp. 209–210, IEEE Press, 2013.

[215] M. Scheeff, J. Pinto, K. Rahardja, S. Snibbe, and R. Tow, *Experiences with sparky, a social robot*, in *Socially Intelligent Agents*, pp. 173–180. Springer, 2002.

[216] C. Bartneck, *Affective expressions of machines*, *emotion* **8** (2000) 489–502.

[217] H. Kozima, M. P. Michalowski, and C. Nakagawa, *Keepon*, *International Journal of Social Robotics* **1** (2009), no. 1 3–18.

[218] E.-S. Jee, Y.-J. Jeong, C. H. Kim, and H. Kobayashi, *Sound design for emotion and intention expression of socially interactive robots*, *Intelligent Service Robotics* **3** (2010), no. 3 199–206.

[219] D.-S. Kwon, Y. K. Kwak, J. C. Park, M. J. Chung, E.-S. Jee, K.-S. Park, H.-R. Kim, Y.-M. Kim, J.-C. Park, E. H. Kim, *et. al.*, *Emotion interaction system for a service robot*, in *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pp. 351–356, IEEE, 2007.

[220] I. Wallis, T. Ingalls, E. Campana, and J. Goodman, *A rule-based generative music system controlled by desired valence and arousal*, in *Proceedings of 8th international sound and music computing conference (SMC)*, pp. 156–157, 2011.

[221] T. Komatsu, S. Yamada, K. Kobayashi, K. Funakoshi, and M. Nakano, *Artificial subtle expressions: intuitive notification methodology of artifacts*, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1941–1944, ACM, 2010.

[222] T. Komatsu, *Toward making humans empathize with artificial agents by means of subtle expressions*, in *Affective Computing and Intelligent Interaction*, pp. 458–465. Springer, 2005.

[223] T. Komatsu and S. Yamada, *How does the agents' appearance affect users' interpretation of the agents' attitudes: Experimental investigation on expressing the same artificial sounds from agents with different appearances*, *Intl. Journal of Human–Computer Interaction* **27** (2011), no. 3 260–279.

[224] K. Allan, *The component functions of the high rise terminal contour in australian declarative sentences*, *Australian Journal of Linguistics* **4** (1984), no. 1 19–32.

[225] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, *Wavenet: A generative model for raw audio*, *arXiv preprint arXiv:1609.03499* (2016).

[226] C. Veaux, J. Yamagishi, K. MacDonald, *et. al.*, *Superseded-cstr vctk corpus: English multi-speaker corpus for cstr voice cloning toolkit*, .

[227] S. Dixon, *Metal performance humanizing robots, returning to nature, and camping about*, *TDR/The Drama Review* **48** (2004), no. 4 15–46.

[228] M. Siegel, C. Breazeal, and M. I. Norton, *Persuasive robotics: The influence of robot gender on human behavior*, in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2563–2568, IEEE, 2009.

[229] S. J. Kessler and W. McKenna, *Toward a theory of gender*, in *The transgender studies reader*, pp. 181–198. Routledge, 2013.

[230] J. Robertson, *Gendering humanoid robots: Robo-sexism in japan*, *Body & Society* **16** (2010), no. 2 1–36.

[231] M. Wright, A. Freed, and A. Momeni, *Opensound control: State of the art 2003.*, in *NIME*, vol. 3, pp. 153–160, 2003.

[232] N. Strauss, *The game*. Canongate Books, 2006.

[233] E. V. Markovik, *The mystery method: How to get beautiful women into bed*, *New York* (2007).

[234] E. Walster, *The effect of self-esteem on romantic liking*, *Journal of Experimental Social Psychology* **1** (1965), no. 2 184–197.

[235] G. H. Gudjonsson and J. F. Sigurdsson, *The relationship of compliance with coping strategies and self-esteem.*, *European Journal of Psychological Assessment* **19** (2003), no. 2 117.

[236] J. A. Hall and M. Canterberry, *Sexism and assertive courtship strategies*, *Sex Roles* **65** (2011), no. 11-12 840–853.

[237] D. L. Strauss, *D. american beuys" i like america & america likes me"*, *Between dog & wolf, Essays on Art and Politics* (1993).

[238] N. Wolf, *The beauty myth: How images of beauty are used against women.* Random House, 1991.

[239] J. Forte, *Women's performance art: Feminism and postmodernism*, *Theatre Journal* **40** (1988), no. 2 217–235.

[240] "Yoko ono, cut piece (1964)." https://www.moma.org/learn/moma_learning/yoko-ono-cut-piece-1964/, journal=MoMA Learning, publisher=MoMA. Accessed: 2019-12-01.

[241] "Valie export tapp und tastkino (tap and touch cinema) 1969 / 1989." https://www.moma.org/collection/works/159727, journal=MoMA Collection, publisher=MoMA. Accessed: 2019-12-01.

[242] M. Torimitsu, "Somehow, i don't feel comfortable: Momoyo torimitsu." https://www.momoyotorimitsu.com/. Accessed: 2019-12-01.

[243] M. Rikić, "Mónica rikić: Experimental games: New media art: Alternative tech." https://monicarikic.com/. Accessed: 2019-12-01.

[244] "Jordan wolfson's (female figure)." https://www.thebroad.org/art/featured-artworks/jordan-wolfsons-female-figure. Accessed: 2019-12-01.

[245] "Peepshow." http://www.artelectronicmedia.com/artwork/peepshow. Accessed: 2019-12-01.

