Cassandra Smith

EM 500

Dr. Cummings

December, 12th 2019

Final Paper: Voice Sensory Sex Dolls

Over the last century emerging media has transformed almost all aspects of human life. From communication and entertainment to news broadcasting and home appliances, almost everything people do is controlled or operated through media technology. Continually there are a few unconventional social spheres that have started to introduce concepts of emerging media technology. One of these social spheres is sexual intercourse and the accompanying sex toy industry. While thousands of battery operated sex devices exist to assist humans there are few products that offer a mediated experience equivalent to other media technology devices that rule our everyday lives. This has since changed when the first artificial intelligence sex robot was manufactured and sold in 2018. However, there are ultimately zero sex toy products that offer mediated user experiences without the extremes of AI and robotics. The goal is to manipulate the already existing sex doll product so that it can offer a new form of entertainment and mediated experience for users without having to be a robot. By adding a voice sensory/virtual assistance application to sex dolls, users can have the pleasures of almost-human connection with a product that is still almost-human.

In the next four years the global sex toy industry is projected to reach $52 million. That is a significantly large number for an entire industry and dimension of social life that has very limited media technology components. As western societies move closer and closer to a state of

open acceptance and comfortability with sex products, they will ultimately become more open to technology innovation. While hundreds of sex toys and products are created for specific and sometimes outlandish needs, kinks, and fetishes, there is a huge gap in the merging of media interactivity and intimacy. This is where voice sensory and virtual assistance in sex dolls comes in. With a user interface similar to that of virtual assistants like Apple's Siri, Amazon's Alexa, and Google Home, voice sensory sex dolls will foster this bridging of human intimacy and technology. Users can push the boundaries of what has been deemed appropriate within the media landscape and be apart of a new and unique experience. One other compelling reason for this design is the role it could potentially play within the emerging social sphere of sex tourism. As sex doll brothels continue to rise up in select countries around the world, adding a more "human-like" component such as speech could allow for heighten attention. If the world is headed this way, we might as well curate it to fit our media needs, wants, and desires.

Sex dolls in all there variations have been around since the mid century with intercourse simulators and artifacts expanding back centuries earlier. Intended for the purpose of simulating intercourse and assisted masturbation, sex dolls' audience for a long time remained early sea voyagers and military men who sought companionship and cleaner alternatives to prostitution. The first "life-like" dolls were made from inflatable plastic material and lacked many of the distinguishable qualities of human partners. Until 1996 when the first anatomical silicone doll was created and released by an ex-halloween mask creator, Matt McMullen. McMullen then went on to create Real Dolls, one of the first and most popular sex doll companies. Throughout the 21st century dozens of companies have started to design, manufacture, and sell both silicone and tpe (thermoplastic elastomer) dolls with movable metal skeletons. Top online companies

include Real Dolls, Sinthetics, Sex Dolls.com, Fair Dolls, and Silicon Wives. Soft skin-like outside, metal wire frame inside, hair, eyes, and skin coloring details all together create a decent human looking sex toy. A toy that is used for the purpose of satisfying either a sexual or companionship need. Any other emotional connection to the product is created strictly through the user's own imagination and perception. This is where voice sensory could help foster those emotional and personal feelings towards the doll and the overall experience itself.

Fit in Media Environment

The two main dimensions of this sex doll design application are Interactivity and Virtual Assistance. Interactivity can be thought about as a process where two separate entities work together and influence each other's choices and overall shape their interpersonal experience. In today's advanced media society interactivity can be between either two people, one person and one media device, or even two media devices. For the sex doll design the specific type of interactivity that will be focused on will be one person and one media device. Academics have debated for decades about what particular user vs media device relationship constitutes as interactivity. As outlined by Sheizaf Rafaeli (1988), Interactivity operates best within a responsiveness model and "requires communicants to response to each other" (Rafaeli, 1988). This perception of interactivity would be supported by the product design, as the voice sensors would require preempting speech patterns and requests before providing a curated response to the user. For the design to work the user has to work as an active participant in the interactive process, another element to responsive interactivity. (Rafaeli, 1988) Another distinguishable characteristic of the design application in terms of its interactive process is that is a different communication than what occurs between two people. Jennifer Stromer-Galley (2014) highlights

this concept by saying, “interactivity face to face is a different phenomenon altogether from interactivity.” (Stromer-Galley, 2014) Essentially this design while attempting to mimic a potential conversation between two sexual human partners will be a completely different type of communication. A type of interactive communication that Stromer-Galley (2014) calls “Interactivity-as-product.” (Stromer-Galley, 2014) Interactivity-as-product looks different than interactivity-as-process because as-process is communication and response between two people and as-product is when technology interface allows for a mediated interaction to take place for users. The voice sensory will work in this way, users speak to the device and then the device will respond, creating a back-and-forth dialogue between user and device.

