"Hey Google, Talk Like Issa": Black Voiced Digital Assistants and the Reshaping of Racial Labor

goldenowens

Blog post published June 5, 2023 in *Sounding Out!*
https://doi.org/10.59350/9cbp8-49610

Abstract

. In summer 2021, sound artist, engineer, musician, and educator **Johann Diedrick** convened a panel at the intersection of racial bias, listening, and AI technology at Pioneerworks in Brooklyn, NY. Diedrick, 2021 Mozilla Creative Media award recipient and creator of such works as *Dark Matters* , is currently working on identifying the origins of racial bias in voice interface systems. *Dark Matters,*

Keywords

Music (Subfield), Digital Economy and Work Transformation (Topic), Sociology and Political Science (Subfield), Advertising, American Studies, Archival, Article, Artificial Intelligence



Recommended citation

goldenowens. (2023, June 5). "Hey Google, Talk Like Issa": Black Voiced Digital Assistants and the Reshaping of Racial Labor. *Sounding Out!*. https://doi.org/10.59350/9cbp8-49610

Copyright

Copyright © goldenowens 2023. Distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



.



In summer 2021, sound artist, engineer, musician, and educator **Johann Diedrick** convened a panel at the intersection of racial bias, listening, and AI technology at Pioneerworks in Brooklyn, NY. Diedrick, 2021 Mozilla Creative Media award recipient and creator of such works as *Dark Matters,* is currently working on identifying the origins of racial bias in voice interface systems. *Dark Matters,* according to Squeaky Wheel, "exposes the absence of Black speech in the datasets used to train voice interface systems in consumer artificial intelligence products such as Alexa and Siri. Utilizing 3D modeling, sound, and storytelling, the project challenges our communities to grapple with racism and inequity through speech and the spoken word, and how AI systems underserve Black communities." And now, he's working with SO! as guest editor for this series (along with ed-in-chief JS!). It kicked off with **Amina Abbas-Nazari's post**, helping us to understand how Speech AI

systems operate from a very limiting set of assumptions about the human voice. Today **Golden Owens** explored what happens when companies sell Black voices along with their Intelligent Virtual Assistants. Tune in for a deep historical dive into the racialized sound of servitude in America. Even though corporations aren't trying to hear this absolutely critical information–or Black users in general–they better listen up. **–JS**

—

In October 2019, Google released an ad for their Google Assistant (GA), an intelligent virtual assistant (IVA) that initially debuted in 2016. As revealed by onscreen text and the video's caption, the ad's announced that the GA would soon have a new celebrity voice. The ten-second promotion includes a soundbite from this unseen celebrity—who states: "You can still call me your Google Assistant. Now I just sound extra fly"— followed by audio of the speaker's laughter, a white screen, the GA logo, and a written question: "Can you guess who it is?"

Consumers quickly speculated about the person behind the voice, with many posting their guesses on Reddit. The earliest comments named Tiffany Haddish, Lizzo, and Issa Rae as prospects, with other users affirming these guesses. These women were considered the most popular contenders: two articles written about the new GA voice cited the Reddit post, with one calling these women Redditors' most popular guesses and the other naming only them as users' desired choices. Those who guessed Rae were proven correct. One day after the ad, Google released a longer promo revealing her as the GA's new voice, including footage of Rae recording responses for the assistant. The ad ends with Rae repeating the "extra fly" line from the initial promo, smiling into the camera.



https://youtu.be/JVjUpoBSewo

Google's addition of Rae as an IVA voice option is one of several recent examples of Black people's voices employed in this manner. Importantly, this trend toward Black-voiced IVAs deviates from the pre-established standard of these digital aides. While there are many voice options available, the default voices for IVAs are white female voices with flat dialects. This shift toward Black American voices is notable not only because of conversations about inclusion—with

some Black users saying they feel more represented by these new voices—but because this influx of Black voices marks a spiritual return to the historical employment of Black people as service-providing, labor-performing entities in the United States, thus subliminally reinforcing historical biases about Black people as uniquely suited for performing this type of work.

