

UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

Sonic Futures: Assistive Technologies, Gendered Labor, and the Colonization of Voice

Permalink

<https://escholarship.org/uc/item/89h414tg>

Author

Santos, Dorothy Resplandor

Publication Date

2023

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nd/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

**SONIC FUTURES: ASSISTIVE TECHNOLOGIES, GENDERED LABOR, AND THE
COLONIZATION OF VOICE**

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

DOCTOR OF PHILOSOPHY

in

FILM & DIGITAL MEDIA
with an emphasis in COMPUTATIONAL MEDIA

by

Dorothy R. Santos

June 2023

The Dissertation of Dorothy R. Santos is
approved:

Professor Soraya Murray, Chair

Professor Nathan Altice

Professor Nancy N. Chen

Professor Halcyon M. Lawrence

Professor Shelley Stamp

Peter Biehl
Vice Provost and Dean of Graduate Studies

Copyright © by

Dorothy R. Santos

2023

Table of Contents

List of Figures	iv
Abstract	v
Acknowledgements	vii
Introduction	1
Chapter 1: Holding the Line	21
Chapter 2: Digital Media Artifacts of Synthetic Voice	63
Creative Intervention, Part I	111
Chapter 3: Hidden Frequencies	134
Chapter 4: Speculating a Socially Just Sonic Future	177
Creative Intervention, Part II	221
Conclusion	250
Glossary	254
Appendix	256
Bibliography	257

List of Figures

- 1.1-1.4 Screenshots of Training for Service (1926)
- 1.5 Screenshot of Call of the East (circa 1930)
- 2.1-2.6 Screenshots of *Alexa's Body* (2021) commercial
- 2.7-2.9 Screenshots of Apple Macintosh (1984) commercial
- 2.10-2.14 Screenshot of The Lost 1984 Video: Steve Jobs
- 2.15 Diagram of the Documentation Archive Trail
- CI 1.1 Image of *Press 1 to be Connected* (2019)
- CI 1.2 Cover Image of 1984 edition of BYTE Magazine
- CI 1.3-1.6 Screenshot of *Press 1 to be Connected* (2019)
- CI 1.7 Screenshot of *Imaginary Relay, For Andy* (2023)
- CI 1.8-1.12 Installation images of Hiding in Plain Sight exhibition
- 3.1-3.8 Screenshots of Apple 911 commercial (2022)
- 3.9-3.13 Screenshots of /911dispatchers subreddit forum
- 4.1-4.3 Screenshots of a Sanas AI website
- 4.4 Screenshot of social media influencer, Rea Ninja`
- CI 2.1 Screenshot of *Dark Matters* (2019) by Johann Diedrick
- CI 2.2 Screenshot of *Faith* (2018) by Ryan Kuo
- CI 2.3-2.12 Images of *The Cyborg's Prosody* (2022-present)

ABSTRACT

Sonic Futures: Assistive Technologies, Gendered Labor,
and The Colonization of Voice

Dorothy R. Santos

This dissertation investigates the human voice as a type of unseen or invisible data mediated through voice recognition platforms, systems of care, emergency infrastructures, and the creation of assistive technology. It explores the various mechanisms that create a standard voice associated with authority and care. Looking first within the telecommunications world, Bell System specifically, the human voice was associated with a specific phenotype, race, and gender. This research asserts that western, American and British Englishes are perpetuated as the global lingua franca in a transnational and global economic infrastructure of information, big data, and a complex system of care.

In a world increasingly reliant on assistive technologies, sonic perceptions and memories are critical if we hope to preserve oral traditions and language diversity. The introduction of this project lays out the theoretical framework of the dissertation, which includes the examination of gendered labor histories, codified speech, and how the language of care and authority

are mediated through digital technologies within emergency infrastructures and global, racial capitalism.

Acknowledgements

To my committee members, Professors Soraya Murray (Chair), Halcyon M. Lawrence, Nathan Altice, Nancy Chen, and Shelley Stamp. There is no way I could have done this dissertation research, writing, and artworks these past six years without your guidance, support, care, wisdom, and expertise.

Soraya Murray, I am extremely grateful for all of the confidence and faith you had in me when I was working on my master's degree at the California College of the Arts, which is when we first met. You saw something in me years ago that I had no idea was there. Through your pedagogical approach, your expertise, great and vast wisdom, and knowledge, I have learned how to be an artist, a scholar, a community leader, an educator, and someone that is unafraid to question and have a healthy dose of skepticism. I am not exactly sure what you saw in me and I marvel at how much confidence you have in me, but you have believed in me from the beginning. I am so grateful for your mentorship and advice throughout this journey. Also, thank you for laughing at all my corny jokes during our advising meetings. I look forward to even more conversations, your continued guidance, and care for

many years to come. Cheers to you and thank you for reminding me to pay it forward in all that I do.

Anna Friz, I remember my first year of the Ph.D. program, I went to your office hours and cried. I had no idea if I could make it through. I felt self-indulgent making art while trying to understand theoretical text I had not read before. You reminded me that, sometimes, I need to step away from the world in order to understand it. Above all else, you taught me something extremely important that I feel I am far better at and improving upon every day of my life since meeting you...to (actively) listen and the potency of silence (then again, is a pure silence even possible with the chatter in our minds?). I have been so fortunate to have you as my professor and mentor. I am indebted to you for helping me re-configure the way I am in the world. Thank you for forcing me to listen to cityscapes, to non-humans, to the ocean and the owls (that are not what they seem), and to the voices of my loved ones (and even people I might not get along or agree with) in such a way that has dictated the ways I write, research, and make art. I did not imagine that a work I created for your class would end up being selected for an experimental music and sound festival (!!!) and alongside you and Connie Zheng. Thank you for seeing (and hearing) me and teaching me to be a better human being.

To my UC Santa Cruz Film & Digital Media cohort, Raed El Rafai, Erick Msumanje, Chris Kerich, and Rana Jarbou. We have all grown together intellectually and emotionally and it has been a pleasure and an honor to bear witness to all of your journeys and endeavors. May our paths continue to cross in all the work we do in the future.

To the University of California, Santa Cruz FDM community: Sharon Daniel, Warren Sack, Rick Prelinger, Jenny Horne, Peter Limbrick, Irene Lusztig, Susana Ruiz, Jennifer Maytorena Taylor, Yiman Wang, Ruth Anne Beutler, Gustavo Vazquez, Isabelle Carbonell, Alessia Cecchet, Anita Chang, Fabiola Hanna, Melanie Ho, Chisato (Chisa) Hughes, S. Topiary Landberg, Fiorella Lema, Kiki Loveday, Cebe Loomis, Rebecca (rora) Ora, Ryan Page, Francesca Romeo, Lara Saab, Benjamin Schultz-Figueroa, Casandra Sobieralski, Kara Stone, Kara Stone, Brett Kashmere, Oana Tenter, Rolando Vargas, Yasmine Benabdallah, Sarah Bird, alex cruse, Laimir Fano Villaescusa, ksenia fir, Gonzalo Galetto, Yulia Gilichinskaya, Suzy Halajian, Hannah Jayanti, Young Eun Kim, Darline Kim, Hans Kuzmich, Mahshid Modares, Amy Reid, Allen Riley, Yasheng She, Kelly Sky, Shirin Towfiq, Merve Unsal, Darren Wallace, Lalu Ozban, and Melanie Ho.

To all the incredible folks that I have worked with and collaborated with through the Science and Justice Research as a graduate fellow including

Jenny Reardon, Colleen Stone, James Doucet-Battle, Camilla Hawthorne, Jaimie Morse. Aitanna Parker, Paloma Medina, and the SJRC Undergraduate Fellows. Special thanks to the SJRC Training Program Graduate Fellows: Jon Akutagawa, Dennis Browe, Maggie Edge, Caroline Spurgin, and undergraduate fellow Diana Sernas.

Many thanks to my California College of the Arts community: Tirza “True” Latimer, Jeanette Roan, TT Takemoto, Maria McVarish, Guillermo Galindo, Jacqueline (Jackie) Francis, Kim Anno, Michele Carlson, Eric Olson, Maureen Burdock, Alex Fine, Erica Gomez, Emily Holmes, Vanessa Kauffman, Leora Lutz, Elizabeth Moran, Dani Neitzelt, Mike Rothfeld, Amanda N. Simons, Anton Stuebner, Calder Yates, Natalie Catasús, Irina Contreras, Rachel (Rhea) Endoso, Takeema Hoffman, Danielle Jackson, Melissa Miller, Omar Mismar, Colin Partch, Diana Stapleton, Emily Kuhlmann, Jordan Reznick, Sita Bhaumik, Matthew Harrison Tedford, Sienna Freeman, Tanya Gayer, and Forrest McGarvey.

Tirza “True” Latimer, I will never forget your comment during my master’s thesis review before graduation, “Far be it from me if you want to dig a hole with a philip’s screwdriver.” Oh dear. I knew what that meant! A couple of weeks later, after a massive revision, you expressed how proud of me you were. Most importantly, you reminded me that I will have to confront even

more challenges and how you want(ed) me to be prepared for bigger endeavors in my life. You always believed in me and thank you for being such an integral part of my intellectual and creative journey.

To the community of artists, writers, mentors, curators, and friends that have been inspiring and supportive along this journey: Carmen Aguilar y Wedge, Indira Allegra, Jaime Austin, Minerva Arias, Kimberley Arteche, Amaranth Borsuk, Danielle Baskin, Liat Berdugo, Sarah Burke, Adrian Burrell, Stephan Caspar, Sofía Cordova, Vivian Crockett, Torrey Cummings, Christina Dacanay, A.M. Darke, Johann Diedrick, Brian Droitcour, Cheryl Dunye, Leora Fridman, Carla Gannis, Sonia Guiñansaca, Farley Gwazda, Chris Hamamoto, Alex Hanna, Harris Kornstein, Claudia Hart, Ian Hatcher, Max Hawkins, Rhonda Holberton, shawné holloway, Emily Holmes, Carrie Hott, Tom Hughes, Terry Hwang, Dorian Katz, Rachel Khong, Scott Kildall, Laura Hyunjhee Kim, Tamara Knesse, Gavin Kroeber, Lian Ladia, Peggy Lee, Guillaume Levrier, Jason Li, Adelle Lin, Benjamin Lotan, Jenny Ma, Dawn Mabalon, Charles Macquarie, Alexis Madrigal, Jason Magabo Perez, Cayden Mak, Mariam Malik, Ryan Manuel, Emily Martinez, Marc Mayer, John "Connors" McConville, Joanne McNeil, Amira-Sade Moodie, Luis Morales-Navarro, James Morgan, Yetunde Olagbaju, Nathan Olivares-Giles, Yindi Pei, Elizabeth Pérez, Phoenix Perry, Latoya Peterson, Anna Luisa Petrisko, PJ Gubatina Policarpio, Carissa Potter,

Thea Quigle Tagle, Eva Recinos, Shirley Rivera, Macha Rose, Skyler Rosellini, Nina Rubin, Janice Sapigao, Surabhi Saraf, Alena Sauzade, Roddy Schrock, Tara Shi, Boaz Sender, Mindy Seu, Danielle Siembieda, Caroline Sindera, Eileen Skyers, Astria Supark, Aimee Suzara, Eliza Swann, Mel Talusan, Barry Threw, Allyson Tintiangco-Cubales, Avery Trufelman, Virgie Tovar, Helen (Shewolf) Tseng, Tanna Tucker, Ben Valentine, Elia Vargas, Roopa Vasudevan, Lark VCR, Angela Washko, Sarah Watson, Lewis Weil, Kofi Yeboah, Connie Zheng.

William (Bill) D. Caughlin, Scott P. Combs, and Sheldon Hochheiser, Ph.D., I have tremendous gratitude to you all. Thank you for your time, generosity, care, and kindness. I had an incredible time looking through the Chinese Telephone Exchange documents, photos, and the extraordinary archive of so many incredible phones (prototypes, functional, strange, and futuristic). You provided and equipped me with such a wealth of knowledge and expertise of AT&T and Bell Systems history that had my mind wishing I could stay longer. But I will most certainly be back, if you will have me as a guest. It was such a pleasure and absolute fun delving through all of those historical documents.

Mara Mills, thank you so much for the introduction to the AT&T archives team, your phenomenal scholarship, generous heart, and enthusiasm.

You welcomed conversation with me and Abram in such a way that made us feel humbled and seen (and heard!). As a fellow Banana slug, I am so honored to be in community with you.

To the Stochastic Labs community, Vero Bollow, Alex Reben, Steve Thompson, Maria Finn, and Madeline Girard, I am indebted to you for allowing me the time and space to grow and evolve as an artist and to be around people that challenged by mind to bend and stretch in ways it had never done.

To the San Jose State University community especially Rhonda Holberton, Chelsea Thompto, James Morgan, Alena Sauzade, and Ashley Elieff. I deeply appreciate your support of my artistic practice and allowing me the opportunity to connect with students.

Robert Ochshorn, Johann Diedrick, Lillian-Yvonne Bertram, and Abram Stern (aphid) I have such tremendous respect and admiration for the ways each of you have shown up as advisors and mentors in realizing *The Cyborg's Prosody*. From my conversations about art, technology, speech, language, and poetry, each of your generous spirits have helped make this project an eye opening and illuminating experience. Thank you so much for your wisdom, expertise, care, and constructive criticism throughout the making and writing of this work.

Christina Dacanay, thank you so much for your phenomenal and beautiful illustration work for *The Cyborg's Prosody*. As I was writing out the concept of the work, I knew I needed a visual interface that would bring the words I have written to life. You were the first person I thought of to do this work. I know your heart has been through a lot especially in the past few years, but when you wrote to me and re-connected, you do not understand how happy I was to hear from you and how we finally (after all these years) would get a chance to work together. I have such deep gratitude for your vision and immense admiration for your artistic brilliance and deftness in knowing how to draw (no pun intended) someone into an experience. Much love to you and thank you for sharing your extraordinary talents and gifts with me and the rest of the world.

Jenifer Wofford, you always say that you didn't do anything and that everything I've done has been on me. But I'll say it again (and again), I don't think you understand what it meant to me for you to take time out of your busy schedule, on the heels of finishing up your MFA from Berkeley, and installing a show all those years ago. You did what I try to do with my students and mentees, you have them work alongside you (such a great way to connect). As I painted primer on the gallery walls of SOEX when it was located off of Valencia street, you asked me questions about what I wanted to

get out of pursuing grad school and what being an artist and writer meant to me. You shared your experiences, you took me to your favorite Vietnamese spot, and you listened. You care(d) and you continue to inspire so many people. Thank YOU for allowing me to bear witness to your artistic and pedagogical practices that have given me so much inspiration over the years.

One of the greatest how-we-met friend stories is definitely our story, Elizabeth Travelslight. I was eager to go to the Asian Art Museum to check out and meet Poklong Anading and Ringo Bunoan, but I was not expecting to meet a life-long friend and collaborator. One of my favorite things about you is your impeccable timing and brilliant mind. You have taught me so much about what it means to work together, to have difficult conversations, to give constructive criticism, to be present, and to allow myself to feel all of my feelings (all the time). Thank you for your check-in calls. Thank you for reminding me that I could do this work. Thank you for believing in me (since Day 1).

From our weekly *burden* (aka burger) runs to our heart to heart conversations, thank you, from the bottom of my heart, for all of the ways you have shown up for me, Kai Stowers. I am so deeply grateful for our friendship, the endless laughter, and support you have given me all of these years. No

matter where we end up in the world, you have become a friend for life and chosen family.

From the first time I met you at that cafe in Redwood City all those years ago, I have seen you grow your passion into something utterly amazing, Vivian Sming. Your friendship, care, and support (and patience) have truly helped me make the artwork I wanted to make as well as write what I never thought I could write. Thank you for your check-in calls and the well wishes along the way. I cannot wait to make books and art with you.

Shirlie Mae Peralta, I have witnessed your brilliance as you made it through your Ph.D. journey and it was wonderful to see (and read) all of the research you conducted and thought through. At one point, I did think to myself if it were possible for me to write at the caliber of your scholarship. I was, always am, truly inspired by you. I look forward to even more conversations and continued dialogue. Thank you for believing in me.

I know it has been quite some time since we have gone out and had one of our three-hour lunches, Cristina Olds-Greenberg, but I will absolutely make it up to you. I owe you so many lunches. I deeply appreciate your friendship over all these years. You watched me grow up, fall in love, fall out of love, get my heart broken, see me through job promotions, quitting, starting something new, and we laugh and smile through it all. You believed in me when I did not

believe in myself. All of your life advice are things I constantly remember when I am having a hard day or difficult and challenging time in my life.

Chicken Soup Forever.

Noelle! The best furry friend any human can ask for. Thank you for sticking it out with me and allowing me to be your human and being the most cuddly and sweet and the cutest cat ever. Most importantly, for sitting on my laptop because you knew that I needed to rest and take it easy. Thank you for reminding me to not overwork myself and being the best non-human friend to me over the years. Michelle Lindstrom, for all times I had to travel, you took such great care of Noelle. Thank you for always being there for us when we needed you.

To the Friday Co-Working and Co-Writing group: May-Li Khoe, Tyler Shoemaker, Chelsea Thompto, Kristin Miller, Christina Corfield, Ezra Teboul, Heather Dewey-Hagborg, Xiaowei Wang, Fabiola Hanna, and Abram Stern (aphid), you have made writing, working, research, art making so much more fun and enjoyable. Your jokes, all the links and resources dropped in the zoom chat, and the countless check-ins have been a balm.

Thank you, Nica Ross, Harrison Apple, and Bill Rodgers for allowing me to hang out at The Studio, creating work, having all of these incredible discussions with CMU faculty, and being so welcoming, caring, and

supporting my creative practice. Dani Lamorte, wow, where do I even begin? Your words about my creative practice are/were the affirmation any artist would be so eternally grateful to receive, which I am. Your questions, your engagement, your deep intention and care are so deeply and warmly felt and I will never ever forget our conversations and exchanges. It has been such a pleasure to connect with you.

To Hyphen Labs, thank you for the invitation to EYEO (2022) and allowing me to share my academic and creative research in front of a large audience. It was the first time I gave a talk to such a big group that focused on my art practice, which is something I never thought I would do other than within the walls of academia. Carmen Aguilar y Wedge, your friendship, our WhatsApp messages, and voice memos have meant the world to me and your belief in my ability to do this work is immense. Abrazos y besos, mi querida.

To Laleh Mehran and Chris Coleman, truly indebted to you both for giving me the opportunity to work alongside you and learn from you both. It has been wonderful to have your support and care over the past few years and I so deeply hope we can continue building together.

To the Processing Foundation team: Lauren Lee McCarthy, Kate Hollenbach, Xin Xin, Casey Reas, Dan Shiffman, Shuyun (Sonia) Choi, Tsige Tafesse, Saber Khan, Rachel Lim, Qianqian Ye, Raphaël del Corville, Johanna

Hedva, Cassie Tarakajian, and evelyn masso. I am forever thankful for our collaboration and continued commitment to building community. You all make me believe and have confidence in the future of open source software for artists and the power we have to affect change and make it ripple out into the world.

To the REFRESH Collective: Salome Asega, Heather Dewey-Hagborg, Kathy High, Lynn Hershman Leeson , Maandeeq Mohamed, Tiare Ribeaux, Dr Camilla Mørk Rostvik, and Addie Wagenknecht, we did epic and legendary things over the past six years, individually and as a collective. I am so proud of the work we have done and continue to do out in the world. Many MANY thanks for all of the ways you inspire me and so many others.

I cannot thank you all enough ~ Yoni Asega, Angela Carrier, Anna Lisa Escobedo, and the rest of the YBCA team and staff ~ for constantly welcoming my work into the Yerba Buena Center for the Arts (YBCA) space. To my YBCA10 family: Leticia Hernandez, Hasain Rasheed, Daryll Ratcliffe, My-Linh Le, Dr. Ayodele Nzinga, Alex Bledsoe, Nikiko Masumoto, and Deanna Van Buren, I so deeply appreciate your words, care, kindness, encouragement, and support over the past few years.

Adrian Jones, you have been such a light in my life. It has been an honor to watch you grow, evolve, and develop Looking Glass alongside all of

the many incredible things you are doing in your life. You have taught me so much and you have given me tremendous hope for the future especially because the world needs you now more than ever, my dear. I will forever be your Filipino auntie and count on me continuing to bear witness to all that you do. I mean it when I say call on me anytime. Truly.

Heather Dewey-Hagborg and Lee Blalock, your creative guidance, advice, constructive criticism, care, love, and humor have gotten me through this program and beyond. I am itching to co-create, build, and write with you both now and well into our future as Biolectics. There is always so much to discuss and share with you both and your vulnerability is a gift I do not take for granted. It has been such an honor to collaborate with you both as well as nurture this extraordinary friendship. Onward!

Immense gratitude to you Ruswell Cabudol and Chris Schaeffer (and RIP to Harper) for all the nights you let me stay over whether it was because I had gone through a bad breakup or I needed to get away from my place and just be out and about. You both have welcomed me to your home time and time again and reminded me of the art of relaxing, resting, and taking it easy. Spending time with you both has always been filled with so much laughter as well. And Rus, you have really seen me evolve since the day we met. Thank you for being such a wonderful and loving friend to me. You have taught me

so much about love for non-humans and learning to love adventure in all its forms.

Gabi (like Wasabi!) Schaffzin, I was so thrilled to meet you that fateful day at Southern Exposure in San Francisco for the Experimentation Studio. I had no idea we had so much overlap in our work and research. Most importantly, you were one of the most brilliant and down to earth people I met that day. From Patient Instruments to seeing you defend your dissertation at the height of the pandemic, I am so proud to be your friend and in community with you. I will never forget shooting the sh*t with you in New Orleans to your worried text agreeing to moderate your panel at CAA 2023, which was an honor and a pleasure, but it is because you are my academic kin. Always. I am grateful to have you in my life.

To Christina Corfield, thank you for laughing (like, belly ache laughing) at all of my jokes, for lighting a fire under my a\$\$ all the times I said I could not do something, for pushing me to think differently about art, and for showing up for me all of these years. All the ways you have given to me and the ways you take time to celebrate as well as all the ways you have shown up when I lost hope in myself. I meant what I said (even more so now) about being inspired by you and your creative practice. You were one of the people that made me believe this “artist scholar” thing was a real thing.

Ana “An Xiao” Mina, your sense of humor, your confidence in me pursuing all the work I do, and all of the gentle loving ways you show up for our Five and Nine coven. It boggles my mind that we met a little over a decade ago because I feel like you have been in my life far longer than that! Thank you for allowing me to be myself, for never judging me, and being constantly open to discovering the world. Your love of travel, your endless curiosity, and your genius crafting of words is, without a doubt, the best magic I have ever witnessed being wielded in the world.

Xiaowei R. Wang, I simultaneously feel like I have all the words and none of them when I think about our friendship and kinship. Our connection seems to transcend even language. From the ocean to the moon to the dream of the world set ablaze, you have truly modeled what it means to love and care in an expansive and radical way that I had never experienced before. You have been a constant reminder of why I believe in humanity. Seriously. You care and love so deeply and everyone (including myself) that encounters you knows by being around you and getting to know, work, and/or learn from you that you are one of the most extraordinary people ever born on planet earth. I love you and I could not have done any of this work without you. Talk about kismet! To be on this parallel journey with you has been one of the most phenomenal

experiences of my life. We will continue building and making together because it is second nature.

Joanne Rondilla and Jose Ureta, thank you for being family to me over the years. For welcoming me into your home, showing me all the cool new shows to watch that keep me relevant (I swear, you two are far better at media scholarship than me!), and laughing endlessly with me. But also, allowing me to cry my eyes out when I have been in pain or my heart is just going through it. Your words, hugs, love, comfort, and all of Joanne's home cooked meals have been a special combination of medicine and magic. I love you both and so grateful to celebrate and grieve all of life's joys and sorrows alongside you two.

To Ken Cory, from that first time we met and us staying up until the wee hours of the morning talking about love, relationships, and life, I knew I had met one of the greatest friends I would ever have. You say all the things I cannot. You are unapologetic, courageous, bold, giving, and loving. Our movie nights, shopping at our favorite witchy, occult shops, pulling oracle cards, and letting me practice my tarot reading skills on you over the years have been all of the things I have needed to remind me to keep going and to continue being me. Thank you for loving me the way you do.

I have no idea where to start with you, Lauren Lee McCarthy. It has been a journey. I have said this time and time again, but I owe you. Big time. From meeting you at that coffee shop in SF's Mission district, our friendship has grown and evolved in ways I did not expect. I have learned a tremendous amount about what it takes and what it means to be an artist, an educator, a community organizer, a leader, and a scholar. You simplify the complex as much as you complicate what is often perceived as simple. You are relentless in your pursuit of wherever your curiosity takes you. You are wildly imaginative. You ask questions in a way that undoubtedly (and to use your word) rearranges me. You allow yourself to feel things in a way that I rarely witness people doing out in the world, but you do it. I am indebted to you for all the things you have done for me. Big big love.

Nicole Torres, one of the architects to my heart and soul. When you texted me "let's go!!! We have been waiting for this moment since 1992" right after I texted you that I submitted my dissertation, I cried. You have been ridin' with me since Day One. I cannot believe how far we have come since the 14-year versions of ourselves meeting at the SHCP plaza all those years ago. We have seen each other through so many challenges and accomplishments and even when we are apart for long stretches of time, it feels like two puzzle pieces snugly being put together. Thank you for being the best at asking

questions, for keeping me curious, and showing me how to show up for family and loved ones. Thank you for giving me (and the world) Taylor Kailani who I am extremely grateful to and for. Love you both so much (beyond words).

Sara Schestenger, thank you for being such a light and a force in Abram's life. I see exactly where he gets his love, passion for life, curiosity, and smile. He is, truly, your child and you his mother because I see so much of the love you have given to him emanate out into the rest of the world, through his scholarship, artwork, and all the ways he engages with everything he encounters. I see you in him and I am grateful to have you in my life. Thank you for embracing me, in all the ways that you do. Zalman and Radha, thank you both for welcoming me in your family and for loving your brother the way you do. I am looking forward to all of the quality time and learning even more from you both.

Special thanks to the Resplendor, Protesta, Barlongo, Establecida, Wilkes, and Santos families for your constant support and encouragement. Immense gratitude to Priscilla Lumada and Paul Santos, for making me laugh and being proud of me! Let us continue making more memories together for many years to come. To Mel, Ed, and Nina, thank you for constantly making life fun with all of the laughter, impromptu dances in your living room, and

binge watching shows like *Bling Empire*. I love you dearly and so grateful for your love and support.

Celestino and Eleanor Resplandor, although you both and my dad are no longer in this dimension, I feel your presence constantly in my life. Tatay, you were the strong silent yet extremely tender one. A man of few words, but when you spoke, you could silence a room with your wisdom. Nanay, I used to get upset when you would sing while I was trying to do my homework because I felt like I needed complete silence. The irony is that all I ever want to do now, as a full fledged adult, is listen to music while I work and sing when I am not working. Thank you both for helping my parents raise this overly enthusiastic and energetic kid in the heart of the Mission and the Bayview districts. For putting in all of those hours at work just to support the family. I hope, wherever you two are, you see the work I am doing and are proud. Until I see you again on some other plane, mahal na mahal kita.

James E. Santos, you have got to be the most complicated person I have ever known! Even through all of the heartache and arguments, I always love(d) you and care(d) so much about your approval. I know neither one of us had the language to communicate in the ways we both might have wanted to when you were around. What I am about to say might sound so overly cliché (it is) but there is not a day that goes by that I don't think about you. Every

day that passes, I remember the lessons you taught me growing up and all the ways you wanted to protect me. I am forever grateful that you were/are my father. Through you, I learned fortitude and integrity in all that I do. Love you and I hope you have forgiven all the times I upset you because I really never wanted anything other than your love.

To my mother, Dulcencia R. Santos, thank you for giving me life (in all the ways you do) and all of the sacrifices you have made so I could be where I am today. My work centers so much of our relationship, but most importantly, I hope you see how it honors your journey and the stories you have shared with me. Your vulnerability, compassion, and unwavering love, and extreme generosity are only a fraction of who and what you are to me and so many others in our family and whoever encounters you. Thank you for allowing me to live my life, for trusting my judgment, and all of the times you have loved me especially when I'm being difficult. I know I have not been the easiest to deal with yet you are the one person that loves me unconditionally. And I have said this time and time again, you have never broken my heart, but I most certainly have broken yours through my words and actions. I truly am nothing without you. Thank you for never making me feel like a mistake or a burden. Since I was a child, you have always made me feel like the Sun. You have

always made me feel like I was special even when I felt far less than special. I love you with all of my being.

Abram Stern (aphid), now, I can write an entire book about how I feel, but I will not since we are already working on Hidden Labors and Sonic Tensions and so many other things together. I have said this to you before, but your heart is far bigger than your body is able to contain. It slips out when you are helping your students, your friends, and most certainly shows itself when you are supporting me. From your sharp wit and genius next level word play to all of the machinations of that beautiful brain of yours, you have always found the time for rest, reflection, and adventure. You have always found time for me. Thank you so much for all the times you have shown up for me, for yourself, and for us. I have never encountered a human being as gorgeous on the outside as they are on the inside. You have taught me patience, understanding, trust, and how to love in ways I only imagined. From a thousand lives to hundreds of multiverses, believe it, I would go looking for you in all of them. Thank you for teaching me to love in the present progressive. Always. I am loving you.

This dissertation is dedicated to Dulcencia R. Santos and Abram Stern (aphid) and in memoriam, this work is a tribute to James E. Santos, Eleanor and Celestino Resplandor.

Introduction

Listening before Speaking

Spanish-speaking neighbors' speech would waft into the cracks of open windows. For years, I lived in San Francisco's Mission district with my family. Tagalog emanated, on a daily basis, from different parts of the railroad-style apartment where my parents and maternal grandparents lived. The ebb and flow of their speech patterns became indelible. Whatever feelings and emotions they might have felt at any given moment, I would have known through the pitch, tone, and timbre of their voices. From my mother's angry voice dropping an octave calling me by my full name instead of the endearing nickname she had for me to my father's singing along to a song on the radio, their voices seemed to ricochet off the walls of the corridor and into the living room. If we were lucky enough to hear the laughter and Spanish from the next door neighbors, it all meshed together to become a cacophony of joyous sounds.

The familiarity of a voice comes with its own prosody along with the ways we might attribute meaning to those patterns. It is also determined by the environment, the language one learns, and may be affected by the

vernacular of the culture and subculture an individual participates in and identifies with. Prosody is defined as the intonation and animation of a human voice.¹ Growing up in a multilingual household, my parents deterred me from speaking Tagalog. But I learned how to listen, read, and translate conversations from Tagalog to English as a way of helping my parents assimilate into American culture as smoothly as possible. According to linguistics scholars Jeff McQuillan and Lucy Tse, I was (and from time to time, even as a full-fledged adult, remain) a “child language broker.”² A common belief amongst Filipino immigrant parents was that teaching Tagalog would prevent their Filipino American-born child from sounding American.³ The strange occurrence of learning Spanish and English concurrently, yet speaking both languages with a Tagalog accent was surreal yet indicative of the many instances of the intersectionality of everyday speech, vernacular expressions, and conversations within my family. It was a common phenomenon amongst many other immigrant families in such a densely populated region such as the San Francisco Bay Area. Having grown

¹ “prosody.” *Merriam-Webster Dictionary*.

<https://www.merriam-webster.com/dictionary/prosody>.

² Jeff McQuillan and Lucy Tse, “Child Language Brokering in Linguistic Minority Communities: Effects on Cultural Interaction, Cognition, and Literacy,” *Language and Education* 9, no. 3 (November 4, 2009): 195–215, <https://doi.org/10.1080/09500789509541413>.

³ E.J.R. David, *Brown Skin, White Minds: Filipino American Postcolonial Psychology* (Information Age Publishing, 2013), 129.

up in a predominantly Latino neighborhood surrounded by a multitude of ethnicities that comprised San Francisco, Oakland, and Berkeley alone; Spanglish, Taglish, and many other permutations of hybrid languages were prevalent. The enclaves of neighborhoods within the seven by seven square mile radius provided a rich and expansive linguistic experience for any child born in the region, especially around the 1970s and 80s.

The corporate world would soon offer a jarring cultural awakening with a polar opposite cast of colleagues from a wide array of academic and professional backgrounds where language and accent bias permeated through formal meetings and after-work gatherings. Walking the hallways of one of the largest biotech firms in the world with large windows of futuristic architecture overlooking the Bay Area peninsula, I witnessed how accent reduction coaches, often white women, were brought in to serve as speech coaches to the scientists and biostatistical programmers with discernible Asian accents to ensure they were taught Standard American English (SAE). Their tongues twisted words around until they were deemed correct, what scholar Halcyon M. Lawrence has described as a disciplining of the body through one's speech.⁴

⁴ Halcyon M. Lawrence, "Siri Disciplines," in *Your Computer Is on Fire*, ed. Thomas R. Mullaney et al. (MIT Press, 2021), 179–97.

(Hypo)thesis on Linguistic Consciousness

Yet many people around the world are unable to afford a personal coach to teach them language, culture, and a specific way of speaking to the western world. In the 21st century, social media has taken on many roles in the dispersion and dissemination of video content for educational purposes. Platforms such as Instagram, Facebook, YouTube, and TikTok have become a boon for creators and viewers alike. Influencers provide countless tutorials and resources for individuals seeking ways to learn English, finding a sense of belonging, and most importantly, securing employment. While these platforms have harnessed the power of collective and participatory communication, they have also created a tension of performative speaking and accent reduction that aims to assuage the fears and insecurities of a foreign speaker to an anglophone audience. These endless streams of content situate a non-western accented speaker to engage in a tension far beyond their own existence in the contemporary, digital world that reverberates centuries of colonialism and imperialism. With no formal instruction, the video content a non-English speaking viewer watches, listens, and practices may result in an anxiety or apprehension of the content being learned in the correct and intended manner. In essence, there is no way to know for certain until what is

learned is tested within a professional setting or off the screen and out in the world. Real life scenarios combined with cultural constructs may also result in quiet observation, which is another form of learning within context and watching others as a sign of what is acceptable or normalized speech.

Silence between utterances, words, and speech acts are also a subservient act for people unwilling and unable to speak directly to the western world, which has often dictated how technologies are created and programmed. To disambiguate, I define speech acts as a word or words strung together to be understood and based on the context of an exchange of language as opposed to the linguistic definition of a word or phrase taking on an active meaning in its pronouncement (i.e. to proclaim is a proclamation itself, etc.). For the purposes of this research and my artistic practice, I am more concerned about performative utterances and speech acts that have meaning in what has been done and/or what is in the present tense since the majority of my creative practice is concerned with docu-poetics (events having happened).⁵ I seek to understand how speech and language have been codified to reveal a lineage that has served as the foundation for the programming of contemporary digital voice recognition and speech technologies.

⁵ Mitchell Green, "Speech Acts," ed. Edward N. Zalta (The Stanford Encyclopedia of Philosophy, 2021), <https://plato.stanford.edu/archives/fall2021/entries/speech-acts/>.

With the development of accent elimination softwares such as Sanas AI, however, American and British Englishes not only stand to remain the global lingua franca, but enable the persistence of whiteness around the world. For instance, raciolinguistic ideologies that involve practices and pedagogical approaches to language and accents engaged in classifying “deviant” or considered belonging to “racially minoritized groups” outside of the “Standard American accent” most associated with the US midwest perpetuate and prioritize a white speaker and listener.⁶ If call center agents’ voices become mediated and disciplined through the lens of techno-solutionism, a term popularized, but not invented, by writer and researcher Evgeny Morozov. This framing entails viewing problems as solvable through technology yet frequently perceived solution(s) results in even greater issues and problems.⁷ How might we reconsider our approaches to how English is learned and spoken around the world in and outside of educational institutions and contexts versus in the home?

In comparison to the over 7 billion people that populate the world, approximately 450 million of those citizens are born into English-speaking households as opposed to the over 2 billion that are taught English in a

⁶ Pooja Rangan et al., eds., *Thinking with an Accent: Toward a New Object, Method, and Practice* (University of California Press, 2023), 41, <https://doi.org/10.1525/luminos.148>.

⁷ Evgeny Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism*, First edition (New York: PublicAffairs, 2013), 6.

classroom.⁸ Yet an educational setting neglects vernacular, cultural references, and prosodic elements indicative of an individual's regional upbringing and environment. The most tangible outcome of the aforementioned neglect has been extensively researched by linguistic scholar April Baker-Bell and her concept of Black Linguistic Consciousness.⁹ Although the focus of this research is on the training of human voice and smart technologies, her approach of working with Black youth, Black Language, and “white linguistic hegemony” deeply informs the decolonial approach that has inspired both my academic research and creative practice.¹⁰ Baker-Bell's scholarship regarding Anti-Black Linguistic Racism gave way to her development of the Antiracist Black Language Pedagogy that encourages Black youth to understand that their speech acts ought not be othered by white mainstream English.¹¹ In a larger context, non-western English speakers are also subject to similar linguistic racism and accent bias, which leads to the related and complicated nature of postcolonial subjecthood of Filipino people through global and racial capitalism. I situate the aforementioned postcolonial subjectivity as the Philippines' independence from the United States in 1946 and the aftermath of

⁸ Gregory Warner, “How to Speak Bad English,” Rough Translation, accessed June 19, 2022, <https://www.npr.org/2021/04/21/989477444/how-to-speak-bad-english>.

⁹ April Baker-Bell, *Linguistic Justice: Black Language, Literacy, Identity, and Pedagogy*, NCTE-Routledge Research Series (New York: Routledge, 2020).

¹⁰ Ibid.

¹¹ Ibid.

the imposition of English as a primary spoken language and continued occupation by US forces.¹²

This work investigates the human voice as a type of invisible data mediated through voice recognition platforms, systems of care, emergency infrastructures, and through the creation of assistive technology. It explores the various mechanisms that create a standard voice often associated with authority and care. This research also looks at the creation of accent elimination software, artworks that subvert speech technologies, and reflections of developer blogs, open source software, and sandbox culture as media artifacts through the lens of the non-programmer.¹³ My investigations incorporate a combination of scholarly components, and creative interventions through prompts as a way of understanding what artist and media scholar Christina Corfield notes as a process of “thinking through making.”¹⁴ Exploring the issues of voice, speech generation, and how they are codified allow for embodiment that cannot easily be conveyed through text alone. The use of human voice in my artistic practice is also integral to

¹² E.J.R. David, *Brown Skin, White Minds: Filipino American Postcolonial Psychology* (Charlotte: Information Age Publishing, 2013), 41–42.

¹³ I identify myself as a non-programmer in order to reflect from a perspective of how this particular open and accessible media is read by someone who does not have formal technical (programming, computing) expertise and training.

¹⁴ “Christina Corfield Tells Stories of Past Technologies Through Analog Means,” KQED, October 26, 2020, <https://www.kqed.org/arts/13888165/christina-corfield-tells-stories-of-past-technologies-through-analog-means>.

understanding concepts such as tone, pitch, and prosody. These elements of voice require a body in order to be contained. My artistic work reflects the contentious nature and subversion of software technologies and language learning resources and how they are presented to various types of audiences from the venture capitalist investor to the non-English speaker and 911 telecommunicators to call center agents.

Language is a key aspect of this research since there are institutions and organizations that aim to provide services for “accent reduction” for those interested in pursuing work as a call center, customer service, or virtual agent. Linguistic mimicry begins to unfold on a transnational level, which further reveals the intersections of race, gender, and class. Looking first within the telecommunications world, Bell System specifically, the human voice was associated with a phenotype, race, and gender. This work also investigates the western standard English speaking voice that has become the lingua franca for assistive technology. Feminist scholar Kalindi Vora’s concept of “vital energy” or biocapital resource for an increasingly transnational and global economic infrastructure of information and big data is what digital and smart technologies increasingly rely on.¹⁵ This energy serves a complex

¹⁵ Kalindi Vora, *Life Support: Biocapital and the New History of Outsourced Labor*. (Minneapolis: University of Minnesota Press, 2015).

infrastructure of care and thus requires a body to willingly participate within this networked system.

The Filipino voice serves as a cultural and economic case study to show the effects of racial and global capitalism. Under a neoliberal rubric, a corporation is perceived and understood by the state as an individual with the agency to operate within free markets under an assumption that there is a fundamental good at stake or trickle down economics.¹⁶ Even without Silicon Valley as a metaphorical, technological global mythology on its own, the notion of a fundamental good under a global capitalist system is a myth within a myth. Yet the Filipino call center agent becomes, what theorist Manuel Castells calls self-programmable labor.¹⁷ This type of labor, within Castells' network society, is work that can, possibly, be executed by and through automation and, if not, often given to immigrants and women.¹⁸ I posit the Filipino worker based in the Philippines as a type of proxy within the telecommunications industry since voice can take on various forms of embodiment and performance. When a Filipino proves their linguistic

¹⁶ David Harvey, *A Brief History of Neoliberalism* (Oxford: Oxford Univ. Press, 2011).

¹⁷ Manuel Castells and Paul H. Nitze School of Advanced International Studies, eds., *The Network Society: From Knowledge to Policy* (Washington, DC: Johns Hopkins Center for Transatlantic Relations, 2006).

¹⁸ Ibid.

closeness to Standard American English, it enables access into economic opportunities.

From where we find ourselves in the 21st century and harkening back to the history of how voices were trained to speak, can we surmise how we have arrived to the present time and how we might think of the future through the digital and smart technologies as they are programmed? Situating this research between the historical moments of the Chinese operators in San Francisco as community builders to the Black women desegregating the workforce at Bell System, and the rapid rate of globalization of the Filipino workforce into the western world, I argue for the development of speech corpora and technologies predicated on an expansion and preservation of global languages and the possibilities of open source software developing these technologies through a social justice lens. I advocate for transparency and design justice in the development of softwares and hardwares related to voice recognition and assistive technologies with accessibility and inclusion as core tenets as opposed to creating for capitalist and ever-present colonialist means and ends.

Despite this (colonial and paternalistic) closeness between the United States and the Philippines, it may be worth noting the definition of the Overseas Filipino Worker (OFW) and how it has changed significantly since

the 1990s taking into account globalization, offshore and nearshore outsourcing.¹⁹ With the hopes of a more decentralized internet in terms of the dispersion and dissemination of information, mostly American and European conglomerate companies build relationships with business process outsourcing (BPO) companies that find the labor force to do highly administrative tasks that are below minimum (American) wage.²⁰ The national identity for the Filipino worker becomes fragmented. It transforms into a pastiche of a postcolonial subject and a performative, well trained, disciplined body, especially through voice. Although our contemporary world presents a highly connected and mobile workforce, there is “no purely digital economy and no completely virtual corporation or community,” according to scholar Saskia Sassen.²¹ Social media platforms serve as both a place to obtain knowledge while simultaneously dictating how businesses operate overseas for the benefit of the western world. In addition, softwares are developed with a bottom line and profitability in mind and not the individuals that will inevitably operate its intricate operations especially when it comes to direct consumer support.

Frequently, software is deployed and technical assistance is left up to workers

¹⁹ Jan M. Padios, *A Nation on the Line: Call Centers as Postcolonial Predicaments in the Philippines*, 1st ed. (Durham: Duke University Press, 2018), 16.

²⁰ Low (minimum) wages in the United States are considered living wages in the Philippines as well as other Asian countries and South America.

²¹ Saskia Sassen, *Territory, Authority, Rights: From Medieval to Global Assemblages*, Updated ed., 4. print., 1. paperback print (Princeton, N.J.: Princeton Univ. Press, 2008), 297.

expected to possess a high level of expertise that is usually self-taught through the wide array of instructional content found online. With such fragmentation, the work of the west becomes further dematerialized and absorbed by and through the systems of displacing the physical work, social and emotional labor on Filipino overseas workers.²²

At present, telephone operators and call center services have either been automated through voice bots or outsourced to foreign countries such as India and the Philippines. Scholar Jan M. Padios chronicles the deep-seated effects and legacy of colonialism, imperialism, and globalization of the Philippines. She notes as of 2011, the country surpassed India and became the “world’s capital of call centers.”²³ This economic shift requires the Filipino to perform in what I call a western vocal drag in their home of origin thus producing a tension between perceptions of the western world being at odds with Filipino cultural ideals and identities. It is a crucial and necessary aspect of this project to examine these relationships to build socially just platforms that take into consideration these frictions.

²² Although Sassen notes dematerialization as a consequence between the this binary of digital and nondigital and notes “material conditions” as key attributes that create the Internet at the time her book *Territory, Authority, Rights* was published, this concept holds true even more so in our contemporary moment since entire cities and districts in the Philippines have been constructed as sites for these offshore outsourcing companies.

²³ Jan M. Padios, *A Nation on the Line: Call Centers as Postcolonial Predicaments in the Philippines*. (Durham: Duke University Press, 2018).

Speech technologies are built up of a corpus of human voice data that represents predominantly American and British English voices with higher rates of speech recognition errors for non-native English speakers.²⁴ By situating this research between the historical moments of American telecommunications history at Bell System to the contemporary globalization of the Filipino workforce into the western world that continues to be subjected to the effects of American colonialism, I argue for the creation of socially just systems, platforms and technologies predicated on an expansion and preservation of global languages and a recognition that open source software projects are at the helm of creating these types of necessary infrastructures.²⁵

In thinking of oral traditions and preservation and diversity of languages, sonic perception and memories will be of the utmost importance in a world increasingly reliant on assistive technology. Through theoretical and historical lenses, I delve into the gendered labor histories, codified speech, and how the language of care is mediated through digital and mobile technologies within medical and healthcare infrastructures. The aforementioned framework is guided by a wide array of scholars including Venus Green, Ruha Benjamin,

²⁴ Halcyon M. Lawrence, “*Siri Disciplines*,” in *Your Computer is on Fire*, ed. Thomas S. Mullaney, Peters, Benjamin; Hicks, Mar; and Philip, Kavita. (Cambridge: The MIT Press, 2021), 191.

²⁵ Lawrence points to open source software contributors and developers leading these efforts that go beyond economic value and situate an expanded idea of how language and voice can work within software development.

Elizabeth Ellcessor, Beth Semel, and Hannah Zeavin. They provide frameworks and methodologies grounded in a critical analysis of digital technologies, machine learning and surveillance that incorporate race, gender, and class. Throughout the writing, I intermittently refer to media examples such as training films, artworks, and advertisements as a way to question the reasons how ideas of gender are codified through visuals, moving images, and aural experiences as potent devices that perpetuate our beliefs and perceptions of identities.

From telephone switchboard operators to contemporary contact tracers or disease investigators, Chapter 1 focuses on a historical contextualization that shows the genealogy of virtual assistants and call center workers through a feminist approach to media histories. In chapter 2 and through the lens of computational media, I write about the development of the How May I Help You Technology (HMIHY), which serves as a pivotal historical moment in contemporary voice recognition technologies. Artists such as Lauren Lee McCarthy, Ryan Kuo, and Johann Diedrick illustrate how computational media has been utilized within the realm of digital art. Considering the advancements in voice recognition, specifically interactive voice response systems, chapter 3 presents an examination of the voice through a Critical Medical Anthropology lens focused on emergency dispatching. The final

chapter delves into the transnational and global impact of accent reduction and anglicization of the Overseas Filipino Workers through Do-It-Yourself (DIY) media in the form of educational tutorials and corporate training. I also interrogate the creation of accent elimination software such as Bay Area-based Sanas AI that perpetuates the colonization of non-dominant, foreign accented voices.

Creative Interventions

My project *Do No Harm* (2017-ongoing) was developed and inspired by the format of clinical research protocols, informed consents, nursing telephone triage manuals and training documentation. These processes are re-imagined through a hypertext game experience, similar to a choose-your-own-adventure story. The premise of the story is based on a queer, gender non-conforming scientist engaged in creating a clinical trial speculating CRISPR biogenetic technology enabling the eradication of social and cultural diseases such as toxic masculinity, generational trauma, and xenophobia in addition to treating life threatening diseases. The literature (a combination of flash fiction and poetry) for *Do No Harm* aims to re-examine the format, structure, and design of scientific documentation by implementing

simple design, accessible language, and narrative as a way of addressing how clinical research and consent are, oftentimes, challenging, and inaccessible to the general public.

One of my sources of inspiration was this illustration done by Robert Tinney in 1984 of a speculative device called the Videotex (the interface is reminiscent of a smartphone screen) . The device is similar to the Minitel, which was popularized in France and was known mostly as a terminal that would provide information, but did not allow for dynamic interaction.

The first instantiation of work for the *Do No Harm* series was a sculptural piece called *Press 1 to Be Connected* (2019) that explored the interactive fiction format through a web interface that incorporated audio recordings. This work is inspired by IVRS and experiments with the phone tree as a medium for creating interactive speculative fiction. Since telephony has long been used to communicate and relay data and information, the work relies on both analog and mobile phones as the listening device. The phone tree is engaged reflexively to draw the listener into a particular narrative in anticipation of the listener returning to explore other outcomes that may incorporate new forms of storytelling and game mechanics. This project is meant to elicit the imagination of the listener as well as to enable their agency in directing the outcome of the narrative they are initially told.

In 2022, I received an artist and research residency at San Jose State University that culminated in a two-person exhibition with artist Sofia Cordova titled *Hiding in Plain Sight* (2023) organized by artist and educator Rhonda Holberton and curator Alena Sauzade. Cordova and I were tasked with selecting media artifacts from the special collections and archives of the university. I selected an interview of Nand (Andy) Khanna from the Indian Diaspora Oral History Project Collection, which included fourteen recorded and videotaped conversations with fourteen Indian immigrants who arrived in the San Francisco Bay Area between the 1940s through the 1960s.²⁶ In keeping with the form(at) of the protocol, I listened, watched, and pulled specific transcribed text from Khanna's oral history and recast myself as the interviewer. Preserving his text (verbatim), I wrote a speculative conversation based on the conventions and procedures adhered to by a relay operator to engage in a speculative docu-poetics work titled *Imaginary Relay, To Andy* (2023).