One of the main concerns from academics and general population alike is can this specific type of media interaction be a genuine replacement to human communication? The answer to this mostly lies within the specific needs of the user of the media product. If the sex toy is already acting as a good enough replacement for real human intimacy then the additive speaking feature is presumed to work in the same way. That is why the virtual assistance component is vital to making the design functional, relatable, and believable to the experience.

When offhandedly thinking about speech there are key attributes that can make or break conversations. These attributes include tone, diction, infliction, pauses, and repetition, and together shape communication processes both directly and indirectly. When it comes to the virtual assistants that people use on their media devices they are all basically the same design female, direct, collective, informative, and mostly lacking emotion. They are designed purposively like this because they are task-oriented and meant to be helpful and all knowing. In their 2005 book *Wired for Speech*, authors Clifford Nass and Scott Brave explain how human’s

preferences for speech and communication with each other does in fact cross over to our media technology. (Nass, et al 2005) What is important about understanding these speech preferences is that they will likely inflict emotional responses from users, especially when the speech revolves around intimacy like this design does. When it comes to media devices and speech interfaces using "I" Nass et al explain that there are four different possibilities of how this can occur.

1. Both recorded-speech and synthetic speech voice interfaces should say "I"
2. Recorded-speech voice user interfaces should say "I." Synthetic-speech voices user interfaces should not say "I"
3. Neither recorded-speech nor synthetic-speech voice user interfaces should say "I"
4. Recorded-speech voice user interfaces should not say "I." Synthetic-speech voice user interfaces should say "I"

Nass et al details that all of these possibilities contain both pros and cons, with the decision resulting in the design and functions of the interface. (Nass, et al 2005) Media users ultimately will feel some sort of way when a fake voice makes personal claims as compared to when a very real voice in a not real object makes the same remarks. For this particular design "I" and other personal remarks and phrases are essential to the interface and tasks it is designed for. With mediated intimacy as the goal, it would be strange if the voice design was not also depicted as personally motivating. Another design function that will be taken into consideration is the "who" behind the voice. In her 2015 article Jessie Hemple depicts that there are specific reasons why most virtual assistants are synthetic female voices with the main reason being sexism. (Hemple, 2015) Alexa, Siri, and Google home are all female sounding and as virtual assistance is implanted into most new media technology, the trend is continuing. Hemple states, "In the short

term, female voices will likely remain more commonplace, because of both cultural bias and the role technology plays in our lives." (Hemple, 2015) As users we are actively aware of this design practice and the indirect effects it has on perceived women's roles, but we also know that it is pleasing to listen to and interact with. When considering this design interface it will have the customizability range of Apple's Siri, with language preferences and the option for male or female. But like Siri, the default mode will be a women's voice because most sex doll products are in fact replicated female forms and are purchased and used by males.

The last component to making the voice interface is finding the balance between real recorded voice and synthetic voice. Most voice assistants rely on a combination of both recordings and synthetic duplications for the devices. Voice actors with consistent tone and rhythm will record common phrases and words and then have those recordings translated into all the possible text and sounds that our devices would say. As outlined in a CNN Business article by Jessica Ravitz, "These snippets were then synthesized in a process called concatenation that builds words, sentences, paragraphs." (Ravitz, 2013) This process, concatenation, will be applied to this design, but with very different voice actors. The voice actors behind virtual task assistants are selected based off of the service the device gives. In the case of the voice sensory sex dolls it would not make sense for them to sound automated and aloof like Siri and Alexa. This is why the voice acting will come form adult film stars who have the experience in talking in the particular experience that the device would be used with. Adult film stars in this sense can record all the expected sounds, phrases, and noises associated with intimacy and intercourse. This will allow for a more relatable and authentic experience to take place for the device users.

By analyzing the distinctive qualities of both virtual assistance and interactivity we develop a better understanding of the framework that all media technology is designed within. Some media devices can't come off as too human while others operate best when afforded the emotions that humans equate with one another. When it comes to developing this particular design, it makes sense to have a mediated experience that mimics real life situations and falls into that blend of realistic and synthetic. After all the doll products themselves look as human as they possible can be, the voice operation system should work the same way.