[246] G. Walker, "Kinetics." http://www.gileswalker.org/kinetics. Accessed: 2019-12-01.

[247] T. D. Conley and L. R. Ramsey, *Killing us softly? investigating portrayals of women and men in contemporary magazine advertisements*, *Psychology of Women Quarterly* **35** (2011), no. 3 469–478.

[248] E. Nimpstch, "Jordan wolfson: Art and objecthood." https://www.riotmaterial.com/jordan-wolfson-art-and-objecthood/, Nov, 2018. Accessed: 2019-12-01.

[249] A. D. Cheok, K. Karunanayaka, and E. Y. Zhang, *Human-robot love and sex relationships*, *Robot Ethics: From Autonomous Cars to Artificial Intelligence,(Eds P Lin, K Abney, R Jenkins)* (2017) 193–220.

[250] M. Kang, *Building the sex machine: The subversive potential of the female robot*, *Intertexts* **9** (2005), no. 1 5.

[251] C. Paul, *Challenges for a ubiquitous museum: From the white cube to the black box and beyond*, *New Media in the White Cube and Beyond* (2008) 53–75.

[252] M.-S. Kim and M. Bresnahan, *Cognitive basis of gender communication: A cross-cultural investigation of perceived constraints in requesting*, *Communication Quarterly* **44** (1996), no. 1 53–69.

[253] M. J. Williams and L. Z. Tiedens, *The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior.*, *Psychological Bulletin* **142** (2016), no. 2 165.

[254] M. S. Graziano and D. F. Cooke, *Parieto-frontal interactions, personal space, and defensive behavior*, *Neuropsychologia* **44** (2006), no. 6 845–859.

[255] V. Lehmann, E. M. Huis, A. J. Vingerhoets, *et. al.*, *The human and animal baby schema effect: Correlates of individual differences*, *Behavioural processes* **94** (2013) 99–108.

[256] M. Borgi, I. Cogliati-Dezza, V. Brelsford, K. Meints, and F. Cirulli, *Baby schema in human and animal faces induces cuteness perception and gaze allocation in children*, *Frontiers in psychology* **5** (2014) 411.

[257] L. Beck and E. A. Madresh, *Romantic partners and four-legged friends: An extension of attachment theory to relationships with pets*, *Anthrozoös* **21** (2008), no. 1 43–56.

[258] E. K. Crawford, N. L. Worsham, and E. R. Swinehart, *Benefits derived from companion animals, and the use of the term "attachment"*, *Anthrozoös* **19** (2006), no. 2 98–112.

[259] A. Yokoyama, *The possibility of the psychiatric treatment with a robot as an intervention-from the viewpoint of animal therapy*, in *Joint 1st Int. Conf. SCIS ISIS, Tsukuba, Japan*, 2002.

[260] J. Panksepp and J. Burgdorf, *Laughing rats? playful tickling arouses high frequency ultrasonic chirping in young rodents*, *Toward a science of consciousness III* (1999) 231–244.

[261] C. Triscoli, I. Croy, H. Olausson, and U. Sailer, *Touch between romantic partners: Being stroked is more pleasant than stroking and decelerates heart rate*, *Physiology & behavior* **177** (2017) 169–175.

[262] C. Schmied, X. Boivin, S. Scala, and S. Waiblinger, *Effect of previous stroking on reactions to a veterinary procedure: Behaviour and heart rate of dairy cows*, *Interaction studies* **11** (2010), no. 3 467–481.

[263] S.-L. Gan, *Squeezables: tactile and expressive interfaces for children of all ages.* PhD thesis, Massachusetts Institute of Technology, 1998.

[264] H. Perner-Wilson, L. Buechley, and M. Satomi, *Handcrafting textile interfaces from a kit-of-no-parts*, in *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, pp. 61–68, ACM, 2011.

[265] A. Flagg, D. Tam, K. MacLean, and R. Flagg, *Conductive fur sensing for a gesture-aware furry robot*, in *2012 IEEE Haptics Symposium (HAPTICS)*, pp. 99–104, IEEE, 2012.

[266] P. J. Boylan, *Running a museum: a practical handbook*. International Council of Museums, ICOM, 2004.

[267] J. G. Hanhardt, K. Hakuta, and N. J. Paik, *Nam June Paik: Global Visionary*. Smithsonian American Art Museum, 2012.

[268] S. Cook, *Immateriality and its discontents: An overview of main models and issues for curating new media*, *New media in the white cube and beyond* (2008) 26–49.

[269] C. Roberts, G. Wakefield, and M. Wright, *The web browser as synthesizer and interface.*, in *NIME*, pp. 313–318, Citeseer, 2013.

[270] M. H. Wan Rosli, K. Yerkes, M. Wright, T. Wood, H. Wolfe, C. Roberts, A. Haron, and F. R. Estrada, *Ensemble feedback instruments*, in *Proceedings of the international conference on New Interfaces for Musical Expression*, pp. 144–149, The School of Music and the Center for Computation and Technology (CCT . . . , 2015.

[271] "When you're ready project." https://whenyoureready.org/, 2019. Accessed: 2019-08-01.

[272] C. Danescu-Niculescu-Mizil and L. Lee, *Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs*, in *Proceedings of the 2nd Workshop on Cognitive Modeling and Computational Linguistics*, pp. 76–87, Association for Computational Linguistics, 2011.

[273] E. O. Kirkegaard and J. D. Bjerrekær, *The okcupid dataset: A very large public dataset of dating site users*, *Open Differential Psychology* **46** (2016).