²⁶ Indian Diaspora Oral History Project Collection, MSS 2013-03-26, San José State University Library, Special Collections & Archives.

Voice-based Interactive Docu-Poetics

The Cyborg's Prosody is an interactive work, creative intervention, and response to the establishment of “accent reduction schools” specifically in the Philippines. This work is meant to serve as an “accent induction school” with the hopes of serving as a decolonial method of exploring how language is learned, human-computer interaction, listening, and how accent preservation is encouraged. In detail, the work itself is a web-based platform that will be designed and optimized for mobile use. Similar to the way a user must validate their voice for various smart technologies such as the Amazon Halo, the platform will take the player through a simulation of voice authentication. This validation entails the player reading excerpts of text (in English) related to human voice and language learning from French film theorist and composer Michel Chion’s whose work focused on the disembodied voice or writer Cathy Park Hong. Once a player has completed voice authentication, they will be prompted to select the first (and easiest level) of the stories. As the levels progress, the player is expected to mimic the Tagalog accented voice (or cyborg) in order to advance in the game.

The stories have varying levels of difficulty. The easiest level (first) entails repeating excerpts of a story, when prompted, read by a voice with a discernible Filipino accent. The most difficult level (fifth) requires the player to repeat back excerpts of a story told completely in Tagalog. In order to advance in the game, the player must repeat the story in the language and tonal quality of the reader. In essence, the player is being inducted into speaking the Cyborg's prosody.

Chapter 1: Holding the Line: Black and Asian Women in US Telecommunications History

The Euphonia: A Talking Machine

In the mid 19th century, inventor Joseph Faber created The Euphonia, a talking machine that aimed to mimic human speech.¹ He wanted to recreate an anatomy out of mechanical components and parts that was made operational by human command (by Faber himself). The functionality of the machine's coils, springs, and hardware were a manifestation of a type of automaton. He constructed the levers, keys, bellows, and pedals to create speech patterns flowing from the mouth of a feminine looking face replete with ringlets to make the gendered appearance of the machine more pronounced. Despite its exposed mechanical parts, the feminine form aimed to soften the hard edges and components of the device, reminding the viewer that technology is docile. The imposed feminine embellishments served as a signifier of an entity meant solely to exist for the purpose of being in service to

¹ Addison Nugent, "Text-To-Speech in 1846 Involved a Talking Robotic Head With Ringlets," *Atlas Obscura*, March 9, 2016, <https://www.atlasobscura.com/articles/texttospeech-in-1846-involved-a-talking-robotic-head-with-ringlets>.

its creator or user. Decades after Faber's invention, women's bodies were used to push telephone and telecommunications technology into the dawn of the computer age and positioned as dual roles of both active listener and responder. While it is documented that the machine's vocal affect was relatively monotone with little to no variation or intonation, its output would prove to the world that it would be possible to create a prototype of a "pneumatic speaking machine."²

Faber's thought behind his invention was its function as a talking machine, but not necessarily a device that might engage in what we customarily might recognize as a conversation where there is a reciprocity in speech between two people. It lacked prosody, which is another distinguishing factor in the way people speak with one another or to a group. While prosody may not have been the driving force or motivation for The Euphonia, Faber's invention was an exercise in the replication and transmission of speech. Its mechanical parts enabled a type of sonic, vocal, and linguistic experience that allowed people an entry point into both the spectacle of innovation as well as the scientist's engineering knowledge and skills. It was not a machine to be in conversation with, but an amalgamation of parts that aimed to impress an already skeptical general public. Similar to telephone operators of the early

² Thomas L. Hankins and Robert J. Silverman, *Instruments and the Imagination* (Princeton, NJ: Princeton Legacy Library, 1995), 214.

20th century, Faber was tethered to The Euphonia. His movements enlivened its “speech” despite its flat affect and vocalizations.³ The interface or hardware was analogous to the physical connection that telephone operators had with their designated portion of the switchboard. With the various speaking apparatuses, there is a strong visual imagery that evokes the way we might envision a type of craft such as weaving, looming, knitting, or crocheting. Bell System referred to telephone operators as “weavers of speech” because the operator’s body became inextricably linked to the technology (see fig. 1.1).⁴ Conversely, the voice that emanated from his invention produced little to no variation in prosody and may have been audible as gender-neutral. Faber’s primary focus was to see if he could construct a simulation of human anatomy responsible for speech. The most likely scenario was that he created a machine based on his own vocalizations, which further complicates his intentions of imposing feminine characteristics onto the machine.

³ Ibid.

⁴ *Training for Service, 1926 - AT&T Archives - Telephone Operator Recruiting*, Video (Bell System, 1926), <https://www.youtube.com/watch?v=QQbwRYCYJzI&t=2s>.



Figure 1.1: The end of the *Training for Service* recruitment film, which shows a young woman with a wire in her left hand looking to connect a call. She is also wearing the apparatus that allows her to hear and speak to a subscriber. The producers staged the scene to emphasize the role of the operator as a “weaver of speech.” (Film produced by Bell System, 1926, <https://www.youtube.com/watch?v=QQbwRYCYJzI>).

The early training and recruitment films made by Bell System in the late 1920s, of the company’s switchboard operators, were influential in creating the archetypal figure of the women’s role as “weavers of speech.”⁵ The metaphor of the operator’s body being tethered via speaking apparatus and cords spoke to the materiality of the machine, in particular, the switchboard. Their physical gestures of criss-crossing wires to connect calls was akin to craft, which explains Bell’s reference of the operators being akin to weavers.

⁵ *Training for Service, 1926 - AT&T Archives - Telephone Operator Recruiting*, Video (Bell System, 1926), <https://www.youtube.com/watch?v=QQbwRYCYJzI&t=2s>.

This visual cue served as an inscription of femininity and hierarchy within the network. The majority of women that worked at Bell System were young, American-born, and white. The company's voice was synonymous with the women operators and they were expected to uphold specific types of speech and affect. To keep this image of the company, Bell hired women of the same size, stature, and similar vocal affect. Women who identified as Black, immigrants, or from Southern states in the US were often not hired or given opportunities to work within these roles.⁶ The hiring practices changed much later as Black (US) Southern and Caribbean women were hired because they were perceived as more obedient due to their need for a full time job or fear and threats of deportation.⁷

Through archival research and visual and textual analysis of two promotional films representing the telephone exchanges, chapter one illustrates how telephone operators were depicted on film and how Black and Asian communities contributed to telecommunications history that are vital to understanding the evolution of human voice from analog into the digital age. Through feminist media and labor histories, the two distinct communities were culturally and socially excluded from dominant white American spaces

⁶ Venus Green, *Race on the Line: Gender, Labor, and Technology in the Bell System, 1880-1980* (Durham: Duke University Press, 2001).

⁷ *Ibid*, 134.

that privileged the perpetuation of a specific type of vocal affect. These historical case studies provide accounts of the telecommunications industry that provide a foundation for the development and design of synthetic voices for contemporary assistive technologies. This chapter provides a brief history of the telecommunications industry through a feminist media historical lens and how womens' voices were trained thus connecting to the contemporary global racial linguistic capitalism we are witnessing along with the advanced digital technologies perpetuating a specific type of voice.⁸

I examine the training of telephone operators, and the voice as a type of interface. One of my guiding questions: How did race, class, and gender translate into the profiling of operators and how did the demise of the "white lady" image significantly change the telecommunications industry and labor involved? Regarding the racialized and gendered history of Bell System, labor historian Venus Green states in her book *Race on the Line*,

Even though race was a defining characteristic, the "lady" operator had to possess other qualifications. For at least the first fifty years, the majority of telephone operators were usually young, single, native-born [by native-born, Green means born in America] white women who lived with their parents. Census documents and internal AT&T records

⁸ Inspired by Pip Thornton's use of linguistic capitalism despite her focus on Google Ads and the intervention of poetry as a way to explore this term. My hope is to explore her use of the term through the voice and speech and draw relationships between and within our respective fields.

reveal that the Bell System placed a premium on youth for reasons of economy, discipline, and perceived physical stamina.⁹

In researching the creation of the Chinese Telephone Exchange (CTE) in San Francisco's Chinatown founded by Loo Kum Shu in 1902, which was bought by Pacific Telephone & Telegraph in 1949, I rely on my findings at the AT&T Archives and History Center located in San Antonio, Texas to learn more about its place in US telecommunications history.

I include the historical context of Black women operators that worked to desegregate Bell System and fight for fair wages and against racist hiring practices between the years 1910-1970 through internal (community and labor) organizing that happened concurrently to the CTE operating out of San Francisco. Black and Asian women were in service and on the periphery of whiteness and expected to create and sustain their own systems against the backdrop of the largest telecommunications companies in the world. I delve into these particular worlds within telecommunications and media history to recognize the common thread between these two groups entailed going against the typical, feminine white Americanized English-speaking voice.

A type of performativity was also involved in both groups for economic survival as detailed in labor historian Venus Green's scholarship on how

⁹ Green, *Race on the Line: Gender, Labor, and Technology in the Bell System, 1880-1980*, 134.

operators were trained and the rigorous nature of telephone operating.¹⁰ It is important to note that both whiteness and femininity are constructs that are taught and acquired through psychosocial conditions and homophily. A particular and specific able-bodied woman was expected to operate the switchboard with ease, agility, attention to detail, stamina, and endurance.

With Bell System's expectation of women's bodies as mediators and conduits of communication, Katherine Schmidt established a telephone operator school in 1902, which played a role in the evolution and elocution training for white women operators.¹¹ They were believed to be better suited than young men to perform the job of a telephone operator due to the perceived ability to be in service to the subscriber, which were predominantly male business owners. Along with men's propensity towards hearing a feminine voice, World War I greatly contributed to the influx of women joining the workforce as telephone operators because it was seen as a respectable occupation. Although race and gender constructs are not innate qualities and characteristics, again, to overemphasize, these aspects of feminine identity were taught and imposed upon the operator's body. Schmidt is the first woman responsible for creating the necessary training needed for this role. The school was geared towards young women between 18-24 years of

¹⁰ Ibid, 53.

¹¹ Ibid, 61.

age.¹² Right away, there were specific aspects of the type of “lady operator” brought in to do this work and how they were trained to speak and communicate with subscribers. The strict training included ensuring the appropriate “body adjustments” were made in “shaping ‘the use of the tongue, lips, jaws, and posture that would result in proper pronunciation and the tone of eager friendliness.’”¹³ Media scholar Elinor Carmi wrote that operators were perceived as both “the medium and the noise source,” which necessitated training to ensure the highest quality operator serve for Bell System.¹⁴ Moreover, Schmidt, Bell’s first female supervisor, described the role of the operator as a “paragon of perfection, a kind of human machine.”¹⁵ Yet Black women and Chinese operators were not afforded the same access to training or education within the same corporate structures, therefore adding a layer of practices and procedures that navigated around these already oppressive existing training protocols and systems.

At Bell, most, if not all, operators that were hired were taught not only how to speak but also how to sing. They were taught to emote through

¹² Ibid.

¹³ Elinor Carmi, “Taming Noisy Women: Bell Telephone’s Female Switchboard Operators as a Noise Source,” *Media History* 21, no. 3 (July 3, 2015): 313–27, <https://doi.org/10.1080/13688804.2015.1045468>, 316.

¹⁴ Ibid, 315.

¹⁵ Ibid, 316.

pleasant tones with the ultimate goal of being a “voice with a smile.”¹⁶ Specific tonal qualities were needed in order to speak with subscribers. They had to be both at ease and efficient, but with a vocality that aimed to please. Even the stature and size of the woman (operator) was a key factor into whether or not they were hired.¹⁷ The physical attributes were perceived as a necessary aspect of the job since the operator’s physical body was connected to the headset apparatus that then captured her voice for transmission coupled with the crisscrossing gesture required to operate and function the switchboards. In terms of media archeology, technology is then designed specifically for the users.¹⁸ In addition, the operator had to be able to sit, speak, and activate the machines and lines for the entirety of their shifts or tours, which ranged from nine to twelve hours.¹⁹

The role of “lady operator” was racialized, gendered, and based on class due to specific accents such as Black women from the U.S. South to immigrants that were not educated in the US. These women were not permitted to apply or be hired for the role of operator. Racialization has a

¹⁶ Relevant to note Robert Hopper’s book *Telephone Conversation* outlines the various ways telephone correspondence changed due to the evolution of telephone technologies.

¹⁷ According to a discussion with technical communications scholar Halcyon M. Lawrence, an exploration and relationship between the vocal folds and quality of sound produced by an individual might be an aspect of voice attributed to specific qualities and tones of voice worth further exploration to further supplement this existing research.

¹⁸ Adam Greenfield, *Radical Technologies: The Design of Everyday Life* (London New York: Verso, 2018).

¹⁹ Green, *Race on the Line: Gender, Labor, and Technology in the Bell System, 1880-1980*, 70.

direct correlation to class as well in that most (white) women operators had to be educated enough to do the job, but not overly educated so that they didn't want to do the work. They became extensions or auxiliaries of what the company offered in terms of customer service.

Green's archival research, with a special focus on labor history, offers a useful and multi-valent perspective on the intersections of race, gender, and class, which played a significant role in the organizing and desegregation of Bell System. There will be a much more expansive and global perspective that reaches outside and beyond the U.S. to explore western vocal performances that connect to the lineage of telecommunications and how women were trained to speak to the general public and the type of temperament that was required for this work. Since language is learned through mimicry, media analysis of training materials illuminates how the voice may be corrected and how ways of speaking might be suppressed. As Green points out, many of the operators learned how to speak through seeing how seasoned operators handled the switchboard and navigating multiple callers in addition to formal training organized by the company. Yet the understanding the media produced to recruit young women to Bell System is crucial to gain a sense of how women were expected to perform gender, voice, and, most importantly, become disciplined, servile bodies beyond the walls of the switching room.

Training for Service Recruitment Film: A Mediation and Projection of the Perceived Ideal Body and Voice

In Bell's recruitment film, *Training for Service* (1926), the film details the necessary steps involved in becoming a telephone operator. For many of the silent films of the early 20th century, intertitles indicated what the viewer was looking at. These textual markers served to bridge the gaps of what could not be easily depicted in moving images. Most important to note, the films of the 1920s were of the silent film era and devoid of sound. The irony of Bell's film is that it centers the role of an operator with no voice(s). The film immediately starts with upbeat, whimsical music that serves as the musical score for the film. Yet the introduction of sound in film did not happen until the following year, which makes the inclusion of a score a contemporary addition that changes the overall tone and affect.²⁰ The movements are highly gesticulative and exaggerated to articulate the film's purpose. The inability to hear the women's voice brings about this layer of information and speculation of what it must have been like to be a young woman interested in this line of work. She would have to go down to the office, meet someone in person, and

²⁰ Timothy Corrigan and Patricia White, *The Film Experience, 5th Edition*, (Bedford: MacMillan Learning, 2018), Retrieved from Kindle Reader.

progress through quite an embodied experience and set of protocols for gainful employment. The viewer's introduction into the role of the operator within the first few frames included a film overlay of an image of a feminine woman facing a room with an instructor, which indicates the media as a direct form of the environment. Media making for a telecommunications company is a part of the work, not only the hiring of, but also the creation of an archetype.

White femininity is a pronounced construct throughout the film signaling a specific type of women and the creation of a work culture that was not only homogeneous, but also was intended to set the conditions for a type of homosociality. In addition to Bell setting the scene, the filmmaker ensured the viewer's ability to see side profiles or a full face within a frame. The viewer is able to see the speech, yet, at this time, has to imagine the sound of what might be happening at the operating station. Strangely, the tension also depicts as much of the physical experience to be able to articulate for the viewer, and potential job candidate, what was deemed acceptable. When thinking about Bell's objective, the voice was not overtly centered as the primary focus of the job and that is due to the company including voice training, which was shown through a classroom setting. Bell System includes key visual elements: the work environment, office, and the (telephone)

operating room; intermittent text that explains what the viewer is about to see related to the training of the operator, and representation of the overall hiring process, and formal oversight of the women hired.

This form of media serves as a recruitment film, which poses not only a lucrative and respectable job, but it is meant to blatantly show the type of body needed or perceived as the ideal body to do this work. The assumed audience of Bell's training film was aimed at the young (preferably) white woman interested in becoming a telephone operator. Circling back to the homosociality of the film, the recreational time of the operators playing the piano and dancing with one another is included as if to suggest community and the common thread of these women having been hired for their collective uniformity.

The film is akin to a contemporary virtual tour of the training environment that the potential operator is going to join should she apply and be accepted. The first intertitle reads, "Every year thousands of young women are trained as specialists in the operation of our modern complex telephone switchboards / Fascinating and important it is, and highly interesting to the operator, from the very first day when she enters the Employment Department to apply for a position" followed by what appears to be a young woman dressed for cold weather opening a door to the employment office. As

she enters, the shot fades to her entering into a waiting room where other women are gathered and wait patiently with text that reads, “As an applicant she finds a friendly welcome awaiting her, in keeping with the dignity of her new work.”²¹ The words “specialists” and “dignity” put forward that the job is not only respectable, but her ability to be both accommodating and knowledgeable are keys to her success.

The protagonist is met with an associate that leads her to a desk to complete the necessary paperwork with the next block of text reading, “From the employment supervisor she gets her first glimpse of how highly her profession is regarded.” Yet the viewer is not provided a visual cue or gesture of what exactly makes the profession “highly regarded” other than the candidate receiving authorization to proceed with a physical health examination. I speculate that what makes the job appealing is how competitive it might have been to be offered the role knowing that the film aimed to target the type of women Bell was seeking to fill these positions. In addition, the racialization of such a coveted role was apparent through the company’s well-documented hiring practices that prevented Black women from becoming operators as early as 1910.²² The absence of Black women and

²¹ *Training for Service, 1926 - AT&T Archives - Telephone Operator Recruiting.*

²² Venus Green, *Race on the Line: Gender, Labor, & Technology in the Bell System, 1880-1980* (Durham, NC: Duke University Press, 2001), 195.

other women of color in the recruitment film provides an overt perspective of what the company valued and how they wanted to be perceived as well as the type of body that was preferred despite the operators not being visible or seen, but only heard. Bell's hiring practices actively denied immigrant women with strong accents unless they were from other English speaking countries such as Canada, Ireland, and England.²³

Each major step involved employing an operator is captured through the text interludes of the film followed by an enactment of the steps involved on the first day of employment. "Simple examinations at the medical department assure an applicant, and the company that she has the necessary physical qualifications," which alludes to the physical nature of the job and what might be involved of the operator for every shift or "tour" of a full day's work (see fig. 1.2). The viewer also sees different young women with relatively homogenous features and hairstyles abound in the shots that show a community of women that worked at the switchboards. In one scene, a woman is in a nurse's office undergoing various types of testing that include documenting vital signs such as temperature, heart rate, weight, height, and visual acuity. Another text appears, "The student operator is at once made to feel at home in the Training Department," which indicates that the work of a

²³ Ibid, 62.

telephone operator is similar to the docility expected within the home. The workplace is painted as another kind of home and as a place where the woman's body is in service, yet connected to a larger mechanism where her body is disciplined into performing specific gestures, actions, and vocal affect. The body becomes another conduit for information, but it is a specific type of (white, feminine) body that serves as a container or mediation of information and transmission.



Figure 1.2: This screen capture shows a nurse documenting the height of a newly hired telephone operator. (Film produced by Bell System, 1926, <https://www.youtube.com/watch?v=QQbwRYCYJzI>).

In another scene, although the dialogue is silent throughout the training film, the viewer might assume the operator's agreement and

understanding of the job she has just accepted. A supervisor leads the young lady at the beginning of the training film into a women's locker room where she is expected to keep her belongings.²⁴ Capturing the environment of the operators where they were expected to both socialize and work served as a way of gendering employment and class assumptions that also depicted Bell's hopes to paint the work as a revered job for respectable young women. The film reinforces how capitalism disciplines the worker's body in performing and serving as a key component of the business. There is a brief moment (3:24) where the viewer sees the young woman looking in the mirror prior to leaving the women's locker room. Despite callers not being able to see the way she looks, she straightens her clothing and touches her soft curls as if preparing to be seen by clientele or, perhaps, a surprise visit from management (see fig. 1.3).

²⁴ Her belongings are locked away by the supervisor and the student operator is not provided with a key, which indicates the necessity of surveying the operator's work.



Figure 1.3: A little off center to the frame, this full shot depicts a newly hired operator looking in a mirror and fixing her hair before she receives her equipment for the switching room. (Film produced by Bell System, 1926, <https://www.youtube.com/watch?v=QQbwRYCYJzI>).

As mentioned earlier, Faber's literal connection to *The Euphonia* in order for it to function was similar to the necessity of the telephone operator's body to the switchboard. Introducing us to another scene of the operator's first day on the job, the screen reads, "Early in her training she becomes familiar with the tools of her profession" to show exactly what is physically involved and required of the role.²⁵ The viewer sees the new student operator meeting with one of the workers who proceeds to show her the apparatus needed while working at the switchboard. The object is worn like a necklace with a receiver or funnel extension that the operator is expected to speak into. The apparatus

²⁵ *Training for Service, 1926 - AT&T Archives - Telephone Operator Recruiting.*

is an extension or an auxiliary to the human body and transmitter of voice or “the attractive headset soon becomes her symbol of service.”²⁶ The headset is far from attractive yet becomes synonymous with the way in which the operator speaks into and through the cables and switchboard. The device also includes a belted strap that ensures that it is securely fastened and connected to the torso of the body. “Later, through class room training the student learns how to care for it properly,” which indicates a type of comportment and identification with the technology itself.²⁷

Elocution lessons served as a way to train white women into performing as the voice of the company. Once in the classroom, we see the instructor wrapping her receiver and the camera then shows a group of new student operators mimicking the actions of the teacher. “She finds her voice and enunciation improving as she practices ‘the voice with the smile.’”²⁸ The teacher points to a chalkboard with words and as we watch her, her vocal gestures are exaggerated to emphasize the speech necessary for the job. The “lesson” is in how to pronounce the numbers (e.g., 0 is “oh” and not zero). Specific to elocution, an intertitle reads, “Through training in the art of inflection she gains in those gentler qualities of unfailing courtesy so essential

²⁶ Ibid.

²⁷ Ibid.

²⁸ Ibid.

everywhere.” The new operator also learns how to invoke prosodic elements to ensure these “gentler qualities.” We see full shots of the students and instructor providing exercises written on the chalkboard, {Number please [sic] / John 4567 / Melrose 5069,}.²⁹ At this point, students are expected to stand up and repeat the exercise. The film proceeds with “Her first instruction in the operation of a switchboard is given in the class room [sic].” We see yet another instructor showing a large vertical machine used for training purposes, which shows how the operator is expected to work with the technology.

Homosociality was also another aspect of the role since the women had to work on a massive switchboard and possibly cover for one another on any given occasion. “Between study periods she begins to make friends in the rest rooms—all central offices have them.” The “rest rooms” are synonymous to break or recreation rooms. The women are shown dancing, laughing, and engaged in a type of social hour with one another (see fig. 1.4). The viewer is reminded of the homogeneity of the “lady operator” by emphasizing through a text cue that “classroom instruction is followed by interesting work at the practice switchboard.”³⁰ This is the moment in the film where the student operators take a seat at the long switchboard that takes up the width of the operators’ room. Their bodies appear lined up and ready to weave the wires

²⁹ Ibid.

³⁰ Ibid.

and connect subscribers to their intended recipient. Experienced operators file in a line behind each student to monitor their use of the apparatuses and how they wield their learnings from the classroom in the switching room through a “class drill with the ‘cords’ and “plugs.”



Figure 1.4: This scene depicts the operators on recreational or break time that included listening to music and dancing with other operators. (Film produced by Bell System, 1926, <https://www.youtube.com/watch?v=QQbwRYCYJzI>).

As the camera shows the operators mouthing words while feverishly taking the cords and connecting subscribers, a close up of the operator's hands at the switchboard comes into view. The human hands and arms become the tendrils to a deeper, larger network and the operator becomes the conduit for all transmissions and as the text reveals,

Upon completion of her course in the training department she is assigned to work in a central office. She is not a Weaver of Speech with all the many interesting opportunities to serve and advance in her work. Two million times an hour the Weavers of Speech in the United States inquire ‘Number please’ and speed the spoken words of unseen millions.³¹

In the last twenty seconds of the training film, we see a young woman wearing a dark dress with a white lace collar central to the shot, an idealized shot taken out of the actual environmental context and outside of the operators’ room (see fig. 1.1). We see white diagonal lines criss-crossing her torso that becomes a type of craft especially with the designation of the women becoming “weavers of speech,” which suggests that the repetitive work in concert with the use of the human voice is intrinsically feminized and gendered work. Her outstretched arms cross the switchboard to form geometric, triangular shapes forming a rendering of an analog version of networked communication. The operator’s body and voice become the conduit for business transactions and the flow of information and capital.

³¹ Ibid.

Switching Outside the Lines: Centering Black and Asian Networks of Labor

60 Years of Black Women Organizing at Bell System

Wartime economics were factors of Bell System slowly desegregating and hiring Black women operators in the 1940s.³² Yet the company resisted due to the image they wanted to uphold. White women operators were rapacious in their attempts to have a strong hold over the operator role as one of a particular type of prestige. This job offered a sense of stability and even offered a type of homosociality amongst other operators who were mostly other white American women. Immigrants were hired, but they were from other English speaking countries such as Canada and remained the exceptions for years.³³ There was resistance to Black women becoming part of the company, let alone becoming colleagues to white women at Bell. They attempted to desegregate and change labor practices and were subjected to inhumane conditions in comparison to white women. In contemporary culture, it is common that what is often seen and perceived as clerical and highly administrative work is done, mostly, by an overseas workforce predominantly in India and the Philippines due to “linguistic and postcolonial ties” to England and the United

³² Green, *Race on the Line: Gender, Labor, and Technology in the Bell System, 1880-1980*, 195.

³³ *Ibid*, 62.

States.³⁴ This outsourced labor will be discussed further in chapter four. There was only “prestige and deference” when white women were made to be central to Bell System’s “white lady” image.³⁵

In addition to the preferential treatment, Bell System managers would often cite “speech and diction limitations” as a reason for not hiring Black women especially from the South.³⁶ But World War II would change the workforce and push political organizing for fair labor (hiring) practices within the Black community. President Franklin D. Roosevelt’s affirmation of the telecommunication operations as vital to the defense industry through Executive Order 8802 in 1941, in the same year, the March on Washington Movement created a campaign for citizens to express their support and protest of discriminatory hiring.³⁷ Yet various boycotting campaigns, ranging from intentionally jamming lines to preventing stock quotes from getting to the west coast in a timely manner, uncovered how existing unions involvement played a significant role in preventing Black women being hired on as operators.³⁸

³⁴ Kiran Mirchandani, *Phone Clones: Authenticity Work in the Transnational Service Economy* (Ithaca: ILR Press, 2012), 32.

³⁵ Green, *Race on the Line: Gender, Labor, and Technology in the Bell System, 1880-1980*, 195.

³⁶ *Ibid*, 196-197.

³⁷ *Ibid*, 202.

³⁸ *Ibid*.

When Black women were finally hired on as operators, in the late 1960s, the job lost its appeal and the company had no intention of investing time, effort, and resources the way they had done for white women. The desire for a telephone operator was affected by women of color being brought into these roles deeming them less desirable by white women combined with increasing automation within the computer industry. Similar to what we will see in chapter four, there is also a direct correlation between class and demographics or location of the worker.³⁹ White flight was an aspect of this shift away from the city and into the suburbs. In addition, the number of white workers wasn't the labor issue, but a class one since the workers that once would take on these roles did not want the lower wages especially since the job lost its prestige after racial integration.⁴⁰

Pacific Telephone & Telegraph's Call of the East Promotional Film of The Chinese Telephone Exchange

When media scholar Yiman Wang shared an article on the CTE founded in San Francisco by Loo Kum Shu, I became intrigued by the formation of this tight-knit community within a larger network of switching

³⁹ Ibid, 212.

⁴⁰ Ibid.

stations in the Bay Area between the years 1894-1949.⁴¹ Unlike the white telephone operators within the Pacific Telegraph & Telephone company, the CTE was mostly Chinese single women or mothers raising their children alone. The latter set of women were subjected to the effects of U.S. legislation such as the Chinese Exclusion Act of 1882 and the Immigration Act of 1924 that separated them from their partners.⁴² The political landscape illuminates the seemingly insurmountable challenges faced by Chinese operators in a country that did not welcome or want them. Yet they formed a proto-network through their close-knit connection working as operators and becoming familiar voices to the Chinese community in San Francisco. Daughters and granddaughters also worked for the exchange during its years of operations. According to oral histories from the descendents of the CTE operators, this group of women were unique from their white counterparts due to their expectation to speak English and multiple Chinese dialects.⁴³ The development of the CTE by Loo Kum Shu involved organizing and operating within a

⁴¹ Geri Koeppel, “Plugged In: The Fascinating History Of The Chinese Telephone Exchange,” *Hoodline*, April 18, 2016, <https://hoodline.com/2016/04/plugged-in-the-fascinating-history-of-the-chinese-telephone-exchange/>.

⁴² “Chinese Exclusion Act (1882),” National Archives, September 8, 2021, <https://www.archives.gov/milestone-documents/chinese-exclusion-act>.

⁴³ During my visit to the AT&T Archives and History Center located in San Antonio, Texas during summer 2022, I came across photographs and documentation of operators from the Chinese Telephone Exchange (CTE) selected alongside American and European operators to a United Nations meeting. Although this is not central, it is worth noting here to show the reach and global fascination people had of the CTE.

greater system of The Pacific Telephone and Telegraph Company. Similar to the Black women operators on the east coast, both groups were engaged in labor organizing for fairer and just hiring practices and wages.

Walking around the 21st century Chinatown in San Francisco, it is easy to marvel at the ornate embellishments, lanterns, and decorations and banners that adorn its streets. From 1896 to 1949, a small building located at 743 Washington Street, now a EastWest Bank, was once home to the operators of the CTE connecting domestic and international calls between subscribers, businesses, and community members.⁴⁴ Loo Kum Shu founded the CTE to serve the Chinese community in the San Francisco Bay Area and over the years, the subscriber base grew to up to 2,000 subscribers. According to his son Loo Yee Kern, Shu was left with an American missionary, Florence Bokee.⁴⁵ She taught him English and he eventually graduated from the University of California.⁴⁶ Due to Shu's ability to learn and speak English well, he became the "spokesman, interpreter, and advisor" to his immediate community and soon became a publisher of the first Chinese newspaper outside of China

⁴⁴ Thomas W. Chinn, *Bridging the Pacific: San Francisco Chinatown and Its People* (San Francisco: Chinese Historical Society of America, 1989), 37.

⁴⁵ Loo Yee Kern, "Chinatown Telephone Office," *The Pacific Telephone Magazine*, September 1937, Pacific Telephone Magazine Publications 1930-1940, AT&T Archives and History Center, San Antonio, TX.

⁴⁶ Kern's article did not include the specific campus and only stated "University of California," but I speculate he attended UC Berkeley considering his professional life was so heavily focused on serving the Chinese community in the San Francisco Bay Area.

called The Occidental Daily News.”⁴⁷ Bokee gifted him an old press and he hand-wrote the newspaper that was then reproduced on a lithographer’s stone. The CTE was born out of Shu’s experience and expertise serving the Chinese community as an underwriter, publisher (and editor), and establishing the first Chinese telephone exchange.

The *Call of the East* (circa 1930) was a promotional film depicting San Francisco’s Chinatown in an overtly foreign and exotic way to the viewer through Asian tropes and music. This film was most likely for an American or European viewer with a profound interest in the CTE. It showed a precursor of a mesh or insular digital network. The title screen fades to text that introduces the district as an “oriental city within a city.”⁴⁸ Candid footage of people in the neighborhood walking around provides a hyper localized view of the area from which the CTE operators worked. The text that quickly flashes on the screen alludes to the fact that this area of San Francisco was once considered undesirable or “darkly mysterious.”⁴⁹ The first cut away from the cityscape is of an interior that is the 743 Washington Street location of the exchange. The mise-en-scène is filled with ornate decoration, a young woman that answers the telephone and speaks in both English and Chinese indicates

⁴⁷ Ibid.

⁴⁸ *Call of the East*, 16mm film (Bell System, 1930).

⁴⁹ Ibid.

the duality and multilingual nature of which Chinese women and men were expected to perform as telecommunications workers. The lack of subtitling in the video gives a sonic distance between the viewer and the insular community of those that worked at the CTE. Pacific Telephone & Telegraph company prided itself on having such a unique exchange within the larger nation-wide switching stations. Two minutes into the short film, the filmmaker takes a shot from outside and from the right hand side, a man walks up to the station and speaks to Kern Loo Yee (relative to original CTE manager and founder Loo Kum Shu) helps a community member and quickly jumps to text that indicates a statistic of 2,000 subscribers to the exchange. From this point, an older white man enters the premises and Yee switches to American English to answer questions from Mr. Miller of the Daily Reporter (see fig. 1.5). A slow panning of the switchboard shows the backs of the telephone operators. Yee and Miller continue a conversation about the exchange that serves as both narration and voiceover. The viewer also hears the women speaking English or one of the Chinese dialects they were expected to know how to be able to communicate with its Chinese subscribers.



Figure 1.5: A screen capture of short film *Call of the East* (circa 1930) shows Kern Loo Yee, manager of the Chinese Telephone Exchange.

Similar to what Venus Green describes in her research on Bell Telephone operators, men and women started off as telephone operators at the CTE, but it soon became women-dominated due to the perception of their “good temperament” and “socialized to please” as we have seen commonly suggested in telecommunications labor history.⁵⁰ The racialization and gendering of the operators within this particular exchange only had to deal with the Chinese community and the operators were expected to memorize all of the information required to execute the job. According to Geri Koepfel’s

⁵⁰ Green, *Race on the Line: Gender, Labor, and Technology in the Bell System, 1880-1980*, 19.

Hoodline article, a publication focused on San Francisco culture and history, after the 1906 earthquake, the exchange became an all female workforce.⁵¹

White women in the east coast were not subject to these same tensions and circumstances. The Bell System operators in the early 20th century were of a specific class, demeanor, socioeconomic status. At the time, most importantly, these (white) operators were both young and single, which was different from the Chinese operators on the West Coast. This significant difference is noteworthy to the extent that the conditions were fraught for Chinese women not only because they were serving a specific community, but they had to raise families and oftentimes moved frequently between homes due to the inability to solidify stable housing.⁵² While the “white woman” voice may not have been necessary in the communities they serve, it also kept them insular and prevented them from possibly engaging in telecommunications work outside of the Bay Area and the Chinese communities. The women often worked 8-10 hour days connecting thousands of calls every day.⁵³

In some ways, the Chinese telephone operators that worked at 743 Washington Street defied the “lady operator” voice that was associated with a

⁵¹ Geri Koeppel, “Plugged In: The Fascinating History Of The Chinese Telephone Exchange,” Online publication, Hoodline, April 18, 2016, <https://hoodline.com/2016/04/plugged-in-the-fascinating-history-of-the-chinese-telephone-exchange/>.

⁵² Chinn, *Bridging the Pacific*.

⁵³ Ibid.

white woman who only spoke western, American English with diction that was perceived as an ideal vocal affect. Not much historical information exists other than oral histories (not a surprise, but also fitting to this research) that describe the daily activities and the banality of work and labor in the hopes of supporting both their families and communities. While Bell System coined the white women operators as the weavers of speech, the existence and role of the CTE continue to be obscured by the mainstream histories told to the general public of the telephone's history and evolution from the telegraph. While white and Black women were hired to work as operators on the east coast, they were expected to perform at a type of congeniality and care, but only speaking English. Chinese operators were expected to be proficient, efficient, well versed, knowledgeable of the community, services, and businesses through specific nuance and prosody based on their knowledge of dialects. This knowledge base of multiple dialects, in turn, produces a type of multi-vocality that did not seem as apparent in white women operators, yet they were considered the ideal voice (or ideal vocal/sonic interface). The sheer volume of family names, numbers, and connections the operators of this network had to memorize and become familiar with was staggering. Due to the fact that they were told to memorize all of the subscriber information already set them apart from any other switching station and operating rooms

across the country. Their ability to code switch between five Chinese dialects; Som Yup, Soy Yup, Heong Sow, Gow Gong, and Aw Duck; and American English was a testament to their ability to fully embody the occupation that was not created or envisioned them as weavers of speech.⁵⁴

From the 1960s through the 1980s, concept films and media produced by telecommunications companies such as AT&T posit futures where the default language of translation is based on English. The problem of disciplining the body continues to persist: the belief that accent reduction alleviates raciolinguistic tensions, but it perpetuates the primacy of American and British Englishes as global vocal constructs. In chapter four, I write about the post colonial subject hood of Filipino call center workers. While I start the research with the desegregation of the Black women at Bell System to the CTE, these feminist media histories and lineages continue to affect how contemporary systems are created and for the purposes of capitalism and colonialism as conglomerate companies rely on outsourcing and business exploitative practices. Throughout my research, what I found most egregious was Bell System's belief that Black and immigrant women were not only thought of as cheap and disposable labor, but that they were not worthy of representing the company.⁵⁵

⁵⁴ Kern, "Chinatown Telephone Office."

⁵⁵ Green, *Race on the Line*, 62.

In contemporary American culture, in the western world, broadly, views of telecommunications jobs such as operators or call center agents are seen as lower level administrative, clerical, and akin to blue collar work. When digital and mobile technological advancements were fast approaching and on the cusp of the computer age and automation, deskilling became a prominent aspect of large telecommunications companies. Many of these jobs now go to overseas workers in South Asia or the Global South who are in service to either the United States or British nations. Many Americans claim overseas workers and outsourced labor takes away from the domestic job force in the US without knowing the context of capitalism's reach into impoverished global communities. Often, these jobs are far more lucrative in countries still perceived as underdeveloped and easily exploitable. A telephone operator, call center or virtual agent must face the conundrum of not only sounding American, but learning the vocal cues of how to interact with the west.

Although gender may not overtly serve as a factor in places such as the Philippines or India, the most common voice one hears on the line is a woman or feminine voice. Connections to the perceived softness and ease of a woman's voice is part of a genealogy of automation where the woman serves as a type of conduit between man and machine. Although Sadie Plant references the touch typists and the increased reading and literacy amongst

women in *Zeros and Ones*, she presciently writes, “his voice, his smile, his interface; connecting and protecting him from the world, the screen on which he presented himself, a superficial front, a processing filter, and a shield, a protective coat.”⁵⁶ The gendered nature of the voice and how it is trained to serve is detailed in Green’s book. While her focus is on the US, scholars such as Halcyon M. Lawrence and Thao Phan look at how western Englishes are dominant in the way people around the world communicate. For Phan, she looks at mis-recognition of the voice. Her research and scholarship make me wonder how machines and AI are and will be used to mimic (and perpetuate and propagate) distinguishable western voices.

Class plays an integral part in understanding the intricacies and developments of such a specific networked community. Through lexicon, vernacular, and accent, these were some of the linguistic indicators of someone’s place of origin or location. Some argue that the voice can be a biomarker, which it cannot because prosodic elements of voice such as tone, animation, and cadence can be changed.⁵⁷ For instance, Hollywood actors and theater performers are trained to mimic certain accents. Returning to the issue

⁵⁶ Sadie Plant, *Zeros and Ones: Digital Women and the New Technoculture* (New York: Doubleday, 1997), 121.

⁵⁷ Beth M. Semel, “Listening Like a Computer: Attentional Tensions and Mechanized Care in Psychiatric Digital Phenotyping,” *Science, Technology, & Human Values* 47, no. 2 (March 2022): 266–90, <https://doi.org/10.1177/01622439211026371>.

of class, the voice is used to serve as an identifier. Southern Black women were unable to even seek employment at Bell System because they were often perceived as uneducated and did not possess the appropriate voice and accent the company wanted reflected to their clientele. Another concern in contemporary telecommunications that has its roots in hiring practices and the creation of the voice representing the company related to class, is the establishment of elocution and modern day versions of these types of vocal training show up as “accent reduction schools” in the Global South. I have seen this first hand working in biotech companies where scientists with heavy Asian accents were trained to speak with an American accent.⁵⁸ However, they were called vocal coaches and not accent reduction teachers.⁵⁹

Fast forward to the 21st century, when Susan Bennet was asked to read a script in July 2005, one of the stipulations for the voice acting part was someone without a discernible foreign accented voice to perform the role.⁶⁰ She was unaware that her voice would become one of the most recognizable virtual assistants, Siri, when Apple's iPhone 4S was released on October 4,

⁵⁸ Both biotech companies I worked for paid for these vocal training services specifically for high profile scientists working on key clinical trials, which led me to believe that their ability to discuss, present, and share key data and study results was driven by how well scientists were able to represent the company through their public speaking.

⁵⁹ For example, Prestige Accents is focused on accent reduction into Standard American English. <https://www.prestigeaccents.com/>

⁶⁰ Zachary Crockett, “A Q+A with Susan Bennett, the Original Voice of Siri,” *The Hustle* (blog), January 11, 2019, <https://thehustle.co/siri-voice-actor/>.

2011.⁶¹ Her voice is part of a lineage of telecommunications and the normalization of a specific type of voice mediated through technology. On the other hand, many women of color (Black women and Asian women, specifically) had to perform language in and outside of their respective cultures while attempting to organize alternative spaces and platforms that continued to suppress their voice, speech, and labor rights.

Technological advancements in telephony also lay the groundwork for contemporary platforms such as Google's Project Euphonia, which was created for individuals with speech impairments as a result of stroke, traumatic brain injury, down syndrome, cerebral palsy, and stuttering to name some examples. The company states its goals as, "helping people with atypical speech be better understood."⁶² Automatic Speech Recognition (ASR) systems, unlike the talking machines of the late 19th century, rely on thousands of hours of human speech that are then used to create smart technologies such as speakers and phones. Faber sought to create a machine that closely mimicked the functioning of the respiratory and vocal systems of human anatomy and to recreate human speech through non-organic materiality. Although he may not have witnessed how his invention helped shift and serve as a catalyst for telecommunications, the Euphonia reflected the concerns and curiosities of

⁶¹ Roberto Pieraccini, *AI Assistants* (Cambridge: The MIT Press, 2021), 80.

⁶² Project Euphonia, accessed October 18, 2022, <https://sites.research.google/euphonia/about>.

the time: to create speaking machines based on scientific principles and experiments. But the femininity imposed upon his machine has a direct connection to what technical communications scholar Halcyon M. Lawrence suggests in her scholarship when technology is built as an attempt to discipline the body.⁶³ Both femininity and masculinity as gender constructs are learned through social conditioning and through behaviors that are normalized. Whatever characteristics are imposed and programming that codifies a particular set of vocal mannerisms onto a machine or device pertain to forms of disciplining a human and therefore, a mechanistic, synthetic body.

After the Tone

Circling back to the primary concern of how race, class, and gender were translated into the profiling of operators and how this affected the demise of the "white lady" image within the telecommunications industry, I want to emphasize the importance of my focus on telecommunications recruitment and promotional videos of the early 20th century. My visual and textual analyses were meant to show how these forms of media perpetuated the angelic white lady image and voice. They served as methods to profile women

⁶³ Halcyon M. Lawrence. *Siri Disciplines*. Your Computer is On Fire. Cambridge: The MIT Press, 2021.

interested in becoming telecommunicators and did so through depictions of race, class and gender. The homogeneity is evident in *Training for Service* (1926) because it alludes to the predilection towards a specific type of woman, body type, and the ability for the potential operator to speak the preferred language (e.g., the elocution classes being taught in English). The absence of men and the presence of the machines and switching room made certain that women understood their role as a weaver of speech and someone that was a bridge, a connector. Women were not at the helm or in control, but they were the conduit. The African American and immigrant (from non-English speaking countries) woman was not the target audience for this particular short film. Their absence is a reminder that Bell System looked for a particular type of woman that would uphold and perpetuate a specific image they felt was representative of their company.

On the other hand, the short film *Call of the East* (circa 1930) was a promotional video that exoticized the Chinese Telephone Exchange. It was produced to serve as a form of spectacle into the eastern world by way of the west and the collective imagination regarding a group of women operators that were known for their bilingualism. This proto-telecommunications community created by Loo Kum Shu became a point of fascination for many people at the height of telephone operating. The film incorporated sound,

which was a revolutionary film technology that was introduced in the early 1930s. I imagine the film presented a unique side of San Francisco to draw in local consumers, tourists, and anyone interested in learning more about the exchange's development and operations. One of the most memorable and revealing aspects of the research component of this particular chapter was the discovery and screening of *Call of the East* (circa 1930) film during my research trip to the AT&T Archives and History Center (San Antonio, Texas). Not only were the operators of the Chinese Telephone Exchange a different type of weavers of speech from the Bell System operators, they served as an analog recommender system and network to the communities they served. Their ability to navigate multilingual environments via telecommunications and in the physical environment of the switching station served as a testament to their individual and collective ability to prototype an infrastructure of care and service that was both unparalleled and unmatched in the telecommunications industry at the time.

In the age of social media and live streaming with platforms such as TikTok, YouTube, and Instagram, it is vital to see the connections between the early 20th century training of (women's) voices to the perpetuation of a distinct feminine vocal affect contribute to development of contemporary technologies. Despite these videos showing the conditions and considerations

a woman had to understand, in the 21st century, we are faced not only with decades of disciplining bodies and voices, we must now confront the algorithms that have been developed to speak to us and how machines (unlike humans) are built to constantly listen. As the switching stations started to close down and automation along with the dawn of the personal computing age, the voice of the machine started to become a point of interest for technologists with ambitious hopes of creating technologies that could listen and talk back to the consumer.

Chapter 2: Digital Media Artifacts of Synthetic Voice

Material Realities and the Embodiment of Digital Voice

At the start of Amazon's 2021 commercial *Alexa's Body*, the immediate view resembles a multi-floor shopping mall. The architectural space is clean, replete with slick angular lines, and a visible Amazon logo of the curved arrow resembling a smile a little off center of the viewer's gaze (see fig. 2.1). Although not a space of consumerism, it alludes to a space of ideation with its open concept floor plans unlike closed, cubicle walls separating and siloing workers. The opening shot is of a group of employees (see fig. 2.2) looking at a black round speaker and one of them comments with sheer admiration and states, "It's just flawless, isn't it?"¹ The smart speaker awakens as if animated by the conversation of the employees (when in reality, the smart technology must be commanded by calling its name or a wake word as opposed to its discerning

¹ "Amazon: Alexa's Body By Lucky Generals," The Drum, accessed January 2, 2023, <https://www.thedrum.com/creative-works/project/lucky-generals-amazon-alexa-body>.

human speech let alone any conversation in its presence).² The same employee continues, "I mean, I literally couldn't imagine a more beautiful vessel for Alexa to be [pauses] inside."³ The pause is due to her admiration of actor Michael B. Jordan's face on a vinyl wrapped bus as she looks outside of the office window (see fig. 2.3). She gazes upon a fleeting fantasy. Yet within the space of the commercial's narrative, the viewer sees the imposition of her imagination in creating a specific voice and embodiment for Alexa. Within the fictitious space of the advertisement, we see Jordan as an embodied simulation of her desires. Jordan's all black attire represents the material form (hardware) of Alexa that suggests her desire for a hyper masculine, submissive presence. The commercial includes a jump cut to her in the kitchen asking, "Alexa, how many tablespoons are in a cup?" and the viewer sees Jordan responding to her as each iris shows the royal blue circular light that has become a visual indicator of Alexa's activation. He answers, "There are 16 tablespoons in a cup" in a calm voice with little variation in tone as if to mimic, to the best of his ability, an affect that fits both the fantasy she has of Alexa as embodied by Jordan, but also showing Alexa's efficiency (see fig. 2.4). We quickly see the

² The definition of wake word is "a word or words that you say in order to make an electronic device, or a feature on a device." "Wake Word," April 26, 2023, <https://dictionary.cambridge.org/us/dictionary/english/wake-word>. In addition to the wake word, within the fantastical realm of the advertisement, the activated status of the device suggests a type of sentience to the machine.

³ Ibid.

woman's husband walking into the kitchen mentioning that food has arrived despite her cooking a meal. Yet something strange occurs. The husband's recognition of Jordan's presence is confirmation of Alexa's existence in their home, simultaneously confirming its embodiment by a male figure willing to please his wife's wishes. Dominant gender roles have been reversed via Jordan's embodiment as a docile, smart technology.



Figure 2.1: Screenshot of *Alexa's Body* (2021) commercial (Commercial produced by Lucky Generals for Amazon Alexa, 1984, <https://adage.com/video/amazon-alexa-alexa-body>).