Target Audience

While sex toys are designed to be available to all kinds of people regardless of age, gender, and sexuality, most devices can be targeted to a specific audience in mine. When it comes to sex dolls that target audience has a tendency to be single straight cis gendered men. A more distinctive attribute of this audience is that they are more likely to be wealthy and often lack both social and intimacy skills. Over the last couple years as silicone and tpe dolls have become more popular, news and entertainment companies have looked into the kind of people that use sex dolls. Vice, Vox, Forbes, and Buzzfeed, just to name a few, have produced articles and short documentary style videos about this increasingly cultural phenomena. These works have all highlighted that sex dolls are manly used by men who crave an intimate attachment and either don't know how to have one with a real person or simply favor the one that can be had with a fake person. Some sex therapists have even said that sex dolls can be very beneficial to individuals who struggle with making interpersonal relationships, and these particular products allow for more intimate and meaningful options as compared to other sex toys. In their 2018 paper N. Döring and S. Pöschl explain that sex doll owners have reported that "sex dolls can

provide a lot of sexual and emotional satisfaction, create feelings of comfort, peace and even love." (Döring, et al 2018) While this research indicates a strong desire and need for a sex doll product, it should be known that only a very small percentage of the population actually uses sex dolls. Döring et al (2018) found in their national online survey that 9% of men have used a sex doll and only 2% of women have. These numbers could change with the additive feature of voice sensory virtual assistance and possibly help in changing the stigma surrounding sex dolls and sex toys as a whole.

The User Experience

When it comes to actually using this product there are essentially two different circumstances in which the product would be used. The first and also the obvious situation would be personal use in which an individual would buy a sex doll and use it at their own discretion and creating whatever relationship they want with it. The second situation is completely reliant on the emergence of sex doll brothels in Europe and Japan, thanks to the sex doll company Lumi Dolls. Lumi Dolls has begun operating sex doll brothels for patrons who either want to specifically use dolls or want the clean and protective reassurance that comes with dolls and not with people. With sex doll brothels users are guarantied safety from STIs, law enforcement, discrete use, and a unique experience. As general sex doll use becomes more popular it is projected that the sex tourism industry will grow to. Regardless of these two different situations in which sex doll use would take place, the user experience with the voice sensory virtual assistance would be the same. A user would have a doll, touch, feel, and manipulate it however they want. And then proceed with engaging in conversation with the device. An example of how this would look is indicated below:

User: “Do you know what I want to do to you?

Doll: : “What do you want to do to me?”

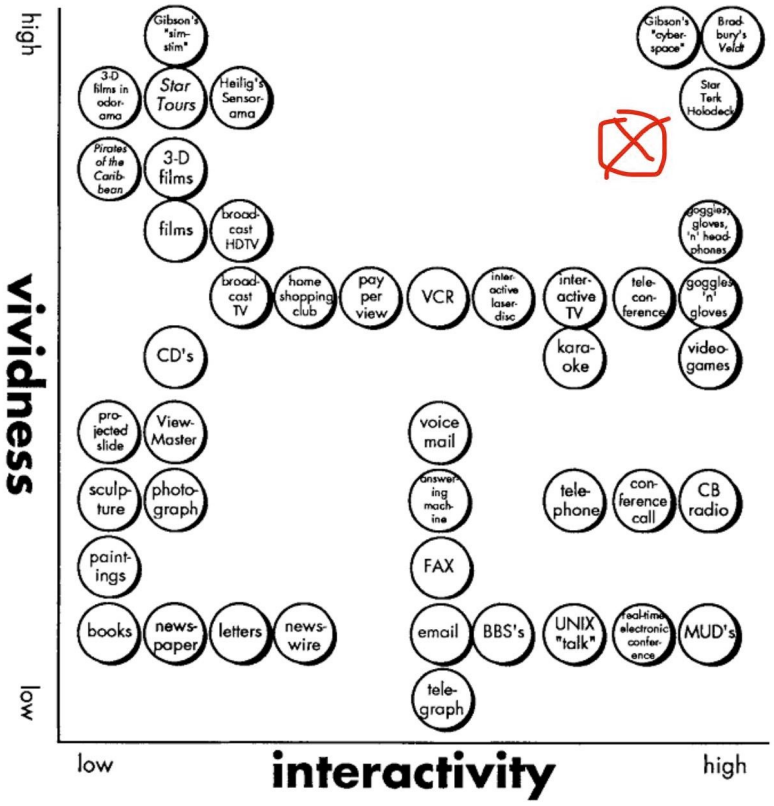
User: “I wanna touch your body”