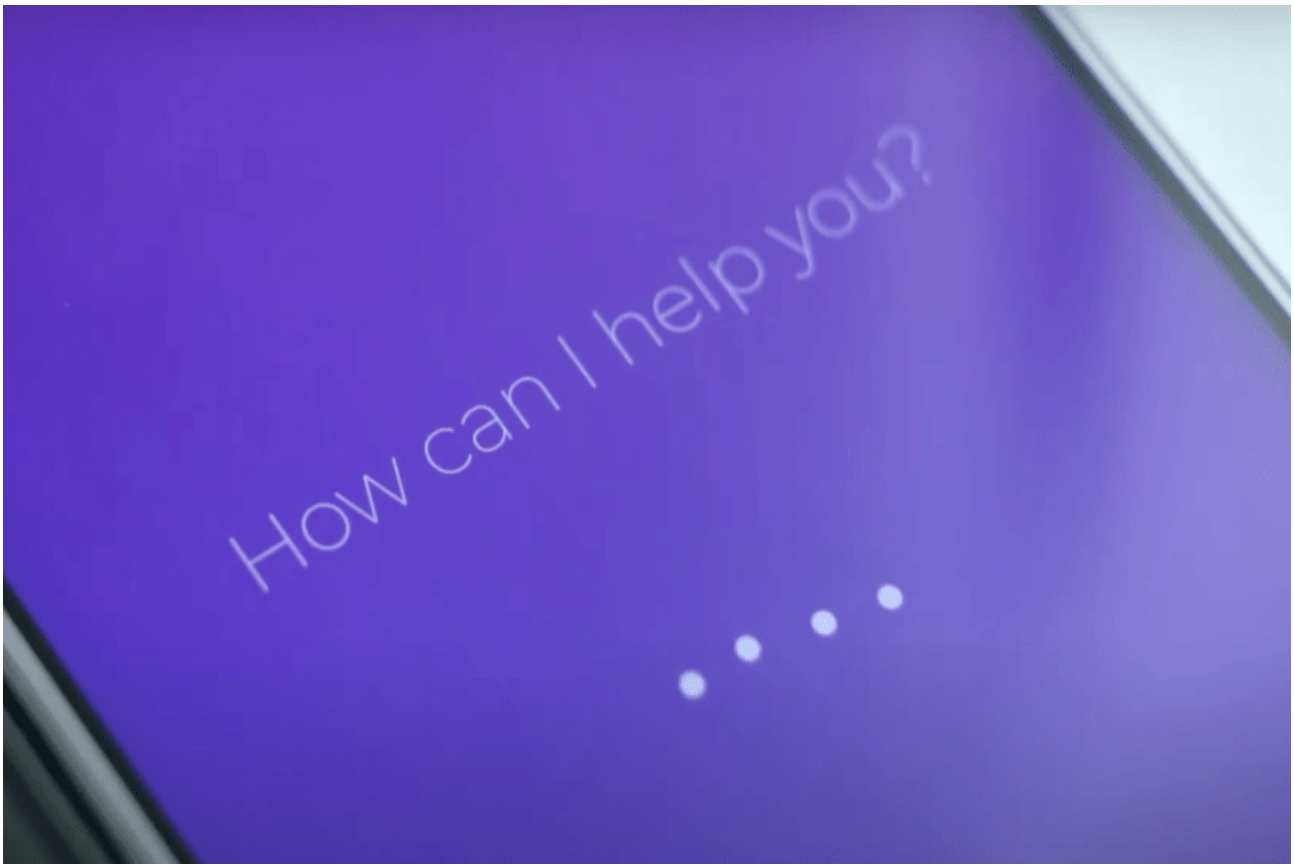
Marketed as labor-saving devices, IVAs are programmed to assist with cooking and grocery shopping, transmit messages and reminders, and provide entertainment, among other tasks. Since the late 2010s they have also been able to operate other technologies within users' homes: Alexa, for example, can control Roomba robotic vacuums; IVA-compatible smart plugs or smart home devices enable IVAs to control lights, locks, thermostats, and other such apparatuses. Behaviorally, IVAs are designed and expected to be on-call at all times, but not to speak or act out of turn—with programmers often directed to ensure these aides are relatable, reliable, trustworthy, and unobtrusive.



Round Grey Speaker On Brown Board, gadget, google assistant, google home (public domain)

Far from operating in a vacuum, IVAs eerily evoke the presence of and parameters set for enslaved workers and domestic servants in the U.S.—many of whom have historically been Black American women. Like IVAs, Black women servants cooked, cleaned, entertained children, and otherwise served their (predominantly white) employers, themselves operating as labor-saving devices through their performance of these labors. Employers similarly expected these women to be ever-available, occupy specific areas of the home, and obey all requests and demands—and were unsettled if not infuriated when maids did not behave according to their expectations.

White women being the default voices of IVAs has somewhat obfuscated the degree to which these aides have re-embodied and replaced the Black servants who once predominantly executed this work, but incorporating Black voices into these roles removes this veil, symbolically re-implementing Black people as labor-performing entities by having them operate as the virtual assistants who now perform much of the labor Black workers historically performed. Enabling Black people to be used as IVAs thus re-aligns Black beings with the performance of service and labor.