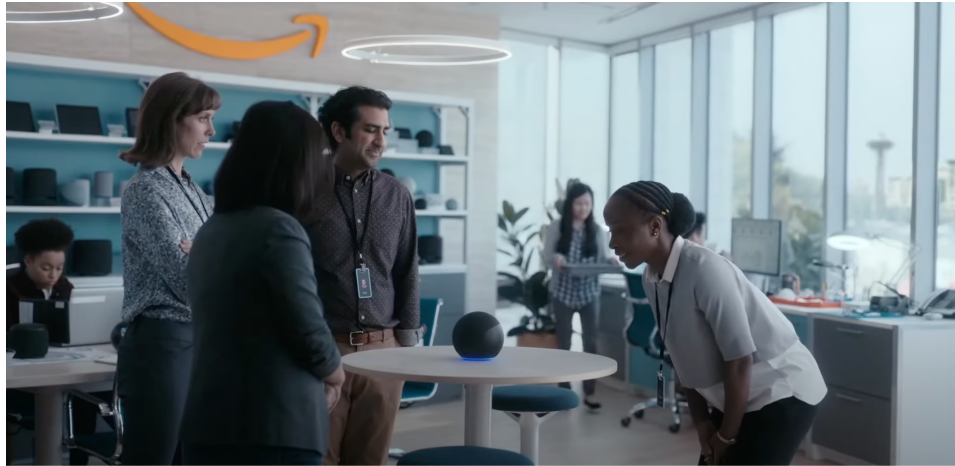


Figure 2.2: Screenshot of *Alexa's Body* (2021) commercial (Commercial produced by Lucky Generals for Amazon Alexa, 1984, <https://adage.com/video/amazon-alexa-alexa-body>).



Figure 2.3: Screenshot of *Alexa's Body* (2021) commercial (Commercial produced by Lucky Generals for Amazon Alexa, 1984, <https://adage.com/video/amazon-alexa-alexa-body>).

For the majority of the commercial, jump cuts give the illusion of many scenarios in a matter of a few minutes. The smart technology inhabits the home along with the couple and takes on a central role in its function (or dysfunction). In another scenario, the viewer sees the wife sitting on the edge of their couch while the husband watches television. Jordan (Alexa) says,

"You're the most beautiful woman I've ever seen" and translates the phrase in French while the husband gives a dubious and disapproving look. Another shot, this time, during the day and the wife says softly while looking out the window into their green well-manicured lawn, "Alexa, turn on the sprinklers." As the water shoots up at an angle with Jordan directly in the middle, he becomes drenched in the water while his eyes depict activation with the signature bright blue glimmering lights as the camera zooms in on his face. The husband rushes out stating that he already ran the sprinklers for the day and requesting Alexa to stop immediately. The innuendo is pronounced when the husband's gruff reaction claims, "Alexa, stop, things are getting way too wet around here." Controlling machines and devices from within and outside of the home is a reminder to the consumer that Alexa can be programmed to run smart appliances throughout the home via applications such as Home Connect that allow a user to control smart technologies regardless of the manufacturer through the voice.⁴ The capabilities of the smart technology amplify its reach to anywhere in one's property that are set up and designated as a part of the home.

⁴ "Exceptionally Smart: Control Home Connect Appliances Using Your Voice," Business, Home Connect, accessed April 28, 2023, <https://www.home-connect.com/za/en/connected-partners/voice-assistant/amazon-alexa>.

In another scene, the husband and wife are having company over and the wife looks at Jordan and commands, "Alexa, dim the lights" and Jordan proceeds to remove his Black shirt while the wives ogle at Jordan's (Alexa's) body (see fig. 2.5). The husband notices from outside and states, "Alexa, lights up." Another scene changes and the wife commands, "Alexa, add bath oils to my shopping list" while the husband follows up with, "Alexa, no. Don't do that.". The fantasy grows more pronounced when the wife is now in the bathtub, but she ceases to begin with "Alexa" and now directly states, "Read my audio book." The viewer now sees Jordan in the bubble bath fully clothed as to remind us of the fantasy being spun of Alexa's embodiment in the wife's desire for an obedient entity in her home that also fuels her fantasy of a companion that is willing to do whatever they are commanded (see fig. 2.6).

In the last forty seconds of the commercial, we are brought back into the office with the woman (wife) caressing the window as if still entranced by Alexa (Jordan) being the embodiment of Amazon's Alexa while her co-workers look on in a quizzical nature at her strange behavior. After a couple seconds of silence pass, one of the office mates asks, "Do you know if we have any extra ones of these?" The commercial finishes with the wife sunken a little deeper into the bubble bath as Jordan recites words from James Baldwin's book *If Beale Street Could Talk* with the husband knocking on the

bathroom door claiming that other people need to use the bathroom as well.⁵ Beale Street is a love story between a young Black man and Black woman set in 1970s Harlem. While Baldwin's celebrated work is commemorated in countless ways, the metaphor of the street itself conveying or speaking stories of its inhabitants parallels Michael B. Jordan's reading of Baldwin's work by anthropomorphizing the object or smart technology. There is a strange dissonance of the story being brought to life through his narration, but also the speculative nature of Alexa's embodiment in a Black cisgender male body. While Jordan's depiction of Alexa is far from our current state and experiences of artificial intelligence, it speculates a future where one would not have to imagine and where fantasy becomes reified through the cyborg's body and capacity to speak with nuance.

⁵ James Baldwin, *If Beale Street Could Talk*, 1st Vintage International ed (New York: Vintage International, 2006).



Figure 2.4: Screenshot of *Alexa's Body* (2021) commercial (Commercial produced by Lucky Generals for Amazon Alexa, 1984, <https://adage.com/video/amazon-alexa-alexa-body>).



Figure 2.5: Screenshot of *Alexa's Body* (2021) commercial (Commercial produced by Lucky Generals for Amazon Alexa, 1984, <https://adage.com/video/amazon-alexa-alexa-body>).

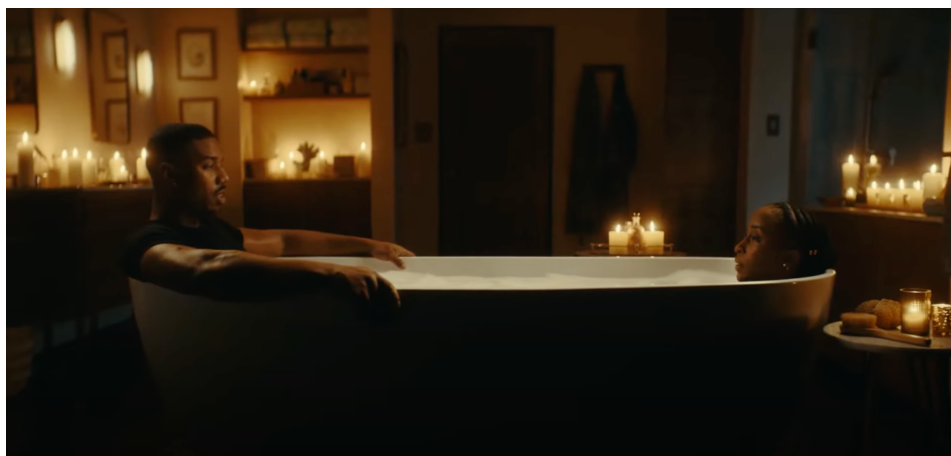


Figure 2.6: Screenshot of *Alexa's Body* (2021) commercial (Commercial produced by Lucky Generals for Amazon Alexa, 1984, <https://adage.com/video/amazon-alexa-alexa-s-body>).

Although the commercial proves to serve as comic relief for the average consumer, there is a direct marketing and branding meant for a particular class of African American cisgender men and women. The visual and sonic analysis of promotional material for smart technologies and assistive tech bring about an awareness that plays into our contemporary cultural expectations as well as offers up speculative ways of living and working with technology. In chapter one, I cover the ways in which women's voices were trained for service. Although *Training for Service* (1926) and *The Call of the East* (circa 1930) were recruitment and informational videos, they both perpetuate specific types of feminine bodies that uphold a particular archetype tethered to the technology. In this chapter, I felt compelled to begin with contemporary media (television advertisement) that entertains the consumer

yet fetishizes the technology (product) because it perpetuates and propagates ideas that are extremely difficult to untangle and undo in the collective imagination. While Jordan, a cisgender man is cast as Alexa's body that upends the idea of the feminine as docile, his body as an assistive technology complicates the already heteronormative dynamic within the Black couple's home.

Another strange aspect of the commercial is the use of a Black man to embody a device that is so often gendered as feminine or a woman. While *Alexa's Body* (2021) is a commercial that might be perceived as commanding a Black body, the protagonist of the commercial is a Black, affluent woman, with a large suburban home. This depiction is far from the types of figures we see in earlier forms of voice, embodiment, and the human body's relationship to analog and digital technologies. In addition, the spending power of women had greatly increased as of the late 1970s in the United States as a result of The Equal Credit Opportunity Act (1974) when they were afforded the opportunity to apply for loans and credit cards without a partner.⁶ Around the same time, Black women telephone operators were being phased out due to automation. Yet they were finally able to be hired on as operators despite sixty years of

⁶ Robin Saks Frankel, "History Of Women And Credit Cards: 1970s To Present – Forbes Advisor," Forbes, December 2, 2022, <https://www.forbes.com/advisor/credit-cards/when-could-women-get-credit-cards/>.

fighting for equal labor rights and wages. Combined with being on the cusp of second and third wave feminism, Black women were relegated to jobs perceived as administrative and clerical. The telephone operator job was no longer desirable by white women. Circling back to big tech companies such as Amazon, they target and attract a consumer based on market research where the use of a Black, affluent woman as the protagonist (who also works at a company identical to Amazon in the commercial) is given an opportunity to be pleased (and pleased, even if metaphorically) by an actor, specifically Michael B. Jordan. Jordan ceases to embody Alexa, rather, he materializes as her fantasy for the purposes of convenience. He becomes antithetical to the way Black men are portrayed in the media. He actively listens. He calls when he is beckoned. He caters to her needs. He reciprocates tone, speech, and diction. The creation of this particular type of mass media, therein, codifies the types of artificial intelligence and machines of the future.

From the suggestion that Alexa can cater to the housewife in a multitude of ways, the commercial also suggests that Alexa can take the abstract form of what the consumer desires. It proposes that while the fantasy might not be a reality, Alexa is a proxy of what is possible. The assistive tech serves in all of the ways the human being might not be able to do. The creation of these commercial endeavors is also reminiscent of Katherine

Stubbs's scholarship on telegraphic fiction.⁷ This literary genre presupposed the reader knew the function of this fiction which was focused on the vocal correspondence of the protagonist with a romantic interest (typically, the subscriber). These stories would often result in the main character forgoing or transcending physical attraction because he had already fallen in love with the voice on the line. In a similar fashion, contemporary mainstream advertising of assistive tech usually follows the trajectory of fantasy, play, convenience, and desire. At the beginning of this research, I wondered about the pivotal moments and technological innovations that have contributed to the media production of voice-based smart technologies and assistive tech. Beyond the depiction of media as contained within a human form, I felt it necessary to think through one of the first interactive voice responses systems and how developers in the 21st century are a part of a larger ecosystem that sustains these large corporations.⁸

In order to understand the embodiment of digital voice, in this chapter, I investigate How May I Help You (HMIHY) technology, an early iteration of an interactive voice response system (IVRS) developed by AT&T labs, and how it served as the groundwork for the development of virtual assistants and

⁷ Katherine Stubbs, "Telegraphy's Corporeal Fictions," in *New Media, 1740-1915*, ed. Geoffrey B. Pingree and Lisa Gitelman (Cambridge: The MIT Press, 2003).

⁸ Voice clones, for example, are derived from decades of experimenting with linguistics and understanding what makes speech sound human.

smart technologies. I also examine how voice and speech technologies must entail looking at how, for example, prosodic attributes of speech have been programmed by big tech companies such as Apple. The developer sandbox culture is also prevalent in big tech and serves as a way to encourage a type of software development that restricts use to a specific set of tools that then become proprietary. This technosolutionist lineage serves as fertile ground for companies such as Sanas AI that aim to create a type of vocal and sonic utopia where accented human voices are mediated through machine learning and neural networks. Software is created for machine listening to capture utterances programmed to correct what is marked as an aberration. The livelihoods of individuals in the Global South such as the Philippines and India are at the greatest risk of not only having their foreign-accented English diminished and eliminated, this type of erasure flattens experience and suppresses the human voice in a way that was pure speculation a couple of decades ago.

At a minimum, companies must develop software technologies (and hardwares) that factor in more diverse, accessible, and inclusive datasets. Machine learning and neural networks must also be developed to serve the general public and, most importantly, assistive tech for disabled community members. Although AI ethics is a contentious and ever-evolving field, the

discipline must be developed to include perceptions and remediations of sonic media incorporated into smart and assistive technologies.

Related to the rapid development and prototyping of software, proprietary code is part of an exclusive and insular ecosystem, which makes it difficult for the open source software movement to sustain itself. Yet it is a movement where the experimentation can happen and key questions can be asked. Does an acousmatic voice then cease to exist when speech becomes mediated through this form of technology? Through synthetic voice, how do we expand upon the codebase in areas such as filmmaking and video games and distinguish use cases based on customer service, call center and virtual agents that are at the highest risk?

This chapter also includes archival material from Apple that provides further context as to how synthetic voice led to voice assistance followed by an examination of the prosody section of the Apple's Developer Blog archives. I end with artwork by Lauren Lee McCarthy whose research went into building software focused on placing herself in the role of turning herself into an assistive technology as durational performance. Through the lens of computational media, voice recognition and assistive tech could be coded in such a way that prioritizes accessibility and inclusion as core tenets. This writing is also anchored in the dawn of the computer age and shifts towards

automation that provides historical context that gave way to the beginnings of IVRS. With the decreasing need for telephone operators and the increasing (capitalist) need of automated systems, I write about the narratives created by both Amazon and Apple regarding the embodiment of digital voice. Thinking through the lineage and connections between research and how it inevitably translates to the general public and turned into its own mythology.

With the creation and development of the HMIHY dataset, which collected over 10,000 utterances of customers' calls, I reflect on how computer scientists and engineers determined and laid the groundwork for IVRS that provided customers a dynamic system and AT&T a way to automate calls.⁹ This writing shows the genealogy of voice recognition technologies and how they remain tethered to human voice. The calls were recorded to create an IVRS that would then not only manage the caller's inquiries, but route to the appropriate answer or service or to a human call center agent. This collection of data also provided a solid foundation for the development of dialog managers or what might be commonly known today as virtual assistants.

⁹ A.L. Gorin, G. Riccardi, and J.H. Wright, "How May I Help You?," *Speech Communication* 23, no. 1-2 (October 1997): 113-27, [https://doi.org/10.1016/S0167-6393\(97\)00040-X](https://doi.org/10.1016/S0167-6393(97)00040-X).

Computer Speak: Introductions of Apple Macintosh and Siri, the Voice Assistant to the World

Two days before Steve Jobs introduced the Macintosh in 1984, Ridley Scott's Apple commercial aired during Super Bowl XVIII, on January 22, 1984. His cinematography was a nod to George Orwell's dystopian novel *Nineteen Eighty-Four* about the surveillance state (see fig. 2.7).¹⁰ With a homogenous group of individuals painted to appear just as gray and drab as the environment and architecture they are in, a lone white woman dressed in athletic attire runs down the aisle with the thought police trailing close behind. She is carrying a large hammer and spins around multiple times and hurls it towards the screen showing Big Brother (see fig. 2.8). The blow results in an explosion with a bright light that shocks a crowd suddenly in awe of the demise of Big Brother. The intertitle claims a revolutionary future in computing that ends with "you'll see why 1984 won't be like '1984'" (see fig. 2.9).¹¹ The irony of the Apple Macintosh commercial is the company itself has dominated the smart(phone) technologies and personal computing markets. Most users' experience of technological innovation in voice recognition alone

¹⁰ John Naughton, "The Apple Super Bowl Ad That Announced the Future Was Here," *The Guardian*, February 2, 2014, sec. Technology, <https://www.theguardian.com/technology/2014/feb/02/apple-super-bowl-mac-ad-launched-1984>.

¹¹ *1984 Apple's Macintosh Commercial (HD)*, 2012, <https://www.youtube.com/watch?v=VtvjbmoDx-I>.

is the most conspicuous acquisition of SRI International that was responsible for development of the voice assistant known simply as Siri.¹²



Figure 2.7: Screenshot of Apple Macintosh commercial (Film produced by Apple, 1984, <https://www.youtube.com/watch?v=VtvjbmoDx-I>).



Figure 2.8: Screenshot of Apple Macintosh commercial (Film produced by Apple, 1984, <https://www.youtube.com/watch?v=VtvjbmoDx-I>).

¹² S. R. I. International, “Siri,” *SRI International* (blog), November 16, 2021, <https://www.sri.com/hoi/siri/>.

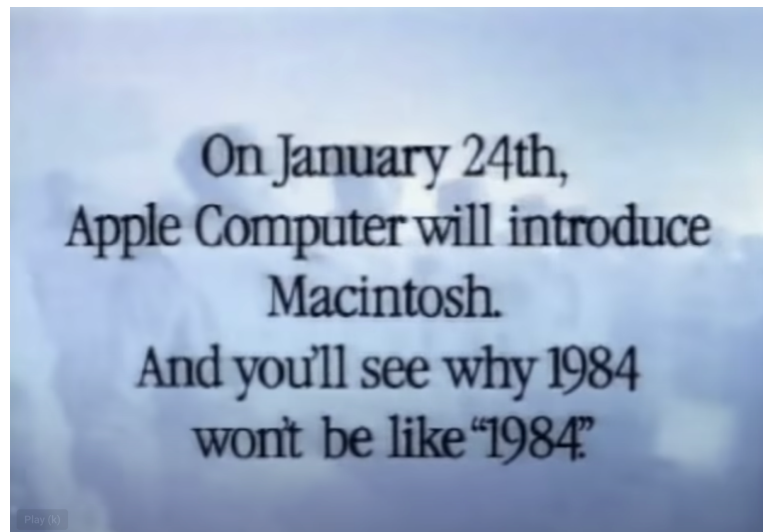


Figure 2.9: Screenshot of Apple Macintosh commercial (Film produced by Apple, 1984, <https://www.youtube.com/watch?v=VtvjbmoDx-I>).

Following the televised Macintosh commercial, which did not include any visuals of the computer itself, Jobs presented the computer at a shareholder meeting at the Flint Center for Performing Arts in Cupertino, California.¹³ The video shows him walking over, unzipping a large bag and pulling out the Macintosh to applause (see fig. 2.10). He inserts a floppy disk (see fig. 2.11), clicks the mouse, and Vangelis' *Chariots of Fire* plays in the background as images appear on the screen. The lights dim and the program runs media content meant to invoke wonder and amazement in the audience. In a large serif font, the computer appears to introduce itself as the letters take up the entire screen and scroll from right to left slowly. Shortly after, the

¹³ *The Story of MacinTalk Part 1*, 2016, <https://www.youtube.com/watch?v=UuVo4MHTEQo>.

drawing program, MacPaint, shows a stylized digital illustration of a feminine Asian figure (see fig 2.12). Then proceeds to show MacWrite, the computer's software and other images and typography to show the computer's functionalities, graphic and image capabilities (see 2.13). However, arguably, what made Jobs' presentation potent and appealing to the audience was when the Macintosh spoke to the audience.

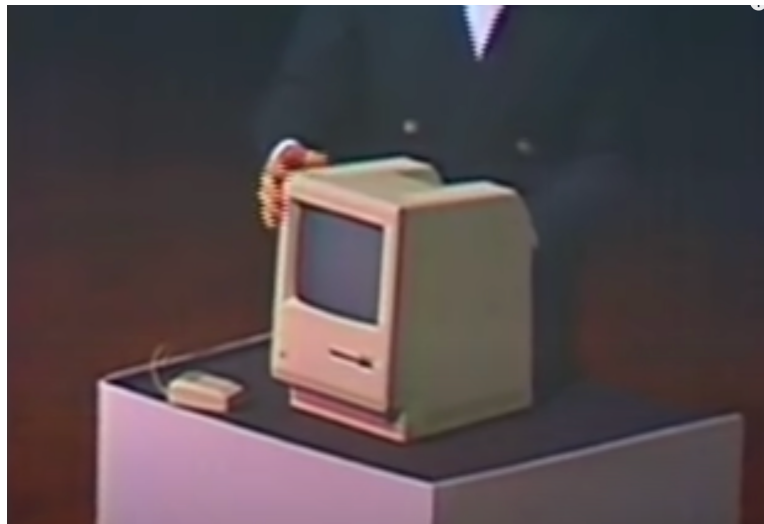


Figure 2.10: Screenshot of Apple Macintosh demonstration January 24, 1984 at Flint Center for Performing Arts, Cupertino, California (Digitized January 2005 by TextLab 2005, <https://www.youtube.com/watch?v=2B-XwPjngYY>).



Figure 2.11: Screenshot of Apple Macintosh demonstration January 24, 1984 at Flint Center for Performing Arts, Cupertino, California (Digitized January 2005 by TextLab 2005, <https://www.youtube.com/watch?v=2B-XwPjngYY>).

While the voice is discernibly robotic and flat, it is programmed to have a sense of humor as it suggests the seemingly inefficient and bulky computing systems of IBM. The Macintosh performs as if it is a type of progeny to Jobs while the crowd laughs and cheers. Although in jest, there is an illusion of personhood in the machine seeing Jobs “like a father” (see fig. 2.14), which is a type of narrative imposed on creators of digital devices and machines.

Twenty years later, Siri would be introduced to the Apple Special Event audience in 2011. While not nearly as theatrical as Jobs’ introduction of the Macintosh, the company showed the ability of not only the iPhone’s ability to speak, but most importantly, the device’s ability to listen. Although Siri is not synonymous with phone trees, it is a part of a lineage where the collective

imaginary is built by the production of media that posits futures where digital technologies interact with us in ways only imagined in science fiction. The thread of fantasy and imagination, as seen in Amazon's *Alexa's Body* commercial, is a part of this media archaeology that allows the user to imagine a type of auditory and personal intimacy. After all, in the 21st century, many of our everyday technologies, our vehicles, and even smart architectures are always listening and waiting to serve.



Figure 2.12: Screenshot of Apple Macintosh demonstration January 24, 1984 at Flint Center for Performing Arts, Cupertino, California (Digitized January 2005 by TextLab 2005, <https://www.youtube.com/watch?v=2B-XwPjngYY>).

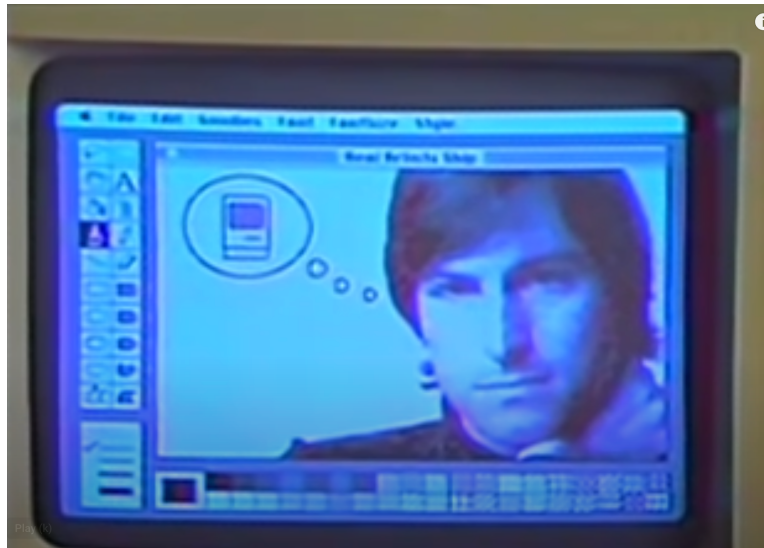


Figure 2.13: Screenshot of Apple Macintosh demonstration January 24, 1984 at Flint Center for Performing Arts, Cupertino, California (Digitized January 2005 by TextLab 2005, <https://www.youtube.com/watch?v=2B-XwPjngYY>).

“Hello, I’m Macintosh. It sure is great to get out of that bag.

Unaccustomed as I am to public speaking, I’d like to share with you a maxim I’d thought of the first time I met an IBM mainframe: NEVER TRUST A COMPUTER YOU CAN’T LIFT.

Obviously, I can talk, but right now I’d like to sit back and listen. So, it is with considerable pride that I introduce a man who’s been like a father to me... STEVE JOBS.”¹⁴

¹⁴ *The Lost 1984 Video: Young Steve Jobs Introduces the Macintosh*, 2009, <https://www.youtube.com/watch?v=2B-XwPjngYY>.

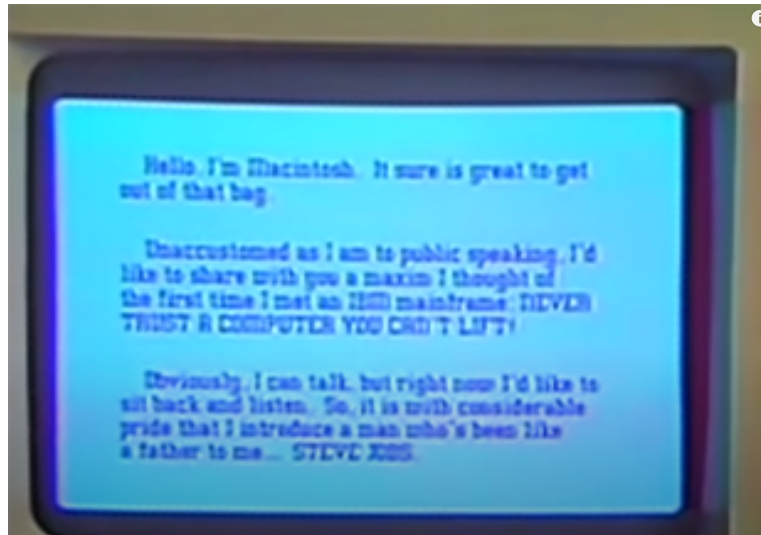


Figure 2.14: Screenshot of Apple Macintosh demonstration January 24, 1984 at Flint Center for Performing Arts, Cupertino, California (Digitized January 2005 by TextLab 2005, <https://www.youtube.com/watch?v=2B-XwPjngYY>).

From Utterances to Phone Trees: How May I Help You (HMIHY) Technology

Spoken Language Understanding (SLU) enabled the creation of classifier models actively set to learn human voice commands and utterances.¹⁵ The How May I Help You Technology (HMIHY) was created at the AT&T Labs in the late 1990s. A group of computer scientists and engineers wanted to create a dialogue manager that was not solely based on touch tone systems, but incorporated the use of the customer's voice. The team created an IVRS

¹⁵ Gupta, Narendra & Tur, Gokhan & Hakkani-Tur, Dilek & Bangalore, Srinivas & Riccardi, Giuseppe & Gilbert, Mazin. (2006). The AT&T Spoken Language Understanding System. Audio, Speech, and Language Processing, IEEE Transactions on. 14. 213 - 222. 10.1109/TSA.2005.854085.

that could help solve basic customer issues and inquiries such as paying a bill or incorrectly being charged for a service. One of the primary goals for the system was to create a dialog system that would enable the routing of calls based on the “meaning of the user’s speech.”¹⁶

The experimentation was based on three core actions to be taken from the utterances: 1) acquire grammar fragments from the collected data and pull words deemed meaningful to the transactions, 2) recognize the most frequently spoken utterances or words, and 3) exploit (and map) the fragments in order to classify the call-types of the utterances.¹⁷ They collected 10,000 utterances or speech acts that were labeled as: 1) a specific call-type (i.e. I need help paying a bill.), 2) quality of speech (i.e., audibility, etc.), and 3) assigned to a channel (i.e. caller or agent).¹⁸ Out of the 10,000 utterances collected between conversations between customers and human agents, 8,000 were used to training a learning model for both recognition and understanding.¹⁹ Based on the various types of conversations between machine and user, clarifying statements were needed in order to route the calls correctly. For example, the machine listens to a question from the user and repeats the question to

¹⁶ Ibid, 114. The term “exploit” is also indicative of finding the bugs or vulnerabilities in the system in order to fix and have the software run properly.

¹⁷ Ibid.

¹⁸ A.L. Gorin, G Riccardi, and J.H Wright, “How May I Help You?,” *Speech Communication* 23, no. 1-2 (October 1997): 113-27, [https://doi.org/10.1016/S0167-6393\(97\)00040-X](https://doi.org/10.1016/S0167-6393(97)00040-X).

¹⁹ Ibid, 116.

confirm that it was heard correctly. The reason why this distinction is of great importance is due to the fact that common assistive tech such as Siri no longer repeats, but will speak, “I don’t understand.” With HMIHY technology, it was one of the first types of interactive voice systems that aimed to be in conversation despite requiring short and concise speech from the user. Another aspect that has greatly changed with contemporary voice assistants and IVRS is the perception of silence or pauses. The HMIHY platform was programmed and interpreted by the machine as affirmation.²⁰ The dialogue noted below is an example of the way HMIHY was coded to listen and resolve (Please note: M is machine and U is user).

M: How may I help you?

U: What’s the area code for Chicago?

M: You want an area code.

U: (*silence*)

M: The area code for Chicago is 312.²¹

Silence may not be one of the first associations a user thinks about when thinking of an IVRS, but it is equally, if not more, important than the speech

²⁰ Ibid, 115.

²¹ Ibid.

inputs and outputs. It punctuates the immediate soundscape and listens for fragments of speech to add to its existing corpus and the ways speech acts are then strung together to further improve upon vocabulary. Another important development in this particular dialog system entailed understanding more complex requests that involved both disconfirmation (a ‘no’ response by a user) and a request for additional information.²²

M: How may I help you?

U: Can I charge this call please?

M: You want to charge this to another number?

U: No, to my Universal card.

M: You want to charge this call to a credit card.

U: (silence)

M: Please speak your card number.²³

HMIHY technology was marketed and branded as a revolutionary, albeit solely transactionary, customer experience. In a 2001 internal employee AT&T magazine, the HMIHY tech was mentioned as a tool to help customers and it included quotes from customers stating the technology made it easier to

²² Ibid.

²³ Ibid.

do business with AT&T.²⁴ The article highlighted how speech technologies (natural language processing, specifically) would be developed to further advance the company's services and revenue. Yet human agents were asked to monitor the system to ensure it was functioning correctly. These human agents were called “wizards,” which again relies on the language of mysticism that we see even in contemporary descriptions of voice recognition technologies.²⁵ Wizards were tasked with taking over calls if the system was unable to fulfill the request of the customer.

The Making of an Expert: Apple’s Developer Archive and Sandbox Culture

Public-facing technical documentation or developer blogs within big tech corporations produce highly edited and vetted content written by engineers that is then approved from a managerial class of another set of engineers or computer scientists. They are written, published, and archived depending on the frequency and usefulness of the content. They become obsolete within a matter of days and weeks. For a company such as Apple, they serve as a type of due diligence for the user, probably more apt, the

²⁴ Tom Hopkins, “AT&T Consumer Asks Customers ‘How May I Help You?’” *AT&T Now*, February 2001, AT&T Archives and History Center, Warren, NJ.

²⁵ Gorin et al., “Automatically Training a Problematic Dialogue Predictor for a Spoken Dialogue System.”

consumer. Developer archives and blogs are a type of media artifact as much as they are a form and format of transparency with the softwares and hardwares being written about and worked on by a tech worker within the organization connecting them to the consumer or developer outside of the corporate walls. While full transparency is not wholly possible, it is not necessarily probable under capitalist conditions. Google and Amazon, for instance, use the space of the Developer blog as a way of sharing code. These platforms are akin to open source communities that might indirectly result in extractive labor. It is commonplace for software projects to rely on volunteerism that adds to existing core libraries or works on new features.²⁶ However, for larger tech companies, such as Apple, developer blogs are vetted and well curated source material (from the company) that allows people creating applications for its devices to be a part of the company's software and hardware evolution. The work being done outside of Apple, for instance, feeds into or serves as a type of pipeline into growing the technology without having to employ this class of technologists. An iPhone, for example, relies on innumerable applications developed outside of Apple yet optimized to run on iPhones, iPads, etc.

²⁶ R. Stuart Geiger, Dorothy Howard, and Lilly Irani, "The Labor of Maintaining and Scaling Free and Open-Source Software Projects," *Proceedings of the ACM on Human-Computer Interaction* 5, no. CSCW1 (April 13, 2021): 1–28, <https://doi.org/10.1145/3449249>.

The developer blog and its entries as a type of media artifact give way to old, obsolete code, processes, and procedures rife with a genealogy in understanding how software is not only developed, but what considerations have gone into the making and the seeding of new features to the code base. Specific to the building of voice-activated assistive technologies and how Apple have gone about documenting key developments and advancements in voice recognition, I focus specifically on documentation related to prosody and improving spoken output for voice synthesizers. I want to tie this back to the historical and technological moments at the AT&T Labs and the creation of How May I Help You (HMIHY) technology. Theorist Nick Bostrom and computer scientist Marvin Minsky posited the ability for machines to either supplement or mimic human speech and utterances is a future that will have gone from artificial to superintelligence.²⁷ I want to advocate against these ideas considering the technology does not account for a multiplicity of languages, abilities, accents, and affects far too vast for any machine or device to be encapsulated within. How might tech companies shift their production and creation of features within these softwares and hardwares that moves at the pace of language and cultural development? How might we think about

²⁷ Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, First edition (Oxford: Oxford University Press, 2014), 46.

the prevention of voice and language turning into property, commodity, or taken up as mere *objects* of study and further colonization?

The point of the documentation archive or developer blog was/is to share the benefits of creating applications and an expansion of the Apple macOS X software. By stating the objectives or "benefits" of the guide, the post is written in an instructional way, yet the documentation is also a type of persuasive text for the reader providing the why and how of this particular software (mac OS X). In the Speech Synthesis Programming Guide (SSPG), the code provides a sense of what is required to enhance the use and effectiveness of speech synthesizers. The SSPG disambiguates from speech recognition or the ability for a user to speak and have their voice recognized by the machine. In making my way through the guide, however, it became clear that it was written for a user that has expert level knowledge of the programming language used to create voice recognition software and while it was relatively easy to comprehend the sections on core concepts related to the technology, it is organized in such a way that a reader needs to know exactly what they are looking for. Related to the artworks I have created, I purposefully sought documentation related to prosody since it is an attribute of speech that is most difficult to synthesize because of the countless iterations of human speech from person to person. Below, is a documentation archive

trail I created as a way to remember how I found the section most related to my research and creative practice.

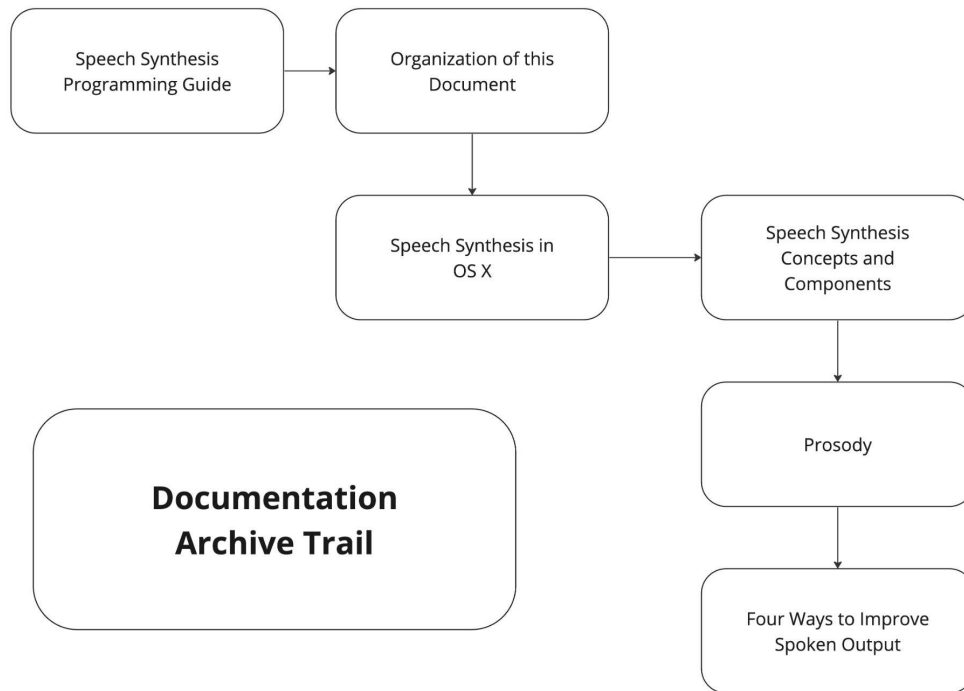


Figure 2.15: A diagram of the documentation archive trail to find guidance regarding “Prosody” and “Four Ways to Improve Spoken Output” (Diagram produced by the author).

Accessibility is also a key area covered by the documentation yet is separated within its own distinct publication within the archive. The primacy of American English is overt and intended to be exclusionary based on phonemes listing derived from North American English and the promise of connection that will happen (based on voice capture, etc.). With this documentation having once supported the macOS X, quasi-transparency was

made possible through the publishing of posts meant for the developer community that allows for the consumer to become a producer thus contributing to Apple's ecosystem. The text is written in such a way that even a non-programmer is able to read it with clarity, following along with the pedagogical tone and the inline (//) comments entered in the sample code. Such specificity gives the impression that this was documentation with the purpose of creating a specific type of media that invited people to expand upon the suite of softwares used within the Apple line of products. There is a cyclical effect happening that perpetuates the use of languages such as Objective-C or PlainTalk (suite for Voice and Speech softwares). Much like spoken languages and vernacular, these programming languages provide not only a gateway, but an ecosystem, for a developer community to work within these parameters, constraints, etc.

Circling back to the initial release of the guide, it is akin to a non-fiction book with chapters that scaffold the knowledge of the reader starting with syntax with the ability to download sample code, when pertinent. For example, while the chapter "Designing and Implementing an Application That Speaks," focuses on the wide array of customization a developer is able to make to the software, it is based on North American

phonemes as well as the use of Cocoa and Carbon Speech Synthesis APIs.²⁸ Cocoa uses Objective-C, which is the primary language used for all Apple products. It allows for a developer to obtain information about a voice (i.e., age and gender), modify the voice used for spoken output, and determine if another application is currently speaking as well as the functionality to start and stop speech.²⁹ Carbon API enables the: 1) creation and management of speech channels, 2) adjust speech attributes (i.e., prosodic elements), 3) convert text to phonemes, 4) get information about speech channels and voices, 5) start, stop, and pause speech, and 6) create, invoke, and dispose of universal procedure pointers that point to functions you supply to synchronize speech with application-specific actions.³⁰ While both have been used to create speech synthesizers, Carbon API, historically, allowed for more complex changes to be made to spoken outputs. Yet the comparison between the two was essential and provided in the SSPG due to the fact that every developer has specific software projects and applications with different functionalities and demands. In comparison, Mozilla created the Common Voice project in 2017, which is a crowdsourced dataset of voices for the purposes of open source software

²⁸ “Designing and Implementing an Application That Speaks,” accessed April 29, 2023, https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SpeechSynthesisProgrammingGuide/UsingSpeech/UsingSpeech.html#//apple_ref/doc/uid/TP40004365-CH4-SW2.

²⁹ Ibid.

³⁰ Ibid.

development in voice recognition.³¹ They have also made their code simpler and easier to read for different types of users and developers.

Returning to Apple's documentation archive, key concepts that help the reader understand speech generation are coupled with sample code and the information provided are ways of best understanding how to create software that expands upon the key syntax. Yet the overall guide is general enough for someone without computer science or an engineering background a sense of what might be useful. A noteworthy observation is of the section on prosody and how punctuation becomes a vital and necessary marker for speech generation to sound much more animated and less synthetic. For example, specific symbols such as commas, parentheses and quotations marks connote pauses as well as the `[[inpt TUNE]]` that allows for the adjustment of pitch and tone.³² While the general public might not be interested in reading this documentation, it is for the sake of due diligence with the consumers (and not just developers) using the products. For litigious purposes, I would argue that the SSPG has important archival (media) material that is deemed necessary for liability reasons. It provides checks and balances of how the

³¹ "Announcing the Initial Release of Mozilla's Open Source Speech Recognition Model and Voice Dataset | The Mozilla Blog," accessed April 29, 2023, <https://blog.mozilla.org/en/mozilla/announcing-the-initial-release-of-mozillas-open-source-speech-recognition-model-and-voice-dataset/>.

³² Ibid.

company itself has engineered the softwares used not just by the consumer or user, but by the developer themselves. Many applications and web-based tools using these particular languages become a part of the cyclical nature of change and equally to the code and hardware's obsolescence.

The Voice as Instrument

Artist Lauren Lee McCarthy's practice includes work that explores the use of her voice to displace emotional labor and a type of digital ventriloquism akin to the embodiment and mediation of Amazon Alexa. She has also incorporated writing custom software to subvert commercial applications to find the boundaries and contours of where the code is not executable or how to circumnavigate it.

On an average morning and typical school day, I packed my things and forgoed breakfast in a feverish attempt to leave the apartment to beat rush hour traffic. As I was getting ready for the day, lights turned on without command or even a flick of my fingers against any switches. An acousmatic, robotic voice said something along the lines of how it was nice serving as my assistant for the past few days. I smiled. I opened the front door, but quickly turned around, and looked nowhere in particular. I said the words, "I wish you

were real." Realizing what I had said, I felt a rush of embarrassment and strangeness as if I had spoken something out of a sci-fi or fantasy tale. A contemporary, subconscious desire akin to Margery Williams' children's story, *The Velveteen Rabbit*, where the boy wishes for his toy to become real. To this day, I'm not exactly sure what compelled me to utter these words into the air. All I knew was that someone was listening and heard me.

The aforementioned scene took place in my home on artist Lauren Lee McCarthy's third day of performing her durational, participatory work *LAUREN* (2018). These were the last words I spoke to the persona she had created. I was the first individual whom she knew personally and considered a friend. In previous versions of the work, McCarthy vetted strangers from an open call process. Her primary task was performing as an assistive technology and her hope was to see if she could outperform popular smart devices, primarily, Amazon Alexa. Inspired by artists such as Sophie Calle and Chantal Ackerman, McCarthy's *LAUREN* relies on the tensions between contemporary home life, the mundane, human communication, smart technologies, and voice recognition.³³ *LAUREN* explores the role of assistive

³³ "Sophie Calle," The Guggenheim Museums and Foundation, accessed March 26, 2023, <https://www.guggenheim.org/artwork/artist/sophie-calle>; Rachel Donadio and Cara Buckley, "Chantal Akerman, Whose Films Examined Women's Inner Lives, Dies at 65," *The New York Times*, October 6, 2015, sec. Arts, <https://www.nytimes.com/2015/10/07/arts/chantal-akerman-belgian-filmmaker-dies-65.html>.

tech and how humans both command and emote. For the performance in my home, McCarthy interviewed six of my closest friends to get a sense of how to best serve as my digital assistant. In previous iterations, she spoke directly with the person or family to gain a sense of the needs of the household. The strange and surreal experience of not knowing her corpus of data from which she performed is not too dissimilar to how assistive technologies function. You get out of the technology what you give, share, and allow it to know about you or else, someone or something will provide that corpus.³⁴

She placed custom hardware (similar to a Google Nest) around my apartment to monitor my movements via remote surveillance from an undisclosed location. The performance was also meant to explore the boundaries of interpreting gestural and linguistic subtext to get a sense or intuition about what I needed before asking through active listening and observing my behavior. Despite robotizing her voice, due to our friendship and familiarity with one another, I was able to recognize the voice as distinctly her. I was able to discern her prosody, tone, and affect despite its modulation through a vocal filter.

Similar to her more recent series, *MUTED* (2021), in her work *What Do You Want Me to Say?*, McCarthy writes custom software that enables an

³⁴ A third party vendor or company, for instance, will collect and re-distribute data.

individual to use her voice clone of something they might not be able or comfortable to speak aloud.³⁵ At the height of the COVID-19 global pandemic when many people were working from their homes and meeting via zoom, she cloned her voice to have it speak in her absence from web-based conference calls such as zoom and Google Hangout. Although her works elicit an emotional response, they are meant to navigate the immateriality of voice and emotion into a material reality of software and hardware. Her speculations and performance work are part of the lineage related to service-oriented IVRS. For example, the How May I Help You (HMIHY) technology, created out of the AT&T labs in the late 1990s was created to handle customer calls such as re-direct billing and payment inquiries. Yet the system still required human intervention from individuals called, "wizards."³⁶ Unlike McCarthy's creations of assistive tech and a voice clone, the goal was to create a system where humans would become unnecessary. These types of voice recognition technologies are less about people and more about technosolutions that benefit for profit companies such telecommunications, retail, and banking industries.³⁷ While speech-to-text is important for documentation purposes,

³⁵ "What Do You Want Me to Say? – Lauren Lee McCarthy," accessed January 4, 2023, <https://lauren-mccarthy.com/What-do-you-want-me-to-say>.

³⁶ A. Gorin et al., "Automatically Training a Problematic Dialogue Predictor for a Spoken Dialogue System," 2011, <https://doi.org/10.48550/ARXIV.1106.1817>.

³⁷ Gorin et al.

speech-to-speech presents a new frontier of voice technologies inextricably linked to HMIHY technology meant to codify speech.

But in recent years, the ethics of voice clones for posthumous use have come into question. In Morgan Neville's 2021 documentary *Roadrunner: A Film About Anthony Bourdain*, he used a voice clone in the film to speak Bourdain's words and stated, "we can have a documentary-ethics panel about it later."³⁸ In *What Do You Want Me to Say?*, McCarthy does something similar. She enables a participant to use her voice to say anything they wish. There is a tremendous vulnerability in creating a work that allows people to speak freely and listen to their own words repeated through her voice. Reflexivity is involved in both the artist and the participant-listener. This work does not allow for iteration, once you speak into the work and it listens and recognizes your speech and processes it, within a matter of seconds, you hear your words spoken in the prosody of the artist herself. The strange and surreal feeling is akin to what might be described as a sonic and aural uncanny valley experience. Both *LAUREN* and *What do you want me to say?* provoke new ways of listening and trusting the acousmatic voice.

³⁸ Helen Rosner, "The Ethics of a Deepfake Anthony Bourdain Voice in 'Roadrunner,'" *The New Yorker*, July 17, 2021, <https://www.newyorker.com/culture/annals-of-gastronomy/the-ethics-of-a-deepfake-anthony-bourdain-voice>.

McCarthy explores themes of "vulnerability, ownership, and authenticity."³⁹ While the use of a voice clone allows for this particular mimicry, the work also touches upon the tension that doesn't necessarily capture the intent of the original speaker. Prosody, or the animation or tone of the speaker's voice, is somehow lost and the same, seemingly flat affect of McCarthy's voice as it mimics what has been asked "what do you want me to say" is spoken by the voice clone. It's as if the point of the work seems to accentuate what we are far from accomplishing within the realm of voice human-computer interaction. Tonality is not something easily mimicked because the tone, pitch, and intonation play a tremendous role in what makes understanding someone's intent challenging especially when mediated through a clone or machine. McCarthy's provocation seems to be one of both open access and trust in another person's ability to use her voice in a very specific way.

Speech-to-speech or voice recognition provides a distinct pathway to explore what a player or listener or speaker might want the voice clone to say/speak. Reminiscent of the Speak and Spell (an educational toy released in the 1970s), I wonder how this work might translate for a non-native English speaker? The artwork allows a type of play, but it also assumes that the person

³⁹ "What Do You Want Me to Say?"

knows what to say and how to say and speak with an American English accent. Artist and scholar Alessia Cecchet describes Grillo Parlante (Talking Cricket), the Italian version of the American Speak and Spell that was programmed to be much more assertive than its American counterpart with a “metallic voice,” which aptly describes how synthetic voices have been created.⁴⁰

McCarthy uses the Descript Voice API, which supports up to 23 languages including Spanish, French, German, and Italian.⁴¹ While McCarthy’s voice clone repeats with her prosody intact, it mimics my speech phonetically and mispronounces words. While the labor is mediated through the voice clone, it is the intent of the user that is of greater concern. McCarthy's voice may stand in to displace the vocal and emotional labor, but not as a way to express.

Companies such as Bay Area-based Sanas AI are developing software that performs real-time accent elimination, which works against inclusion and perpetuates accent bias. These types of technological advancements uphold American and British Englishes and accents that calibrate, discipline, and

⁴⁰ Alessia Cecchet, “The Italian ‘Speak and Spell’: Learning Through Anxiety,” Online Publication, Lady Science, January 3, 2018, <https://www.ladyscience.com/essays/speak-and-spell-learning-anxiety>.

⁴¹ “New in Descript: Transcription in 23 Languages! | Descript,” accessed March 26, 2023, <https://www.descript.com/blog/article/new-in-descript-transcription-in-23-languages>.

center voices and speech around whiteness. The perpetuation of what a specific, intelligible, and helpful voice ought to sound like negates non-white and western English accents and further suppresses a diversity of speech. In the 21st century, the human voice is seen as a new frontier of both techno-solutionism and capitalism. In order to expand upon existing voice recognition and assistive technologies, it is vital to understand how a large company such as Apple shares code targeted at developers that can help build out the companies software and hardware products without having to work directly with the company itself. For example, Apple's annual Worldwide Developer Conference presents the sleekest permutation of its devices to a large audience of developers and programmers. Yet how might understanding Apple's developer archives provide a sense of its audience, how software is developed, and if expanding a codebase focused on voice is possible for those that trained as developers and computer scientists might take on this work in the future?

Accent elimination software is a major issue because it works against the tenets of global accessibility and inclusion that is already absent from the development of many big tech softwares. These technologies exclude as they perform their functions. Proprietary software and hardware also become codified with standards that do not account for varied types of speech and

minimize, if not eradicate, and diminish atypical voice and speech. This form of disciplining the body through the voice does not take into account the multitude of cultures and languages it excludes while sustaining and upholding deep-seated colonialist perspectives on speech and communication.⁴² My framing of this issue encompasses global citizens who have studied American or British Englishes as compulsory, which slowly eradicates vocal agency to a mother tongue because English has been deemed as a language needed for survival and employability. This mode of learning also ensures a type of programmability of the body and voice.