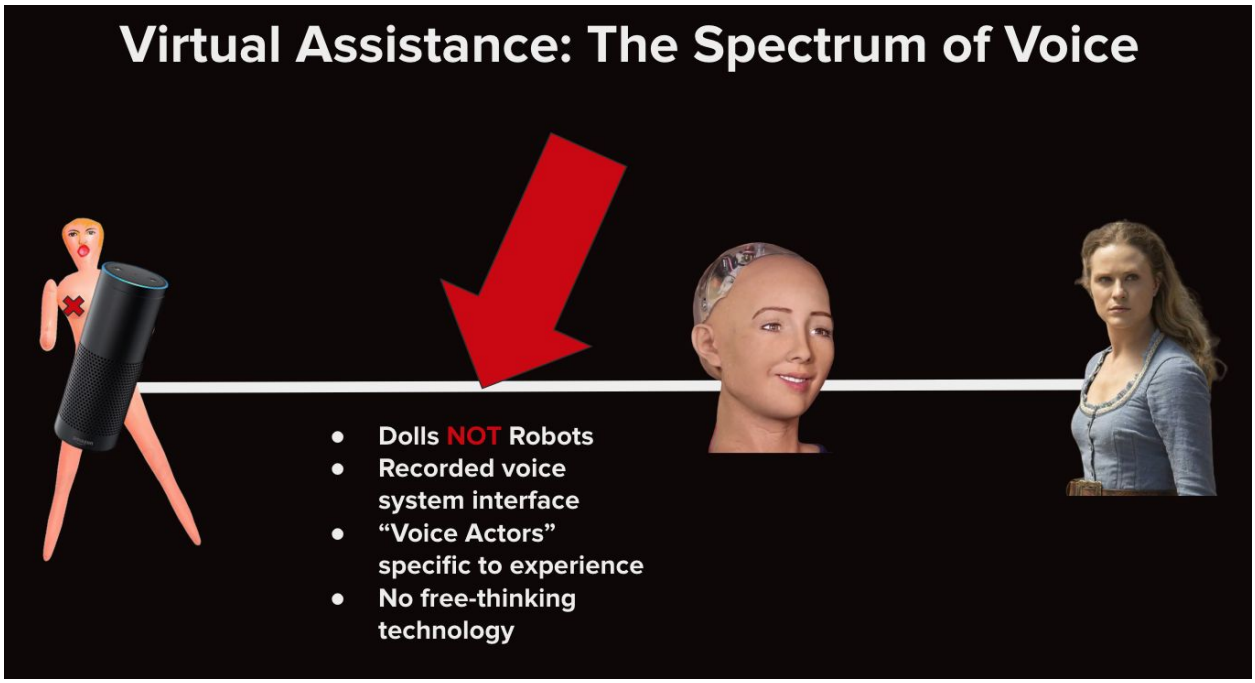
Doll: “Then touch me”

User: “Like that?”

Doll: “Yes. Just like that”

Users would simply talk out loud to the doll and using a microphone and speaker system the doll would respond back accordingly. Questions would be met with scripted answers that are programed to respond as accurately as possible. For example when using Siri, if a user asked about the weather she would only respond with answers about the weather and not about anything else. This device would work in the same way. If a user was talking about a specific spot on the doll, like touching the breasts the doll would have an answer that correlates to the breasts and not another part of the body. This specific program ensures that the mediated experience is as close to a real life intimacy as possible. This whole process is a pretty distinct example of Interactivity-as-product. (Stromer-Galley, 2014) Users speak and the media technology listens, correlates an accurate and helpful answer, and then responses back. With the notion being this is a fake doll that looks like a real person and has the voice of a real person and helping to create a replication of what real intimacy looks, feels, and sounds like. When compared to the diagram by Jonathan Steur (1992) about vividness and interactivity, this design when coupled with the sex doll product, would fall somewhere around the same level of high interactivity and high vividness. It is placed there because touch, hearing, sound, and sight are all senses that are stimulated throughout the user experience.


high
vividness
low
low
interactivity
high
Gibson's "sim-stim"
3-D films in odor-ama
Star Tours
Heilig's Sensor-ama
Pirates of the Carib-bean
3-D films
films
broad-cast HDTV
broad-cast TV
home shopping club
pay per view
VCR
inter-active laser-disc
inter-active TV
tele-con-ference
goggles 'n' gloves
video-games
kara-oke
CD's
pro-jected slide
View-Master
sculp-ture
photo-graph
paint-ings
books
news-paper
letters
news-wire
voice mail
answer-ing mach-ine
FAX
email
BBS's
tele-graph
tele-phone
con-ference call
CB radio
UNIX "talk"
real-time electronic confer-ence
MUD's
Gibson's "cyber-space"
Brad-bury's Veldt
Star Trek Holodeck
goggles, gloves, 'n' head-phones