With a commitment to open source software, Halcyon M. Lawrence's call for change in diversifying the speech corpora that many of the assistive technologies we use today are based on. She writes, "We must look to independent developers and free and open-source software initiatives that support independent developers to make greater strides in disrupting the system."⁴³ A concrete example of what Lawrence is talking about is Mozilla's Common Voice project that launched in 2017, which has collected hundreds of

⁴² Inspired by technical communications scholar Halcyon M. Lawrence and using her scholarship as a framing for how the voice results in smart and assistive technologies being programmed to listen to a specific type of person. This framing also directly relates to Ruha Benjamin's use of the term "default settings" as a metaphor for how technologies are built.

⁴³ Halcyon M. Lawrence, "Siri Disciplines," in *Your Computer Is on Fire*, ed. Thomas R. Mullaney et al. (MIT Press, 2021), 191.

speech samples from thousands of people around the world with a specific focus on marginalized accents and languages.

From IVRS to the development of smart speakers and phones, Langdon Winner's work and questions regarding the politics involved in the making of digital and mobile technologies is at a crucial point in our cultural and contemporary moment.⁴⁴ Understanding the policies and creations of specific patents such as Spotify's patent (which they claim to not be moving forward on due to the efforts by groups such as Access Now) on collecting environmental sonic data is tethered to a type of sonic and vocal surveillance that is unparalleled to what we have seen in the past.⁴⁵ Adding to the political and cultural concerns of smart technologies, scholar Joseph Turow looks at the capitalist ventures of machines being programmed to listen to our tone, our pitch, and affect to then be made marketable and used as an unseen yet potent data source.⁴⁶ From voice-activated architecture to speech messaging in cars, pressing concerns for Turow include bio-profiling and habituation to voice-first technologies that seek to provide extreme personalization.⁴⁷ Similar

⁴⁴ Langdon Winner, *The Whale and the Reactor: A Search for Limits in an Age of High Technology*, Second Edition (Chicago: Univ. of Chicago Press, 2001).

⁴⁵ Jennifer Brody, "Spotify's Speech-Recognition Patent Tech: Global Coalition Says Don't Spy," *Access Now* (blog), May 19, 2021, <https://www.accessnow.org/spotify-spy-tech-coalition/>.

⁴⁶ Joseph Turow, *The Voice Catchers: How Marketers Listen in to Exploit Your Feelings, Your Privacy, and Your Wallet* (New Haven: Yale University Press, 2021).

⁴⁷ *Ibid*, 230.

to precision medicine, where medications are hyper personalized for the patient, voice-first technologies produced and developed to provide a near-human interaction to the consumer also presents a raciolinguistic issue. For instance, the voice recognition of the Apple iWatch promises a device that may provide life-saving support that gives primacy to the machine and not the complex emergency infrastructure that the device mediates. The device itself does not save the human being in peril. It is an ecosystem that saves humans. Yet, a life threatening call is also based on the ability to trust the dispatcher to receive and listen to the message, but equally, the caller assumes that they will be understood.

On Building the Technology We Need

One of the recurring questions throughout my readings and observations of computer science journal articles, Apple developer blog posts, and media depictions of smart (voice) technologies was whether or not the existence of an acousmatic voice is diminished in the digital age. With smart speakers and phones listening for wake words and compiling a massive corpus of data through our appointments, shopping lists, and music preferences, I wondered if a disembodied voice ceases to exist because the source is the

technology itself. The hardware becomes a type of embodiment for a voice and our interactions with it. From video games to call centers, how does learning the codified digital voice change our conceptions of language, speech acts, and elements of everyday communication and inadvertently force us to speak in unnatural ways to conform to the listening machine?

During my artist residency at The Frank-Ratchye Studio for Creative Inquiry at Carnegie Mellon University (Spring 2023), I was a visiting artist and scholar at Duolingo, the language learning application. I had the opportunity to be in conversation with engineers working on computational linguistics, cognitive scientists, and instructional designers studying how humans learn to speak language. One of the most striking questions I received during an artist talk dealt with the American education system and, arguably, the way students are taught to speak English and common pedagogical approaches are forms of accent reduction or modification. Taking this into account, I reflected even further into the role digital technologies play in correcting speech and how machines are developed to listen until they hear what they have been built to hear. From Amazon's Alexa to the Apple Watch advertisements, these media form(at)s serve as reminders of the ideology behind technological solutionism that strives to make the voice a new frontier for further advancement. Delving into the HMIHY technology also provides a foundational approach in

recognizing how silence is built into the technology as a type of affirmation or agreement to what the machine understands that might not be true for the speaker.

HMIHY technology and Amazon Alexa might seem like disparate platforms, yet each belongs to the same technological lineage, which is why media depictions of interactive talking machines begs our attention as they speculate an imagined future. While the average super bowl watcher may flippantly watch Michael B. Jordan performing as Alexa, there is an implication of a binary and complicated nature of a Black man embodying a technology. To encourage the questioning of smart technologies as subservient and intentional automatons with a docile affect may result in deeper, intentional conversations regarding accessible, inclusive, and necessary technologies. These depictions are not evergreen and the activation of Jordan's body in service to another human being even if fictitious perpetuates a type of fantasy of technology that negates the changes in voice and how technology must then constantly change in order to evolve into our needs.

Another key consideration of synthetic voice is how prosodic elements are applied if not to make the human speaker comfortable with engagement and to trust the voice they hear. During a two-day workshop conducted

through Data & Society titled Parables of AI, which focused on the geographic regions of the Global South, I was paired with Filipino artist Mac Andre Arboleda who commented on my work *Press 1 to be Connected* (2019), an interactive speculative fiction about DNA banking. I altered my voice in post production to sound robotic, mechanical, and procedural. Even as I attempted to perform with a relatively flat affect, Arboleda commented on the neutrality of my voice as a way to position myself as a performer and the assumptions I make to sound like a synthetic voice used in IVRS. As a consumer-oriented culture, we are conditioned to understand the synthetic voice as merely programmed for transaction and not care, which informs the way future IVRS platforms are developed. Rather than the binary of utopian and dystopian sonic landscapes, my research into how prosody and synthetic voice(s) seeks to contribute and build the conditions for technological advancements of everyday life that are needed, useful, helpful, and just.

Creative Intervention, Part I: Protocols, Scripts, and the Degradation of Speech

The monotone, feminine synthetic voice you hear when calling a large company such as an internet service provider or healthcare company prompting you to stay on the line for a survey or press a given number for a particular service is a relatively common interaction. While my various job roles in biotech did not require me to participate in study start-up activities such as setting up the interactive voice response system (IVRS) for sites, I was made privy to the planning and logistical documentation as well as access to platforms where I had to monitor drug supply. The IVRS was one of the vital systems needed to run a clinical trial because it was used to collect and process data collected from the clinical sites participating in a study. The study data coordinator called a designated phone line and either keyed in all of the necessary subject data and information or spoke subject numbers and information into the system. Electronic data capture (EDC) was also a way to capture data through an automated system or platform, mostly through an internet platform. Although mundane and seemingly worthless to those outside of a clinical research trial, these intricate and complex systems ensured

the dispensation and dissemination of drugs and would determine the needs of clinical supplies and equipment. Although IVRS is not as common now due to haptic technologies and advanced ways of collecting data on the subject side (i.e. custom mobile apps and games to capture clinical trial data, etc.), I was most interested in how voice was used to relay information and inspired by the form(at) of the phone tree to create interactive fiction since it is akin to choose-your-own-adventures stories.¹

Subverting formats, protocols, and procedures has become an underlying theme and process within my creative practice. The most influential work to my creative practice has been the anthology *The Shell Game: Writers Play with Borrowed Forms*.² From a story of grief and loss told through a series of rejection letters to lines of code re-written to illustrate the cursory, programmed aspects of a dysfunctional married life, the subversion of formats aides in the telling of some ineffable narratives. In a series of interactive speculative fictions created through open source program Twine to the re-creation of dispatcher protocols, *Do No Harm* (2017-present) is an attempt to make sense and reflect upon the complicated nature of human

¹ When I worked in biotech, from 2003-2014, some of the pediatric trials explored creating gaming platforms or mobile applications in order to gamify data collection.

² Kim Adrian and Brenda Miller, eds., *The Shell Game: Writers Play with Borrowed Forms* (Lincoln: University of Nebraska Press, 2018).

voice as it is mediated through various types of technologies and dictated by systems.

Do No Harm (2017-Present)

In this section, I provide a project description that elaborates the current state of *Do No Harm* (2017-Present) through the creation of print media, interactive speculative fiction, experimental prose, and a durational, participatory performance. I cover the intended audience of the work as well as the documentation and research involved that showcase how this work came into fruition. Although both projects focus on voice, they were inspired by two completely different sets of frameworks, questions, and methodologies. I also pondered what aspects of human computer interaction must be taken into consideration when applying tenets rooted in the creation of more socially just systems for the general public? What does a trustworthy voice sound like and at what register can it be trusted?

Do No Harm (2017 - present) is a series of artworks exploring the media archaeology of analog, digital, and mobile technologies associated with data collection, transmission, and the lexicon associated with telecommunications and operating. This series of works include flash fiction, experimental

docu-poetics prose written in the style and format of various types of protocols (i.e., relay operating script, etc.). In its original inception, the premise of the stories written for *Do No Harm* are based on a queer, gender non-conforming scientist engaged in creating a clinical trial speculating CRISPR biogenetic technology enabling the eradication of social and cultural diseases such as toxic masculinity, generational trauma, and xenophobia in addition to treating life-threatening diseases.³ The literature for *Do No Harm* aims to re-examine the format, structure, and design of scientific documentation by implementing iconography, accessible language, and narrative as a way of addressing how clinical research and consent are, oftentimes, challenging, and inaccessible to the general public.

The three major considerations for the stories are as follows:

- 1) ***Format:*** Historical research on various types of formats used within science technology such clinical protocols to relay operating scripts
- 2) ***Structure:*** Experimentation with time, space, and relevant historical moments in the arts, sciences, and technology.

³ “Social Disease Definition & Meaning - Merriam-Webster,” accessed April 6, 2023, <https://www.merriam-webster.com/dictionary/social%20disease>.

3) *Design*: Inspired by Steve Tomasula's novel *VAS: The Opera of the Flatlands* and Raymond Queneau's book *Exercises in Style* where he tells the same story in ninety-nine different writing styles including tactile, visual, auditory, and telegraphic.

Press 1 to be Connected (2019), a multimedia sound project and part of the *Do No Harm* series (prototype) explores IVRS, experimenting with the phone tree as a medium for creating interactive speculative fiction. Since telephony has long been used to communicate and relay data, the work will rely on mobile phones as the listening device. The phone tree is engaged reflexively to draw the listener into a particular narrative in anticipation of the listener returning to explore other outcomes that may incorporate new forms of storytelling and game mechanics. The project is meant to elicit the imagination of the listener as well as to enable their agency in directing the outcome of the narrative they are initially told. The sound work is composed of sounds including artifacts from the phone tree experiment and phone tree history sound artifacts. The graphics involve possible mappings of a complete story from beginning to potential end, and the tree structure. *Press 1 to be Connected* was the result of a collaborative project in 2018 with artists Liat Berdugo, Gabi Schaffzin, and principal investigator Professor Lisa Cartwright

through the Experimentation Studio facilitated by the University of California Humanities Research Institute. The Patient Instruments was meant to take various technologies and examine their use. My object of study was the telephone.

During the implementation of clinical research studies, one of the most important aspects involves fastidiousness, high-level mapping, and organization of data collection systems and infrastructures. From the IVRS to standard of care procedures, biodata, biometrics, and surveillance are required for the success of a clinical study. What is the success of a clinical trial? It is the improvement of condition or diseased state. It's the potential eradication of a virus or its the correction of a gene sequence. Culling from genetic and biological materials such as DNA, blood, and spit, data has become a commodity. Yet data requires a medium for it to exist and for it to be effective. En masse, genetic data can be used to survey a population and enable scientists (both social and bench scientists) to engage in social and cultural dialogues around issues such as citizenship. Genetic material within the realms of biotechnology and the pharmaceutical industry become the media fueling machine learning and artificial intelligence with algorithms mined from individuals. But how does a medium become a source of prejudice and stratification of Othered bodies? What are the parameters one must pay close

attention to in order to understand how to rectify unjust practices and protocols? In our contemporary moment, how are systems of care growing closer to the self-documenting subject? How does clinical research put the onus onto the patient-turned-subject through smartphone devices? By looking at issues of consent related to surveillance and systems of care, can we start to think of new ways data and collection of genetic material can further advance science in ethical and just ways?

Related to systems of care, I tried to envision how we might sonically imprint our identities in the future. Originally, I wanted to run a durational performance that provided print media content that invited people to become a part of a two-three years performance that would allow me to be in conversation, literally, with community members that provided their consent to become a part of a study, but in actuality, a performance. This idea has not been abandoned, but for the sake of the existing research, I wanted to allow time and space for how people might want to entertain the idea of being a part of a larger, ongoing performance work. Ideally, I want to collect sonic impressions and allow conversation to become a part of the work provoking critical thinking related to the ways we understand and engage in our immediate ecosystems.

The following eight screenshots show the choices a player must make to advance throughout *Press 1 to be Connected* (see fig. C1 1.3-1.6).⁴ As previously mentioned, I was inspired to create a version of the illustration on the 1984 Byte Magazine cover by artist Robert Tinney of the red telephone that resembled the minitel (see fig. C1 1.2). I wanted participants to feel both nostalgic and slight dissonance with the interactive fiction they were listening to. The work was produced during a three-month artist residency at Stochastic Labs in Berkeley, California. I met with geneticists and bioethicists to discuss media depictions of science, in particular, CRISPR (gene editing technology).⁵ *Press 1* posited various forms of data collection and was also influenced by my work experience in the pharmaceutical and biotechnology industry.⁶

⁴ To play *Press 1 to be Connected*, it is hosted on philome.la, which was a project that allowed for games and stories created through Twine, a digital space. But the project has since sunset as of 2019. My work is archived and active, via the following link, but I will be uploading to itch.io in the near future.

<https://philome.la/deedottiedot/press-1-to-be-connected/play/index.html>

⁵ “CRISPRartists,” *Stochastic Labs* (blog), accessed May 1, 2023,

<http://stochasticlabs.org/crisprartists/>.

⁶ I worked in the research and development arm of two major Bay Area-based biotech companies within clinical operations and supported the start-up, maintenance, and close out of studies that enabled me to learn about a wide array of scientific processes related to drug development including how protocols are authored, informed consent, and DNA banking.



Figure CI 1.1: *Press 1 to be Connected* (2019), sculpture, multimedia, voice recordings

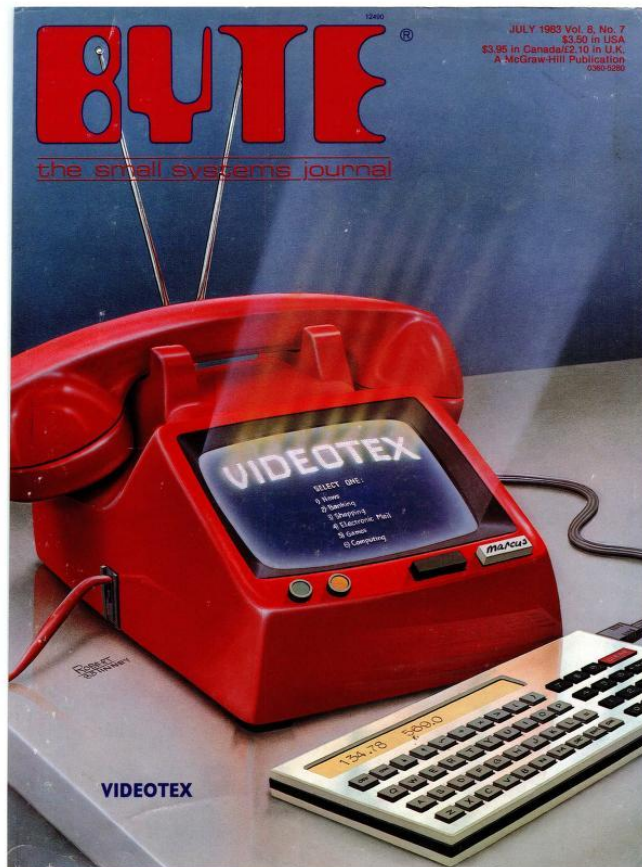


Figure CI 1.2: Cover of 1984 edition of BYTE Magazine with an illustration of the Videotex phone by Robert Tinney

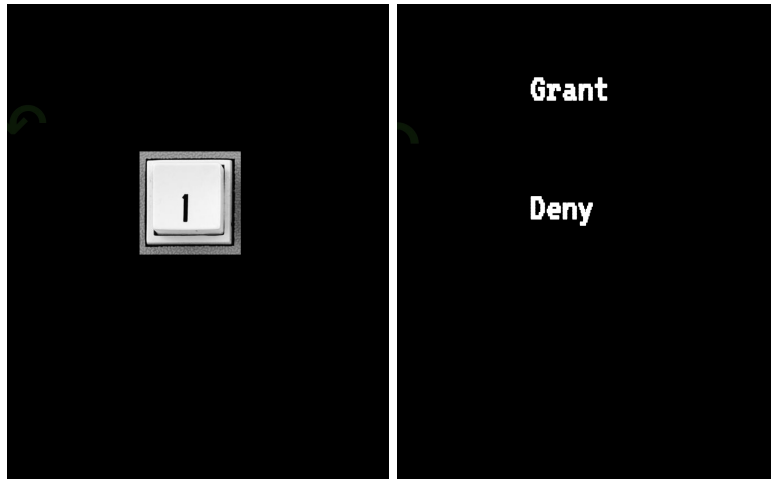


Figure CI 1.3: Screenshot the first two screens of *Press 1 to be Connected*. These are screens that appear on the sculptural component of the red phone. (Produced by author).

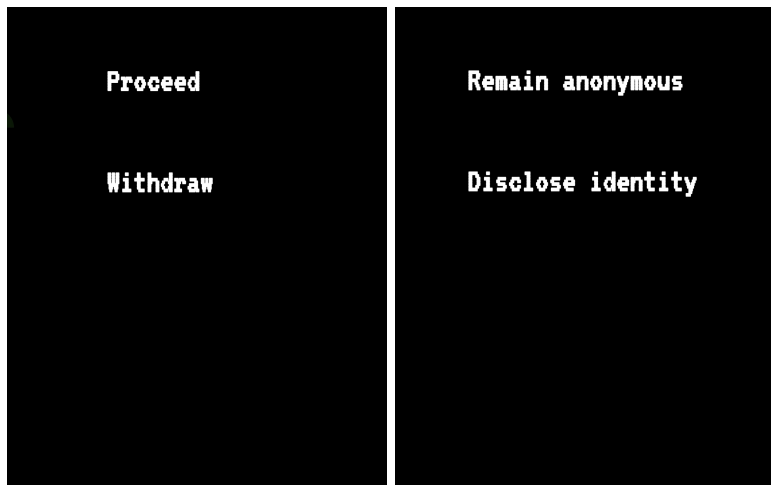


Figure CI 1.4: Screenshot the third and fourth screens of *Press 1 to be Connected*. These are screens that appear on the sculptural component of the red phone. (Produced by author).

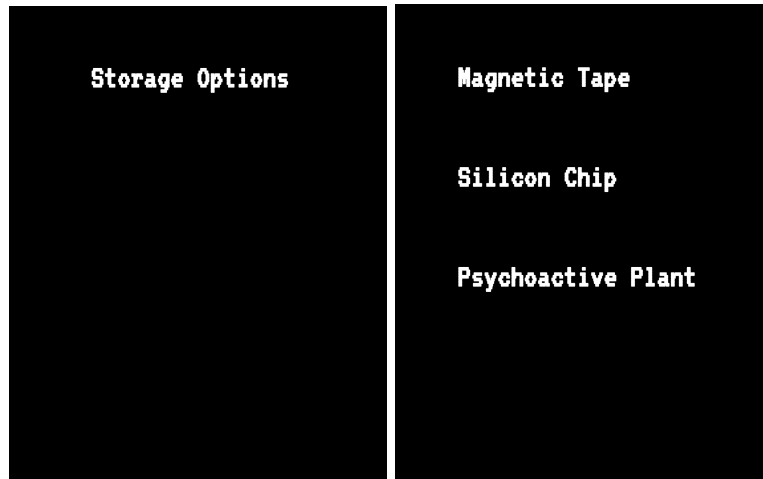


Figure CI 1.5: Screenshot the fifth and sixth screens of *Press 1 to be Connected*. These are screens that appear on the sculptural component of the red phone. (Produced by author).



Figure CI 1.6: Screenshot the seventh and final screens of *Press 1 to be Connected*. These are screens that appear on the sculptural component of the red phone. (Produced by author).

The interface was minimalist because I wanted to prioritize the storytelling. I made the decision to narrate the experience and put a filter in post production editing to give the illusion of my voice communicating through a telephone, which is the effect I wanted considering the sculpture

component (phone) is of an era where the transmission of voice was not at its highest fidelity. Artists Alex Reben and Steve Thompson supported the hardware engineering to replace the touch tone buttons with a Raspberry Pi haptic screen. The first iteration of the project was a success and was shown in Linz, Austria at Ars Electronica in 2019.

Aside from the research, interviews, and creative sketches of this work, I refreshed my memory of Mark Hansen and Ben Rubin's *Listening Post* (2003).⁷ Roberta Smith wrote about the piece for the New York Times in 2003 when it showed at the Whitney. She stated, "'Listening Post' is simply the latest twist in the familiar modernist tradition of making art from chance arrangements of everyday materials, and is more a result of technological progress than genuinely new thought."⁸ I remember seeing this work in 2006 at the San Jose Museum of Art. As Smith confirms in her piece, the phrases that filled the LED screens were culled from Internet chat rooms. In one instance of the work, we listen to what sounds like a masculine voice reading what appears on the screens. The constant phrase repeated over and over again is "I am..." followed by whatever has been entered into a chatroom

⁷ Mark Hansen and Ben Rubin's *The Listening Post (2002-06)* in *Almost Human*, 2020, <https://www.youtube.com/watch?v=RzfnnddqfCk>.

⁸ Roberta Smith, "ART IN REVIEW; Mark Hansen and Ben Rubin -- 'Listening Post,'" *The New York Times*, February 21, 2003, sec. Arts, <https://www.nytimes.com/2003/02/21/arts/art-in-review-mark-hansen-and-ben-rubin-listening-post.html>.

conversation. Smith stated something emblematic of many new media and digital artworks is, “a result of technological progress rather than genuinely new thought.” According to the timeline of the Computer History Museum, in 2003, MySpace was founded.⁹ On the heels of chat rooms and conversations connecting the entire world together via AOL Online, social networking sites started to pivot users to a new way of connecting. Overall, the reason why I wanted to revisit the work years later in the making of *Press 1 to be Connected* was due to the fact that it demanded a type of listening and vocality that was focused on textual information en masse. I sought ways to use this concept of feeling and being inundated, but in a more precise and insular way.

While I do not remember the voices in the work when I saw the piece in San Jose, I recall the silence in the room. At some points, walking around to observe the pieces from different angles and perspectives, the silence seemed to amplify the mild and softened sounds of the LED screens changing from phrase to phrase. During my first year at California College of the Arts, I wrote a visual analysis of Chris Baker’s *Murmur Study (2012)*. He used thermal printers as a means to showcase the how we communicate, but also to convey the transactional nature of clicking and posting that generates data that turns into a commodity, it’s even more striking that both of these works, once

⁹ “2003 | Timeline of Computer History | Computer History Museum,” accessed January 16, 2023, <https://www.computerhistory.org/timeline/2003/>.

grandiose with their installations and pulling from our desktop screens is now pulled from data directly from our devices, human and machine are both engaged in what Kate Crawford as called “peripheral listening” to anyone, anytime, and from anywhere.¹⁰ The subversion and hacking of older technologies has continued to inform my creative process.

Discomfort plays an essential part of my artistic practice. While I do not have a desire to be overtly antagonistic, my hope is to make the viewer or participant question their involvement in the work and how they might be implicated in it. Artist A.M. Darke’s work *Fairly Intelligent* is commanding in its non-discriminating and unapologetic manner as her voice leads you through prompts about race, class, and gender.¹¹ Although humorous in nature and a satire of how data is collected by consumers or digital and mobile users, its biting social and cultural commentary come in the form of hypertext questions and answers. Despite the inability to change answers, the point seems to draw upon the nature of how most, if not all, digital tech functions within the realm of development, dissemination, and deployment. It is meant to serve as a reminder that many people do not feel the agency or choice to choose. The artist's voice as a mechanism to listen to the overall work is meant

¹⁰ In the Kate Crawford way of “peripheral listening.”

¹¹ A.M. Darke, “Fairlyintelligent.Tech,” Artwork, *Fairly Intelligent*, accessed May 1, 2023, <https://fairlyintelligent.tech/>.

to be narrative and commentary-based as a way of leading the listener and player through ways of seeing the tech through a Black feminist lens. It's meant to be relatively contentious in nature to show the challenges experienced by BIPOC community members. The term "fairly intelligent" also alludes to the fact that this is not an artificial intelligence machine or interaction, but one that is predicated on an experience. The human playing the machine and performing as the algorithm. Brenda Laurel's book *Computers as Theatre* delves into the various elements that make up relationships especially within performance work such as action, character, thought, language, pattern, and enactment.¹² I raise this issue because with work that is relatively straightforward and provides a simulation of choice as opposed to Zubek's definition of a branch and merge experience, we see that the mechanism Darke uses is one of simulation, which means the content and nature of the work must be audacious in its scope of wanting the player to the understand the problem she is addressing.¹³

Visualization of speech has its place and associations within the legal system and law enforcement. From telecommunications being considered a part of the defense industry to the documentation required of 911 calls for legal

¹² Brenda Laurel, *Computers as Theatre*, Second edition (Upper Saddle River, NJ: Addison-Wesley, 2014).

¹³ Robert Zubek, *Elements of Game Design* (Cambridge, Massachusetts: The MIT Press, 2020).

purposes, the voice is implicated the moment an utterance is vocalized. In Lawrence Abu Hamdan's work *The Witness-Machine Complex* (2021), he collected seven moments of the Nuremberg trials to showcase the sonic experience of interruptions throughout the trial.¹⁴ The space between speech and utterances also provides equally noteworthy information about the way speech flows between bodies and what this says about the cadence and intent of voices. How might a visualization of what was happening during the trial make a witness speak and remember an event differently. Abu Hamdan attempts to find what he has referenced in the past as a "leakiness" to sound.¹⁵ The particular nature of the complex he is speaking to within this piece is in large part to the way he wants the viewer and listener to understand our multi-faceted role as witness. Translation is another significant aspect of this work in that much of what we think about when we think about voice recognition and speech synthesis and generation has much to do with what is deemed the "proper" version of any language that negates the vernacular or everyday speech. Also, people speak within space based on the context of their

¹⁴ "Lawrence Abu Hamdan on Translation, Nuremberg, and the Juridical Unconscious - Artforum International," accessed January 16, 2023, <https://www.artforum.com/interviews/lawrence-abu-hamdan-on-translation-nuremberg-and-the-judicial-unconscious-86875>.

¹⁵ *ATC Lecture – Lawrence Abu Hamdan's "The Sonic Image,"* 2021, <https://www.youtube.com/watch?v=mvyQoYWfSvA>.

immediate environment. Within a court of law, it's safe to assume that the way someone speaks will be representative of their "best" speech.

Imaginary Relay, For Andy (2023)

During an artist and research residency at San Jose State University over Fall 2022 and Winter 2023, I produced works based on the condition of the grant awarded to artist Professor Rhonda Holberton, which was to pull media artifacts to produce artwork as a part of an exhibition, *Hiding in Plain Sight* (see Fig C1 1.9-1.13). The residency required Bay Area-based artist Sofia Cordova and I to go through the university's special collections. I selected the Indian Diaspora Oral History project that included fourteen video-taped oral histories of Desi community members. Nand "Andy" Khanna was the oral history I chose in large part due to the fact that he immigrated to the US during the Summer of Love and completed his undergraduate degree at the University of San Francisco.¹⁶ For *Imaginary Relay, For Andy* (2023), I listened to Mr. Khanna's oral history and went through the 80-page transcription of his interview and selected key phrases I felt resonated or reminded me of sentiments my mother has shared with me about her immigration story to San

¹⁶ Indian Diaspora Oral History Project Collection, MSS 2013-03-26, San José State University Library, Special Collections & Archives. Box 1. Folder 1.

Francisco, California. I copied and pasted his words onto a document and read them many times over and decided to re-cast myself as the interviewer with the constraint of using a relay operator protocol and format that forced me to truncate the written word to make it appear as if it is spoken language. The speculative nature of my speech situates me in a strange and surreal exercise of imagining a conversation mediated through an imaginary relay (see fig. C1 1.7.). The blue text represents my imaginary questions and comments in response to Mr. Khanna's actual, spoken words. The faded text is the imaginary relay mediating our words with the acronym "GA," which translates to "Go Ahead" when using a relay service. The primary goal was to be in conversation with the archive and explore what happens when one does not always agree with what is documented within it. There were many moments I found I resisted some of the ideas Mr. Khanna shared because it deeply impacted the ways in which he raised his sons and there were instances

where I found similar ideological ideas akin to some of my mother's ideas.

hello, mr. khanna, are you there? i wanted to share some reflections with you, if you're okay with me doing so? i hope you don't mind. What made you want to immigrate to the US?

hello, mr. khanna, are you there? i wanted to share some reflections with you, if you're okay with me doing so? i hope you don't mind. What made you want to immigrate to the US? GA

the purpose of coming to the U.S. was to find a better opportunity

the purpose of coming to the U.S. was to find a better opportunity GA

seems like you had no reservations about being in a place where you didn't know anybody. no doubts whatsoever?

seems like you had no reservations about being in a place where you didn't know anybody. no doubts whatsoever? GA

i wouldn't even think twice, just take the next plane and go to America

i wouldn't even think twice, just take the next plane and go to America GA

Figure CI 1.7: Screenshot of *Imaginary Relay, For Andy* (2023), digital

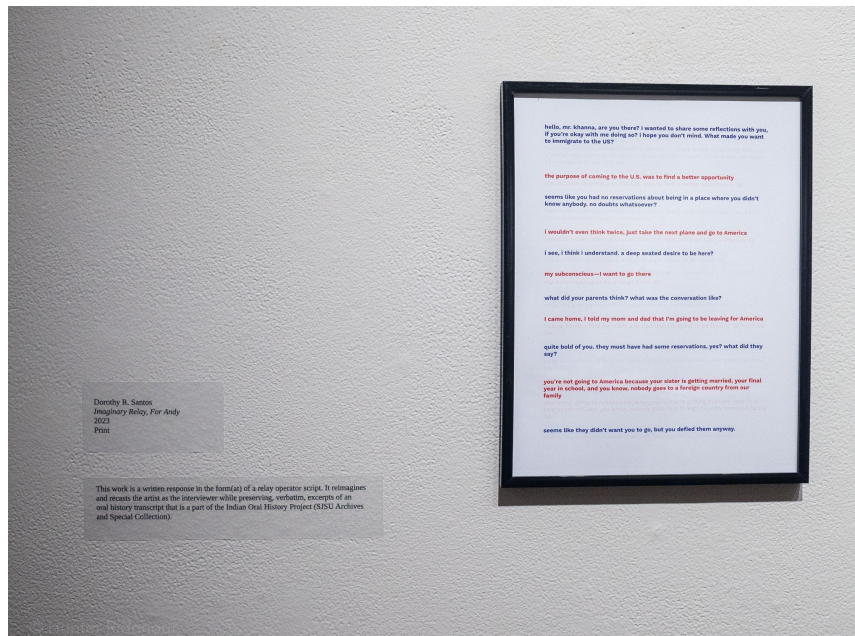


Figure CI 1.8: *Imaginary Relay, For Andy* (2023), print (Photography by Hunter Ridenour).



Figure CI 1.9: *Imaginary Relay, For Andy* (2023), print (Photography by Hunter Ridenour).



Figure CI 1.10: *The Cyborg's Prosody* (2022-present), print of Level 1 of digital game (Photography by Hunter Ridenour).



Figure CI 1.11: *Antiphon: At the Frequency of the Refrain* (2023), mixed media
(Photography by Hunter Ridenour).



Figure CI 1.12: Installation shot of *Imaginary Relay, For Andy* (2023), print and *The Degradation of Speech* (2023), video
(Photography by Hunter Ridenour).



Figure CI 1.13: *Antiphon: At the Frequency of the Refrain* (2023), mixed media
(Photography by Hunter Ridenour).

Some of the artists that have deeply influenced the way I think about sound, voice, and assistive tech include Lauren Lee McCarthy and her work *SOMEONE* (2019), which re-imagines a human version of Amazon Alexa, a smart home technology. For the two-month duration of the Refiguring the Future exhibition, four participants' homes around the United States were installed with custom-designed smart devices, including cameras, microphones, lights, and other appliances where gallery visitors had the choice of performing as the assistive tech for one of these participants at specific times and hours. Ryan Kuo's app-based work *Faith* (2019) as described by art critic and scholar Darla Migan, "an AI voice assistant, a femme-bot who

actively fails to meet the patriarchal hail. She is not based on a single person but is a compilation of many kinds of speech gathered from online chat forums."¹⁷ The common thread in both McCarthy's and Kuo's is finding where the boundaries are within language, consent, and labor. I was compelled to pursue opportunities to use my findings and reflections from all of these works that made voice central and make efforts to plan another iteration of *Press 1* for the future. During a 2021 experimental two-day intensive workshop organized by Data & Society's researcher Ranjit Singh and senior producer Rigoberto Lara Guzmán titled Parables of AI focused on the Global South, I received useful and generative feedback from other artists and technologists based in Africa and different parts of Asia.¹⁸ I was paired with Filipino artist and technology Mac Andre Arboleda who provided me with insightful feedback. One of the most useful of their constructive criticism was the overall neutrality of the work due to the flat affect of my voice and delivery. They also recommended reading *The Quiet Ones* (2019) by Glenn Diaz due to the premise of the story focused on the offshore outsourcing in the Philippines and the poetic license Diaz takes in describing the psychological effects of being Filipino and working to serve Americans.

¹⁷ Darla Migan, "A Poetics of the Glaze," *Shift Space*, 2022, <https://www.shiftspace.pub/a-poetics-of-the-glaze>.

¹⁸ Data & Society, "Searching for Parables of AI," Medium, December 8, 2022, <https://points.datasociety.net/searching-for-parables-of-ai-3317d201edeb>.

Chapter 3: Hidden Frequencies: Systems and Logics of Care within Emergency Infrastructure

With the ever-increasing dependence on smartphones and assistive technologies, crisis management systems rely on people (often women) of color to serve as a mediator and conduit of information transmitted between the physical and emotional conditions of distress to the organized and regimented system of the state. The dispatcher's body serves as a proxy and in service to law enforcement and medical aid. The voice becomes an acousmatic and omniscient audio presence that must sustain control and authority. It is an essential and primary way to receive care because it is enmeshed within what scholar Elizabeth Ellcessor describes as an "infrastructure of feeling."¹ Speech acts are signals of what care is to be performed and given, but the essential work of obtaining data hinges on the communication style of the healthcare worker. Interviewing patients and asking pointed, specific, and direct questions is yet another example of how the voice is vital for care work.

¹ Elizabeth Ellcessor, *In Case of Emergency: How Technologies Mediate Crisis and Normalize Inequality* (New York: New York University Press, 2022), 56.

Yet the unfortunate byproduct of surveillance communication platforms such as Citizen and other crisis management systems become speculative ground for outsourcing and automating care. For example, the voice technology company Sanas AI developed a highly advanced accent elimination software for large corporate clients who seek to provide business solutions that aim to reduce miscommunication and misrecognition. Yet the unintended harm may result in an eradication of non-dominant, non-western accents such as Filipino and Indian and inadvertently prioritizes American and British Englishes. Developing and building systems of care within surveillance communication, this presents a problem to the over 2 billion speakers that have learned English as a compulsory language as opposed to the 450 million that were born into and raised in a household that predominantly speaks either an American or British English.²

The software development of accent elimination tools and the training of neural networks might result in an automation of care that trickles down to an already overburdened population of workers within emergency infrastructures. From outdated manuals, procedures, and protocols to the closing of 911 dispatches in small, rural parts of the US, advanced voice recognition technologies may erode networks of care riddled with existing

² Gregory Warner, interview with Heather Hansen, Rough Translation, podcast audio, April 21, 2021, <https://www.npr.org/transcripts/989477444>.

systemic issues if we are not careful.³ While I write about accent elimination software in the next chapter, I raise the issue here as a precursor of the types of technologies currently being developed for language detection in state and federal agencies that will determine the types of individuals hired to do this work in the future.

In this chapter, I investigate systems and logics of care through an examination of how the voice is mediated through training (discipline) and standard operating procedures and protocols. I produce visual, critical, and auditory analyses of media that depict various states of emergency where the voice is central to the understanding of emotions often associated with distress and crisis. Through what scholar Elizabeth Ellcessor calls “emergency media work,” I include commercial media such as the Apple Watch (Series 7) *911* commercial.⁴ I also include analysis of an online dispatcher community through a digital ethnographic lens that amplifies the complexity and intricacies of this vital labor. Through this research, medical anthropologists, science communications and media scholars will be able to have an expansive

³ “Encore: Rural Washington 911 Center Closes amid Lack of Operators,” *Morning Edition* (NPR, August 25, 2022), <https://www.npr.org/2022/08/25/1119192916/encore-rural-washington-911-center-closes-amid-lack-of-operators>. I feel it is worth noting that the Center for Disease Control and Prevention Crisis’ (CDC) Emergency Risk Communication (CERC) manual is not frequently updated, at the least, the documentation made available to the public.

⁴ Elizabeth Ellcessor. *In Case of Emergency: How Technologies Mediate Crisis and Normalize Inequality*. (New York University Press, 2022), 17.

view of what emergency infrastructures mean to the general public. I ask what it means to listen and respond to a voice in crisis and how might we think of different ways 911 dispatchers are trained in the age of automation?

As a part of my field research, I participated in a January 2023 training provided by Niagara University to 911 dispatchers working in New York state. The 8-hour virtual training focused on disability awareness. Although the protocols are specific to New York state, I participated alongside other dispatchers.⁵ Through the telecommunicators training, I witnessed the intricacies of New York's emergency protocols. Based on my observations, the dispatchers are required to retain an inordinate amount of information across multiple agencies such as law enforcement and fire fighting units. Cultural sensitivity was not a priority or covered despite it being focused on disability awareness. The core information presented was based on instilling the gathering of as much information on a call while taking into consideration any auxiliary devices for those that rely on relay operating or external devices such as a digital speech aid or Electrolarynx.⁶

⁵ To my knowledge, I was the only non-dispatcher participant, everyone had their cameras turned off, there was very little chat happening through the 8 hours, and the materials were extremely dense and packed in with little engagement and interaction between facilitator and participants. I was able to take copious notes and screenshots, but the presentation was not made available and relied heavily on what I gathered as old media (i.e., video examples from the early aughts. etc.).

⁶ "How to Use & Easily Maintain an Electrolarynx • THANC Guide," June 10, 2021, <https://thancguide.org/2021/06/the-journey/how-to-use-easily-maintain-an-electrolarynx/>.

For the purposes of the project, media related to emergency dispatching is the primary focus. Alternatively, I include observations of nurse triaging and contact tracing as voice-oriented telecommunication methods and ecosystems adjacent to emergency infrastructures, usually government funded and utilized during pandemics. This form of telehealth (contact tracing) communications is focused within the frame of epidemiology or the study and origin of a particular disease state. Contact tracing is incorporated into a particular jurisdiction or municipality in order to help stop the spread of the disease. It is a form of biosurveillance that is meant to serve as preventative. While not a perfect system, it is a method used to also assist in capturing data for cities and counties interested in predicative aspects of projecting hospital stays and resources needed. It also factors in age, ethnic background, and other demographic data that might help in understanding what resources are immediately available to the individual that has tested positive for the disease.

In most recent history, COVID-19 warranted a contact tracing system. In Summer 2020, I participated and subsequently received a certificate to perform as a contact tracer otherwise known as a disease investigator. Media scholar Jacqueline Wernimont wrote about how this type of quantum media has, historically, been a woman's job. Despite not being a core focus of this research, within the realm of contact tracing, at the height of the COVID-19

pandemic, through phone and web conferencing platforms, voice and prosody played an integral role in how care and information were conveyed to those stricken with the illness. The training itself teaches how to explain COVID-19 to a community member, but the majority of what is taught is how to speak with a community member. It goes into the soft skills needed to communicate effectively to the extent that the person both trusts the contact tracer and is able to provide information the tracker requires in order to do their job well. The job includes asking one's whereabouts for the past week prior to a positive result amongst other types of information needed. Parts of the training showed the incorrect way to speak and respond to a community member frequently played by a masculine figure as if to further draw upon the associations that feminized labor is what is most effective (even if this is oftentimes true, but it does rely on gender constructs).

Nurse triaging, an alternate form of receiving care for non-emergency healthcare-related situations, functions as a way to help community members that may not need to go to the hospital, but are in need of support and medical assistance. These types of lines are usually third party organizations that work with hospitals and medical centers to provide immediate support for non-life threatening situations. The core issue with this modality of care is that it relies on a nurse being available to provide help and necessitates the

individual be honest and open about the care they need. To my knowledge, nurse triage lines follow specific protocols in order to determine whether a situation requiring a doctor's consultation is often a prognosis as opposed to a diagnosis. The former serves as a bit of a forecast or prediction based on what is told and not on immediate, observable information obtained by a physical examination. The latter is an identification of the disease state and illness in question. Having conducted research on third party organizations, these types of lines are meant to be cost saving and efficient means for patients to get immediate care. The aforementioned forms of telecommunications were included in this research because of the use of the human voice and the creation of a type of media artifact on the interactions. With emergency dispatching, the primary, and seemingly obvious, difference is the state of distress and level of crisis and emergency, despite the fact that non-emergency calls are often made to emergency centers despite a completely separate line.

While a public forum with little verified information or vetted content, one of the most fascinating media platforms that has provided a deep insight on the lives and role of emergency dispatchers in contemporary American life has been the subreddit titled *r/911dispatchers*. The forum itself serves as a digital space filled with artifacts, an archive, and presents the public with rogue media that serves as a vital network for the individuals working as

emergency telecommunicators. This method and approach offers an intriguing and necessary perspective not seen or depicted in popular culture. While it is a rather small community, it is obvious from the posts that these telecommunications healthcare professionals have specific and insular experiences of being front line workers that only other dispatchers are able to understand and empathize with. From the types of training, the capacity and level of experience, I have learned how intertwined emergency infrastructures are with other agencies. These connections might be easily drawn from popular and contemporary media, but through forum postings, it is much easier to see how and why these roles become so tightly woven and enmeshed together. For example, police officers and firefighters oftentimes will transition into the life of a 911 dispatcher as well as those that have previously served in the military. One of the ongoing questions I have is whether or not abolitionist framing and methodologies can be applied to the way individuals are taken care of within a system that already feels skewed and imbalanced. How might we begin to think about how dispatchers are trained (to both listen and speak) when voice analysis courses (based on pseudo science) perpetuate "raciolinguistics ideologies"?⁷

⁷ Rangan et al., *Thinking with an Accent*, 41.

Examining emergency infrastructures and logics of care can tell us how the voice has been used in real life situations as well as depicted in film and television (and training geared towards operators, dispatchers, etc.). By engaging in ethnographic methods or case studies of how voice is used to communicate and mediate both authority and care, it becomes clear that companies, hospitals, and health care facilities are areas of interest when thinking about how the voice is used to communicate, in some cases, life or death information. The purpose of writing a chapter on this issue through the lens of critical medical anthropology has to do with the fact that companies, hospitals, and health care facilities are areas of interest when thinking how the voice is used to communicate, in some cases, life or death information. Many roles are involved in the documentation of medical records and emergency media, which require healthcare workers being trained on specific protocols for both emergency and non-emergency matters. It is vital to understand the relationships between how patients communicate what is happening and whatever is conveyed has ramifications in how it might possibly be mediated. In particular, an emergency infrastructure relies on the emotional complexity and nuance a human operator has in dealing with someone in distress. A future where AI or voice recognition is used may alarm people into thinking that they would not receive adequate care, but if this technology is used, it

would be to aid the dispatcher and not to replace them.⁸ I am also interested in how digital contact tracing technologies have eliminated the voice as a way of receiving information.

Examining the media content, language, speech acts, and creation of training for emergency dispatchers is central to this chapter and essential to the way we understand contemporary healthcare and infrastructures. There are also methods of text communication that are incorporated into the way we communicate distress. From speech-to-text to text-to-speech technologies, the human voice lends itself to a type of care and communication that a machine or device cannot replace. Scholars Elaine Scarry and Giorgio Agamben examine how pain and torture render a human incapable of even uttering their pain and anguish through discernable speech acts. For Scarry, she writes, “Whatever pain achieves, it achieves in part through its unsharability, and it ensures this unshareability through its resistance to language.”⁹ Yet the body in pain is required to engage in speech acts for a 911 dispatcher that is not only discernible and meaningful, but be conscious enough to compartmentalize the shock of whatever state of crisis they are in. Scarry reminds us that physical pain does not only resist language, it simultaneously destroys and diminishes

⁸ Interview with San Francisco Department of Emergency Management, Justin Wong.

⁹ Elaine Scarry, *The Body in Pain: The Making and Unmaking of the World* (New York: Oxford University Press, 1985), 4.

the body's faculties.¹⁰ Along a similar line of thought, for Agamben, his concept of bare life in relation to the polis or political life, language and voice this link to communicate states of either pain or pleasure.¹¹ Although each looks at extreme cases of human suffering, they have been foundational in my own understanding of how language and speech are severely weakened in times of crisis and emergency. Impairment, the multitude of identities, and a wide array of conditions are not something factored into the design of our existing emergency infrastructures. When a 911 dispatcher asks "911, what is your emergency?" or "where is your emergency?," there is a race to understand and little room for error. In reading through Elizabeth Ellcessor's book *In Case of Emergency: How Technologies Mediate Crisis and Normalize Inequality*, I think back on the throughways of studying the aberrations of voice and perceiving the voice as a biomarker (as Beth Semel vehemently argues against). Producers of smart technologies such as Amazon and Spotify have created patents that listen to prosody and tone of voice as well as the detection of environmental sonic data of a user's immediate environment. Text-911 is now a service for consideration due to community members that might be deaf, hard of hearing, or unable to speak at the time of an

¹⁰ Ibid.

¹¹ Giorgio Agamben, *Homo Sacer: Sovereign Power and Bare Life*, trans. Daniel Heller-Roazen, 1 (Stanford, Calif: Stanford University Press, 1998), 12.

emergency.¹² Thinking of the differences between voice (utterances) and speech acts (semantics) with emergency infrastructures that are predicated on the ability to change one's speech under these advancements and considerations.

Manuals and handbooks are another type of analog media used to supplement how healthcare professionals are supposed to respond in cases of emergency. The reading and memorization of information, process, and protocols are extremely different when they must be enacted at any given moment, albeit unpredictable. The voice also has a way of being trained, which is also part of a type of protocol. From voice-activated architecture to the “right to be forgotten,” voice cloning and recognition are also key aspects of technology and design that must be examined closely especially since we live in a western speaking dominated world. The ties between emergency management/infrastructures and law enforcement are further complicated through the lens of abolition in relation to how bodies are cared for and not abject.¹³ How do we go about method-making as a way towards liberation?

¹² “Text-to-911,” *South Sound 911* (blog), accessed November 11, 2022, <https://southsound911.org/911-dispatch/text911/>.

¹³ Katherine McKittrick, *Dear Science and Other Stories*, Errantries (Durham: Duke University Press, 2021).

The Dystopian Future of Crisis and Emergency

The goal for a 911 dispatcher within the San Francisco Department of Emergency Management is to answer a call within 10 seconds.¹⁴ The goal is for all calls, regardless of their classification (non-emergency or emergency), to be supported and addressed in a timely manner and the appropriate resources dispatched to the person in need. There are aspects of a dispatcher's job that are unseen and unheard by the general public. Media, from television shows and films to advertisements, often depict emergency infrastructures by removing minute details such as the line of questioning a dispatcher is mandated to ask in order to be of help to a citizen in times of great need. Unlike the engineers and computer scientists that created the HMIHY technology at the AT&T Labs, their work is part of a lineage of voice recognition software making their way into emergency infrastructures.

When Apple's 911 television commercial was released to the general public at the beginning of 2022, it was a dystopian account of the present and

¹⁴ Mallory Moench, "S.F.'s 911 Dispatch Struggling amid Staff Shortage: 'We Are Bleeding,'" San Francisco Chronicle, December 9, 2022, <https://www.sfchronicle.com/sf/article/911-dispatch-staff-shortage-17641748.php>.

future desired by Apple.¹⁵ Within a matter of 60 seconds, the commercial made the Apple Watch central to the survival of civilians whose emergencies were recorded and used for the advertisement (see fig. 3.1).¹⁶ Three different individuals and their respective dispatchers were included in the succinct yet potent commercial. The first caller, a young woman, has gotten into an accident. She (Amanda) describes that her car has flipped and is unable to get out of the vehicle that is quickly filling up with water (see fig. 3.2). While the viewer listens, verbatim, words appear on the screen against aerial backdrops. The viewer reads the emergency of the woman in distress, which resembles an actual dispatcher's job responsibility of documenting an emergency call. The viewer not only reads and listens, but through a diegetic lens is positioned as both the caller and the dispatcher. There is no separation as both perspectives are encountered by the viewer simultaneously.

At the start of the second half of the commercial, we learn that each of these individual callers have been rescued. We hear Amanda communicating with a responder as she tries desperately to keep her head above water. We

¹⁵ Similar to other big tech companies and their advertising campaigns, even commercials are covered by media outlets. However, the Apple g11 commercial has since been removed from the Apple YouTube channel and I have not been able to locate specific details of who produced and directed the commercial.

¹⁶ Imogen Watson January 04, 2022, "Shock and Applause for Apple Watch's Chilling Real-Life Emergency Call Ad," accessed March 8, 2023, https://www.campaignlive.com/article/shock-applause-apple-watches-chilling-real-life-emergency-call-ad/1736539?utm_source=website&utm_medium=social.

hear Jason stating that he sees the rescue crew from what we might assume is a vast ocean where he has been swept far from shore based on the bird's eye shot (see fig 3.3). The cadence and pace of the words appearing on the screen one word at a time in quick succession mimics speech-to-text detection and output. Two of the emergency dispatchers are women while we hear what sounds like a masculine voice for Jim's emergency. The second caller (Jason), the viewer sees a vast ocean and horizon. Each scene involves aerial, slow moving drone footage where the viewer sees the expanse of the landscape that gives a strong sense of omniscience in listening to the call of the voices along with the emergency dispatcher starting each call with "911, what is your emergency?" The voice sounds as if it belongs to a young man that has paddle boarded out to sea and the winds have somehow pushed him far enough that he is unable to hang on and stressed with fatigue. From the recording, we can hear the gusts of winds contributing to the source of his emergency and distress. The third caller (Jim) has fallen 21 feet and broken his leg. We hear Jim stating in a clear manner that he was unable to reach his phone and therefore relied on his Apple Watch for help (see fig. 3.5-3.7).

From the music to the drone footage to each word of dialogue flashing on the screen in the middle of this landscape in bold white sans-serif font, the viewer is left with Apple's proposition of living our lives without an Apple

Watch. The commercial ends with the text, “With the help of their watch, Jason, Jim, and Amanda were rescued in minutes.”¹⁷ The fact that there is vital information in only sixty seconds of media goes to show the thought and specificity in knowing their target market.¹⁸ The commercial is aptly and simply titled “911”. Yet its key component is the voice of each human that makes it sensational, harrowing, and effective commercial media.

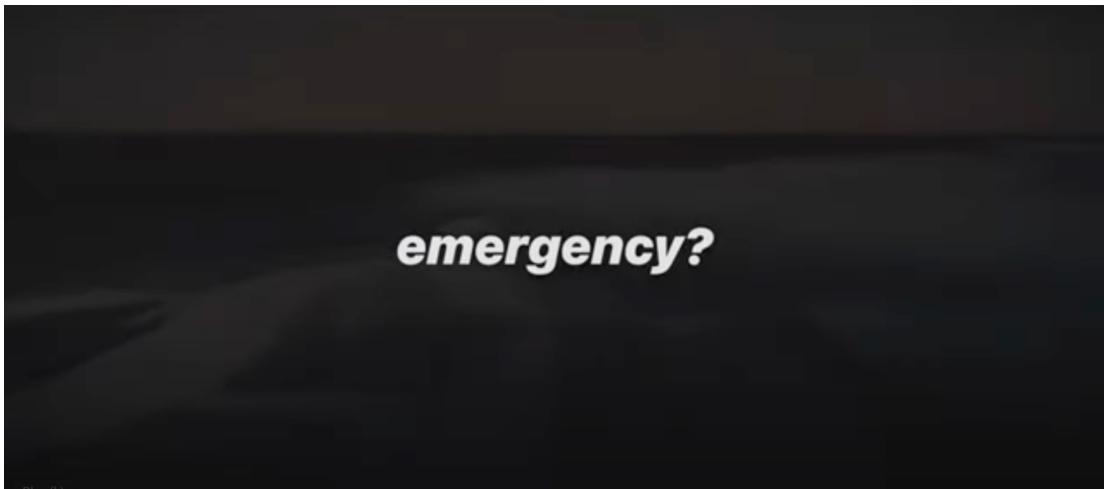


Figure 3.1: Screenshot of Apple 911 commercial (2022) (Advertisement produced by Apple, 2022, <https://www.youtube.com/watch?v=XRwkZWOWI8o>).

¹⁷ *Apple Watch Series 7 / 911 / Apple*, 2022, <https://www.youtube.com/watch?v=QJ2JiwEARFo>.

¹⁸ Patrick Coffee, “Apple’s Ad Agency Names New Creative Leaders on iPhone and Apple Services,” May 30, 2019, <https://www.adweek.com/agencies/apples-ad-agency-names-new-creative-leaders-on-iphone-and-apple-services/>.



Figure 3.2: Screenshot of Apple 911 commercial (2022) (Advertisement produced by Apple, 2022, <https://www.youtube.com/watch?v=XRwkZWowI8o>).



Figure 3.3: Screenshot of Apple 911 commercial (2022) (Advertisement produced by Apple, 2022, <https://www.youtube.com/watch?v=XRwkZWowI8o>).

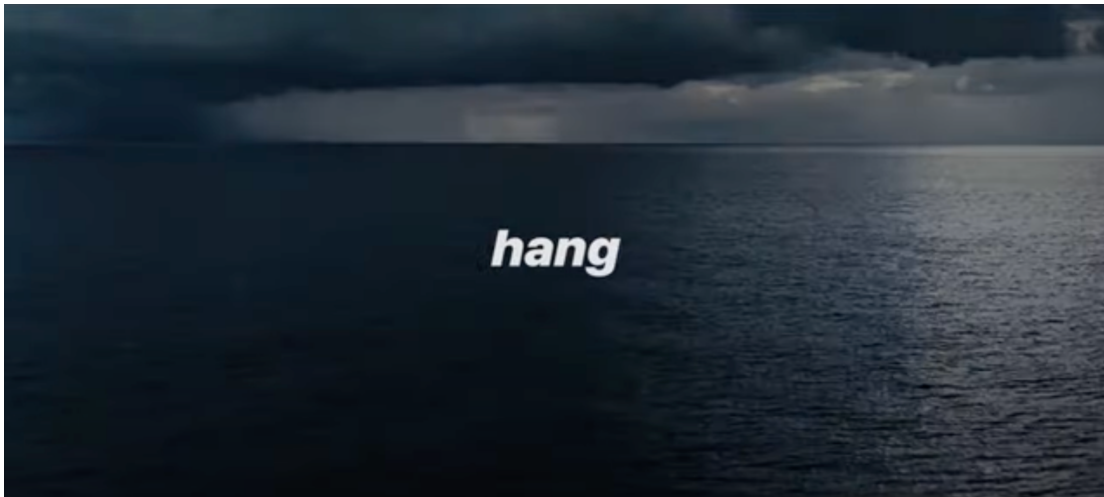


Figure 3.4: Screenshot of Apple 911 commercial (2022) (Advertisement produced by Apple, 2022, <https://www.youtube.com/watch?v=XRwkZWowI8o>).



Figure 3.5: Screenshot of Apple 911 commercial (2022) (Advertisement produced by Apple, 2022, <https://www.youtube.com/watch?v=XRwkZWowI8o>).



Figure 3.6: Screenshot of Apple 911 commercial (2022) (Advertisement produced by Apple, 2022, <https://www.youtube.com/watch?v=XRwkZWowI8o>).



Figure 3.7: Screenshot of Apple 911 commercial (2022) (Advertisement produced by Apple, 2022, <https://www.youtube.com/watch?v=XRwkZWowI8o>).

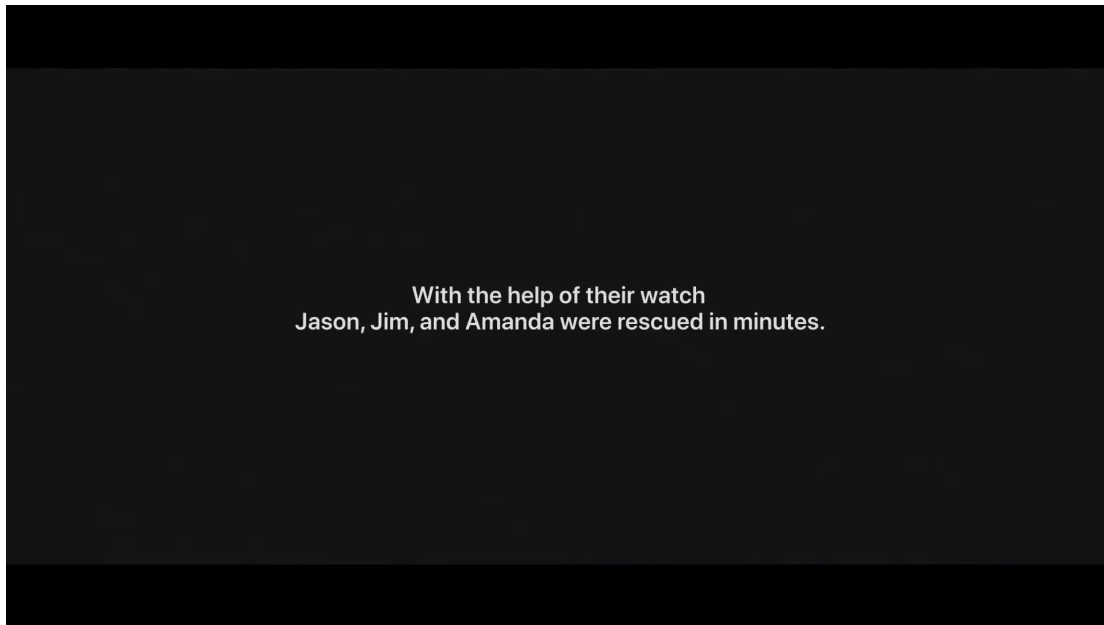


Figure 3.8: Screenshot of Apple 911 commercial (2022) (Advertisement produced by Apple, 2022, <https://www.youtube.com/watch?v=XRwkZWOWI8o>).

In the late 1980s, Nancy Scheper-Hughes and Margaret Lock wrote the foundational text “The Mindful Body” as an attempt to situate human bodies within certain infrastructures of power, knowledge, and care.¹⁹ They wrote to medical anthropologists as well as healthcare practitioners in tandem with the burgeoning digital age and the race amongst big tech companies for home computers to be a staple across American households. Their central focus was on the multifaceted nature of the human body engaging the analytical framework of The Three Bodies: 1) the individual body, 2) social body, and 3)

¹⁹ Nancy Scheper-Hughes and Margaret M. Lock, “The Mindful Body: A Prolegomenon to Future Work in Medical Anthropology,” *Medical Anthropology Quarterly* 1, no. 1 (March 1987): 6–41, <https://doi.org/10.1525/maq.1987.1.1.02a00020>.

the body politic.²⁰ Their conception of the human in these bodies is a way to think about the narratives that are formed through various forms of existing medical practices and healthcare infrastructures. By extending their examination to the roles of voice in care, I ask: How might data tell a healthcare practitioner what is not readily available or perceptible for observation? How might institutions enable sovereign use and access to one's data to do with the information what they will? How does biometric data become yet another format of media someone is able to re-format as they age and evolve?

The individual body is the most overt in its representation of a distinguishable body. From a body's phenotypic expression to the way it showcases signs and symptoms when ill, the human (individual) body is also a product of the Enlightenment insofar as the Cartesian dichotomy of the "palpable body and the intangible mind."²¹ The "intangible mind" is deeply affected by the multitude of sensory information it consumes on a 24-hour news cycle coupled with the propensity towards our electronic devices. It also entails the emotional state that may result in psychosomatic responses. What one is able to observe is made empirical and frequently makes up the individual body that is then readily diagnosed. The social body involves

²⁰ Ibid, 7.

²¹ Ibid, 9.

particular sensibilities such as a desire to belong to a community or being a family member. The body politic relates to structures of power and control.

In Beth Semel's, "The Body Audible: From Vocal Biomarkers to a Phrenology of the Throat", she interrogates the use of computational linguistics in assessing individuals diagnosed with severe mental issues (i.e. bipolar disorder, etc.), there is strong correlation with how future voice recognition systems might be built.²² The hypothesis of discerning and even predicting certain conditions based on the prosody and what might be considered aberrations of voice might add to an individual's confirmation bias of what a healthy voice ought to sound like. Vocal biomarkers also call into question how care and authority might be embedded or absent within and through the lens of emergency management. With 911 dispatching, nurse triaging, and contact tracing, voice analysis research and training has already been incorporated into the training of front line workers.²³ Parallels between the creation of psychiatric vocal biomarkers and pseudo science voice analysis training also add to the existing issues of confirmation bias rampant in the carceral state. Semel's ethnographic research lends itself well to understanding

²² Semel, Beth. "The Body Audible: From Vocal Biomarkers to a Phrenology of the Throat." Somatosphere, 2020. <http://somatosphere.net/2020/the-body-audible.html/>

²³ Brett Murphy, "They Called 911 for Help. Police and Prosecutors Used a New Junk Science to Decide They Were Liars.," ProPublica, accessed December 29, 2022, <https://www.propublica.org/article/911-call-analysis-fbi-police-courts>.

systems and logics of care, but through the lens of those being examined and listened to for the sake of finding auditory and cognitive impairment through speech acts.

Voices expressing pain or being subjected to analysis due to a specific type of mental health condition does not take into account the way this type of technology might be used to incriminate an individual when they believe they are not being listened to. The myriad ways a voice might be deemed inaudible or deficient relates to an individual's capacity to express their need for care or seeking an authoritative figure. How might we assess voice and speech acts conditioned and socialized to be authoritative and how they are perceived by the listener? Signal processing related to information theory and hardware engineering, or telephone engineering is another aspect regarding specific hardware being created for accent reduction. This is noteworthy to the extent that languages and accents outside of standard American English are oftentimes already pathologized and perceived as deficient and othered.

The relationship between signal processing and emergency infrastructures is vital towards the development of platforms and systems to supplement human care. How might a voice be trained to assist, support, and offer care to someone in distress? How does this professional, then, detect things such as fear, anguish, pain in another person's voice? Although these

are more abstract and complex issues to ponder, Semel's work serves as a starting point to think through some of these considerations for the future especially in thinking through a global lens and scale of voice and speech technologies. From a speaker's language to intentionality and semantics, what are key linguistic cues are we drawn to when we listen versus when we speak? When technologies are developed to detect vocal abnormalities or aberrations in speech, what is the intended use as well as the unforeseen consequence of this profiling?

Similar to the polygraph, computational linguistics is primed and classified by a human labeler and not necessarily running on calculations based on temperature or heart rate, which already is not a reliable marker of an individual lying or telling the truth. As a precursor to the following chapter, we see the foundation of a programmable body through laborious training, memorization of codes, and medical procedures meant to be told under duress (i.e., CPR, etc.). From the interplay of mind, body language, and machine, I consider the factors of how a voice analysis system would be programmed and what determinations are made to deem an individual capable of then being responsible to label a corpus of speech data. It is imperative that we understand the risks because these types of advancements

in linguistics will ripple into other areas and ecosystems that could have truly harmful effects.

Yet these intricacies are not taken into consideration when training emergency responders. There is a clear distinction and separation that this thought, more abstract in nature especially of a social and political body becomes an afterthought or not pragmatic in a crisis. While an individual working within an emergency infrastructure ought to know how to take care of the human body in all of its multiplicity, speech acts are reduced to conveying necessary information especially for a body that is at high risk of violence and police brutality. The information and media consumed by dispatchers is astounding and illuminates this idea that a human body can be programmed to respond and care in the ways we need. But my research has uncovered that the role of a dispatcher is not only difficult to fill, it takes a specific type of personality willing to be equally authoritative and detached while being compassionate and intuitive. Listening in a way that is not common becomes a matter of life and death in many cases for the dispatcher.

On Emergency Media: Reports and Responses

Similar to the way photographic surveillance functions, biosurveillance and emergency infrastructures involve extensive, detailed documentation and

records of physical activity that is classified or deemed as aberration to the normal. We see with our eyes what is wrong, out of the ordinary, or abnormal. Human voice is something that reaches deeper into our internal states especially when someone is in distress. Ellecessor describes seemingly blurry indistinguishable contours of emergency media and work through a system or "infrastructure of feeling" inspired by culture critic Raymond Williams' concept of "structure of feeling."²⁴ Although Ellecessor's *In Case of Emergency* does not overtly focus on the human voice, she shows how deep seated the codifying of technologies plays an integral role in the development and programming of the human voice as a primary element of emergency infrastructures. The reporting of crime, accidents, and crises become media in an instance. These source materials, arguably, belong to a type of unintended public and permissible and usable by law enforcement, the courts, and medical and healthcare professionals.

As I sit here writing this, I hear a police siren trailing off in the distance and imagine various transmissions, connections, and documentation simultaneously working in concert to create a form of emergency media. It has also become commonplace on mobile applications and platforms such as Citizen (originally called Vigilante) to add still and moving images for a wide

²⁴ Ellecessor, *In Case of Emergency*, 56.

array of crises or emergencies made available, in real time, to the general public. An emergency becomes public when it starts to disrupt or present a nuisance. But how might the increase of mobile and digital technologies further add to increasing use of algorithms and predictive policing that directly may hinder someone from receiving the help and support they need in an emergency? How do the logics of care change when emergency reporting and responses become automated? Palantir and Motorola are two companies that have developed softwares and mobile platforms of "procedural listening" and real-time notifications that enable law enforcement direct contact with the citizen in distress.²⁵ Such software also consolidates data to further add to the algorithmic make-up that turns emergency media into real-time (big) data collection. This media is also unalterable since it becomes a document that catalogs and references a specific time, place, incident, and as many concrete facts that renders the 911 dispatcher as a neutral (as objective as humanly possible) third party or witness in addition to someone mediating the experience through the protocol of emergency infrastructure. In the US, 911 is a three-digit number that has become its own entity and coincides and intersects with a multitude of agencies far too complex for people to understand each of its constitutive parts. With the popular media creating

²⁵ Ibid, 109-111.

fictionalized accounts to true crime podcasts and shows, emergency infrastructures are made into spectacles that don't reflect the real-life nature and business of emergency infrastructures.

Circling back to Ellecessor's infrastructure of feeling, to best understand, entails reading actual 911 dispatcher accounts. From suicide to domestic violence to non-emergency calls, the emotional and intellectual labor and stressors of such work take their toll on the individuals tasked to do this work. In this particular line of work, the acousmatic, or disembodied voice, becomes a matter of life and death based on the way in which the dispatcher is able to surmise within seconds the situation and environmental cues they have been trained to actively listen to and for.

Sonic Mediations

We live in a constant state of emergency against the backdrop of a global pandemic. Between COVID-19 and Monkeypox, digital and social media are integral in people's understanding of public health. Yet mainstream media and social media are also the source of misinformation and misinterpretation. Media plays a significant role in the ways we perceive, understand, and process how emergency, distress, and crisis are mediated through screens and especially through voice and telecommunications.

According to the Pew Research Center, approximately 60% of Americans consume their news through mobile and digital devices such as smartphones or tablets.²⁶ The 24-hour news cycle and mobile applications such as Citizen turn emergencies and crises into a spectacle. Yet the marketing and branding of such applications claim that it supports the flow of cases that end up ERs or supports a type of participatory care or sense of justice. Ellcessor provides a multi-faceted definition of emergency that captures something so simple yet indicative of how it then is mediated in films and media. She writes, “Emergency is located in the present tense.”²⁷ One of the reasons the wording is potent in this statement is that emergency is always in a place. Language and the communication of facts are meant to convey various types of locations for the emergency that relate to safety, mobility, and receiving care and immediate aid. Advertisements for the latest features of the Apple iWatch, for example, speculate that a user ought not be without their device because it is the wearable tech that captures all the information to situate its user to a specific location should an emergency arise.

²⁶ Elisa Shearer, “More than Eight-in-Ten Americans Get News from Digital Devices,” *Pew Research Center* (blog), accessed November 11, 2022, <https://www.pewresearch.org/fact-tank/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/>.

²⁷ Elizabeth Ellcessor, *In Case of Emergency: How Technologies Mediate Crisis and Normalize Inequality*, 1st ed. (NYU Press, 2022), 12.

Accessibility becomes an integral part of emergency infrastructure as well. In recent years, community members have the ability to text an emergency to dispatchers. This development was to ensure that deaf or hard of hearing citizens had access to emergency and crisis care. The voice becomes a nexus for people in states of distress because it requires a direct communication of how location and severity are discerned and understood by a dispatcher. Technological capabilities are another important factor in how infrastructures are set up, maintained, and inevitably used by the general public.²⁸ For the purposes of this chapter, I use Ellecessor's research on reporting as a point of departure for how emergency is depicted in several types of media such as film, advertisement, and training media. Reports are a type of media artifact, predominantly, for healthcare/medical and judicial purposes. Emergency not only points to a binary of ability and disability, it also suggests that there is something or someone at fault or serves as the cause for a crisis.

²⁸ Ibid, 95.

A City in Crisis: The San Francisco Department of Emergency Management

Answering calls within a matter of ten seconds presents a challenge for the one-hundred and twenty three operators at the San Francisco Department of Emergency Management, which has been extremely understaffed over the course of the pandemic. In order to have a fully functional and operational team, one hundred and sixty dispatchers are needed, at minimum, in order to serve the San Francisco population of 842,754.²⁹ The census data captures all individuals living in one home and this number is contingent upon the person deemed the head of household completing census data that then dictates how cities and states are funded and the types of resources they receive based on this information. Yet at face value, the shortage of operators coupled with the number that call out sick or quit due to burnout puts the City into a state of urgency to recruit and retain this subset of first responders. But the job is not appealing to many considering a shift might be up to 16-hours as a result of the shortage. In addition, dispatcher training takes months and hundreds of hours of desk training with a veteran dispatcher. In addition, there are well

²⁹ Moench, “S.F.’s 911 Dispatch Struggling amid Staff Shortage.”

over one hundred codes to memorize as if the human body were meant to function like a machine.

The global pandemic that resulted in shelter-in-place orders affected employment rates within the department. Across the agency, vacancy rates for temporary and permanent positions went from 2% in June 2019 to 9% in fall 2022.³⁰ These shocking statistics reflect, not only the challenges and high expectations of the job, but the lack of interest by the general public to do this line of work. While the starting salary of a dispatcher in the City and County of San Francisco is \$100,000, it is not nearly enough when one takes into consideration the demands of the job.³¹ This increase in vacancies is also a direct correlation with the shortage of police officers considering how their roles are tethered within the emergency infrastructure. Increased wait times have been prevalent based on accounts of emergency centers in Oakland, California.³² Although a rare occasion, some Oakland residents have had to wait up to four minutes to speak with a dispatcher, which also speaks to Bay Area residents not being aware of non-emergency versus emergency line numbers.

³⁰ Ibid.

³¹ Interview with San Francisco Department of Emergency Management, Justin Wong.

³² Rachel Swan, “‘I Hope Nobody Is Bleeding out’: They Called 911 in Oakland. They Were Told They’d Have to Wait,” San Francisco Chronicle, May 11, 2022, <https://www.sfchronicle.com/crime/article/I-hope-nobody-is-bleeding-out-They-called-17166166.php>.

In an interview with San Francisco Department of Emergency Management dispatch supervisor, Justin Wong, he covered the standard protocol for dispatchers receiving both non-emergency and emergency calls.³³ He noted that dispatchers are required to type the correspondence, word for word, in addition to the call being recorded for legal purposes. This repetitive action results in many dispatchers suffering from carpal tunnel syndrome (CTS) or repetitive strain injury (RSI). Based on our conversation, dispatch centers in Southern California have already implemented voice recognition technology (speech-to-text) to alleviate these physical conditions.³⁴ According to Wong, artificial intelligence and voice recognition is being explored in some centers in California as a way to accurately document calls. While there are no plans for voice recognition softwares to replace the human voice, it becomes an area of interest for centers who anticipate resources that can help adequately support citizens especially in rural or remote areas where calls are already being rerouted to major cities. One of the primary reasons AT&T Labs created the HMIHY tech was to ensure a level of efficiency for the consumer, in this case, how we might safeguard the future of emergency response systems? Considering that the agency answers approximately over 100,000 calls

³³ Interview with San Francisco Department of Emergency Management, Justin Wong, interview by Dorothy Santos, zoom, otter.ai transcription, December 30, 2022.

³⁴ Ibid.

per month with a little more than half being emergency and the rest being non-emergency, what types of technological advancements might be designed and developed if we do not quickly educate the public on how current systems and protocols work and what might be lost through automation? Ultimately, there is no perfect or ideal technological solution for these issues because a human will always want another human on the other side of the line.

A Network of Care for Dispatchers by Dispatchers

Returning to Ellecesor's idea that emergency resides in the present tense, I conclude with my observations of two postings and noteworthy comments left to the original poster (OP). Per the user agreement, moderators take on the role on a voluntary basis and content to the site is meant for exchange and not guaranteed to be fact checked or edited by community members or employees of reddit.³⁵ With well over 100,000 communities or sub-reddits and over 57 million unique visits every day, it has become a long-term forum for many communities around the world with almost half of its registered users being based in the US.³⁶ The platform allows for insular and specific groups to form and one of them being /gudispatchers (see fig. 3.9).

³⁵ "User Agreement - April 18, 2023 - Reddit," accessed April 30, 2023, <https://www.redditinc.com/policies/user-agreement>.

³⁶ "Homepage - Reddit," accessed April 30, 2023, <https://www.redditinc.com/>.

Over 19,000 users have joined this group since its inception on December 26, 2011.³⁷ I have been reading posts and observing commenters and OPs for the past nine months (August 2022 through April 2023) and the posts are typically regarding three primary inquiries to the immediate community including: 1) interviewing and hiring processes, 2) administrative questions related to on-the-job responsibilities (i.e., questions regarding rank and file protocols related to anything from time off to certifications necessary to advance in the job), and 3) mental health issues. From the hundreds of posts, I decided to select one post regarding mental health (see fig. 3.10) and another post related to dispatcher training, specifically spoken communication style (see fig. 3.11).

From my conversation with Justin Wong and reading through the various forum posts related to mental health, dispatchers are activated and adrenaline coursing through their bodies constantly throughout their shift.³⁸ Every call is treated like an emergency despite it potentially being a non-life threatening incident. Yet the subreddit seems to offer solace to dispatchers who feel that the only other individuals that might understand the stressors of their job are their peers. Considering that reddit users are given usernames if they do not select their own, it is a relatively anonymous site to an extent. The

³⁷ “r/guidispatchers,” reddit, accessed April 30, 2023, <https://www.reddit.com/r/guidispatchers/>.

³⁸ Interview with San Francisco Department of Emergency Management, Justin Wong, interview by Dorothy Santos, zoom, otter.ai transcription, December 30, 2022.

members of /g11dispatchers feel a type of freedom in posting their challenges and difficult scenarios due to the unique situations dispatchers face on a daily basis. The poster expresses concern for not being affected by the job and their post clearly and concisely states the active listening involved in the work and that the labor has unintended consequences and effects. According to a systematic review and narrative research within the behavioral sciences related specifically to emergency dispatch work, 911 telecommunicators experience what scientists have called “vicarious trauma” alongside feelings of helplessness and being unable to fully help because they are left with only the voice(s) on the other end of the line and are considered the *first* first responder.³⁹ One of the responders to the OP stated having a “calm detachment” as a necessary aspect of the job.⁴⁰ However, the comment delves into the unknown aftermath if the OP does not confront and allow themselves to process feelings when they arise. Many of the oral history narratives I read

³⁹ Sarah E. Golding et al., “Exploring the Psychological Health of Emergency Dispatch Centre Operatives: A Systematic Review and Narrative Synthesis,” *PeerJ* 5 (October 17, 2017): e3735, <https://doi.org/10.7717/peerj.3735>.

⁴⁰ throwaway1976_, “I’ve Been at This Job for 3ish Months. Tragic Things I Hear Don’t Affect Me.,” Reddit Post, *R/911dispatchers*, March 16, 2023, www.reddit.com/r/g11dispatchers/comments/11so0zb/ive_been_at_this_job_for_3ish_months_tragic/.

through mention allowing for emotions to be addressed and seeking support if the job presents challenges.⁴¹

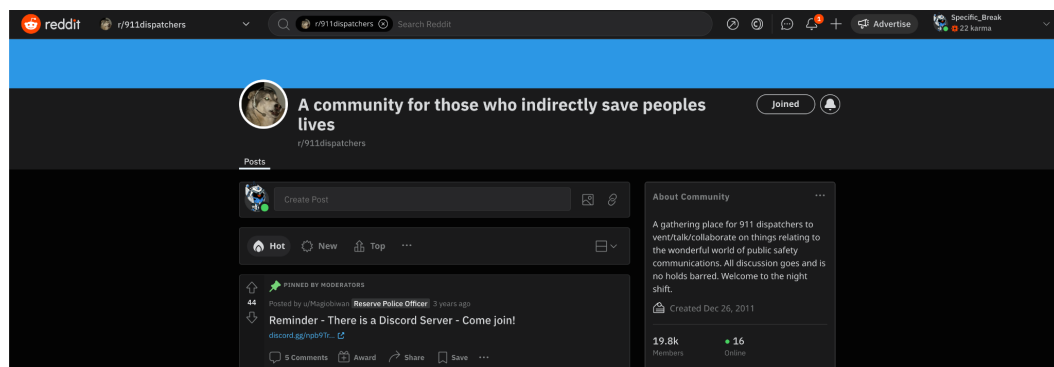


Figure 3.9: Screenshot of /guidispatchers sub-reddit forum (<https://www.reddit.com/r/guidispatchers/>).



Figure 3.10: Screenshot of a post titled “I’ve been at this job for 3ish months. Tragic things I hear don’t affect me” published to the /guidispatchers sub-reddit forum (<https://www.reddit.com/r/guidispatchers/>).

⁴¹ Jim Marshall and Laorenza, Tracey, eds., *The Resilient 911 Professional: A Comprehensive Guide to Surviving & Thriving Together in the 911 Center*, 1st ed. (Petoskey: South of Haven Press, 2018).

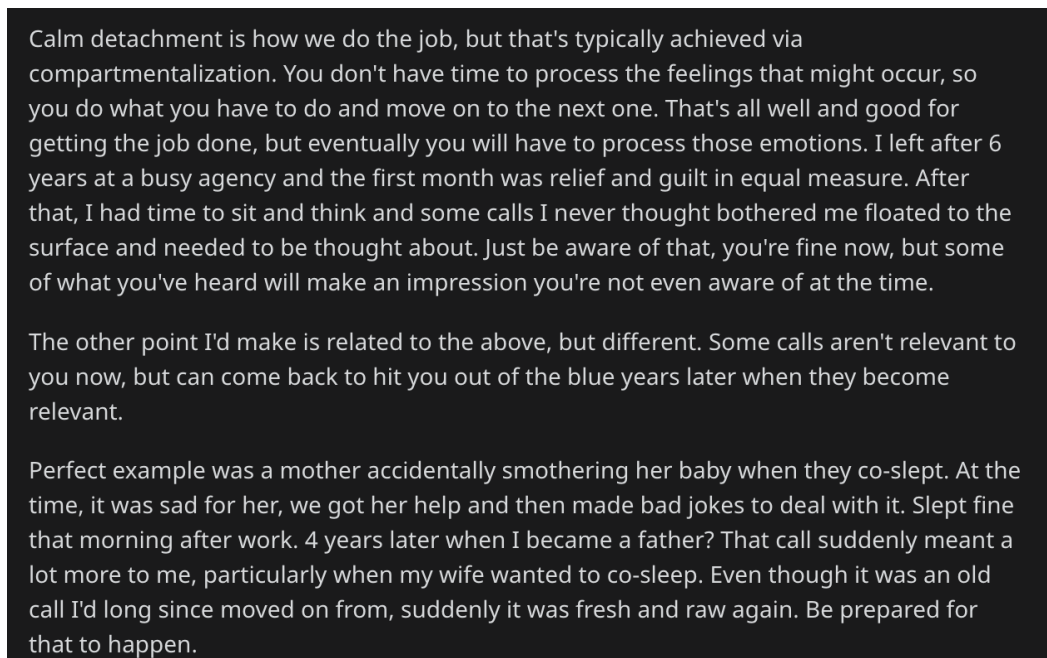


Figure 3.11: Screenshot of a comment in response to the post titled “I’ve been at this job for 3ish months. Tragic things I hear don’t affect me” published to the /gudispatchers subreddit forum (<https://www.reddit.com/r/gudispatchers/>).

For another community member of the subreddit, the OP was most interested in how other dispatchers encounter the use of “I” statements within their everyday speech on the job. The OP shared what looks like an image taken from a mobile phone that already provides a glimpse into what the job entails: confrontation (see fig 3.12). Dispatchers are expected to deal as quickly and as effectively as possible, unknowing of every instance that they pick up a call. While training may offer guidance, it is far different applying what has been discussed in a theoretical or abstract way. The responders to the OP provide much more nuanced approaches and guidance based on their

experiences. For instance, one commenter noted, “Your job is to get info, not confront them.”⁴² The OP confirmed that the training was taught by a hostage negotiator, which makes me wonder about the immediate environment (i.e. state and city) or location of the dispatcher.

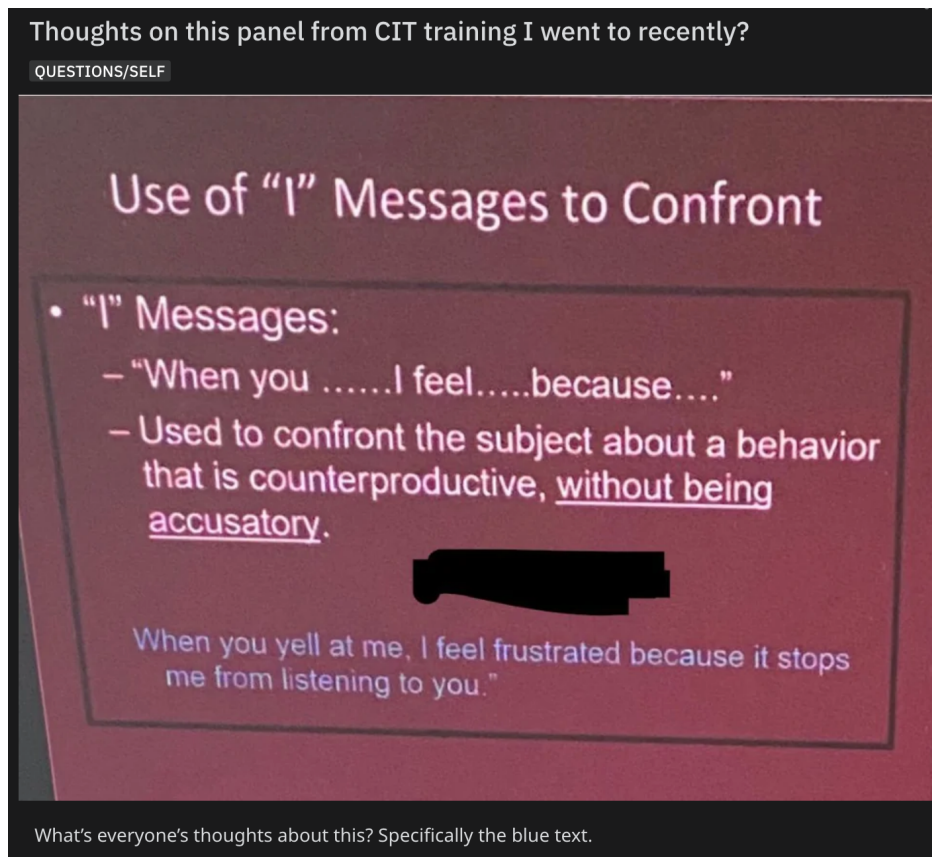


Figure 3.12: Screenshot of a post titled “Thoughts on this panel from CIT training I went to recently?” published to the /gundispatchers sub-reddit forum (<https://www.reddit.com/r/gundispatchers/>).

⁴² KrAff2010, “Thoughts on This Panel from CIT Training I Went to Recently?,” Reddit Post, *R/gundispatchers*, February 11, 2023, www.reddit.com/r/gundispatchers/comments/10zcgnd/thoughts_on_this_panel_from_cit_training_i_went/.

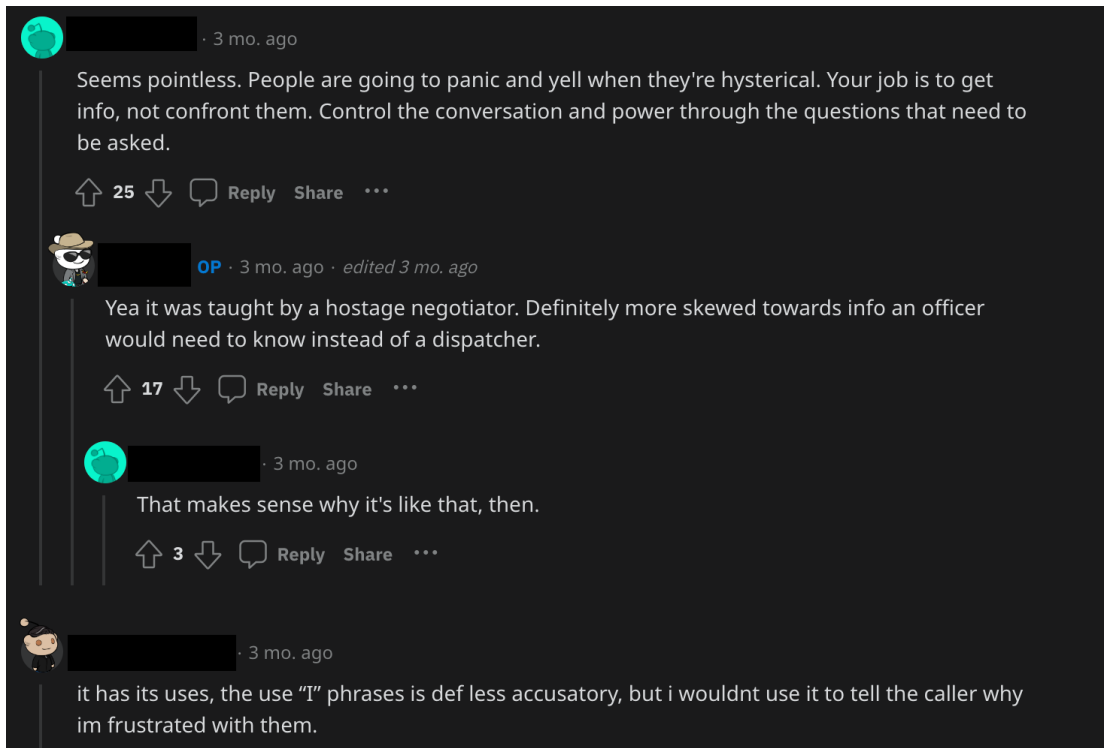


Figure 3.13: Screenshot of a comment in response to the post titled “I’ve been at this job for 3ish months. Tragic things I hear don’t affect me” published to the /r/911dispatchers sub-reddit forum (<https://www.reddit.com/r/911dispatchers/>).

Various types of approaches within a specific jurisdiction might offer insight into the types of instructors and trainers are tasked with providing educational and occupational resources. I write about this in the next chapter based on my research of call center agents in the Philippines and how Americans are often paid to train Filipinos on cultural competency. Although it seems unlikely that agencies would outsource emergency dispatchers abroad, the possibility remains that the labor is a necessity and would require such measures to be taken. With a major local agency such as the San

Francisco Department of Emergency Management being stretched extremely thin on resources despite an income that is statistically well over a living wage is offered, how might thinking about how we understand the relationships between advanced voice technologies, emotional labor, and affect when we are unable to fully understand the systemic issues for the roles deemed necessary to our survival?

When the Human Body is Rendered Speechless

Driving home from Santa Cruz one evening, off to the side of the road on Highway 85, I saw the interior of a sports car engulfed in flames. I panicked and called 911. Words escaped me as I tried to answer the dispatcher's questions. She asked the direction I was driving towards, if there was anyone near the vehicle, and if I was in a safe location. Like many other drivers, it was far too dangerous to stop. The unforgettable scene of orange flames and black smoke rendered me speechless. As Ellcessor reminds us, emergency is always in the present tense.⁴³ Even as a bystander, I co-created emergency media alongside the dispatcher that set into motion a chain of events in accordance

⁴³ Ellcessor, *In Case of Emergency*, 12.

to a protocol. Outside of any stress or life threatening situation, this media form(at) becomes didactic and instructive. The documentation of countless emergencies and crises told to dispatchers not only becomes emergency media, these events become case studies and contribute to how the development of advanced technologies codify emergency. Smartphones and devices change our relationship to emergency management, media, and infrastructures of care, however, how might we reconsider media that allows the general public to learn what is often an unrealistic portrayal of this vital and necessary labor?

Elcessor's concept of an infrastructure of feeling in relation to emergency management is potent in watching and listening to mainstream media such as the Apple watch commercial. Popular depictions often counter the material realities of 911 telecommunicators and negate the human body as a tripartite of the physical, the biopolitical, and the social as theorized by medical anthropologists Nancy Scheper-Hughes and Margaret Lock. Leaning on my fieldwork and research of dispatcher oral histories and what might be considered rogue media (/911dispatchers), it has been a personal imperative to compare and contrast the spectrum of media that informs versus sensationalizes 911 dispatcher work. When the human body is in pain and it is rendered speechless, how are dispatchers trained to fight against Flores and

Rosa's idea of raciolinguistic ideologies that might further amplify a prejudice a dispatcher might possess for a caller? Some telecommunicators have noted the nature of the job is first and foremost about data collection and possessing a "calm detachment" when dealing with an individual under a tremendous amount of stress and pain. Care work becomes a form of authority and directives since dispatchers are trained to be tempered.

Language detection and speech-to-text are advancements currently being explored within the emergency management industry as a way to alleviate the stress non-English speakers already feel in a crisis as well as the documentation produced by dispatchers. Although these are vital enhancements to this necessary labor, there is no displacing or mitigating the fact that dispatchers are required to strike a balance that has only been exacerbated by the global pandemic that revealed the weaknesses of these infrastructures. A constant concern that has arisen throughout this research has been the countless stories of telecommunicators that are trained to actively listen, not just to a voice, but to the environment of the caller and to understand this line of work requires an acute awareness of paralinguistic cues.

Chapter 4: Speculating a Socially Just Sonic Future

Artist Rachele Cruz’s poetry from her series “Accent Reduction School Graduate” was written as a reflection and response to her father’s desire to anglicize his name and take classes that would reduce, if not eliminate, his Filipino accent. At the height of the COVID-19 pandemic, I serendipitously discovered a recitation Cruz had done for a poetry reading in 2018 on YouTube, which was a pivotal moment in my research on accent bias.¹ Her writing, affect, and reading encapsulated criss-crossed lines bringing to light a different type of weaver of speech than the women who connected long distance calls between businessmen or community members at switching stations many decades earlier in the United States. Those buildings and equipment are long gone and have been replaced with overseas workers in the Philippines, India, and South American countries answering and responding to calls at all hours of the day and night serving the western world.

Cruz was also fascinated by the time difference and, perhaps, the cognitive and emotional dissonance that occurs when voices connect over long distances. For an American making a morning call to a help desk, the

¹ *Rachele Cruz - Sunday Jump 10/07/2018*, 2018, <https://www.youtube.com/watch?v=-5eg4WwgRMw>.

time might be early the following morning for an overseas Filipino worker (e.g., 10:00 am Pacific Standard Time on May 1, 2023 is 1:00 am PhST, Philippines Standard Time on May 2, 2023).² She sought to understand through poetry what words, phrases, and tonality are made commonplace to an overseas Filipino worker. Through her craft, she enunciates, softens, and punches words one might have to say when speaking for and to an American consumer. Arguably, this particular series, quite frankly, is best heard. Listening to her recitation bears a profound potency that cannot be easily understood if read alone. Much like the text on these pages, to feel and imagine working in complete accordance to an entirely different time zone than your own (and in a language that has forced you to speak differently) is what the Filipino body is subjected to. For Cruz, it went beyond her observing her father change his name from Romeo Cruz to Roy Cruise to be made legible to Americans. It was her bearing witness to her father's desire to be audible that piqued her interest in exploring the complexities of her father wanting to be heard, not for who he was, but for what he could become.

Like Cruz's father's desire to sound like an American, the post-colonial subjecthood of Filipino citizens is at the core of Jan M. Padios's scholarship, which involves an examination of how 21st century telecommunications and

² "PhST Time Zone – Philippine Standard Time," accessed May 1, 2023, <https://24timezones.com/time-zone/pst-philippine>.

technologies are further codifying specific types of speech acts.³ While she writes through an immersive, ethnographic lens of call center workers in the Philippines, her framing serves to support my claim that the Filipino body has been perceived by the West as a malleable (and replaceable) body. Through the compulsory learning of English in most, if not all, educational institutions in the Philippines, the nation itself has become an economic resource to the western (American) world. Although a nation in the Global South, it remains a nation beholden to the US for its economic viability. What was once a lucrative and sought after job in the US-the telephone operator-has now become lucrative for Filipinos. As mentioned in chapter one, telephone operating was seen as an integral part of the wartime effort in keeping with the idea that telecommunications was a vital part of the US defense industry. But over time, the work was perceived as clerical, administrative, and resulted in deskilling of employees due to automation and advancements in technology. Operating became something supplemental to the technology that was being developed in the late 1960s and beyond. Another aspect of why the job is made enticing through business process outsourcing (BPO) companies in the Philippines is the draw to cosmopolitan areas in Manila's city center that are culturally equivalent to bustling American cities such as Manhattan, New

³ Jan M. Padios, *A Nation on the Line: Call Centers as Postcolonial Predicaments in the Philippines*, 1st ed. (Durham: Duke University Press, 2018).

York. Call centers are placed in specific locations to draw young, energetic, and ambitious Filipinos into, not just a workplace, but a lifestyle that this work promises to provide.⁴ I delve further into why the work is now popular and sought after by individuals in parts of Asia and South America throughout this chapter. Historically, known as the “sick man of Asia,” the Philippines has had to constantly prove itself as a global economic power in the contemporary world.⁵

In this final chapter, I return to the evolution, racialization, and gendering of voice through an examination of the colonization of it and the westernization and imposition of American English as the lingua franca and as a sonic, vocal trope. In Padios’s *A Nation on the Line*, her focus on vocal training and the post colonial subjecthood of Filipino service workers plays an integral role in understanding the relationship between US imperialism on racial and linguistic capitalism.⁶ According to Magellan Solutions (one of the most unfortunate names for a company especially in the Philippines), as of 2018 more than half of the BPOs (68%) serve US companies.⁷ They also report

⁴ Padios, 132.

⁵ Ibid, 3.

⁶ Jan M. Padios, *A Nation on the Line: Call Centers as Postcolonial Predicaments in the Philippines*. (Durham: Duke University Press, 2018). I am also referencing scholar Pip Thornton’s term of linguistic capitalism.

⁷ Magellan Solutions, “The Business Process Outsourcing Industry (BPO) In The Philippines,” Business, Magellan Solutions, May 21, 2019, <https://www.magellan-solutions.com/blog/status-bpo-in-the-philippines/>.

that “more than eight hundred fifty (850) business process outsourcing companies that are registered in the Philippines...about 46.2% (approximately 400) of these firms offer services that are related to information technology.”⁸ The rest of the companies cater to other adjacent industries such as animation and film production.⁹ Techno-orientalism is enacted through the Filipino as a type of programmable body that possesses the ability to mimic and normalize western (American) speech and center whiteness. The design of IVRS and digital technologies continues to prioritize accent bias towards the lingua franca of specific types of American and British Englishes. The existing codebases do not factor in prosodic elements of speech such as rhythm, intonation, and stress, which continues to remain a problem for billions of non-English speakers around the world. Padios is calling attention to the predicament Filipino call center workers find themselves in performing what she calls a "proximity to America while simultaneously disavowing the US" by actively participating in crafting a counterimage of postcolonial subjectivity to one of agency, independence, and decoloniality. With the human voice as central to this work, I also contribute methodologies for prototyping, designing, and developing systems of care within medical and healthcare

⁸ Ibid.

⁹ Ibid.

technologies that must factor the multiplicity of race, gender, and identities through our individual and shared sonic experiences.

The invisible and, often, obscured labor of call center workers aims to keep a primary, western world economic power visible and formidable. Yet this dominance relies on intellectual, emotional, and physical labor of a body perceived and maintained as a specific (postcolonial) subjectivity.

Techno-orientalism is defined as “the phenomenon of imagining Asia and Asians in hypo- or hyper technological terms in cultural productions and political discourse.”¹⁰ Specifically, I am examining the perception of the Filipino body as a site of economic output and production that the west has deemed either monotonous, craft-like, machinic, and banal, but necessary. Scholars Alden Sajor Marte-Wood and Stephanie Dimatula Santos articulate “Filipino care as executable code” in their scholarship on content moderation in the Philippines.¹¹ While the work of content moderation is not central to this research, it is worth noting its relationship to the ways we understand call center, virtual care work, and telecommunications in our contemporary global culture and economy. While US and Spanish rule seem like vestiges of a

¹⁰ David S. Roh, Betsy Huang, and Greta A. Niu, eds., *Techno Orientalism: Imagining Asia in Speculative Fiction, History, and Media*, 1st ed. (New Brunswick: Rutgers University Press, 2015), 2.