Virtual Assistance: The Spectrum of Voice
Dolls NOT Robots
Recorded voice system interface
"Voice Actors" specific to experience
No free-thinking technology

The last key element to the user experience is understanding where this particular design product falls on the spectrum of voice and physical device. We can think of this product and others like it existing on the spectrum of obviously fake looking and sounding to hyper realistic and almost indistinguishable from real humans. On one end you have cheap hard plastic dolls that don't look like real humans and could possible have voice components that are equivalent to that have Siri and Alexa. Next would be this product, more realistic look and feel to the doll body while still remaining a doll, and has the voice operation that has been outlined. Further into the spectrum would be sex robots. What categorizes other products as sex robots and no longer sex dolls is the device's capability to not only speak but to move body parts, vibrate, and be controlled/stylized through smartphone applications. Therefore allowing for more customization that our particular design does not include. On the other end of the spectrum we have a product design that is so realistic that it can not possibly be made and as of currently, only exists in film, novels, and imagination. The most accurate depiction of this would be Michael Crichton and HBO's *Westworld*, where robots that look not only realistic to humans, but have the capability to feel emotion and self learn. It is with this spectrum that there is some level of reassurance to be had in our design tastefully and safely allowing for a small intersection between human intimacy, interaction, and media technology to take place.

Closing & Future Directions

The goal of this product design is to create a new form of sexual interactivity by merging together aspects of virtual assistance technology with the designs of already existing sex toy products. Trough voice sensory assistance attached to silicone lifelike sex dolls, a potential user could partake in a significantly more active and realistic mediated sexual experience. The

reasons for this design come from not only the rise in popularity of sex doll and robot products, but also to give a mediated option that is not as extreme as using sex robots and artificial intelligence. By having the step between just a sex doll and a sex robot that does it all, then freedom of user choice expands.

In terms of future design directions the general thought would be allowing for more personalization of the voice system, with possibly adding in options to record other people's voices and have a more personalized experience and have it be self learning AI. In regards to these design aspects, they have already been addressed in current sex robot models. Sex robots that are currently manuafacuted by Real Dolls have the ability to have customzied tone and attitudes to the voice but not personalzied voice recorderings. This is to have a design quality that differentiates between what is real-life and what is not. Also when given the option to record your own voice users then have the ability to selfishly use and possible abuse the voice of a person they no longer are intimate with or that of a minor. Sex doll comapnies currently do not sell products that resemble minors and while extreme customization is posible, there are safety limitons in place. The final aspect of voice sensory sex dolls that maintains its place as a functioning sex toy product is that becuase it is not a robot, it can never be affored rights. When discussing the place that robots have in our society one of the biggest questions that ultimately always comes up is, can and do robots/AI deserve rights? This becomes a hot topic because depending on the job sector and social sphere that the robot operates in, media users must look to see that actual humans in those jobs should be awarded basic rights before the robots do. When it comes to sex dolls and sex robots joining and possibly replacing human sex workers, one can only hope that it is after human sex work is seen as less dehumanizing like it always has been.

The overall purpose of this design to allow for a sex toy product to exist compfortable in a society that wants to push the boundries of media interactivity while also not compleatly destorying existsing social spheres in which human life security is threatened.

References

Döring, N., & Pöschl, S. (2018). Sex toys, sex dolls, sex robots: Our under-researched bed-fellows. *Sexologies*, *27*(3). doi: 10.1016/j.sexol.2018.05.009

Hempel, J. (2018, June 6). Siri and Cortana Sound Like Ladies Because of Sexism. Retrieved from https://www.wired.com/2015/10/why-siri-cortana-voice-interfaces-sound-female-sexism/

.

Nass, C. & Brave, S. (2005). Wired for speech: How voice activates and advances the human-computer relationship. Cambridge, MA: MIT Press.

Rafaeli, S. (1988). Interactivity from new media to communication. In R. P. Hawkins, J. M. Wiemann & S. Pingree (Eds.), Advancing communication science: Merging mass and interpersonal processes (pp. 110–134). Newbury Park, CA: Sage.

Ravitz, J. (2013, October 15). 'I'm the original voice of Siri'. Retrieved from https://www.cnn.com/2013/10/04/tech/mobile/bennett-siri-iphone-voice/.

Jonathan Steur, "Defining Virtual Reality," *Journal of Communication* 42 (1992): 79-90

Stromer-Galley, J. (2004). Interactivity-as-Product and Interactivity-as-Process. *The Information Society*, *20*(5), 391–394. doi: 10.1080/01972240490508081