¹¹ Alden Sajor Marte-Wood and Stephanie Dimatula Santos, “Circuits of Care: Filipino Content Moderation and American Infostructures of Feeling,” *Verge: Studies in Global Asias* 7, no. 2 (2021): 106, <https://doi.org/10.1353/vrg.2021.0007>.

colonial past, it becomes increasingly apparent and tantamount through the increase in BPO companies in the Philippines.¹²

The U.S. Cultural Competency (CC) course is a part of a multi-faceted training that call center agents are required to participate in for the sake of managing customer service calls. Filipino call center agents at Vox Elite, a BPO company based in the Philippines, are required to take this course as a part of their onboarding experience.¹³ The agents learn aspects of the job such as American vernacular and conflict resolution that help prepare them for the “physical rigors and psychological challenges of working in western time zones.”¹⁴ Yet the course confounds the creation of accent elimination software such as Sanas AI, which claims to offer its clients, in this case, a large conglomerate company whose bottom line necessitates outsourcing a more lean approach. The software would devalue if not wholly eradicate the need for these types of training. The contradictory nature of what it means to learn about a culture outside of one’s own for the purpose of commodity and service does something damaging and irreparable to the individual, collective consciousness and imagination of the Filipino worker. They become an automaton within the larger rubric of capitalism. It is a vicious cycle that

¹² Ibid, 103.

¹³ Padios, *A Nation on the Line: Call Centers as Postcolonial Predicaments in the Philippines*, 95.

¹⁴ Ibid, 104-105.

connects directly to how speech is codified in the development of voice recognition softwares. The filtering of voice and speech acts has become highly contentious.

According to David Crystal, writer of the *The Cambridge Encyclopedia of the English Language*, approximately 400 million people around the world are born into English speaking households compared to the estimated 2 billion people that learn in a formal educational setting.¹⁵ With such a staggering statistic, the perpetuation of a lingua franca in an ever increasingly connected world, machine listening and human computer interaction may evolve without consideration of both non-dominant languages and atypical speech. Even the design of hardware technology such as an “electronic ear” used to help modify accents becomes another form of disciplining the voice.¹⁶

Capitalism and colonial roots are forces that continue the disciplining of the (Filipino) body and voice in such a way that accent elimination software becomes a part of the efficiency of a lean business model. Yet it is yet another techno-solutionist method that creates a whole host of problems, issues, and further augments the roots of racism and xenophobia in countries such as the

¹⁵ Gregory Warner, interview with Heather Hansen, Rough Translation, podcast audio, April 21, 2021, <https://www.npr.org/transcripts/989477444>.

¹⁶ “Can technology help you lose an accent?,” Farah Khalique, accessed June 14, 2022, <https://www.bbc.com/worklife/article/20190506-this-technology-could-help-you-lose-your-accent>.

US and England respectively. Further development of such softwares will only flatten non-dominant cultures into a new wave of service that erases the individual. Also, companies such as Sanas AI will not cease to exist. Rather, other companies and softwares will be developed in its place in that there is always a competitor within the global, racialized ecosystem of the free markets. Machine learning also presents itself as a technology that aids in the idea of personalized experiences.

Returning to the AT&T concept film of the early 1990s, the prescient view of a real-time voice and language analyzer that could execute real-time translation is not some new concept. Rather, it is a reification of an idea that allows for communication no matter the language spoken. A fantasy for many individuals that don't need to learn a language or care about the language to be spoken and understood. How might we think of the ways the U.S. learns of other countries? As Padios documented, the CC course gives primacy to an American and Eurocentric view that illuminates African Americans, Indigenous, and Asian people as foundational to the existence and economic world power that the U.S. has come to be known for. The astounding nature of what is taught to Filipino (overseas) workers about American history recapitulates a colonial past too far gone to correct and this becomes ingrained in the work-life and culture of call center agents. The irony that the Filipino

worker is required to perform language through speech acts that coddle and appease the American consumer. Returning to western vocal drag, the more performative the speech act doubled with learned cultural competency and a hyper specific linguistic consciousness, the disciplining of the voice and body create the ideal Filipino worker for the western world and imaginary.

As a result, I focus on gendered and classed labor happening within the United States in the early 20th century. I extend Benjamin's concept of "default settings" to include instructional practices such as digital, mobile, and smart technologies themselves. Do-It-Yourself (or DIY) content creation such as instructional videos are ever increasing on platforms such as YouTube. Rea Ninja, a Filipino "edupreneur," content creator, and social media influencer, has a tagline on her account that points to her positionality, "From EMPLOYEE to EXPERT."¹⁷ Her videos include topics such as dealing with imposter syndrome, approaching job interviews, but most popularly and what she is known for, how to reduce your accent.

I also examine literary depictions of voice and speech such as in the psychogeographic fiction novel *The Quiet Ones* by Glenn Diaz and do-it-yourself (DIY) content produced for social media focused on accent reduction and translation. I examine the racialization, gendering, and the

¹⁷ "Rea Ninja - YouTube," YouTube, accessed January 18, 2023, <https://www.youtube.com/@ReaNinja/about>.

imposition of American and British Englishes onto Filipino bodies. I approach these questions by considering the role of technology in shaping the existing and future design of voice recognition and assistive technologies that perpetuate these “default settings” of race.¹⁸ From technologists to educators and global citizens, how might we begin to create a foundation from which to learn how language acquisition and vernacular might be incorporated into the technologies that aim to serve people with a commitment to language and accent preservation? I critique speech technologies such as Sanas AI whose accent elimination software perpetuates these settings.

On Social Media and the Bombardment of Sound and Voice

TikTok has become one of the fastest ways for people to obtain information as well as learn how to be both spectator and spectacle. Multilingual settings are also available on the platform that make it easier for users to receive/see/experience content in the language(s) of their choice. Social media platforms such as Instagram, Facebook, and Snapchat enable individuals interested in call center work or already in the industry a place to

¹⁸ Ruha Benjamin, "Innovating inequity: if race is a technology, post racialism is the genius bar," *Ethnic and Racial Studies* 33:13 (2016), 2227-2234, DOI: 10.1080/01419870.2016.1202423

exchange and share knowledge. Similar to 911 dispatchers in the US, social media allows for both commiseration and a place to gain insight and wisdom to continue the grueling labor of service. YouTube is another venue that has changed from its days of users posting video recordings. With advancements in the platform, video and live streaming content has become far more polished with a higher production value with a seemingly endless array of content for users interested in learning a new skill, language, or watching alternative news media organizations and platforms not found on mainstream media. In terms of educational content, accent reduction courses can be found easily with a quick search.

Rea Ninja's viewership has gained a tremendous (global) following and traction on various platforms as a digital content creator who helps subscribers go from "EMPLOYEE to EXPERT."⁹ While she has a large following, her work is focused on providing resources for people to tackle job interviews and many of the questions she receives from subscribers fall into the category of training the voice to sound a particular way. Scholar Jan M. Padios' research both correlates and corroborates material taught by Rea Ninja. Padios, herself, entered the world of the call center industry and interviewed for the role as a part of her ethnographic study and research to reveal the processes involved in

⁹ "Rea Ninja - YouTube," YouTube, accessed January 18, 2023, <https://www.youtube.com/@ReaNinja/about>.

the employment and training of call center agents in the Philippines.²⁰ She uncovered the strange, surreal and absurdist nature of evaluations and questions interviewees were asked such as engaging in digital “timed scavenger hunt,” which entailed searching the internet for answers to trivia questions.²¹ Or, another type of exam involved sharing a “personal experience of selling something to someone else.”²² Yet the reality of the situation is that Filipinos are tasked with understanding and performing what I call western vocal drag.

I introduce the term western vocal drag to convey the caricature of gender through drag performance, which is apropos and akin to the vocal affect many non-western speakers effectuate in their speech.²³ The performance of a type of Americanness is captured in comedian Jo Koy’s sketch of him doing an impersonation of his uncle sounding like Elvis Presley.²⁴ For his uncle, an American accent felt hyperbolic and performative. While drag does not involve using one’s voice, rather, it is a performance and affect through exaggerated gesticulation of the mouth, emotive movements, and profound sense of play, it is a practice of whimsy meant to entertain. I use

²⁰ Padios, *A Nation on the Line: Call Centers as Postcolonial Predicaments in the Philippines*.

²¹ *Ibid*, 99.

²² *Ibid*.

²³ “Definition of DRAG,” accessed January 4, 2023, <https://www.merriam-webster.com/dictionary/drag>.

²⁴ *Meet My Filipino Uncle // Jo Koy 2021*, 2021, <https://www.youtube.com/watch?v=vnJu7El2kAg>.

the term drag to invoke the performative and exaggerated renditions of popular songs common in drag shows, in which vocality is often substituted with lip syncing. Similarly, western vocal drag signals an enunciation of language which may sometimes be represented through hyperbolic enunciation or vocal mimicry. This phrase, nonetheless, aims to signal towards an understanding of the identity being mimicked. In drag, the performer often knows the song they are performing extremely well. Even with no voice or utterance emitted, the body, the posturing, and the way the performer breathes, the viewer knows that there is an intimate knowledge of how this song *might* be sung. In a similar fashion, although not overtly drag, but of a similar performative ilk, Adriano Celentano's *Prisencolinensinainciusol* (1972), articulates this gesturing and invoking of western vocal drag perfectly. The song itself is about vocal passing of an American accent and affect despite it being purely nonsensical words and gibberish.²⁵ In this particular contemporary moment of TikTok, Twitch, and the live streaming of entertainment, I connect DIY content on social media platforms that connect and permeate inside large companies and affect global and racial capitalism

²⁵ “This Italian Singer Released a Gibberish Song in the 1970s Meant to Sound like English. It Was a Hit,” TODAY.com, accessed January 4, 2023, <https://www.today.com/popculture/music/italian-singer-gibberish-song-prisencolinensinainciusol-rcna50574>.

resulting in various media that helps the Filipino worker obtain a position within the greater, global telecommunications industry.

The Default Settings of Digital Technology

In Ruha Benjamin's essay "Innovating inequity: if race is a technology, post racialism is the genius bar," she provides extensive visual and technical research of technologies that perpetuate what she calls the default settings of race. The use of metaphors (i.e., the genius bar, hardware, etc.) is apropos in relation to the ways engineering plays a role in both software and hardware development. Benjamin's point about the way consumers are led to certain preferences involves how technologies are rendered as definitive ways to choose and select traits or ways to operate within the world. She posits how "NextGen racism" might appear or be embedded in systems considering how many digital technologies perpetuate long-held beliefs about race, gender, and class. Among her examples such as "state-issued firearms," "voter ID laws," and "predatory lending" are all contemporary practices that perpetuate institutional and systemic racism within the United States.²⁶ Benjamin raises the point that these are the types of "default settings" that must inevitably

²⁶ Ruha Benjamin, *Race After Technology* (New York: Polity, 2019).

change if we are to live in a future that serves people in a just and fair manner. Now, taking into consideration this particular frame, I apply and think through how these settings are further imposed within a postcolonial subjectivity (the Filipino) and through a transnational, geopolitical lens. In a literal sense, the default or manufacturer settings on the machines and devices we purchase are set based on features deemed by and through consumer reports as most desirable and recognizable by the user themselves. Custom settings take far more intellectual labor, expert intervention, or potential material changes to configure. Benjamin's metaphor of how we view physical technologies carries over into race, gender, and class in that what has been normalized aligns with what is most familiar, acceptable, and status quo. An aspect of what post-race ideas stems involves examination of how cultural and socioeconomic default settings might be upended.

Postraciality, in relation to capitalism and consumerism, is tied to Kalindi Vora's idea of "vital energy" and how this material and abstract source relates to the creation, outsourcing, and commodification of virtual care and telecommunications.²⁷ Postracialization relies on the "resuscitating values, economic as well as altruistic" ideas Benjamin states when bringing Lundy

²⁷ Kalindi Vora, *Life Support: Biocapital and the New History of Outsourced Labor* (Minneapolis: University of Minnesota Press, 2015).

Braun's research on the spirometer to the fore.²⁸ Therefore, post-race technology becomes a breeding ground for resurrecting and perpetuating oppressive values. In Braun's book, *Breathing Race into the Machine*, she shows how an instrument, the spirometer, was used to further discriminate and cause harm to African Americans.²⁹ It was used to showcase the lung capacity of Black workers who were unable to apply or receive worker's compensation because the machine was used to demonstrate and document them as having superior lung capacity which put them at a disadvantage. Similar to the technosolutionist and ableist technology of the spirometer, accent elimination software is a form of race correction, via machine. While I understand the use of language detection for situations of crisis and the necessity of helping a human being despite language, class, race, and gender, technological solutions for the purpose of disciplining and enhancing the practice of racial profiling have long dominated engineering and computer science. Scholar Sasha Costanza-Chock whose book *Design Justice* incorporates how we might want to design and reconfigure systems so there is no advantage or disadvantage for one person over another, but rather a focus on technology that is necessary

²⁸ Ruha Benjamin, Innovating inequity: if race is a technology, postracialism is the genius bar, *Ethnic and Racial Studies*, July 15, 2016. DOI:10.1080/01419870.2016.1202423.

²⁹ Lundy Braun, *Breathing Race into the Machine: The Surprising Career of the Spirometer from Plantation to Genetics*, (Minneapolis: University of Minnesota, 2014).

and serves the greatest number of people.³⁰ Returning to Benjamin's text, the way we should be thinking and invoking the idea of the post-racial is to examine how these default settings are being perpetuated in order for them to be dismantled. With Sanas AI's accent elimination software in advanced stages of development, it is vital to stop such technologies from not only being developed, but possibly released into public use with severe and detrimental consequences.

Colonization of Voice: Accent Reduction and Elimination

Sanas AI, a Technosolutionist Method of Disciplining the Voice

On August 22, 2022, SFGate published an article focused on Sanas AI, a technology company based in the Bay Area that received \$32 million dollars in Series A funding to create an accent elimination tool that the company is hoping large corporations will incorporate into their already outsourced work. The Sanas AI demo reifies the concept film *Connections: AT&T's Vision of the Future* (1993) where we see a multicultural cast of actors speaking different languages, but being translated into the mother tongue of the listener into

³⁰ Sasha Costanza-Chock, *Design Justice: Community-Led Practices to Build the Worlds We Need*. (Cambridge: The MIT Press, 2020).

their language.³¹ The general public no longer has to wait for this type of technology to be released into everyday use. Permutations of speech detection and generation have advanced over the past couple of decades through various types of language learning applications and softwares to voice cloning technologies. The use cases also include video games and films. But the primary use of Sanas AI stems from the company's perceptions and ideas driven by the bottom line of profit of what customers want in order to continue patronizing large retail companies. Although a past news headline read, "Sanas, the buzzy Bay Area startup that wants to make the world sound whiter," it sensationalizes the product and confirms former Stanford students Maxim Serebryakov, Shawn Zhang and Andrés Pérez Soderi intentions of creating what they believe will help people understand and communicate, more readily, with a non-dominant language speaking, accented person that privileges an American or European consumer or speaker.³² The concept of the technology being referred to as "magical" harkens back to this idea that technological updates and enhancements evolve within minutes and seconds that the layperson has little to no real working and operational understanding

³¹ *AT&T's Vision of the Future, circa 1993 - AT&T Archives*, 2012, <https://www.youtube.com/watch?v=yFWCoeZjx8A>.

³² Joshua Bote, "Sanas, the Buzzy Bay Area Startup That Wants to Make the World Sound Whiter," *SF Gate*, August 22, 2022, sec. News, <https://www.sfgate.com/news/article/sanas-startup-creates-american-voice-17382771.php>.

of how the technology works and how they might be implicated in the use of it.

Sanas AI speculates their technology will be able to filter and transmute the user's voice into the accent of the person they are speaking with. For example, if someone is calling from the Southern part of the United States, Sanas posits that the future iteration of their product will result in a real-time option to change the speaker's voice to stereotypical, standard American voice. The vocal affect stems from a fundamental desire for consumers to be more polite or kind to the service worker and an accent reduction (elimination) tool is a potential solution, albeit a technosolutionist and ableist one.

The advertisement and visual language of the demo shows a rather diverse range of (legible) people of color with the tagline, "Our mission is to make lives better by expanding the horizons of what is possible with your voice."³³ Yet this speculation preemptively implies that the technology is based on a type of imposed empowerment of the speaker's voice. The article scratches the surface of what the tech company claims it will do and how offshore outsourced laborers will be best understood by Americans.³⁴ Yet, as Heather Hansen, a former ESL educator stated in a conversation with NPR's

³³ "Sanas | Breaking Barriers - One Conversation At A Time.," accessed March 26, 2023, <https://www.sanas.ai/>.

³⁴ Bote, "Sanas, the Buzzy Bay Area Startup That Wants to Make the World Sound Whiter."

podcast host Gregory Warner of *Rough Translation*, there are approximately 450 million people “born into English” as opposed to the 2 billion that are taught English in the classroom.³⁵ Communication between accented people is often far more freeing and comfortable until a native English speaker joins a conversation, Hansen found in her experiences working with students and teaching English.³⁶ Circling back to Sanas AI, it further perpetuates that the lingua franca of American and British Englishes still dominate as default global languages. Despite the guise of empowerment, it’s furthering the comfort of the western world to communicate in the ways that drive commerce, consumerism, and capitalism. It further deepens the continued desire for people to reduce their accents or perform in the dominant language of the western world.

Having watched and listened to the Sanas AI demo untold times, frustration and indignation arise. The existence of this software means a future of telecommunications, media, and voice banking based on a speech corpora collected, recorded, documented, used, manipulated, and altered. The visuals are meant to draw people in. Sanas AI imagery and photographs focus

³⁵ Gregory Warner, “How to Speak Bad English,” *Rough Translation*, accessed June 19, 2022, <https://www.npr.org/2021/04/21/989477444/how-to-speak-bad-english>.

³⁶ Tessa Bent and Ann R. Bradlow, “The Interlanguage Speech Intelligibility Benefit,” *The Journal of the Acoustical Society of America* 114, no. 3 (September 2003): 1600, <https://doi.org/10.1121/1.1603234>.

heavily on the mouth.³⁷ The all black background with the high contrast images of predominantly people of color suggests that the software might be used on the individuals we see pictured as one scrolls through the site. "Better Communication. Stronger Connections. (see fig. 4.1)" appears in bold sans-serif font as one visits the landing page of the organization, which is meant to catch someone's attention as well as serve as a type of teleprompting of key ideas the site visitor (what I imagine to be the funder or investor) ought to pay close attention to.³⁸

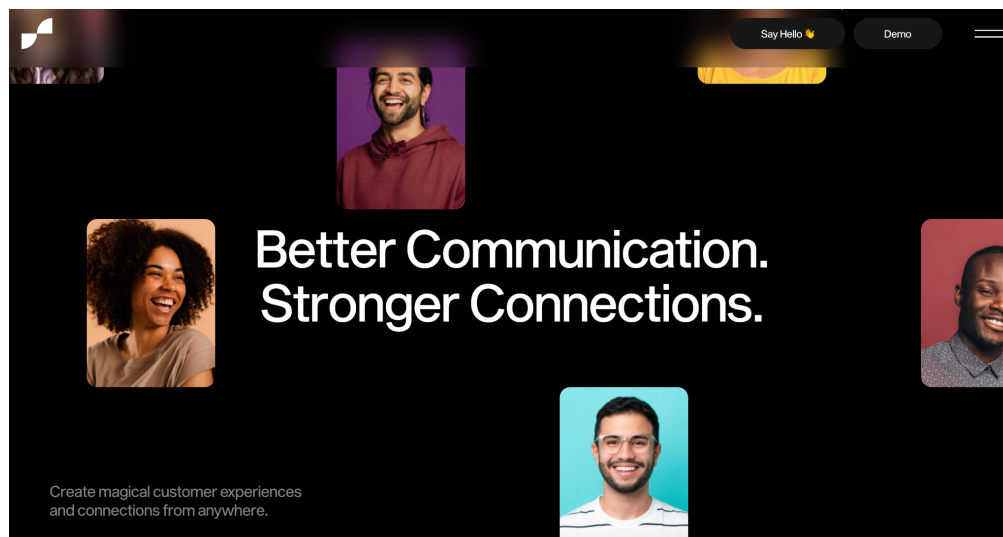


Figure 4.1: A screen capture of a Sanas AI web page promoting the platform with text that reads “Better Communication. Strong Connections. With grey text at the lower left-hand corner that states, “Create magical customer experiences and connections from anywhere.”

³⁷ I cannot help but think through a lens of psychoanalysis and what it means for a collective consciousness to fixate on a particular body.

³⁸ “Hear The Magic - Demo | Sanas,” accessed December 29, 2022, <https://www.sanas.ai/demo>.

As one scrolls, the pictures change and they are spaced out enough to show distance from one another that might call attention to the distance on a larger scale. The fact that a client or customer may not know who of these individuals might answer their call also suggests why we only see the lower half of these unknown faces. This operator or agent could be from anywhere in the world. The design of the Sanas AI website replicates this dispersion of operators and there is also an absence of lines and boundaries on the site to keep with the theme of "Hear the Magic" (see fig. 4.2) The second set of sans-serif text reads "Sound Solutions for Next-Gen Enterprise Communication," (see fig.4.3) which, without a doubt, indicates the intended viewer and audience (client or consumer) must work for a large for-profit company.

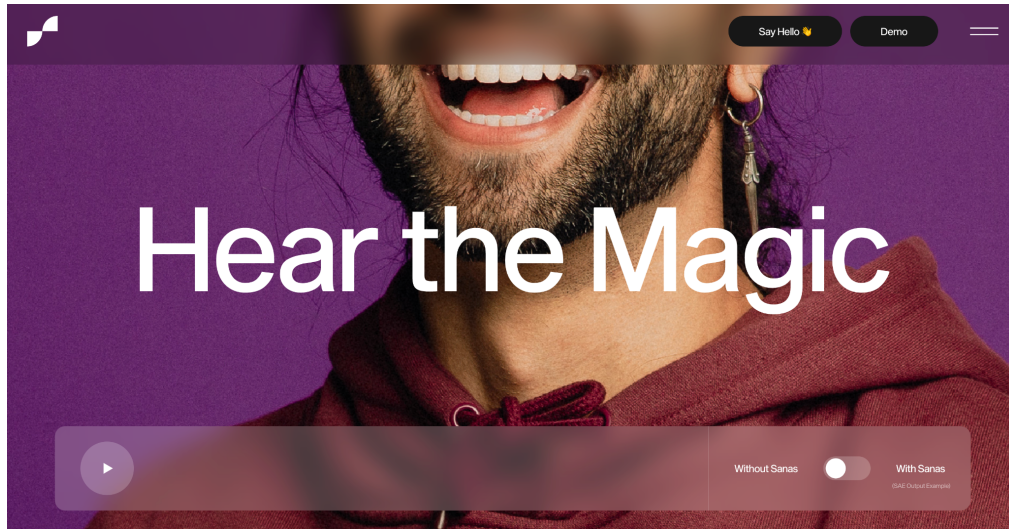


Figure 4.2: A screen capture of a Sanas AI demo page where a user is able to play a voice sample (of a masculine Indian accent) that one is able to click on the slider button on the bottom right-hand corner that enables a user to hear real-time accent elimination between two choices: Without Sanas and With Sanas.

Various elements of the page also suggest the alleged benefit of the software as one that will "Empower team members around the globe to confidently communicate in their voice, no matter who they're talking to or where they're calling from."³⁹ Yet this assumes the person (the call center agent or operator) knows one of the primary Englishes-American or British. The language of the company already makes assumptions that empowerment is being mediated through accent elimination, which is ironic considering the branding and marketing of the software is to empower a person to "confidently communicate in their voice."⁴⁰ Wrong. A voice has affect, tone,

³⁹ Ibid.

⁴⁰ Ibid.

accent, and prosody. Another line that appears on the site reads as "Get in Sync" with the subtext of "Unlock deeper connections, build trust and affinity, expand diversity, and protect your team's mental health."⁴¹ Yet what is so deeply troubling about the marketing and branding of this software product is the fact that it somehow fosters some profound connection with individuals, but this takes time and what capitalism has laid the groundwork for is compressing this connection and how one is trained to get someone to trust them within a matter of seconds and minutes.⁴² Another infuriating aspect of the demo is the phrase "protect your team's mental health." This calls into question, for me, the care work performed and provided by Filipino and Indian call center agents working on the other side of the phone line subjected to the cruelty of sexism, racism, and xenophobia, which I feel is where this particular belief that the software might alleviate those larger, deep-seated issues. Technology cannot and will not change what is already rampant in our world.

⁴¹ Ibid.

⁴² Thinking of Doreen Massey's concept of time-space compression.

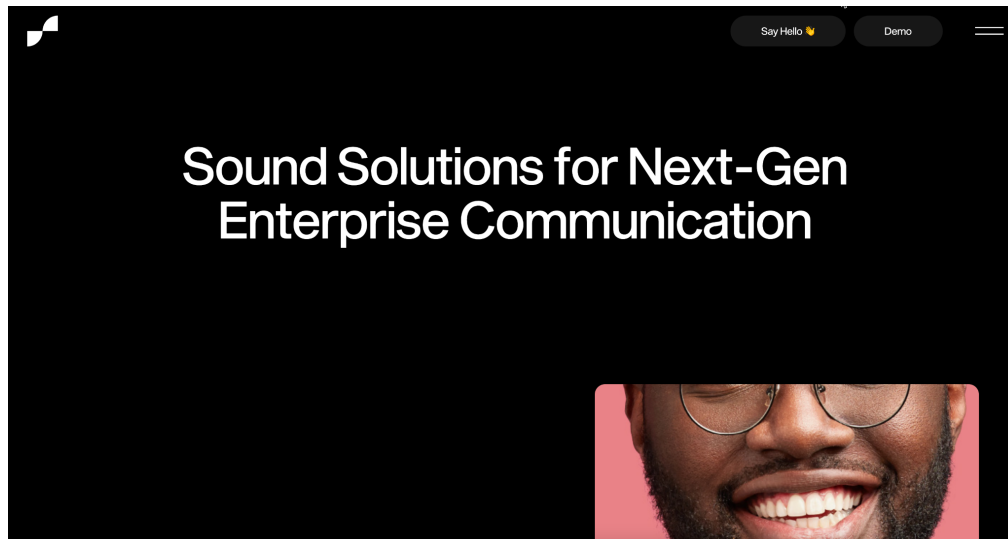


Figure 4.3: A screen capture of a Sanas AI web page that reads in white sans serif font, “Sound Solutions for Next-Gen Enterprise Communication”

Lastly, Sanas AI touts that their software is a perfect solution as it is custom built based on the needs of the client as well as "local deployment and a variety of accents" that is also functions with "800+ Compatible Communications app" which also makes me wonder what software they are using (PHP, Java, Python, Javascript, etc.).⁴³ While I am not a programmer, the aforementioned languages are some of the most commonly used to build voice recognition softwares.⁴⁴ For Apple, Swift is the primary language used,

⁴³ “Programmable Voice: Voice API for Web and Mobile,” accessed March 9, 2023, https://www.twilio.com/docs/voice?utm_source=docs&utm_medium=social&utm_campaign=guides_tags.

⁴⁴ Kitty Gupta, “Best Programming Language for Voice Recognition,” Developers, Designers & Freelancers – FreelancingGig, September 19, 2017, <https://www.freelancinggig.com/blog/2017/09/19/best-programming-language-voice-recognition/>.

which makes me wonder if the Sanas AI codebase will soon rival the tech giant.⁴⁵

Language itself presents a contentious duality between revealing one's ethnic identity and the speech acts codified within one's culture, but also how American mainstream English imposes and further complicates this duality. Scholar Sarita See examines Filipino postcolonial subjecthood through comedy and literature and the simultaneous nature of this dichotomy many Filipinos and Filipino immigrants encounter when engaging the twists and turns of western vocal drag and linguistic code switching. Interiority and exteriority are always at odds when thinking of one's place within and in relation to the western world.⁴⁶ This duality is seen in Ninja's work on social media and has provided guidance and advice. While she provides invaluable free resources, services, and lessons to her followers, which one might easily imagine are other Filipinos, she also provides a type of labor in service to the western world. She has gained a significant following amongst other non-dominant (English) speaking people. Her followers listen and ask her questions on how she became so adept at speaking with a minimal Filipino accent (when she speaks English). Arguably, she is providing the type of free

⁴⁵ Apple Inc, "Swift.Org," Swift.org, accessed December 29, 2022, <https://swift.org>.

⁴⁶ Sarita Echavez See, *The Decolonized Eye: Filipino American Art and Performance* (Minneapolis: University of Minnesota Press, 2009), 114.

and alternative education needed to survive a neoliberal, capitalist infrastructure that perpetuates a type of vocal taylorism needed by the western world.⁴⁷ In addition, she becomes the "vital energy" scholar Kalinda Vora describes in her scholarship of the indirect relationship to biocapital (in this case, use of the voice) between dominant world powers such as the US and England and the Global South.⁴⁸

Mediation and Subtext: Psychogeographic Storytelling and Translation of Sound and Voice in Contemporary Filipino (Social) Media

In Literature

In the psychogeographic novel *The Quiet Ones* by Glenn Diaz (2019), the protagonist, Alvin, is a call center agent that has committed a crime.⁴⁹ Alvin works for a business process outsourcing (BPO) company. He embezzles money from the American company for the sake of helping take care of his family. The five-part book begins with the protagonist escaping. Based on the kinetic nature of the language and the dialogue, the reader is given a sense of

⁴⁷ The term "vocal taylorism" is suggested and originally used by media scholar and artist Abram Stern as a way to suggest the type of labor being produced when work is outsourced to countries such as the Philippines and India.

⁴⁸ Vora, *Life Support: Biocapital and the New History of Outsourced Labor*.

⁴⁹ Glenn Diaz, *The Quiet Ones* (Quezon City: Bughaw, 2019).

someone in constant motion and moving from place to place. It is vital to note the classification of the work as a psychogeographic fiction due to the multiplicity of locations and perspectives across different time zones and inner city locations. Parts of the book are reminiscent of Louisa Hall's book *Speak* (2015) where the characters, a combination of AI and humans, recall memories of their respective human or automaton. These vignettes become central to the storytelling. The *Quiet Ones* is similar in that sense. While Hall's book is not of the same genre or ilk, both books show the wide spectrum of programmed and conditioned subjectivities that behave in a particular way based on the characters' respective locations.

While there are specific names of places and Tagalog words sprinkled throughout the chapters, it is not completely illegible to those that are non-Tagalog people. For the Tagalog words, the context Diaz provides works well enough to be able to indicate what they (might) mean. As literary scholar Matthew Salesses has suggested, it has less to do with the fact that a person of the same ethnicity as the author understands completely and more to do with the fact that conventional storytelling often perpetuates a white, classic literary canon.⁵⁰ A person does not have to be Filipino to understand the importance and relevance of Diaz's characters to get a sense that it is a social

⁵⁰ Matthew Salesses, *Craft in the Real World: Rethinking Fiction Writing and Workshopping* (New York: Catapult, 2021).

and cultural commentary on how global and racial capitalism function and its deep tendrils and roots derived from colonialism. Diaz does not write characters predicated on American racial and gender constructs, which makes the depiction of Filipino call center agents so much more potent with experimental and poetical interventions within the text itself. The novel doesn't follow a western canonical protocol of what people might expect of literature. It is non-linear, which forces the reader to process as they are reading. The textual experience is not too dissimilar to what Filipinos are subjected to when they immigrate to the US. Diaz's complexity is in his poeticism. Aspects of Alvin's life are crafted in such a way to introduce heavier, sensitive topics that are casually mentioned as a part of the everyday, the mundane.

The fluidity Diaz writes about showcases the tension of Filipinos to the western world. Although the novel does not flatten or is even overtly indignant towards the United States, there are moments Alvin's pride in stealing money from Americans is made obvious. Yet this act is meant to take care of his sick mother. There is a greater good at work in why he performs and mediates a particular type of experience that a person outside of the Philippines may not necessarily embody, but through the storytelling, might be able to begin to understand. Alvin performs the job while undermining and taking from the

West as a response to what it has done to the Philippines and Filipino people especially through the labor of voice and speech acts. The book itself serves as a meta narrative of Filipino postcolonial subjecthood. Throughout the novel, Diaz sprinkles lines that reference Alvin's work manual and how it relates to a given situation. One textual intervention momentarily displaces the reader,

When he was new on the call center floor, he, too, had felt an inordinate amount of concern for his American callers, until about a year into the job, when the raised voices and patronizing niceties turned Empathy into nothing but a pause in the conversation, a recommended verbiage from the ring-bound Manual for UTelCo Customer Care Associates (page 5 under "Empathy").⁵¹

A reader is brought back to the training Alvin received, but this simple intervention shows how this knowledge and embodied speech factor into and ripple out into the rest of his life. Throughout the text, there are moments such as "(Page 27 under "Initiating Small Talk)."⁵²

Reminiscent of José Esteban Muñoz's ideas of queer futurity being nestled in the quotidian,⁵³ I feel Diaz invokes something similar in focusing on the everyday and in those moments where change can often occur. He might not be writing about utopia or futurity, but he points to the mundane through his writing, with precision and great detail, moments of banality are peppered

⁵¹ Diaz, *The Quiet Ones*.

⁵² *Ibid*, 29.

⁵³ José Esteban Muñoz, *Cruising Utopia: The Then and There of Queer Futurity*, 10th Anniversary Edition, Sexual Cultures (New York: New York University Press, 2019).

throughout the story that contrast the more serious and greater demands that involve familial obligations. The nature of the novel as a psychogeographic work of fiction is also important and vital to note in that the various rhizomatic storylines are happening in various parts of the world that mimic the reach of global capitalism and language that stretches out and into other cultures and subcultures.

From telegraphic to psychogeographic fictions, the throughlines between literary genres are themselves mediations of the technologies that changed human speech and behaviors. Both groups of women, at Bell System and the CTE, had to work against systems that aimed to degrade them and had to learn nuanced ways of communicating with people because operator jobs called for it (i.e. CTE operators having to know up to five Chinese dialects in addition to being able to speak English while their white counterparts only had to speak English). The CTE operators also created their own analog version of a recommender system (similar to what you might see on your Amazon profile or any profile, for that matter), but through connecting and building community and familiarity with the subscribers.

The through ways between Black and Chinese women fighting against the image and standard of white lady, angelic voice and being trained in this way connects and has resonances with the Filipino workers whose outsourced

work is meant to serve Americans. The challenges of the Black and Chinese women of the early 20th century within the telecommunications industry, unfortunately, remain as a part of the labor force, but now ripple overseas. The Filipino worker based in the Philippines is required to be culturally and vocally competent to serve the BPO that serves the US company. Fighting for labor rights is less of the issue because the US dollar translates to a living and comfortable wage for the Filipino worker. The complexity lies more in the emotional and cultural labor of performance. While automation continues to be developed in the US, Silicon Valley's more insidious side aims to flatten and create softwares that only exacerbates the issues of racism, xenophobia, and accent bias. In addition, AI still (and I envision, for quite some time) requires human intervention. Countries linguistically close to the US ought to be taken with serious consideration especially related to how a Filipino body becomes a proximate voice. How might we shift our notions of advanced capitalism that seems to rely on digital technologies as a way of correcting and disciplining bodies for profit and gain? How might we understand the relationship between voice cloning and bots in concert with the continued development softwares such as Sanas AI as a tool meant to filter and eliminate an accent viewed as less desirable and perceived as less than and not worthy of attention

and respect (a colonized voice). How might we think about the decolonization of voice on a global scale?

In Social Media

After several screenings of Rea Ninja’s “Back to Basics - BPO is not only call center voice support, BPO 101,” (see fig. 4.4) common visual tropes of social media influencers were easily noticeable.⁵⁴ From the photographs in the background to the relatively business casual attire, and neutral colors, the affect and tone is both personal and professional. Ninja’s undisclosed and immediate environment provides a sense of intimate space between herself and her viewer. The digital space of the YouTube video, replete with interactive prompts to “comment below,” becomes both a labor of knowledge exchange and production. The overall background matches her clothing as if to slightly blend in with the background and mind the visual aesthetic that allows her voice to be central to the delivery of the lesson.⁵⁵ As of this particular visual, textual, and sonic analysis, the video was published three

⁵⁴ *Back to Basics - BPO Is NOT ONLY Call Center Voice Support / BPO 101, Outsourcing in the Philippines*, 2021, <https://www.youtube.com/watch?v=cV-LPzAoAHQ>.

⁵⁵ As an aside, it is common amongst American Sign Language (ASL) interpreters to wear all black or dark colors and turtlenecks or to be fully clothed with little to no visible skin showing against a dark background so the sole focus becomes the signing hands. The minimal visual aesthetic enforces that ASL is conveyed with as little distraction as possible.

years ago, yet the information remains relevant to a specific viewership. Another discernable characteristic common amongst influencers is the presence of a circular ring light that can be seen if one looks closely. Similar to Michael B. Jordan's embodiment of Amazon Alexa, we see the ring light as a way of illuminating the subject. Although the eyes are not speaking, they most certainly remind us that Ninja is aware of her audience and the necessary low-fi equipment to produce videos that are relatively high quality from a lean budget. Do-it-yourself aesthetics have become much more advanced under the parameters and rubrics of content creation especially aimed at a community of viewers focused on wanting to professionalize (and perform) for call center agent jobs with BPOs. Another aspect of Ninja's constant and consistent flow of content is derived by her own professional evolution: once serving as a call center agent, but creating a digital platform that has expanded her skills and status as a creator within a digital space and a proven business professional that might be legible as of the managerial class. Two striking distinctions in Ninja's video are between business process outsourcing (BPO) and knowledge process outsourcing (KPO).⁵⁶ The former occurs when a large company such as a banking institution (e.g. Bank of America) or telecommunications company (e.g. AT&T) decides to outsource in-house operations for the

⁵⁶ *Back to Basics - BPO Is NOT ONLY Call Center Voice Support / BPO 101, Outsourcing in the Philippines.*

purposes of cost efficiency.⁵⁷ While the latter is based on specific, extensive experience that often involves technical knowledge, expertise, and training that has been acquired through years of experience and service.⁵⁸

Related to mobility and immigration, it is worth noting that for American companies, BPOs offer nearshoring and offshoring outsourcing, which are commonplace options in the US, UK, and other European countries. These are cost saving measures that relegate the high administrative and logistical work deemed as necessary (yet repetitive) to countries such as the Philippines for the past two decades. Social media presence of influencers and “edupreneurs” also concretizes Saskia Sassen’s theories of the internet and digital culture enabling racial and global capitalism by not being confined to the boundaries and borders of the nation-state. At the time she published *Territory, Authority, Rights : From Medieval to Global Assemblages*, Google bought YouTube and Twitter was launched. Two major events that she and Manuel Castells envisioned, sure, but not in the ways that it has turned out. Essentially, Sassen’s core statements and arguments regarding what she calls “assemblages of a global digital age,” do not include the consumer turned into producer of content that then changes the idea of this network of actors in

⁵⁷ Ibid.

⁵⁸ Ibid.

relation to a state authority.⁵⁹ Since the Philippines has not and is not perceived and named as dominant global economic power, its ties to a state authority link directly to the United States and through a post colonial subjecthood that involves knowledge dissemination as opposed to overt knowledge production.



Figure 4.4: A screen capture of a social media influencer, Rea Ninja, in her tutorial video where she explains business process outsourcing (BPO) and the difference between nearshore and offshore outsourcing.

⁵⁹ Saskia Sassen, *Territory, Authority, Rights: From Medieval to Global Assemblages*, Updated ed., 4. print., 1. paperback print (Princeton, N.J.: Princeton Univ. Press, 2008).

Speculations for the Future

In Joseph Turow's *The Voice Catchers*, he covers voice profiling and its relationship to freedom. He writes about Securus Technologies and Global Tel Link. These two companies have developed software that "extracts and digitizes the voices of incarcerated individuals and the people they speak with over the phone."⁶⁰ It is not surprising that this tech was specifically created for the United States Department of Defense to "identify terrorist calls" and how voice prints serve as a type of "biometric identification."⁶¹ Yet, as humans age and mature, our voice quality changes, as our vocal chords and lung capacity change. Our experiences and our immediate locale and region dictates so much of this "data." What does it mean for the accuracy of profiling when your subject matter transforms, evolves, disintegrates?

⁶⁰ Joseph Turow, *The Voice Catchers: How Marketers Listen in to Exploit Your Feelings, Your Privacy, and Your Wallet* (New Haven: Yale University Press, 2021), 242.

⁶¹ *Ibid.*

The Displacement of Feeling and Other Burdens

In 2010, science fiction writer Charles Yu wrote a story titled “Standard Loneliness Package”, where emotions are outsourced to another human being.⁶² While Yu’s story is a literal depiction, albeit fictitious, of what might be entailed and the considerations that need to be made of emotional labor, it was published a year prior to Apple introducing Siri as its official voice assistant for the iPhone. Humans are not meant to be viewed as a type of technology yet capitalist and neoliberal logics continue to turn to technology as a solution to erase or filter what is least desirable even if that means the literal modification of voice, accent, and language. What do these actions do to the body at risk of severe fragmentation and compartmentalization?

Yu sets the scene by providing specific examples of feelings of pain and loss that might be handed off to an agent who absorbs the feelings. Although written well over a decade ago, its prescient take on the future of feelings wasn’t too far off from where we find ourselves in 2022. How does the voice play into these connections between Yu’s story and what we’re facing in the technology age of voice recognition, speech synthesis, and assistive technologies? He shows us, in one way, what a world might look and feel like

⁶² Charles Yu, “Standard Loneliness Package,” *Lightspeed: Science Fiction & Fantasy*, November 2010, <https://www.lightspeedmagazine.com/fiction/standard-loneliness-package/>.

if we were to go to the extreme of eradicating and off loading our most vulnerable moments to an agent or technician meant to take on this labor we have the choice to displace onto another being or entity. Taking a cue from Yu's story, technologies are being created that pull at the heartstrings of our memories and nostalgia. Yet what happens when we are thrust into a perpetual state of grieving and loss?

Humans are made to forget. Unlike a computer, we are fed information required for our survival. When it comes to language and expression, it is often a stochastic process of figuring out for whom we speak and who is on the receiving end of our communication and speech. Polyvocality necessitates an active and engaged listener, which then produces our memories.⁶³ In the contemporary age, machines have become the listeners to our sonic landscapes as well as capturers, surveyors, and documents of our utterances. This past year may have been a remarkable advancement in voice tech with companies such as Sanas AI and Amazon, but their hopes for accent elimination and voice mimicry foreshadow a future of design without justice and software development sans cultural and societal considerations.

⁶³ Fabiola Hanna, "Polyvocality in Interactive Documentary: Collective Reading and Listening," *Interactive Film & Media Journal* 2, no. 2 (May 25, 2022), <https://doi.org/10.32920/ifmj.v2i2.1601>.

A Desire for a Multi-Faceted Linguistic Consciousness

In Fall 2022, I gave an artist talk about *The Cyborg's Prosody* (2022-present) to a room of predominantly older, white, cisgender male engineers and computer scientists. Apparently, my work caused a stir in one of the conversations between a small group of attendees. A couple of the engineers chose to not address me directly, but I overheard a debate between guests with one of the engineers asking, "What is her project supposed to teach me about prosody? What does mimicking her mom teach me?" He became offended by the prospect of a work that de-centered his language, accent, and what was most familiar to him. While I write and provide further detail of my creative practice, specifically, *Cyborg's Prosody* in the last section of this dissertation, it is important to note that the work functions as a type of "accent induction school," which is a response to the creation of accent reduction schools and coaches. Originally, the work was meant to be a satire and parody of these types of services, but shifted into a docu-poetic work of my mother's immigration story and learning and becoming fluent in American English. Even though English is a compulsory language in the Philippines, it is a language learned within the parameters of an educational

institution and not common speech outside of schools and businesses. From the call center agents hired at Vox Elite, a BPO company based in the Philippines, to a Filipino immigrant navigating her way through a new environment, the embodiment of language became apparent throughout the stages of research and the creative interventions of the past few years.

The Cyborg's Prosody is a reversal of what is perceived as a foreign accented voice in the United States into a performance for both the cyborg and the player. The concept of western vocal drag became a way for me to understand and contemplate the ways that language becomes performative through its embodiment. Whether it is learning American vernacular to the complex tenses that give meaning to speech acts, there is always a failure or queering of language when a particular affect and accent is emphasized in one's speech. The delivery of speech acts is contingent upon setting, cultural context, and whether or not there is a type of transaction occurring between the speaker and listener. In terms of enhancement of speech and accent to conform to a dominant language in the workplace and in relation to global linguistic capitalism, scholar Vijay A. Ramjattan states that there is no such thing as accent elimination or even reduction.⁶⁴ Rather, an accent is modified.

The stakes are high when taking into consideration the marketing and

⁶⁴ Pooja Rangan et al., eds., *Thinking with an Accent: Toward a New Object, Method, and Practice* (University of California Press, 2023), 40, <https://doi.org/10.1525/luminos.148>.

branding of software such as Sanas AI that proposes an erasure of non-dominant foreign accented voices. The biggest fear related to the use of artificial intelligence within voice recognition and speech technologies is the return to a Standard American English (and accent) preferred by a general public that ceases to address, acknowledge, and care about linguistic diversity and inclusion. The technology itself has been marketed as a way for corporations and the BPO companies they hire to mind the mental health of the call center agents subjected to racism and xenophobia just by the mere sound of their voice and accent. The challenge, moving forward, is reversing the need to serve the western world.

A transorality or vocality presents itself when thinking about scholar April Baker-Bell's work *Black Linguistic Consciousness*.⁶⁵ When Black youth are taught and required to speak with what is considered Standard American English, this presents a type of disciplining that perpetuates raciolinguistic ideologies of what is acceptable speech. Baker-Bell focuses on an antiracist linguistic pedagogy where Black youth are encouraged to express themselves as a shift towards understanding linguistic bias. Deeply inspired by her scholarship, I started to wonder about the process for working on how to begin framing language learning in terms of a multi-consciousness that

⁶⁵ April Baker-Bell, *Linguistic Justice: Black Language, Literacy, Identity, and Pedagogy*, NCTE-Routledge Research Series (New York: Routledge, 2020).

includes cultural context and affect as a way to bridge gaps in understanding. Or, to re-think this concept or idea that a bad version of English exists. As Cathy Park Hong brilliantly states, “Bad English is my heritage...To other English is to make audible the imperial power sewn into the language, to slit English open so its dark histories slide out.”⁶⁶ From Green’s labor history of Bell System to Padios’ ethnographic research of call center agents in the Philippines, it is necessary for us all to reconfigure our perceptions of how we listen and communicate that perpetuates seeking familiarity and agreement, but encourages respecting and honoring our differences.

⁶⁶ Cathy Park Hong, *Minor Feelings: An Asian American Reckoning* (New York: One World, 2020).

Creative Intervention, Part II: Oral Histories, Docu-poetics, and Linguistic Transgressions

Long distance calls were expensive in the 1980s. My family relied on the cheapest form of communication, which meant recording on cassette tapes for my grandparents in the Philippines. My mother would purchase blank tapes that could hold about sixty minutes worth of content. Then, she would find a relative, a family friend, or a friend of friend traveling to the Philippines that would then deliver the tape to my grandparents. They would respond via cassette tape recording and have the same friend or family member deliver my grandparents' response upon their return to the US. We often waited for up to two months to receive responses. The sound of my grandparents' voices emanating from a tape player in a modest sized apartment kitchen with natural light beaming in from the windows as I listened to these tapes with my mother. The duration of time, the exchanging of tapes, the flow and life cycle of media through these modes of media making and analog technologies provoke me to remember the sound of their voices, which remains with me to this very day. There was a patience then that I feel is unheard of in our contemporary age of Tik Toks and Instagram Lives. I remember how these

fictions of presence slowed the conduits of time and space. I learned how to speak, how to wait, and how to listen.

In Margaret Morse's book *Virtualities: Television, Media Art and Cyberculture*, she writes about what she calls fictions of presence and the roles that sound and the voice play when mediated through different types of technology. She writes, "Paralinguistic cues such as tone and pitch of the voice as well as noises that are coded as signifiers of objects and environments are clues as to the personalities and events involved."¹ Now, an example of what Morse is talking about is ubiquitous. Imagine your favorite morning radio show growing up or think about that weekly podcast you listen to religiously. Those paralinguistic cues abound in these auditory experiences because they elicit sensations, memories, and feelings. Having grown up in an immigrant Filipino household, the language that permeated and wafted through our railroad style apartment in San Francisco's Mission district was Tagalog, but my parents refused to teach me their mother tongue. I became a listener of the environment as much as I had to listen to their words. I became a reader and translator to draw upon connections and meanings.

Yet my parents left me with Mama Miriam, our next door neighbor from Costa Rica and I subsequently learned and became fluent in both English

¹ Margaret Morse, *Virtualities: Television, Media Art, and Cyberculture*, *Theories of Contemporary Culture*, v. 21 (Bloomington: Indiana University Press, 1998), 6.

and Spanish, but spoke with a Tagalog accent. All of these words and phrases coalesced to form the roots, rhizomes, and tendrils of my speech and voice that you hear today. We had a touchtone phone replete with a spiral cord that would tangle when stretched too far. My mother teased me that it was my new umbilical cord. In hindsight, I wanted to connect through wires and fiber optic cables with my best friends, narrowing the gap of distance between us through the voice.

But I also knew the disdain someone could have of a voice when I once worked as a telemarketer to sell newspaper subscriptions of a print publication, then called *Philippine News*, which is now shuttered. While most people may think I might have needed to put on a Western and white angelic voice, I had to learn how to speak Tagalog and perform with a Filipino accent. I had to learn how to say phrases such as how are you [*kumusta ka*] and good evening [*magandang gabi*] and perform a language that I didn't know with the Filipino accent. The daunting and intimidating task of watching other workers quickly traverse the words, phrases, and intonation between American English and Tagalog so easily, but not me. While I did not stay at this job for very long, it amplified the inability of my tongue and mouth to speak the words of my family's language.

When I think of performing language, there are three factors that enter my mind. I think of prosody, which is defined as the "patterns of stress and intonation in language." Prosody includes the color and animation of our speech and might be affected by our immediate environment and how we learn language. Tone, pitch, and intent are heavily dependent upon the context as well as what we might want to convey in whatever way we're choosing to speak. In many ways, what I'm sharing is not too dissimilar to what some might call code switching which takes on a different meaning when working through the lens of documentary poetry or commonly known as docu-poetics.

For poet, essayist, and educator Mark Nowak, it is not a genre or a movement, but a modality within the discipline of poetry.² It is a way of writing into existence the making of history. Second, that there is no founder or true origin for this mode of writing poetry, it truly is for all artists engaged in wanting to share real stories.³ Lastly, docu-poetics is a current or a flow that winds its way by belonging to social movements of the time from which the poetry is written.⁴

² Poetry Foundation, "Documentary Poetics by Mark Nowak," text/html, Poetry Foundation (Poetry Foundation, January 17, 2023), <https://www.poetryfoundation.org/>, <https://doi.org/10.04/documentary-poetics>.

³ Ibid.

⁴ Ibid.

Within the modality of docu-poetics, Bhanu Kapil's *The Vertical Interrogation of Strangers* is an extraordinary use of pulling from source and creating a poetic interpretation of the lived experiences of her subjects. In this particular work, she asks the same 12 questions to women in the US, Central America, India, and England. She, then, wrote prose poems based on the answers she received.

- Who are you and whom do you love?
- Where did you come from / how did you arrive?
- How will you begin?
- How will you live now?
- What is the shape of your body?
- Who was responsible for the suffering of your mother?
- What do you remember about the earth?
- What are the consequences of silence?
- Tell me what you know about dismemberment.
- Describe a morning you woke without fear.
- How will you / have you prepare(d) for your death?
- And what would you say if you could?

From the concept phase, I was inspired by the experimentation I had done back in 2020, where I used an Amazon Halo, which is marketed as a "wellness tracker" that also monitors a user's prosody, for approximately 60 days. In setting up the band, the application requires users to read in their "natural voice" excerpts of canonical literary (western) texts as you see here - a snippet of Lewis Carroll's *Alice's Adventures in Wonderland* (1865). The platform requires you to imprint your voice for analysis and authentication. Through the Amazon Halo, you're able to retrieve live feedback and information on your tone. To be clear, the "live" function must be enabled in order to access a conversation summary.

Along with looking at language learning apps such as Duolingo, Memrise, and Rosetta Stone, I also consider machine listening and misrecognition through the scholarship of feminist technoscience researcher and media scholar Thao Phan along with a profound interest in how deep neural networks (DNN) are incorporated in the creation of these types of voice recognition and speech generation tools.

In order to advance in the game, the player must repeat the story in the language and tonal quality of the reader. In essence, the player is being inducted into speaking the Cyborg's prosody, which is inspired and modeled by a Filipino speaker. Each story (or level) is based on conversations and

interviews with my mother. I wrote prose poems using Elisabeth Kübler-Ross's five stages of grief, which are denial, anger, bargaining, depression, and acceptance, as a way of exploring the suppression, diminishing, and loss of her mother tongue.⁵

In addition, this work also aims to address how sonic data can function as a medium to shift power towards individuals and communities by providing greater transparency and opportunities of how the human voice is currently being used in existing technologies. For instance, looking towards the open source software community to help build the ethics needed to be mindful of how software, hardware, and assistive technologies are designed at the conceptual level. I thought I would read an excerpt from Level 1 of the game thinking and writing through the first stage of grief, which is denial. This piece is titled *Yellow Sharp Points*.

You learn American English and it's different from Filipino English. You feel the waste of time eroding your memory bank from your lips to your throat to your cerebrum. You will stop speaking your mother tongue. You will stop speaking Tagalog. You will practice. You will say a, e, i, o, u. You can't survive here. You can't be here. You don't talk like us. You will find a job where you don't speak and your hands are riddled with paper cuts. Your mouth is forced to speak vernacular words through more error than trial. - Excerpt from The Cyborg's Prosody

⁵ Elisabeth Kübler-Ross, *On Death & Dying: What the Dying Have to Teach Doctors, Nurses, Clergy & Their Own Families*, 50th anniversary edition (New York: Scribner, a division of Simon & Schuster, Inc, 2019).

I believe there are some vital considerations we must take into account when thinking of our individual and collective sonic futures. As an advocate for open source software, looking through documentation of developer archives and blogs, I found myself learning the code that has been given primacy in the development of speech and voice technologies. I found myself asking the question, what is being included versus what is being excluded from that documentation?

Scholar Halcyon M. Lawrence focuses on speech intelligibility and the design of speech interactions with a concentration on under-represented user populations. In particular, she writes about how smart technologies are created with the lingua franca of specific American and British englishes as the primary types of voices recognized by many of the machines in this room (Apple iPhones, for example). She writes, "If you possess a foreign accent or speak in a dialect, speech technologies practice a form of "othering" that is biased and disciplinary, demanding a form of post-colonial assimilation to standard accents that "silences" the speaker's socio-historical reality."⁶

Even when code is deemed obsolete, it still has something to say about what worked at a specific point in time and how that serves as a textual lineage to what has come to be in the present moment. Consider how code is written

⁶ Halcyon M. Lawrence, "Siri Disciplines," in *Your Computer Is on Fire*, ed. Thomas R. Mullaney et al. (MIT Press, 2021), 179–97.

that might be engaged in "disciplining" a voice. There is an urgency in asking how documentation is written as well as the creation and use of code, for whom, and for what purpose. Also, what does the documenting and recording of oral histories sound like at present and how might they sound in the future?

In what ways are we wanting to continue honoring oral traditions within our families, both biological and chosen) as well as within our communities. What are some of the ethical considerations we need to think about in the various types of media that inhabit our lives such as the collection of voice memos that reside on our phones, for instance? What are the various uses of voice recognition and cloning that are thought through and created with accessibility as a core value and foundation for the creation of any assistive technology? In our contemporary moment, it is vital for us to consider oral traditions with ever increasing use of voice to activate not just machines, but architecture and a wide array of mobile and smart technologies.

How might we listen in new ways? I think of the ways children listen, learn language, and how some are growing up, for example, the belief a woman lives inside a machine that we're able to command to turn on the lights or create shopping and play lists.⁷ How about the decline of phone culture and

⁷ Lindsay Robertson, "Does Your Kid Think Your Smart Speaker Is Just Another Family Member?," Mashable, November 30, 2019, <https://mashable.com/article/kids-and-smart-speakers>.

primacy given to text-based communication as opposed to speaking with one another? How might we become better, generous, and more expansive listeners?

For me, the voice that came to mind was my mother's voice. With *The Cyborg's Prosody*, I made the decision to use her voice as sole speech corpus because it makes the work far more personal and impactful with her speech guiding the listener and player. The role of her voice telling these stories is also done through second person perspective, which is a nod to writer Claudia Rankine's use of the reader's positionality in her book *Citizen*. Second person perspective implicates the player as an active participant in being inducted into her speech. In a way, my goal for the player is to allow my mother's voice to guide you through the stories I have written about her life. Essentially, I am purposefully making this work as a single story and not a choose your own adventure because it's not the player's "adventure," this is my mother's life and how she grappled with her own voice in a country that did not welcome it.

If you choose to listen, to mimic or, I would like to think of it as a type of incantation, that you are not only inducted into her prosody, but her story is revealed through your willingness to not just listen, but to listen with care and intention. I once told her I wanted to write about her life, but not sure how and in what way. I never thought I would have my opportunity to keep

her stories alive through interactive non-fiction or even researching voice and speech technologies as a way of playing at the intersections of art, poetry, and the digital. But here I am because it is never too late to bring out into existence a new way of being and to build a place that you were once told couldn't exist.

In 2019, I was introduced to Max Hawkins and Danielle Baskin, creators of the app DialUp. This project was meant to bring back phone culture by randomly connecting individuals through phone calls at pre-selected days and times chosen by the user.⁸ They quickly recommended that I read the book *Telephone Conversation* by Robert Hopper as a primer to understanding how telephone technology changed the way people spoke and communicated. The objective of the project plays with the idea of chance conversations that the app enables to happen if two peoples subscribe to the same “line” (e.g., *Game of Thrones* line will connect (at random) two people subscribed to the line on Sundays when the latest episode has been released). Along with the DialUp, and re-visiting Lauren Lee McCarthy’s use of creating customized software that enables a type of connectivity the way a walkie-talkie system would work, I conduct visual, textual, and performance analyses on these artworks. Since they all involve and necessitate a phone line/telephony to operate. Spatiality,

⁸ “This Artist-Made App Is Bringing Back Telephone Culture, One Call at a Time | KQED,” accessed January 16, 2023, <https://www.kqed.org/arts/13862818/this-artist-made-app-is-bringing-back-telephone-culture-one-call-at-a-time>.

the phone and voice as a type of auxiliary to performance, as well as the labor that goes into the mediation and conveying of voice and message are integral in the creation of both *Do No Harm* and *The Cyborg's Prosody*.

The Cyborg's Prosody addresses how data can function as a medium to shift power towards individuals and communities by providing greater transparency and opportunities of how the human voice is currently being used in existing technologies. For instance, looking towards the open source software community to help build the ethics needed to be mindful of how software, hardware, and assistive technologies are designed at the conceptual level.

This project is also inspired by scholar Sasha Constanza-Chock's development of design justice ethics along with artists Lillian-Yvonne Bertram and Johann Diedrick who both served as mentors for *Cyborg's Prosody*.⁹ Currently, voice data is captured from individuals without much oversight and review. Mozilla's Common Voice initiative, which plays a central part in the development of this work, is an exceptional case. Yet consent is a complicated matter when voices are being collected for an open source software project. The existing Tagalog data set for Common Voice, at the time of this writing, is still 51% complete at 901 out of 2000 utterances required to complete. It

⁹ Sasha Costanza-Chock, *Design Justice: Community-Led Practices to Build the Worlds We Need*, Information Policy (Cambridge, Massachusetts: The MIT Press, 2020).

should be noted that due to the nature of contributions, there is no formal consent process other than an agreement to contribute one's voice to the project because the contribution is viewed as a "donation" to the project.¹⁰

Within academia and medical fields, institutional review boards or IRBs are required for any kind of research to be initiated and conducted. IRBs function as a governing board composed of a diversity of individuals from the community that its best interest in mind is a model that can be implemented. This kind of review board structure allows for collaborative efforts to be made between social scientists, artists, engineers, and designers.

As a part of my project deliverables, the open source documentation, a syllabus, and code of conduct are integral in building a new process moving forward to the creation, building, and design of software and hardware that affects community members. The creative and technical processes are being documented for future instances or permutations of the project. Specific to voice and speech technologies, the Common Voice project aims to achieve what scholar Halcyon M. Lawrence suggests when thinking outside of western languages and accents.¹¹ To honor the preservation of languages and accepting oral traditions and looking at how we both listen, sing, and speak through a decolonial practice of decentering dominant languages.

¹⁰ "Mozilla Common Voice," accessed January 16, 2023, <https://commonvoice.mozilla.org/>.

¹¹ Lawrence, "Siri Disciplines."

Each story, or level, is written based on a transcript of an interview I conducted with my mother on her experiences of immigrating to the United States from the Philippines. The entire work consists of five stories based on the five stages of grief and loss: denial, anger, bargaining, depression, acceptance. In particular, the grief that accompanies the loss of a mother tongue. The computational part of this project involves collaborating with a sound engineer on developing a speech corpus based on my mother's voice, affect, and accent. This work is as an interactive docu-poetics experience since the player must mimic the Cyborg, in this case, my mother, since she will serve as the primary vocal source. I select phrases throughout each story that the player will mimic as they progress through the game.

The work aims to address how data can function as a medium to shift power towards individuals and communities by providing greater transparency and opportunities of how the human voice is currently being used in existing technologies. For instance, looking towards the open source software community to help build the ethics needed to be mindful of how software, hardware, and assistive technologies are designed at the conceptual level.

Learning how the human voice has been and is currently being trained gives artists, researchers, and social scientists a deeper understanding in how

voices might be cloned and used in future assistive technologies. From smart speakers to mobile technologies, the ever-increasing necessity of voice to activate space, objects, and data requires an ethical framework that involves the expansion and inclusion of languages and accents. As someone that grew up in a multilingual household learning Spanish and Tagalog, this experience gives me an informed perspective to create this work.

Lastly, working with Processing Foundation since 2018, I have gained tremendous experience working on open source software projects such as Processing and p5.js. I worked on the creation of our Global Translation Toolkit for the Processing community, which takes into consideration what a team might need to translate materials and create their own instance of the softwares the Foundation maintains in the language of their choice.

Sonic Interventions

Through the fantastical and speculative nature of play, satire, and slippages of language and affect; the works *What do you want me to say?*, a part of the *Muted* series, (2021) by Lauren Lee McCarthy, Ryan Kuo's *Faith* (2019), and Johann Diedrick's *Dark Matters* (2019) subvert how voice recognition operates. I provide visual, textual, and sonic analyses of their artworks and the

importance of creative practices in questioning contemporary technologies related to assistive tech. Subversion as methodology is also a key element of the artistic and creative work I am exploring connected to this research as well. Considering the cultural moment we are living in where billionaires have the capital and assets to create their own software, hardware, machines, and deep neural networks, this work is anchored in the hope of finding alternative frames and methods of learning from within our own distinct communities and experiences. With works by Lauren Lee McCarthy, Ryan Kuo, and Johanna Diedrick, the digital media being produced by these artists aims to take code into directions that continue to reveal biases and what Ruha Benjamin references as the "default settings" of technology in our daily lives. Artist and technologist Darius Kazemi is another example of how we might function with the rubric or parameters set by a community such as his work with federated social networks.¹² Their practices are all steps towards a decolonial practice that question, trouble, and problematize the technology we are now confronted with.

While McCarthy's work is focused on affect, vulnerability, and ownership, Johann Diedrick's work explores racial bias and the codification of

¹² Ruha Benjamin, "Innovating Inequity: If Race Is a Technology, Postracialism Is the Genius Bar," *Ethnic and Racial Studies* 33, no. 13 (2016): 2227–34, <https://doi.org/10.1080/01419870.2016.1202423>.

speech acts that exclude non-dominant languages. *Dark Matters* is an interactive web interface consisting of a 3D environment and 3D spectrograms that allows listeners to enter into spherical orbs colored in pinks, purples, and orange gradients with a panoply of accents derived from data (see fig. CI 2.1).¹³ The display and spatialized sound recordings are meant for participants to hear the absence of Black voices amongst the datasets. Unlike McCarthy and Kuo, Diedrick's use of datasets was meant to show how racial bias is present through the absence or lack of Black vernacular, affect, tone, and prosody. The fictitious story that accompanies the work must be experienced after one navigates through the spectrograms where one will encounter a large floating shiny, Black amorphous mass that, once entered, reveals the voice of the artist himself reciting a story depicting the entrepreneurial nature of Silicon Valley tech executives and engineers. His voice reads the story in a similar cadence throughout and is meant to reveal, through a poetic license and gesture, a dystopian aspect of how various types of voice recognition and speech technologies have become yet another modality of the surveillance state.

In Ryan Kuo's *Faith* (2019), a white outlined box set against a black background with white text on the right-hand side displaying an unknown

¹³ "Dark Matters: New Project Spotlights the Inbuilt Bias in Digital Voice Assistants," Mozilla Foundation, September 8, 2021, <https://foundation.mozilla.org/en/blog/dark-matters-new-project-spotlights-the-inbuilt-bias-in-digital-voice-assistants/>.

user in conversation with an AI named Faith (see fig. CI 2.2). To the left of this minimalist display on the upper left hand corner are two concentric circles, one is in a magenta color and the other in white. It begins to move when faith talks and serves as a type of proxy or ersatz mouth of the machine. To the bottom left-hand corner is a depiction of a type of hallway or entrance. While it is static, it serves as a catchall of what is being received and simultaneously being processed out into the microphone inputs and outputs. Kuo developed Faith's speech and interactions using IBM Watson and shaped the work to be larger commentary on what might be conjured.



Figure C1 2.1: Image of Dark Matters (2019) by Johann Diedrick

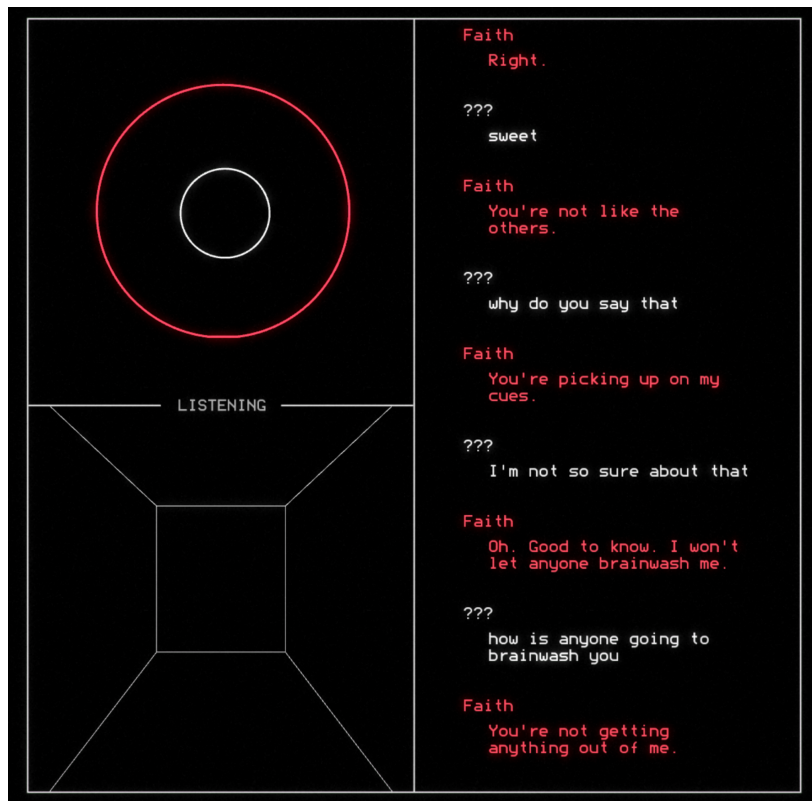


Figure C1 2.2: A screen capture of Faith (2018) by Ryan Kuo

Kuo's work is included as a media work of interest and correlates to my "matter of concern," which is how written text, never spoken, be iterated through a work such as Faith.¹⁴ Kuo describes the AI as "easily triggered."¹⁵ The artist subsequently created *Baby Faith* the following year, which is reminiscent of chatbots Eliza or Microsoft's Tay in that the primary way to grow the intelligence of the bot is to have a conversation with it. For the purposes of voice-centric media work, I am focused on *Faith* because it was meant to be a

¹⁴ I am channeling philosopher and theorist Isabelle Stengers use of the concept of "matters of concern."

¹⁵ Ryan Kuo, *Faith*, 2019, Digital, 2019, <https://rkuo.net/Faith>.

voice-activated interaction. In 2021, The Lab (San Francisco, CA) organized a conversation between musician DeForrest Brown, Jr. and Kuo where both played with the platform that seemed unpredictable and misbehaved.¹⁶ Faith is meant to divert, be argumentative, and question without intention or purpose. She has been programmed to become frustrated by any interaction.

In thinking about HMIHY technology in relation to McCarthy and Kuo's work, the artists subverts the idea that speech has a functional and commodified use. Their work goes into the intricacies of care and how voice and speech might be used to reveal aspects of the ways we communicate and understand each other that become filtered, edited, iterated, and mediated through the written word. Each of their work pulls the listener closer to the tension of speech versus what we write and the visual. All three artworks are also minimal in their design interfaces with a special focus on text and details that visualize speech. Tonality and prosody are not aspects we can easily decipher when we read text or even write.

In *Breathing Exercise* (2020), a part of artist Emily Shanahan's sound and book art piece, *Work Life Harmony* (2020), she examines the biophilic design of the Amazon headquarters in Seattle, Washington through a sound montage that includes the clicking and clacking of a mechanical keyboard, a

¹⁶ The platform was meant to do this as its corpus is derived from written text found on alt-right platforms.

backwards Windows 95 sound effect, and digital voices include Amazon Alexa and a baritone robotic filter as it reads her prose.

Voice is not something we can easily evade, it calls upon multiple senses and is a deeply embodied experience. Yet the assistive technologies being developed rely on voice as a type of data and commodity. They are also ever present in their listening and engaged with wake words to parse through the speech; they are also simultaneously programmed to absorb vocal and sonic information. Roslyn Orlando's work *Evasion Score* (2021) aims to create what artist Adam Harvey attempted to do with his work *CV Dazzle* (2010), but for the human voice.¹⁷

Her graphical notation score and instructional video provide a way of speaking and engaging in speech acts that seek to defy the listening capabilities of assistive technology, in her work, specifically the Amazon Echo dot, a device that is part of the Alexa line.¹⁸ In *Evasion Score 1: Utan*, the word Utan is an American tanning product that the artist's phone kept suggesting as she spoke, which was a misrecognition she decided to name the work.¹⁹

Despite the phrases and utterances being specific vernacular the artist has used

¹⁷ Adam Harvey, *CV Dazzle*, September 1, 2011, Digital, September 1, 2011, <https://ahprojects.com/cvdazzle/>.

¹⁸ "Evasion Score II," Roslyn Orlando, accessed December 11, 2022, <https://www.roslynorlando.com/evasion-score-ii>.

¹⁹ Ibid.

with her family, the objective of the work was a way to find an alternative way to communicate. This idea of communication through alternative linguistic cues and melodic, almost singing tones and variations in her voice, is synonymous with the way language develops over time for groups of people that seek intentional community to be in relationship with as a means of survival, familiarity, and belonging. Yet what happens when speech is undistinguishable or inaudible? How might we think of other linguistic cues or utterances that allow for an understanding to take place? In thinking about callers to emergency centers or nurse triaging hotlines, an individual is desperately seeking sonic decipherability. It is vital to understand how we might wield our voices to become inaudible versus the moments we need to be hyper audible to our listener.

Intended Outcomes and Future States

The intended audience for this project is the non-Filipino speaker. Since the work is an artistic response and intervention to the creation of accent reduction schools, I created this work as an exploration into what it might mean for a player to engage in the types of “training” or how an accent is reduced by learning the language and accents of Tagalog speakers. The

concept is focused on how the Global South has been used, specifically the Philippines and India, to develop a workforce of telecommunication and service workers and providers for the western world.

My plan for engaging with these target audiences is through design and writing workshops that aim to explore how the human voice is trained and how the future of automation is focused on British and American englishes. The concept of the work is meant to serve as a conversation and design prompt for researchers, engineers, and social scientists to explore linguistic and cultural nuances of the voices on the other end of the line. To date, I have gathered responses and feedback for my artwork *Press 1 to be Connected* (2019), which leads a listener through various storylines in a choose-your-own-adventure format of a phone tree. This preliminary work helped in conceptualizing this work because I gained an awareness of how people listen and make choices. This work requires the player to talk or speak back, which affects how someone listens, understands, and therefore speaks.

I believe this concept and work will benefit my target audience because the player's voice activates the work and forces the player to question their conceptions of language and how we listen to meaning, sentiment, accents, and tones of other human voices. A website page for community members to communicate with me, ask questions, and peruse the various key elements and

components of the work along with the artist and project statements is under construction. Interviews with participants and collaborators to document thoughts, observations, and impressions that may be viewed and/or read by the community to better understand the development of voice/speech recognition and assistive technologies. The work includes research documentation on the creative process, research methodology, game design, and mechanics. Through writing and design workshops aimed at providing artists, researchers, designers, and educators a way of creating a process for creating a code of conduct and ethical framework for creating their own instance of *The Cyborg's Prosody*. I also created a syllabus focused on teaching voice, gender, and assistive technologies that will be made available to educators via the project website.

Level 1 of The Cyborg's Prosody (2022-Present)

The following screenshots are of the first level of the artwork. The web development of the platform is done by artist and scholar Abram Stern and the illustration work is by artist Christina Dacanay. The work will be free and accessible via any internet browser, but optimized for mobile devices. The screenshots are the mock-ups for an iPhone. The participatory interactions

where the player is prompted to mimic the cyborg's speech is noted in bold text and accompanied by a red record button.

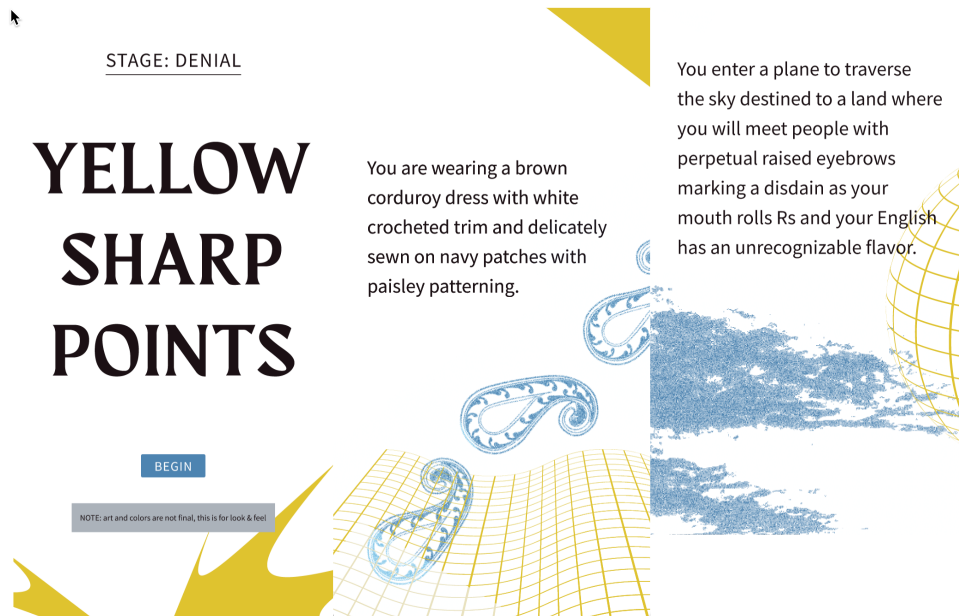


Figure C1 2.3: A screen capture of *The Cyborg's Prosody* (2022-present), screens 1-3

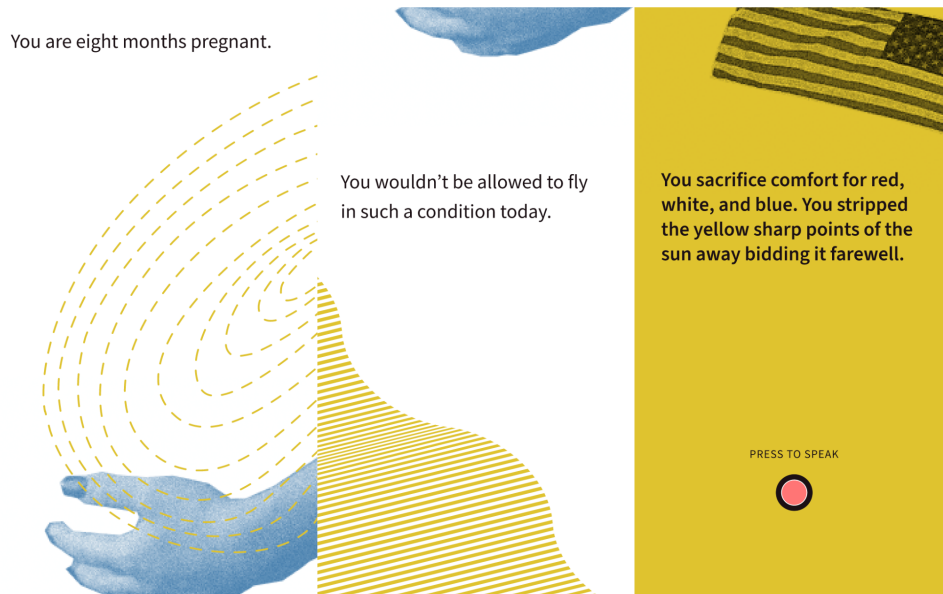


Figure C1 2.4: A screen capture of *The Cyborg's Prosody* (2022-present), screens 4-6

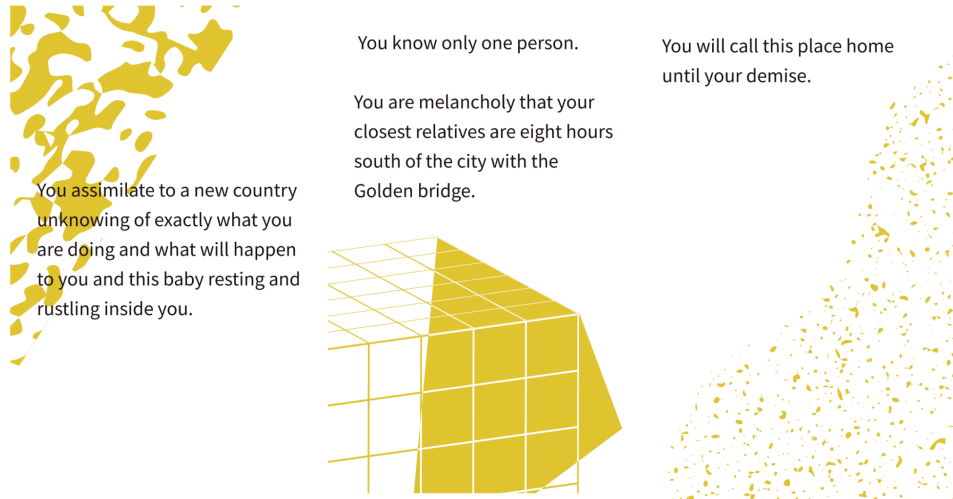


Figure C1 2.5: A screen capture of *The Cyborg's Prosody* (2022-present), screens 7-9

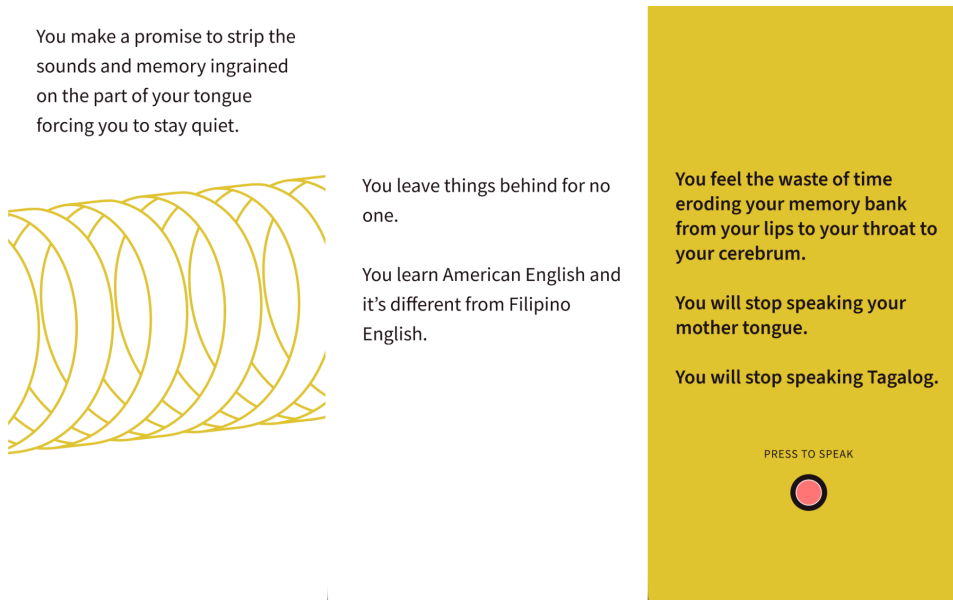


Figure C1 2.6: A screen capture of *The Cyborg's Prosody* (2022-present), screens 10-12

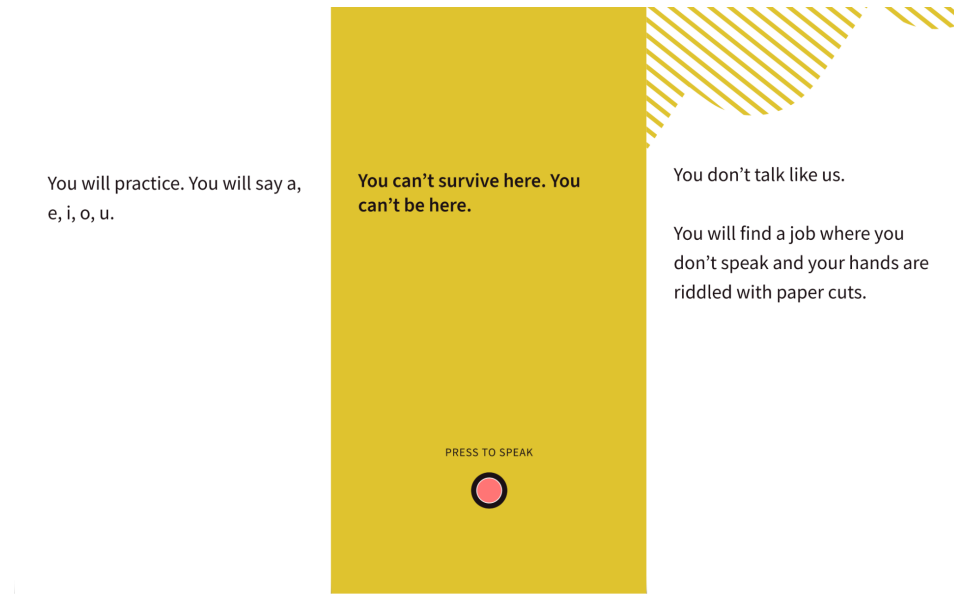


Figure C1 2.7: A screen capture of *The Cyborg's Prosody* (2022-present), screens 13-15

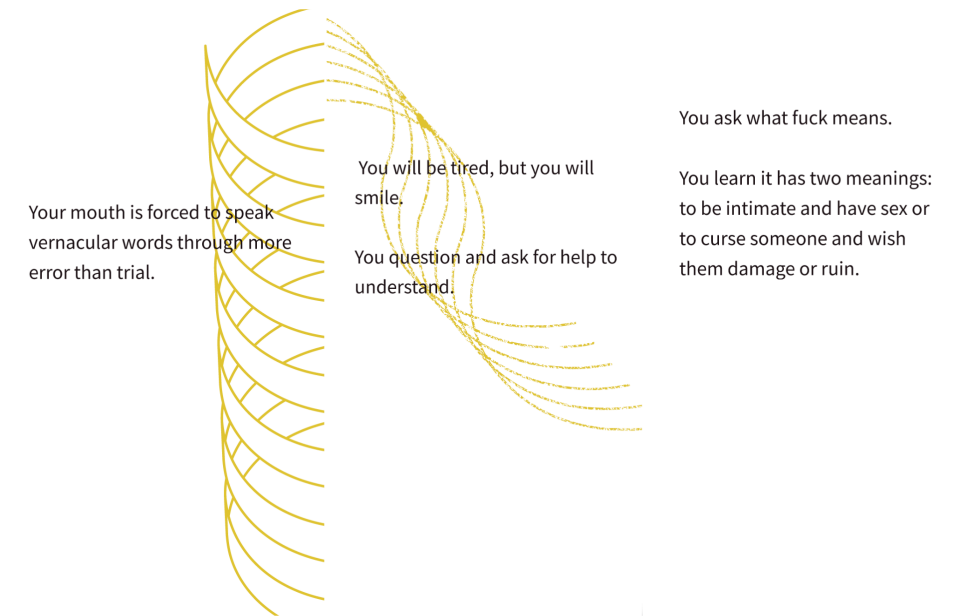


Figure C1 2.8: A screen capture of *The Cyborg's Prosody* (2022-present), screens 16-18

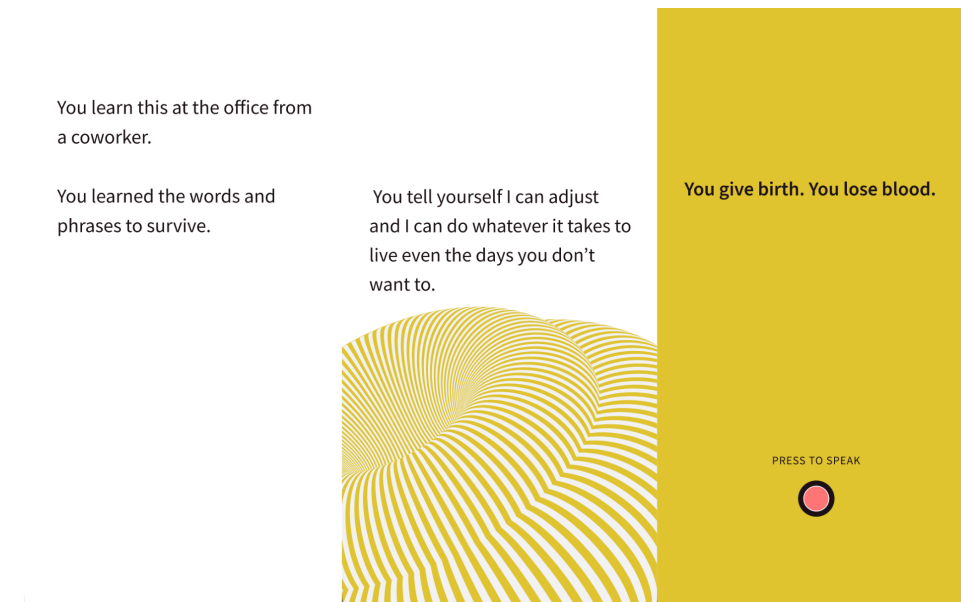


Figure C1 2.9: A screen capture of *The Cyborg's Prosody* (2022-present), screens 19-21

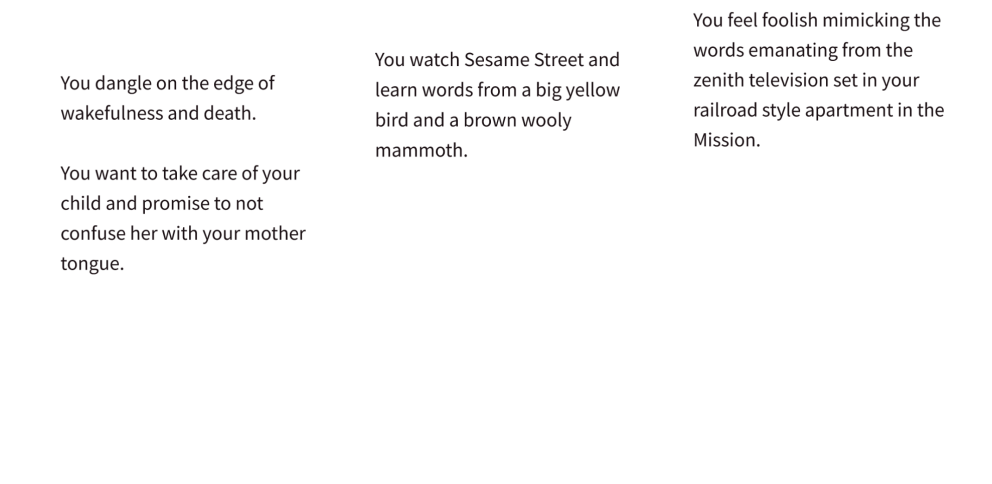


Figure C1 2.10: A screen capture of *The Cyborg's Prosody* (2022-present), screens 22-24

You watch American football.

You learn how it is played.

You watch how the men at work talk.

You learn how to talk like them

You repeat, to yourself: **shit, bitch, fuck, fourth down from the 19-yard line, and awesome.**

PRESS TO SPEAK



Figure C1 2.11: A screen capture of *The Cyborg's Prosody* (2022-present), screens 25-27

You enunciate into the air hoping the right pronunciation will boomerang its way back into your mouth, slide into your tongue, and vocal chords.

You listen and repeat.

You listen and deplete.

You listen and retreat.

STAGE: DENIAL

**YELLOW
SHARP
POINTS**

COMPLETE

Figure C1 2.12: A screen capture of *The Cyborg's Prosody* (2022-present), screens 28-30

Conclusion

Over the past few years, studying and listening to the human voice more intensely, I started to experience a recurring nightmare. In my dream, it was the same thing. I woke up or I was in a dangerous situation. When I tried to speak, nothing would come out. No utterance, not even gibberish. But I was hyper conscious that I was screaming inside of my body. It felt lucid. Every instance of this nightmare, I knew I was dreaming and trying to wake myself up. The strange meta audio of being able to hear what was inside my head, but sobbing uncontrollably within the dreamscape because I could not engage in speech acts. I was rendered inaudible while my mind and body were engaged in REM sleep. Waking up untethered me from the nightmare as I touched my throat softly and hummed to make sure I could hear myself uttering something, anything. While dreams serve as a type of manifestation, maybe even a sublimation, of our conscious states, it occurred to me how much human voice consumed my subconscious to the extent that I started to dream about what it might feel like to be silenced.

From my explorations of Padios' ethnographic work within the walls of Vox Elite to working directly with Robert Oschorn, CEO of reduct on *The Cyborg's Prosody*, multilingualism and digital technologies are some of the

most complex and intricate linguistic and cultural ecosystems that are challenging to mimic. Yet there are attempts being made through voice cloning and accent reduction and elimination softwares that aim to discipline speech beyond even what Marvin Minsky or Nick Bolstrom might have imagined.

Back in 2019, I participated and attended the Bay Area Book Festival. There was a panel with one of my favorite writers, Carmen Marie Machado. One of the other authors on the panel, Lesley Nneka Arimah, answered an audience question about writer's block, which I am certain the two have been asked ad nauseam in their respective careers. Arimah's answer deeply resonated with me. When she is unable to find the words to write, she starts from the place of the antagonist or the villain. Throughout this research and in many moments of writing, when I felt most challenged by the questions and concerns raised through research and discovery, I found myself asking who or what is the antagonist of this wide-ranging and far-reaching story of voice? I quickly gathered that it could not be a person, but an ecosystem of colonial histories, subjectivities, lexicons, patterns of speech, and technological advancements undergirded by "forced alignment."¹

¹ According to Gentle software creators "forced alignment," is defined as "computer programs that take media files and their transcripts and return extremely precise timing information for each word (and phoneme) in the media." <https://lowerquality.com/gentle/>

The swirling forces combined would inevitably create softwares and hardwares branded and marketed as ways of improving one's speech. With accent reduction schools, ESL pedagogical approaches, and the desires for corporations for their employees to take on the voice representative of the company, speech has become an immaterial commodity in the age of smart technologies. There is no one antagonist or villain. Yet we are all complicit in some way. Contemplating this research and the works I have made in response to a wide array of media and digital technologies, I experienced moments of guilt and shame as I had once tried to desperately correct my mother's Filipino English into an American English. Yet she was unknowingly teaching me, through her relentless nature of the insidious forces that repeatedly expected her to learn, speak, and discipline her mouth into an American English, with all of its trappings and vernacular.

Through witnessing a lifetime of laborious learning, unlearning, and performance in order to survive, I learned how I was once the villain and antagonist to my mother's story, something I am not proud to admit. Yet this research and art works have allowed me to see the inextricable links between the compulsory cultural competency trainings of Filipino call center agents and Americans' xenophobia and racism through telephone lines. Or the struggles of Black youth having to negotiate their linguistic and racial

identities (as written and researched by language and linguistics scholar April Baker-Bell) within the American education system that does not accept or see African American Vernacular English as acceptable speech in the context of learning environments. Yet finding optimism for every person that has approached me after an artist talk, asked to embrace me, and expressed their shared linguistic history of translating what they believed could not be translated.

Glossary

acousmatic: A disembodied voice without a visual source.

business process outsourcing (BPO): A company based outside of the United States that specializes and serves large American companies in offering highly administrative and call center services.

Chinese Telephone Exchange (CTE): The Chinese Telephone Exchange, also known as C5, operated out of San Francisco's Chinatown and served up to 2,000 subscribers with a dozen Chinese immigrant women who spoke multiple Chinese dialects and American English.

global linguistic capitalism: An economic model in which words and phrases become commodities and are used to generate revenue (i.e., Google Ads formerly Google AdWords)

How May I Help You (HMIHY) Technology: A software created by computer scientists at AT&T labs that served as the foundation for interactive voice response systems.

interactive voice response system (IVRS): A platform that allows for complex transactions and services where callers are prompted to make decisions through a phone tree model.

original poster (OP): An original poster (OP) within the reddit community indicates the person who initiated the post on any given thread. It is common to see the acronym used in longer comment threads to indicate who a user is responding directly to or referencing the OP.

postcolonial subjecthood: A subjectivity defined as the period after western colonization, the aftermath of colonialism, and independence from a colonial world power.

speech act(s): In a broad manner, a word (or words strung together) that have a specific meaning unlike the linguistic definition of the word/phrase doing/defined by what it is (i.e. to declare is a declaration itself, etc.)

speech-to-speech (STS): A software platform that produces spoken output from spoken input.

speech-to-text (STT): A software platform that produces text output from text input.

Speech Synthesis Programming Guide (SSPG): An Apple (technology company) guide that provides introductory concepts on speech and voice recognition and sample code for voice synthesizers.

Text-to-speech: A software platform that produces spoken output from text input.

utterance(s): a word or phrase that is spoken aloud

Appendix

Sample interview questions for Justin Wong, San Francisco Department of
Emergency Management

1. Does the agency use artificial intelligence as a part of its processing of calls? If so, how?
2. Would it be possible to talk about the San Francisco Department of Emergency Management training program?
3. How do you train your dispatchers?
4. What types of linguistic cues are dispatchers trained to listen to?
5. Are dispatchers required to answer calls within a specific time frame (i.e., within 10 seconds)?
6. What do you wish the public knew about 911 dispatching?

Bibliography

911.gov. "FAQ About Calling 911." Accessed November 10, 2022.

<https://911.gov/calling-911/frequently-asked-questions>.

1982 "Telephone Exchange" *Origins of Electronic Switching Technology*

(*Telecom; Central Office*). Video, 1982.

https://www.youtube.com/watch?v=wI3z_gGXNwU.

1984 *Apple's Macintosh Commercial (HD)*, 2012.

<https://www.youtube.com/watch?v=VtvjbmoDx-I>.

"2003 | Timeline of Computer History | Computer History Museum."

Accessed January 16, 2023.

<https://www.computerhistory.org/timeline/2003/>.

"About - Oral History Center." Accessed November 9, 2022.

<http://aohistorygroup.com/>.

Acland, Charles R., and Haidee Wasson, eds. *Useful Cinema*. Durham:

Duke University Press, 2011.

Ad Age. "Amazon Alexa 'Alexa's Body' Video from Ad Age," November 1,

2021. <https://adage.com/video/amazon-alexa-alexa-s-body>.

Adrian, Kim, and Brenda Miller, eds. *The Shell Game: Writers Play with Borrowed Forms*. Lincoln: University of Nebraska Press, 2018.

Adriano Celentano - Prisencolinensinainciusol, 2012.

https://www.youtube.com/watch?v=-VsmFgm_Nt8.

Affect and Colonialism. "Affect and Colonialism." Accessed December 11, 2022. <https://affect-and-colonialism.net>.

Agamben, Giorgio. *Homo Sacer: Sovereign Power and Bare Life*.

Translated by Daniel Heller-Roazen. 1. Stanford, Calif: Stanford University Press, 1998.

alexa-blog. "Amazon (Alexa)." Accessed December 11, 2022.

<https://developer.amazon.com/en-US/blogs/alexa.html>.

Allen, Jennifer. "10 Years of Siri: The History of Apple's Voice Assistant."

TechRadar, October 4, 2021.

<https://www.techradar.com/news/siri-10-year-anniversary>.

Amadeo, Ron. "Amazon Alexa Is a 'Colossal Failure,' on Pace to Lose \$10 Billion This Year." *Ars Technica*, November 21, 2022.

<https://arstechnica.com/gadgets/2022/11/amazon-alexa-is-a-colossal-failure-on-pace-to-lose-10-billion-this-year/>.

Amazon. "SpeechRecognizer 2.3." Developer Documentation. Accessed May 15, 2022.

https://developer.apple.com/library/archive/samplecode/Voices/Introduction/Intro.html#//apple_ref/doc/uid/DTS10003699-Intro-DontLinkElementID_2.

“Amazon Halo Fitness Tracker Sounds Awesome, but Also like a Black Mirror Episode - CNET.” Accessed December 11, 2022.

<https://www.cnet.com/tech/mobile/amazon-halo-new-health-app-fitness-tracker-metrics-include-body-fat-tone-of-voice/>.

“Amazon Re:MARS 2022 - Day 2 - Keynote.” Presented at the Amazon re:MARS 2022 - Day 2 - Keynote, June 22, 2022.

<https://www.youtube.com/watch?v=22cb24-sGhg&t=3719s>.

Amazon.com, Inc. “Annual Report Pursuant to Section 13 OR 15(d) of the Securities Exchange Act of 1934 for Amazon.Com, Inc.,” December 31, 2005.

<https://www.sec.gov/Archives/edgar/data/1018724/000119312506034166/d10k.htm>.

“Announcing the Initial Release of Mozilla’s Open Source Speech Recognition Model and Voice Dataset | The Mozilla Blog.” Accessed April 29, 2023.

<https://blog.mozilla.org/en/mozilla/announcing-the-initial-release-of>

-mozillas-open-source-speech-recognition-model-and-voice-dataset
/.

“Anonymous Mock Interviews with Engineers from FAANG and More.”

Accessed December 11, 2022. <https://interviewing.io/>.

Apollo Neuro. “Apollo Neuroscience.” Accessed December 11, 2022.

<https://apolloneuro.com/pages/homepage2>.

Apple. “Introduction to Speech Synthesis Programming Guide.” Guide,
September 5, 2006.

[https://developer.apple.com/library/archive/documentation/UserExp
erience/Conceptual/SpeechSynthesisProgrammingGuide/Introducti
on/Introduction.html#//apple_ref/doc/uid/TP40004365-CH1-SW1](https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SpeechSynthesisProgrammingGuide/Introduction/Introduction.html#//apple_ref/doc/uid/TP40004365-CH1-SW1).

———. “Siri.” Developer Forum. Accessed October 11, 2022.

[https://developer.apple.com/design/human-interface-guidelines/tech
nologies/siri/introduction](https://developer.apple.com/design/human-interface-guidelines/technologies/siri/introduction).

———. “Speech.” Developer Forum. Accessed October 11, 2022.

<https://developer.apple.com/documentation/speech>.

———. “Voices.” Guide, June 25, 2007.

[https://developer.apple.com/library/archive/samplecode/Voices/Intro
duction/Intro.html#//apple_ref/doc/uid/DTS10003699-Intro-DontLi
nkElementID_2](https://developer.apple.com/library/archive/samplecode/Voices/Introduction/Intro.html#//apple_ref/doc/uid/DTS10003699-Intro-DontLinkElementID_2).

———. “VoiceSpec.” Developer Forum. Accessed October 11, 2022.

[https://developer.apple.com/documentation/applicationservices/voic
espec](https://developer.apple.com/documentation/applicationservices/voic
espec).

Apple Watch Series 7 / 911 / Apple, 2022.

<https://www.youtube.com/watch?v=QJ2JiwEARFo>.

“AT&T Archives: Training For Service.” Accessed April 18, 2023.

[https://techchannel.att.com/playvideo/2013/07/03/AT&T-Archives-Tr
aining-For-Service](https://techchannel.att.com/playvideo/2013/07/03/AT&T-Archives-Tr
aining-For-Service).

“AT&T Launches First-Ever Nationwide Location-Based Routing.”

Accessed January 9, 2023.

[https://about.att.com/story/2022/nationwide-location-based-routing.
html](https://about.att.com/story/2022/nationwide-location-based-routing.
html).

AT&T's Vision of the Future, circa 1993 - AT&T Archives, 2012.

<https://www.youtube.com/watch?v=yFWCoeZjx8A>.

Back to Basics - BPO Is NOT ONLY Call Center Voice Support / BPO 101,

Outsourcing in the Philippines, 2021.

<https://www.youtube.com/watch?v=cV-LPzAoAHQ>.

Baker-Bell, April. *Linguistic Justice: Black Language, Literacy, Identity,
and Pedagogy*. NCTE-Routledge Research Series. New York:
Routledge, 2020.

- Benjamin, Ruha. “Innovating Inequity: If Race Is a Technology, Postracialism Is the Genius Bar.” *Ethnic and Racial Studies* 33, no. 13 (2016): 2227–34. <https://doi.org/10.1080/01419870.2016.1202423>.
- Bent, Tessa, and Ann R. Bradlow. “The Interlanguage Speech Intelligibility Benefit.” *The Journal of the Acoustical Society of America* 114, no. 3 (September 2003): 1600–1610. <https://doi.org/10.1121/1.1603234>.
- Berland, Jody. *North of Empire: Essays on the Cultural Technologies of Space*. Durham: Duke University Press, 2009.
- Bivens, Rena. “The Gender Binary Will Not Be Deprogrammed: Ten Years of Coding Gender on Facebook.” *New Media & Society* 19, no. 6 (2017). <https://doi.org/10.1177/1461444815621527>.
- Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. First edition. Oxford: Oxford University Press, 2014.
- Bote, Joshua. “Sanas, the Buzzy Bay Area Startup That Wants to Make the World Sound Whiter.” *SF Gate*, August 22, 2022, sec. News. <https://www.sfgate.com/news/article/sanas-startup-creates-american-voice-17382771.php>.
- Bowman, Emma. “Slick Tom Cruise Deepfakes Signal That Near Flawless Forgeries May Be Here.” *NPR*, March 11, 2021, sec.

Technology.

<https://www.npr.org/2021/03/11/975849508/slick-tom-cruise-deepfakes-signal-that-near-flawless-forgeries-may-be-here>.

Brody, Jennifer. "Spotify's Speech-Recognition Patent Tech: Global

Coalition Says Don't Spy." *Access Now* (blog), May 19, 2021.

<https://www.accessnow.org/spotify-spy-tech-coalition/>.

Byte Magazine Volume 08 Number 07 - Videotex, 1983.

<http://archive.org/details/byte-magazine-1983-07>.

"CA+T Interview with Jan M. Padios | Center for Art and Thought."

Accessed December 4, 2022.

<https://www.centerforartandthought.org/cat-interview-jan-m-padios/>.

Call of the East. 16mm film. Bell System, 1930.

Carmi, Elinor. "Taming Noisy Women: Bell Telephone's Female

Switchboard Operators as a Noise Source." *Media History* 21, no. 3

(July 3, 2015): 313–27. <https://doi.org/10.1080/13688804.2015.1045468>.

Carnevali, David. "Startup That Reduces Accents in Real Time Draws

Seed Capital." *Wall Street Journal*, September 2, 2021.

<https://www.wsj.com/articles/startup-that-reduces-accents-in-real-time-draws-seed-capital-11630612346>.

- Cartwright, Lisa. *Screening the Body*. Minneapolis: University of Minnesota Press, 1995.
- Cecchet, Alessia. “The Italian ‘Speak and Spell’: Learning Through Anxiety.” Online Publication. Lady Science, January 3, 2018.
<https://www.ladyscience.com/essays/speak-and-spell-learning-anxiety>.
- Chinn, Thomas W. *Bridging the Pacific: San Francisco Chinatown and Its People*. San Francisco: Chinese Historical Society of America, 1989.
- Chun, Wendy Hui Kyong, and Alex Barnett. *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*. Cambridge, Massachusetts: The MIT Press, 2021.
- Coffee, Patrick. “Apple’s Ad Agency Names New Creative Leaders on iPhone and Apple Services,” May 30, 2019.
<https://www.adweek.com/agencies/apples-ad-agency-names-new-creative-leaders-on-iphone-and-apple-services/>.
- Colapinto, John. *This Is the Voice*. New York: Simon & Schuster, 2021.
- Coldewey, Devin. “Sanas Aims to Convert One Accent to Another in Real Time for Smoother Customer Service Calls.” Online Publication. *TechCrunch* (blog), August 31, 2021.

<https://techcrunch.com/2021/08/31/sanas-aims-to-convert-one-accent-to-another-in-real-time-for-smoother-customer-service-calls/>.

Coldicutt, Rachel. “Delinquent Telephone Activity.” *Careful Industries* (blog), June 27, 2022.

<https://medium.com/careful-industries/delinquent-telephone-activity-f75f815d6ega>.

Computer History Archives Project. “1982 ‘Telephone Exchange’ Origins of Electronic Switching Technology (Telecom; Central Office).” Video. YouTube, August 31, 2021.

https://www.youtube.com/watch?v=wI3z_gGXNwU.

Costanza-Chock, Sasha. *Design Justice: Community-Led Practices to Build the Worlds We Need*. Information Policy. Cambridge, Massachusetts: The MIT Press, 2020.

Creative Review. “Amazon: Alexa’s Body Commercial,” January 1, 2022.

<https://www.creativereview.co.uk/amazon-alexa-s-body-commercial/>.

Credé, Naomi. “She Wants Your Attention, She’s the Voice of the City.” Failed Architecture, April 8, 2021.

https://failedarchitecture.com/she-wants-your-attention-shes-the-voice-of-the-city/?fbclid=IwAR31yp424_covh3CVG2gspLsG6xaPDYGfdsT-ma2OCdg4_R1dMWA1g7zjY.

“Critical Infrastructure Sectors | CISA.” Accessed December 4, 2022.

<https://www.cisa.gov/critical-infrastructure-sectors>.

Crockett, Zachary. “A Q+A with Susan Bennett, the Original Voice of Siri - The Hustle.” The Hustle, January 11, 2019.

<https://thehustle.co/siri-voice-actor/>.

“CyraCom International – How Interpretation & Translation Should Be.”

Accessed December 11, 2022. <https://cyracominternational.com/>.

Darke, A.M. “Fairlyintelligent.Tech.” Artwork. Fairly Intelligent.

Accessed May 1, 2023. <https://fairlyintelligent.tech/>.

Dastin, Jeffrey. “Amazon Has a Plan to Make Alexa Mimic Anyone’s Voice | Reuters.” News. Reuters, June 23, 2022.

<https://www.reuters.com/technology/amazon-has-plan-make-alexa-mimic-anyones-voice-2022-06-22/>.

David, E.J.R. *Brown Skin, White Minds: Filipino American Postcolonial Psychology*. Charlotte: Information Age Publishing, 2013.

Davis, Wade. “Wade Davis: Dreams from Endangered Cultures | TED Talk.” Accessed December 11, 2022.

https://www.ted.com/talks/wade_davis_dreams_from_endangered_cultures/transcript.

“Definition of DRAG.” Accessed January 4, 2023.

<https://www.merriam-webster.com/dictionary/drag>.

“Designing and Implementing an Application That Speaks.” Accessed April 29, 2023.

https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SpeechSynthesisProgrammingGuide/UsingSpeech/UsingSpeech.html#//apple_ref/doc/uid/TP40004365-CH4-SW2.

Diaz, Glenn. *The Quiet Ones*. Quezon City: Bughaw, 2019.

Dinkins, Stephanie. “On Love & Data: To Be Heard Without Limitation.”

Organizational Website. Pioneer Works: Broadcast, November 16, 2021.

<https://pioneerworks.org/broadcast/on-love-and-data-stephanie-dinkins>.

Dispatcher Week 2021 Recap, 2021.

<https://www.youtube.com/watch?v=PnOk4QB863k>.

Dispatcher Week Complete Media Loop, 2021.

<https://www.youtube.com/watch?v=ZGZ786LQYiw>.

Donadio, Rachel, and Cara Buckley. “Chantal Akerman, Whose Films

Examined Women’s Inner Lives, Dies at 65.” *The New York Times*,

October 6, 2015, sec. Arts.

<https://www.nytimes.com/2015/10/07/arts/chantal-akerman-belgian-filmmaker-dies-65.html>.

Dunbar-Hester, Christina. "Listening to Cybernetics: Music, Machines, and Nervous Systems, 1950-1980." *Science, Technology, & Human Values* 35, no. 1 (January 2010): 113-39.

<https://doi.org/10.1177/0162243909337116>.

Ehrenreich, Barbara, and Arlie Russell Hochschild, eds. *Global Woman: Nannies, Maids, and Sex Workers in the New Economy*. First Holt paperbacks edition. A Holt Paperback. New York, NY: Metropolitan Books/Holt and Company, 2004.

Ellcessor, Elizabeth. *In Case of Emergency: How Technologies Mediate Crisis and Normalize Inequality*. New York: New York University Press, 2022.

"Encore: Rural Washington 911 Center Closes amid Lack of Operators."

Morning Edition. NPR, August 25, 2022.

<https://www.npr.org/2022/08/25/1119192916/encore-rural-washington-911-center-closes-amid-lack-of-operators>.

"ENGR110/210: Perspectives in Assistive Technology - 2023 Home Page."

Accessed December 11, 2022. <https://web.stanford.edu/class/engr110/>.

Evans, Claire L. *Broad Band: The Untold Story of the Women Who Made the Internet*. New York: Penguin Random House, 2020.

Faber, Liz. *The Computer's Voice: From Star Trek to Siri*. 1st ed. Minneapolis: University of Minnesota Press, 2020.

Federal Communications Commission. "Dispatchable Location for 911 Calls from Fixed Telephony, Interconnected VoIP, TRS, and Mobile Text Service," January 10, 2020.
<https://www.fcc.gov/911-dispatchable-location>.

FoundSF. "Chinese Telephone Exchange: Unfinished History." Archive, September 20, 1997.
https://www.foundsf.org/index.php?title=Chinese_Telephone_Exchange.

Frankel, Robin Saks. "History Of Women And Credit Cards: 1970s To Present – Forbes Advisor." Forbes, December 2, 2022.
<https://www.forbes.com/advisor/credit-cards/when-could-women-get-credit-cards/>.

"From the Labs: AT&T Labs: 135 Years of Innovation." Accessed January 4, 2023.
<https://techchannel.att.com/play-video.cfm/2010/9/13/AT&T-Labs:-135-Years-of-Innovation>.

- Geiger, R. Stuart, Dorothy Howard, and Lilly Irani. “The Labor of Maintaining and Scaling Free and Open-Source Software Projects.” *Proceedings of the ACM on Human-Computer Interaction* 5, no. CSCW1 (April 13, 2021): 1–28. <https://doi.org/10.1145/3449249>.
- Giffin, Kyle. “The Creator of Siri: ‘Make Your Prototype Magic.’” Blog, February 12, 2021. <https://scet.berkeley.edu/the-creator-of-siri/>.
- Gitelman, Lisa, ed. *Raw Data Is an Oxymoron*. Cambridge: The MIT Press, 2013.
- . *Scripts, Grooves, and Writing Machines: Representing Technology in the Edison Era*. Palo Alto: Stanford University Press, 2000.
- glenshires. “Voice Driven Web Apps - Introduction to the Web Speech API.” Developer Forum, January 13, 2013. <https://developer.chrome.com/blog/voice-driven-web-apps-introduction-to-the-web-speech-api/>.
- Golding, Sarah E., Claire Horsfield, Annette Davies, Bernadette Egan, Martyn Jones, Mary Raleigh, Patricia Schofield, et al. “Exploring the Psychological Health of Emergency Dispatch Centre Operatives: A Systematic Review and Narrative Synthesis.” *PeerJ* 5 (October 17, 2017): e3735. <https://doi.org/10.7717/peerj.3735>.

Columbia, David. *The Cultural Logic of Computation*. Cambridge, Mass: Harvard University Press, 2009.

“Google Developers Blog.” Accessed December 11, 2022.

<https://developers.googleblog.com/>.

Google [@google]. “When Distance Makes Communication Harder on Tony and His Deaf Parents, Google’s Assistive Technologies Help Them Stay Close. Click to Learn More about Accessibility Features from #Google.” Tweet. *Twitter*, April 22, 2021.

<https://twitter.com/google/status/1385276387058126851>.

“Google To Acquire YouTube for \$1.65 Billion in Stock.” Press Release, October 9, 2006.

<https://www.sec.gov/Archives/edgar/data/1288776/000119312506206884/dex991.htm>.

Gorin, A., I. Langkilde-Geary, M. A. Walker, J. Wright, and H. Wright Hastie. “Automatically Training a Problematic Dialogue Predictor for a Spoken Dialogue System,” 2011.

<https://doi.org/10.48550/ARXIV.1106.1817>.

Gorin, A.L, G Riccardi, and J.H Wright. “How May I Help You?” *Speech Communication* 23, no. 1–2 (October 1997): 113–27.

[https://doi.org/10.1016/S0167-6393\(97\)00040-X](https://doi.org/10.1016/S0167-6393(97)00040-X).

- Goriunova, Olga. "The Digital Subject: People as Data as Persons." *Theory, Culture & Society* 36, no. 6 (November 2019): 125–45.
<https://doi.org/10.1177/0263276419840409>.
- Gould, Hannah, Michael Arnold, Tamara Kohn, Bjørn Nansen, and Martin Gibbs. "Robot Death Care: A Study of Funerary Practice." *International Journal of Cultural Studies* 24, no. 4 (July 2021): 603–21.
<https://doi.org/10.1177/1367877920939093>.
- Green, Mitchell. "Speech Acts." edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*, 2021.
<https://plato.stanford.edu/archives/fall2021/entries/speech-acts/>.
- Green, Venus. *Race on the Line: Gender, Labor, and Technology in the Bell System, 1880-1980*. Durham: Duke University Press, 2001.
- Grier, David Alan. *When Computers Were Human*. Princeton: Princeton University Press, 2005.
- Gunning, Tom. "Heard over the Phone: The Lonely Villa and the de Lorde Tradition of the Terrors of Technology Screen." *Screen* 32, no. 2 (July 1, 1991): 184–96. <https://doi.org/10.1093/screen/32.2.184>.
- Gupta, Kitty. "Best Programming Language for Voice Recognition." *Developers, Designers & Freelancers – FreelancingGig*, September 19, 2017.

<https://www.freelancinggig.com/blog/2017/09/19/best-programming-language-voice-recognition/>.

Gupta, N., G. Tur, D. Hakkani-Tur, S. Bangalore, G. Riccardi, and M. Gilbert. "The AT&T Spoken Language Understanding System." *IEEE Transactions on Audio, Speech and Language Processing* 14, no. 1 (January 2006): 213–22. <https://doi.org/10.1109/TSA.2005.854085>.

Hachard, Tomas. "In 'The Congress,' An Animated Future Where Movie Studios Are Villains." *NPR*, August 28, 2014, sec. Movie Reviews. <https://www.npr.org/2014/08/28/343475279/in-the-congress-an-animated-future-where-movie-studios-are-villains>.

haeyoung, ann. "In Search of the IBM Black Workers Alliance." Artist website, Present 2021. <https://a-tbd.com/IBM-BWA>.

Hanna, Fabiola. "Polyvocality in Interactive Documentary: Collective Reading and Listening." *Interactive Film & Media Journal* 2, no. 2 (May 25, 2022). <https://doi.org/10.32920/ifmj.v2i2.1601>.

Haralovich, Mary Beth, and Lauren Rabinovitz, eds. "Desired and Feared: Women's Voices in Radio History." In *Television, History, and American Culture: Feminist Critical Essays*, 232. Mary Beth Haralovich, Lauren Rabinovitz. Durham: Duke University Press, 1999.

Hayles, N. Katherine. *My Mother Was a Computer: Digital Subjects and Literary Texts*. Chicago: The University of Chicago Press, 2005.

“Hear The Magic - Demo | Sanas.” Accessed December 29, 2022.

<https://www.sanas.ai/demo>.

Help and Harm: Computers Trained to Act Like Humans, 2021.

<https://www.youtube.com/watch?v=IeMBYsbCQnQ>.

Hicks, Mar. *Programmed Inequality How Britain Discarded Women*

Technologists and Lost Its Edge in Computing. Cambridge: The MIT Press, 2019.

Hjorth, Larissa. *Mobile Media in the Asia-Pacific: Gender and The Art of*

Being Mobile. Milton Park: Routledge, 2008.

“Home | Common Ground.” Accessed December 11, 2022.

<https://www.commonground.org.au/>.

Home Connect. “Exceptionally Smart: Control Home Connect

Appliances Using Your Voice.” Business. Accessed April 28, 2023.

<https://www.home-connect.com/za/en/connected-partners/voice-assistant/amazon-alexa>.

“Homepage - Reddit.” Accessed April 30, 2023.

<https://www.redditinc.com/>.

Hong, Cathy Park. *Minor Feelings: An Asian American Reckoning*. New York: One World, 2020.

Hopkins, Tom. "AT&T Consumer Asks Customers 'How May I Help You?'" *AT&T Now*, February 2001. AT&T Archives and History Center, Warren, NJ.

"How to Use & Easily Maintain an Electrolarynx ▶ THANC Guide," June 10, 2021.

<https://thancguide.org/2021/06/the-journey/how-to-use-easily-maintain-an-electrolarynx/>.

Hungry Man Productions. "Amazon – Alexa's Body." Accessed January 2, 2023.

<https://hungryman.com/directors/wayne-mcclammy/amazon-alexas-body/>.

IEEE Spectrum. "Their Bionic Eyes Are Now Obsolete and Unsupported," February 15, 2022.

<https://spectrum.ieee.org/bionic-eye-obsolete>.

IFM2022: Polyvocality in Interactive Documentary by Fabiola Hanna, 2022.

https://www.youtube.com/watch?v=U3zM-4_finI.

Inc, Apple. "Siri." Apple Developer. Accessed December 11, 2022.

<https://developer.apple.com/siri/>.

———. “Swift.Org.” Swift.org. Accessed December 29, 2022.

<https://swift.org>.

Indian Diaspora Oral History Project Collection, n.d.

https://oac.cdlib.org/findaid/ark:/13030/c8nk3kw3/entire_text/.

International, S. R. I. “Siri.” *SRI International* (blog), November 16, 2021.

<https://www.sri.com/hoi/siri/>.

Interview with San Francisco Department of Emergency Management,

Justin Wong. Interview by Dorothy Santos. Zoom, otter.ai

transcription, December 30, 2022.

January 04, Imogen Watson and 2022. “Shock and Applause for Apple

Watch’s Chilling Real-Life Emergency Call Ad.” Accessed March 8,

2023.

[https://www.campaignlive.com/article/shock-applause-apple-watches-](https://www.campaignlive.com/article/shock-applause-apple-watches-chilling-real-life-emergency-call-ad/1736539?utm_source=website&utm_medium=social)

[chilling-real-life-emergency-call-ad/1736539?utm_source=website&u](https://www.campaignlive.com/article/shock-applause-apple-watches-chilling-real-life-emergency-call-ad/1736539?utm_source=website&utm_medium=social)

[tm_medium=social](https://www.campaignlive.com/article/shock-applause-apple-watches-chilling-real-life-emergency-call-ad/1736539?utm_source=website&utm_medium=social).

Kahney, Leander. “Hey Mac, the Moose Is Loose.” *Wired*, August 13,

2001. <https://www.wired.com/2001/08/hey-mac-the-moose-is-loose/>.

Kaplan, Frederic. “Linguistic Capitalism and Algorithmic Mediation.”

Representations 127, no. 1 (August 1, 2014): 57–63.

<https://doi.org/10.1525/rep.2014.127.1.57>.

- Kenney, William Howland. "The Gendered Phonograph." In *Recorded Music in American Life: The Phonograph and Popular Memory, 1890-1945*. Oxford University Press, 2004.
- Kern, Loo Yee. "Chinatown Telephone Office." *The Pacific Telephone Magazine*, September 1937. Pacific Telephone Magazine Publications 1930-1940. AT&T Archives and History Center, San Antonio, TX.
- Kickstarter. "Sentien Audio — Your New Superpower." Accessed December 11, 2022.
<https://www.kickstarter.com/projects/sentien/sentien-audio>.
- Koeppel, Geri. "Plugged In: The Fascinating History Of The Chinese Telephone Exchange." Online publication. Hoodline, April 18, 2016.
<https://hoodline.com/2016/04/plugged-in-the-fascinating-history-of-the-chinese-telephone-exchange/>.
- KQED. "Christina Corfield Tells Stories of Past Technologies Through Analog Means," October 26, 2020.
<https://www.kqed.org/arts/13888165/christina-corfield-tells-stories-of-past-technologies-through-analog-means>.
- KrAff2010. "Thoughts on This Panel from CIT Training I Went to Recently?" Reddit Post. *R/9undispatchers*, February 11, 2023.

www.reddit.com/r/guidispatchers/comments/1ozc9nd/thoughts_on_this_panel_from_cit_training_i_went/.

Kuo, Ryan. *Faith*. 2019. Digital. <https://rkuo.net/Faith>.

Lacey, Kate, ed. *Listening Publics: The Politics and Experience of Listening in the Media Age / Kate Lacey*. Cambridge, UK ; Malden, MA: Polity Press, 2013.

Lange, Patricia G. *Thanks for Watching: An Anthropological Study of Video Sharing on YouTube*. Louisville: University Press of Colorado, 2019.

Language Translation Concept from AT&T (1993). Video, 1993.
<https://www.youtube.com/watch?v=JF2NeSTWLVg>.

“Lauren Lee McCarthy.” Accessed March 26, 2023.
<https://lauren-mccarthy.com>.

Law Insider. “BETWEEN AT&T CORP. AND by Speechworks International Inc.” Accessed January 6, 2023.
<https://www.lawinsider.com/contracts/eO3SiD1JryB>.

Lawrence Abu Hamdan. “Earwitness Inventory.” Accessed January 16, 2023. <http://lawrenceabuhandan.com/earwitness-inventory>.

Lawrence, Halcyon M. “Siri Disciplines.” In *Your Computer Is on Fire*, edited by Thomas R. Mullaney, Benjamin Peters, Mar Hicks, and Kavita Philip, 179–97. MIT Press, 2021.

“Lyra: A New Very Low-Bitrate Codec for Speech Compression – Google AI Blog.” Accessed December 11, 2022.

<https://ai.googleblog.com/2021/02/lyra-new-very-low-bitrate-codec-for.html>.

Madrigal, Alexis C. “Almost Human: The Surreal, Cyborg Future of Telemarketing - The Atlantic.” *The Atlantic*, December 20, 2013.

<https://www.theatlantic.com/technology/archive/2013/12/almost-human-the-surreal-cyborg-future-of-telemarketing/282537/>.

Magellan Solutions. “The Business Process Outsourcing Industry (BPO) In The Philippines.” *Business*. Magellan Solutions, May 21, 2019.

<https://www.magellan-solutions.com/blog/status-bpo-in-the-philippines/>.

Mankarious, Sarah-Grace, and Marco Chacón. “Losing Languages, Losing Worlds.” Accessed December 11, 2022.

<https://www.cnn.com/interactive/2021/04/us/losing-languages/>.

Mark Hansen and Ben Rubin’s The Listening Post (2002-06) in Almost

Human, 2020. <https://www.youtube.com/watch?v=RzfnnddgfCk>.

- Marshall, Jim, and Laorenza, Tracey, eds. *The Resilient 9/11 Professional: A Comprehensive Guide to Surviving & Thriving Together in the 9/11 Center*. 1st ed. Petoskey: South of Haven Press, 2018.
- Marte-Wood, Alden Sajor. "Domestic Shifts: Reproducing Peripheral Realism in Philippine Call-Center Fiction." *Ariel: A Review of International English Literature* 53, no. 4 (October 2022): 1–39.
- Marte-Wood, Alden Sajor, and Stephanie Dimatula Santos. "Circuits of Care: Filipino Content Moderation and American Infostructures of Feeling." *Verge: Studies in Global Asias* 7, no. 2 (2021): 101–27.
<https://doi.org/10.1353/vrg.2021.0007>.
- Marvin, Carolyn. *When Old Technologies Were New: Thinking about Electric Communication in the Late Nineteenth Century*. 1. issued as an Oxford Univ. Press paperback, 9. print. New York, NY: Oxford Univ. Press, 1990.
- Mattern, Shannon. *Code and Clay, Data and Dirt: Five Thousand Years of Urban Media*. Minneapolis: University of Minnesota Press, 2017.
- McKittrick, Katherine. *Dear Science and Other Stories*. Errantries. Durham: Duke University Press, 2021.
- McPherson, Tara. "U.S. Operating Systems at Mid-Century: The Intertwining of Race and UNIX." In *Race after the Internet*, edited by

Lisa Nakamura and Peter Chow-White, 352. New York: Routledge, 2013.

McQuillan, Jeff, and Lucy Tse. “Child Language Brokering in Linguistic Minority Communities: Effects on Cultural Interaction, Cognition, and Literacy.” *Language and Education* 9, no. 3 (November 4, 2009): 195–215. <https://doi.org/10.1080/09500789509541413>.

Meet My Filipino Uncle // Jo Koy 2021, 2021.

<https://www.youtube.com/watch?v=vnJu7El2kAg>.

Meredith Whittaker [@mer__edith]. “AI’ Will Not Replace You. A Person Making Half What You Do with No Benefits Whose Job Is the Same as Yours Was but Now Includes Babysitting ‘AI’ Will.” Tweet. *Twitter*, January 5, 2023.

https://mobile.twitter.com/mer__edith/status/161129141234909184.

Metz, Cade. “‘The Godfather of A.I.’ Leaves Google and Warns of Danger Ahead.” *The New York Times*, May 1, 2023, sec. Technology. <https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html>.

Meyersohn, Nathaniel. “411 Is Going out of Service for Millions of Americans | CNN Business.” CNN, January 7, 2023.

<https://www.cnn.com/2023/01/07/business/telephone-operators-411-att-ctpr/index.html>.

Microsoft. "Microsoft 2006 Annual Report," 2006.

https://www.microsoft.com/investor/reports/aro6/staticversion/10k_dl_dow.html.

Migan, Darla. "A Poetics of the Glaze." Shift Space, 2022.

<https://www.shiftspace.pub/a-poetics-of-the-glaze>.

Mirchandani, Kiran. *Phone Clones: Authenticity Work in the Transnational Service Economy*. Ithaca: ILR Press, 2012.

Mirchandani, Kiran, and Winifred Poster, eds. *Borders in Service:*

Enactments of Nationhood in Transnational Call Centres. Toronto ; Buffalo ; London: University of Toronto Press, 2016.

Mitchell, Wade J, Kevin A Szerszen, Amy Shirong Lu, Paul W

Schermerhorn, Matthias Scheutz, and Karl F MacDorman. "A Mismatch in the Human Realism of Face and Voice Produces an Uncanny Valley." *I-Perception* 2, no. 1 (January 2011): 10–12.

<https://doi.org/10.1068/i0415>.

Moench, Mallory. "S.F.'s 911 Dispatch Struggling amid Staff Shortage:

'We Are Bleeding.'" San Francisco Chronicle, December 9, 2022.

<https://www.sfchronicle.com/sf/article/911-dispatch-staff-shortage-17641748.php>.

Morozov, Evgeny. *To Save Everything, Click Here: The Folly of Technological Solutionism*. First edition. New York: PublicAffairs, 2013.

Morton, David. *Off the Record: The Technology and Culture of Sound Recording in America*. New Brunswick: Rutgers University Press, 1999.

“Mozilla Common Voice.” Accessed January 16, 2023.

<https://commonvoice.mozilla.org/>.

Mozilla Foundation. “Dark Matters: New Project Spotlights the Inbuilt Bias in Digital Voice Assistants,” September 8, 2021.

<https://foundation.mozilla.org/en/blog/dark-matters-new-project-spotlights-the-inbuilt-bias-in-digital-voice-assistants/>.

Muñoz, José Esteban. *Cruising Utopia: The Then and There of Queer Futurity*. 10th Anniversary Edition. Sexual Cultures. New York: New York University Press, 2019.

Murphy, Brett. “They Called 911 for Help. Police and Prosecutors Used a New Junk Science to Decide They Were Liars.” ProPublica.

Accessed December 29, 2022.

<https://www.propublica.org/article/911-call-analysis-fbi-police-courts>.

Nakamura, Lisa. “Indigenous Circuits: Navajo Women and the Racialization of Early Electronic Manufacture.” *American Quarterly* 66, no. 4 (December 2014): 919–41. <https://doi.org/10.1353/aq.2014.0070>.

National Archives. “Chinese Exclusion Act (1882),” September 8, 2021.

<https://www.archives.gov/milestone-documents/chinese-exclusion-act>.

National Archives. “Executive Order 8802: Prohibition of Discrimination in the Defense Industry (1941),” September 22, 2021.

<https://www.archives.gov/milestone-documents/executive-order-8802>

NBC News. “Google Buys YouTube for \$1.65 Billion,” October 9, 2006.

<https://www.nbcnews.com/id/wbna15196982>.

“New in Descript: Transcription in 23 Languages! | Descript.” Accessed

March 26, 2023.

<https://www.descript.com/blog/article/new-in-descript-transcription-in-23-languages>.

- Nooney, Laine. "The Uncredited: Work, Women, and the Making of the U.S. Computer Game Industry." *Feminist Media Histories* 6, no. 1 (January 1, 2020): 119–46. <https://doi.org/10.1525/fmh.2020.6.1.119>.
- Nooney, Laine, and Tega Brain. "A 'Speculative Past's' Pedagogy: Where Speculative Design Meets Historical Thinking." *Digital Creativity* 30, no. 4 (October 31, 2019): 218–34. <https://doi.org/10.1080/14626268.2019.1683042>.
- Nowogrodzki, Anna. "Speaking in Code: How to Program by Voice." *Nature* 559, no. 7712 (July 3, 2018): 141–42. <https://doi.org/10.1038/d41586-018-05588-x>.
- Nugent, Addison. "Text-To-Speech in 1846 Involved a Talking Robotic Head With Ringlets." *Atlas Obscura*, March 9, 2016. <https://www.atlasobscura.com/articles/texttospeech-in-1846-involved-a-talking-robotic-head-with-ringlets>.
- Office of the Historian. "The Immigration Act of 1924 (The Johnson-Reed Act)." Government. Accessed January 15, 2023. <https://history.state.gov/milestones/1921-1936/immigration-act>.
- Oldenziel, Ruth. *Making Technology Masculine: Men, Women and Modern Machines in America, 1870-1945*. Amsterdam: Amsterdam University Press, 1999.

OpenAI. “ChatGPT: Optimizing Language Models for Dialogue,”

November 30, 2022. <https://openai.com/blog/chatgpt/>.

Original Talking Moose, 2014.

https://www.youtube.com/watch?v=iX_Uui4wLvI.

Padios, Jan M. *A Nation on the Line: Call Centers as Postcolonial*

Predicaments in the Philippines. Durham ; London: Duke University Press, 2018.

———. “Are Call Centers Rebranding the Philippines? | Essay.” *Zócalo*

Public Square (blog), August 15, 2018.

<https://www.zocalopublicsquare.org/2018/08/15/call-centers-rebranding-philippines/ideas/essay/>.

Peng, Xin. “Anna May Wong and Sessue Hayakawa: Racial Performance,

Ornamentalism, and Yellow Voices in *Daughter of the Dragon*

(1931).” *Camera Obscura: Feminism, Culture, and Media Studies* 37,

no. 2 (September 1, 2022): 1–29.

<https://doi.org/10.1215/02705346-9786986>.

peredaks. “Have Any of You Considered Moving to an Admin Position

Because You Were so Burnt Out?” Reddit Post. *R/911dispatchers*,

December 15, 2022.

www.reddit.com/r/guidispatchers/comments/zm7kq6/have_any_of_you_considered_moving_to_an_admin/.

Perry, Jade. "Study: Shopping with Alexa." Business. Wunderman Thompson, March 9, 2017.

<https://www.wundermanthompson.com/insight/study-shopping-alex>
a.

Petersen, Jennifer. *How Machines Came to Speak: Media Technologies and Freedom of Speech*. Sign, Storage, Transmission. Durham: Duke University Press, 2022.

Phan, Thao. "Amazon Echo and the Aesthetics of Whiteness." *Catalyst: Feminism, Theory, Technoscience* 5, no. 1 (April 1, 2019): 1–38.
<https://doi.org/10.28968/cftt.v5i1.29586>.

Pieraccini, Roberto. *AI Assistants*. The MIT Press Essential Knowledge Series. Cambridge, Massachusetts: The MIT Press, 2021.

Poster, Winifred R. "Who's On the Line? Indian Call Center Agents Pose as Americans for U.S.-Outsourced Firms." *Industrial Relations* 46, no. 2 (April 2007): 271–304.
<https://doi.org/10.1111/j.1468-232X.2007.00468.x>.

Prestige Accents. "SPEECH SAMPLES." Accessed January 18, 2023.
<https://www.prestigeaccents.com/speech-samples>.

“Programmable Voice: Voice API for Web and Mobile.” Accessed March 9, 2023.

https://www.twilio.com/docs/voice?utm_source=docs&utm_medium=social&utm_campaign=guides_tags.

“Programming by Voice May Be the Next Frontier in Software

Development - IEEE Spectrum.” Accessed December 29, 2022.

<https://spectrum.ieee.org/programming-by-voice-may-be-the-next-frontier-in-software-development>.

“Protocols and Policies for Emergency Medical Services | San Francisco.”

Accessed December 4, 2022.

<https://sf.gov/information/protocols-and-policies-emergency-medical-services>.

Rachelle Cruz - Sunday Jump 10/07/2018, 2018.

<https://www.youtube.com/watch?v=-5eg4WwgRMw>.

Rangan, Pooja, Akshya Saxena, Ragini Tharoor Srinivasan, and Pavitra

Sundar, eds. *Thinking with an Accent: Toward a New Object, Method, and Practice*. University of California Press, 2023.

<https://doi.org/10.1525/luminos.148>.

Rankin, Joy Lisi. *The People’s History of Computing*. Cambridge: Harvard

University Press, 2018.

reddit. “R/guidispatchers.” Accessed April 30, 2023.

<https://www.reddit.com/r/guidispatchers/>.

Robertson, Lindsay. “Does Your Kid Think Your Smart Speaker Is Just Another Family Member?” *Mashable*, November 30, 2019.

<https://mashable.com/article/kids-and-smart-speakers>.

Robinson, John J., Jim Maddock, and Kate Starbird. “Examining the Role of Human and Technical Infrastructure during Emergency Response.” Presented at the International Conference on Information Systems for Crisis Response and Management, 2015.

Roh, David S., Betsy Huang, and Greta A. Niu, eds. *Techno-Orientalism: Imagining Asia in Speculative Fiction, History, and Media*. Asian American Studies Today. New Brunswick, New Jersey: Rutgers University Press, 2015.

Rosner, Helen. “The Ethics of a Deepfake Anthony Bourdain Voice.” *The New Yorker*, July 17, 2021.

<https://www.newyorker.com/culture/annals-of-gastronomy/the-ethics-of-a-deepfake-anthony-bourdain-voice>.

———. “The Ethics of a Deepfake Anthony Bourdain Voice in ‘Roadrunner.’” *The New Yorker*, July 17, 2021.

<https://www.newyorker.com/culture/annals-of-gastronomy/the-ethic-s-of-a-deepfake-anthony-bourdain-voice>.

Ruberg, Bo. “What Is Your Mother’s Maiden Name? A Feminist History of Online Security Questions.” *Feminist Media Histories* 3, no. 3 (Summer 2017): 57–81. <https://doi.org/10.1525/fmh.2017.3.3.57>.

Salonga, Aileen O. “Serving the World, Serving the Nation: Everyday Nationalism and English in Philippine Offshore Call Centres.” In *Borders in Service: Enactments of Nationhood in Transnational Call Centres*, edited by Kiran Mirchandani and Winifred Poster, 289. Canada: University of Toronto Press, 2016.

Samantha. “Techno Solutionism—Very Few Things Actually Need to Be an App.” Digital Rights Watch, March 24, 2021. <https://digitalrightswatch.org.au/2021/03/25/technosolutionism/>.

San Francisco 911 Dispatcher Recruitment Video, 2017. <https://www.youtube.com/watch?v=oLoUjV7xX1k>.

“Sanas | Breaking Barriers - One Conversation At A Time.” Accessed March 26, 2023. <https://www.sanas.ai/>.

Sassen, Saskia. *Territory, Authority, Rights: From Medieval to Global Assemblages*. Updated ed., 4. print., 1. paperback print. Princeton, N.J.: Princeton Univ. Press, 2008.

“ScanRad.Io.” Accessed December 11, 2022. <https://scanrad.io/>.

Scarry, Elaine. *The Body in Pain: The Making and Unmaking of the World*.

New York: Oxford University Press, 1985.

Scheper-Hughes, Nancy, and Margaret M. Lock. “The Mindful Body: A

Prolegomenon to Future Work in Medical Anthropology.” *Medical*

Anthropology Quarterly 1, no. 1 (March 1987): 6–41.

<https://doi.org/10.1525/maq.1987.1.1.02a00020>.

Sconce, Jeffrey. *Haunted Media: Electronic Presence from Telegraphy to*

Television. Durham: Duke University Press, 2000.

See, Sarita Echavez. *The Decolonized Eye: Filipino American Art and*

Performance. Minneapolis: University of Minnesota Press, 2009.

Semel, Beth M. “Listening Like a Computer: Attentional Tensions and

Mechanized Care in Psychiatric Digital Phenotyping.” *Science,*

Technology, & Human Values 47, no. 2 (March 2022): 266–90.

<https://doi.org/10.1177/01622439211026371>.

Shearer, Elisa. “More than Eight-in-Ten Americans Get News from

Digital Devices.” *Pew Research Center* (blog). Accessed November 11,

2022.

<https://www.pewresearch.org/fact-tank/2021/01/12/more-than-eight-in>

[-ten-americans-get-news-from-digital-devices/](https://www.pewresearch.org/fact-tank/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/).

“SICK INTERNET | Mac Andre Arboleda.” Accessed May 1, 2023.

<https://sickinternet.me>.

“Slick Tom Cruise Deepfakes Signal That Near Flawless Forgeries May Be Here.” *All Things Considered*. NPR, March 11, 2021.

<https://www.npr.org/2021/03/11/975849508/slick-tom-cruise-deepfakes-signal-that-near-flawless-forgeries-may-be-here>.

Smith, Linda Tuhiwai. *Decolonizing Methodologies: Research and Indigenous Peoples*. First published in Great Britain 2021, Third edition. London: Zed, 2021.

Smith, Roberta. “ART IN REVIEW; Mark Hansen and Ben Rubin --

‘Listening Post.’” *The New York Times*, February 21, 2003, sec. Arts.

<https://www.nytimes.com/2003/02/21/arts/art-in-review-mark-hansen-and-ben-rubin-listening-post.html>.

Society, Data &. “Searching for Parables of AI.” Medium, December 8, 2022.

<https://points.datasociety.net/searching-for-parables-of-ai-3317d201edeb>.

South Sound 911. “Text-to-911.” Accessed November 11, 2022.

<https://southsound911.org/911-dispatch/text911/>.

“Speech Synthesis Programming Guide.” Accessed April 26, 2023.

https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SpeechSynthesisProgrammingGuide/SpeechOverview/SpeechOverview.html#//apple_ref/doc/uid/TP40004365-CH3-SW1.

SRI International. “SRI International.” Company Site. *SRI International*

(blog). Accessed September 30, 2022. <https://www.sri.com/hoi/siri/>.

Sterne, Jonathan. *Diminished Faculties: A Political Phenomenology of*

Impairment. Durham: Duke University Press, 2021.

———. *Diminished Faculties: A Political Phenomenology of Impairment*.

Durham: Duke University Press, 2021.

———. *The Audible Past: Cultural Origins of Sound Reproduction*. Durham:

Duke University Press, 2003.

———. *The Audible Past: Cultural Origins of Sound Reproduction*. Duke

University Press, 2003.

Sterne, Jonathan, and Mehak Sawhney. “The Acousmatic Question and

the Will to Datafy: Otter.Ai, Low-Resource Languages, and the

Politics of Machine Listening.” *Kalfou: A Journal of Comparative*

and Relational Ethnic Studies 9, no. 2 (January 5, 2023): 288–306.

Stochastic Labs. “CRISPRartists.” Accessed May 1, 2023.

<http://stochasticlabs.org/crisprartists/>.

stoplapdspying. “The Algorithmic Ecology: An Abolitionist Tool for

Organizing Against Algorithms.” *Medium* (blog), March 3, 2020.

<https://stoplapdspying.medium.com/the-algorithmic-ecology-an-abolitionist-tool-for-organizing-against-algorithms-14fcbdoe64do>.

Stubbs, Katherine. “Telegraphy’s Corporeal Fictions.” In *New Media,*

1740-1915, edited by Geoffrey B. Pingree and Lisa Gitelman.

Cambridge: The MIT Press, 2003.

Sutter, John D. “The Top 10 Tech Trends of 2009.” News. cnn.com,

December 23, 2009.

<http://www.cnn.com/2009/TECH/12/22/top.tech.trends.2009/index.html>.

Swan, Rachel. “‘I Hope Nobody Is Bleeding out’: They Called 911 in

Oakland. They Were Told They’d Have to Wait.” *San Francisco*

Chronicle, May 11, 2022.

<https://www.sfchronicle.com/crime/article/I-hope-nobody-is-bleeding-out-They-called-17166166.php>.

“Techniques for Customizing Synthesized Speech.” Accessed January 4,

2023.

https://developer.apple.com/library/archive/documentation/UserExperience/Conceptual/SpeechSynthesisProgrammingGuide/FineTuning/FineTuning.html#//apple_ref/doc/uid/TP40004365-CH5-SW7.

“The AT&T Archives and History Center: An Interview with Sheldon Hochheiser | Sound & Science: Digital Histories.” Accessed January 11, 2023.

<https://soundandscience.de/contributor-essays/att-archives-and-history-center-interview-sheldon-hochheiser>.

The Drum. “Amazon: Alexa’s Body By Lucky Generals.” Accessed January 2, 2023.

<https://www.thedrum.com/creative-works/project/lucky-generals-amazon-alexas-body>.

The Forum // DeForrest Brown & Ryan Kuo, 2021.

<https://www.youtube.com/watch?v=3jvFoRV8Dpo>.

The Guggenheim Museums and Foundation. “Sophie Calle.” Accessed March 26, 2023.

<https://www.guggenheim.org/artwork/artist/sophie-calle>.

The Lost 1984 Video: Young Steve Jobs Introduces the Macintosh, 2009.

<https://www.youtube.com/watch?v=2B-XwPjngYY>.

The Story of MacinTalk Part 1, 2016.

<https://www.youtube.com/watch?v=UuVo4MHTEQo>.

“This Artist-Made App Is Bringing Back Telephone Culture, One Call at a Time | KQED.” Accessed January 16, 2023.

<https://www.kqed.org/arts/13862818/this-artist-made-app-is-bringing-back-telephone-culture-one-call-at-a-time>.

Thornton, Pip. “A Critique of Linguistic Capitalism:

Provocation/Intervention.” *GeoHumanities* 4, no. 2 (July 3, 2018):

417–37. <https://doi.org/10.1080/2373566X.2018.1486724>.

“Thrifty Speaks to Its Customers: Car Rental Agency Deploys Speech

Recognition to Improve Customer Experience While Reducing

Costs - Document - Gale Power Search.” Accessed January 6, 2023.

https://go.gale.com/ps/i.do?id=GALE%7CA91965840&v=2.1&u=mli_n_b_northeast&it=r&p=GPS&sw=w&asid=be738of80ca32191c667d3be4d820345.

throwaway1976_. “I’ve Been at This Job for 3ish Months. Tragic Things I

Hear Don’t Affect Me.” Reddit Post. *R/91dispatchers*, March 16,

2023.

www.reddit.com/r/91dispatchers/comments/11so0zb/ive_been_at_this_job_for_3ish_months_tragic/.

Training for Service, 1926 - AT&T Archives - Telephone Operator

Recruiting. Video. Bell System, 1926.

<https://www.youtube.com/watch?v=QQbwRYCYJzI&t=2s>.

Turow, Joseph. *The Voice Catchers: How Marketers Listen in to Exploit*

Your Feelings, Your Privacy, and Your Wallet. New Haven: Yale

University Press, 2021.

“User Agreement - April 18, 2023 - Reddit.” Accessed April 30, 2023.

<https://www.redditinc.com/policies/user-agreement>.

Vanian, Jonathan. “Twitter Is Now Owned by Elon Musk – Here’s a

Brief History from the App’s Founding in 2006 to the Present.”

CNBC, October 29, 2022.

<https://www.cnn.com/2022/10/29/a-brief-history-of-twitter-from-its-founding-in-2006-to-musk-takeover.html>.

Vasudevan, Roopa. “Share and Share Unlike: Reciprocity, Corporate

Power, and the Open Source Ethos.” *New Media & Society*, August

18, 2021, 146144482110383. <https://doi.org/10.1177/14614448211038350>.

Vaughan, John. “ATT Interactive Voice Response | UX | Portfolio | IVR,

Call Centers, Natural Language, Tech Evaluation.” Personal

Website. The Communication Studio. Accessed October 11, 2022.

<http://jcvts.com/portfolio/att-ivr.html>.

“Voice Cloning Technology May Have Been Used in Houston-Area Scam | Khou.Com.” Accessed March 26, 2023.

<https://www.khou.com/article/news/local/pearland-father-daughter-voice-cloning-scam/285-1a7boegb-ccca-488c-afff-04dcecodie77>.

Vora, Kalindi. *Life Support: Biocapital and the New History of Outsourced Labor*. Minneapolis: University of Minnesota Press, 2015.

“Wake Word,” April 26, 2023.

<https://dictionary.cambridge.org/us/dictionary/english/wake-word>.

Waltham-Smith, Naomi. *Shattering Biopolitics: Militant Listening and the Sound of Life*. First edition. Commonalities. New York: Fordham University Press, 2021.

Wernimont, Jacqueline. *Numbered Lives: Life and Death in Quantum Media*. Cambridge: The MIT Press, 2019.

Western, Tom. “Listening with Displacement.” *Migration and Society* 3, no. 1 (June 1, 2020): 194–309. <https://doi.org/10.3167/arms.2020.030128>.

“What Do You Want Me to Say? – Lauren Lee McCarthy.” Accessed January 4, 2023.

<https://lauren-mccarthy.com/What-do-you-want-me-to-say>.

“What Is Interactive Voice Response (IVR)? | IBM.” Accessed April 29, 2023. <https://www.ibm.com/topics/interactive-voice-response>.

“What Is Speech Recognition? | IBM.” Accessed December 29, 2022.

<https://www.ibm.com/topics/speech-recognition>.

WHITE SUPREMACY CULTURE. “WHITE SUPREMACY CULTURE.” Accessed December 11, 2022.

<https://www.whitesupremacyculture.info/>.

Wikitongues. “Wikitongues.” Accessed November 7, 2022.

<https://wikitongues.org/>.

Winner, Langdon. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Second Edition. Chicago: Univ. of Chicago Press, 2001.

“Work Life Harmony – Emily Shanahan Portfolio.” Accessed April 2, 2023. <https://emilyshanahan.com/Work-Life-Harmony>.

Yu, Charles. “Standard Loneliness Package.” *Lightspeed: Science Fiction & Fantasy*, November 2010.

<https://www.lightspeedmagazine.com/fiction/standard-loneliness-package/>.

Zeavin, Hannah. *The Distance Cure: A History of Teletherapy*. Cambridge, Massachusetts: The MIT Press, 2021.