Still image from Mozilla Explains: Bias in AI Training Data

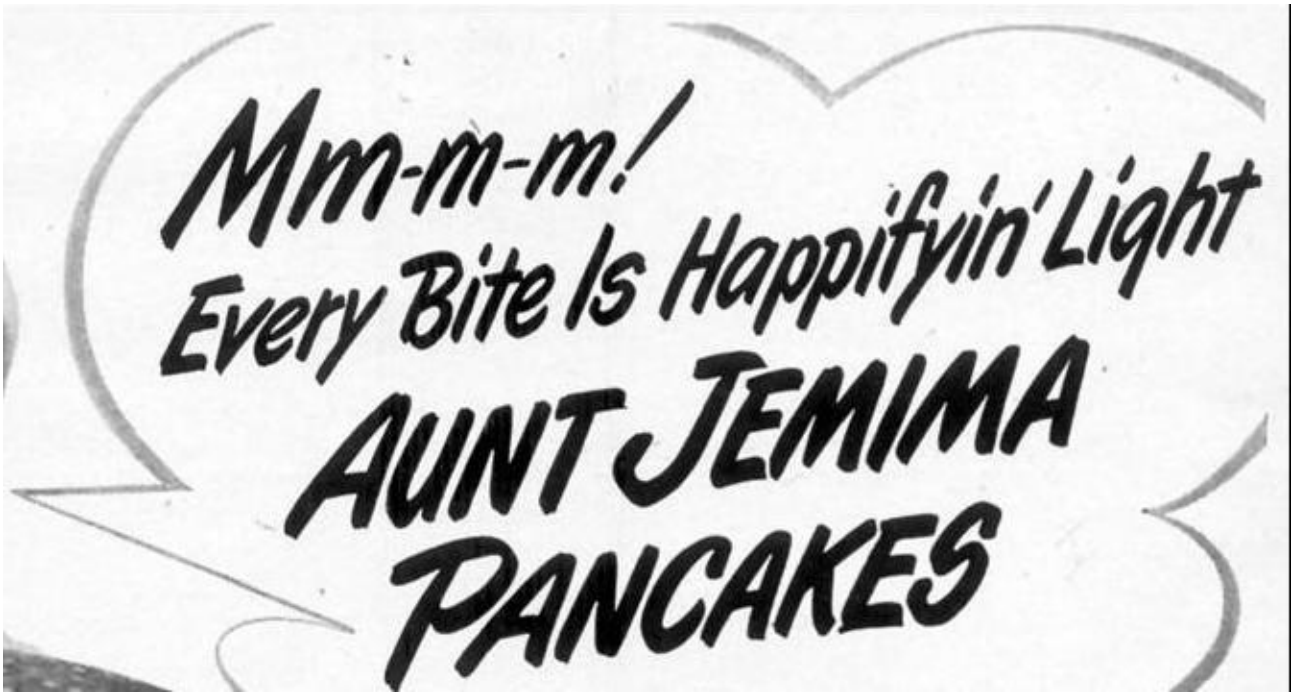
While Black women were far from the only demographic conscripted into domestic labor, by the 1920s they comprised a "permanent pool of servants" throughout the country, due largely to the egress of white American and immigrant women from domestic service into fields that excluded Black women (183). Black women's prominence in domestic service was heavily reflected in early U.S. media, which overwhelmingly portrayed domestic servants not just as Black women, but as Black Mammies—domestic servant archetypes originally created to promote the myth that Black women "were contented, even happy as slaves." Characters like *Gone with the Wind*'s "Mammy" pulled both from then-current associations of Black women with domestic labor and from white nostalgia for the Antebellum era, and specifically for the archetypal Mammy—marking Black women as idealized labor-performing domestics operating in service of white employers. These on-screen servants were "always around when the boss needed them...[and] always ready to lend a helping hand when times were tough" (36). Historian Donald Bogle dubbed this era of Hollywood the "Age of the Negro Servant," referenced in this reel from the *New York Times*.

—-

.

I

Cinema and television merely built from years of audible racism on the radio—America's most prominent form of in-home entertainment in the first half of the 20th century—where Black actors also played largely servant and maid roles that demanded they speak in "distorted dialect, exaggerated intonation, rhythmic speech cadences, and particular musical instruments" in order to appear at all (143). This white-contrived portrayal of Black people is known as "Blackvoice," and essentially functions as "the minstrel show boiled down to pure aurality" (14). These performances allowed familiar ideals of and narratives about Blackness to be communicated and recirculated on a national scale, even without the presence of Black bodies. Labor-performing Black characters like Beulah, Molasses and January, Aunt Jemima, and Amos and Andy were prominent in the Golden Age of Radio, all initially voiced by white actors. In fact, Aunt Jemima's print advertising was just as dependent on stereotypical representations of her voice as it was on visual "Mammy" imagery.



Close up of Aunt Jemima advertising appearing in Woman's Day in 1948.

When Black actors broke through white exclusion on the airwaves, many took over roles once voiced by white men and/or were forced by white radio producers and scriptwriters to "'talk as white people believed Negroes talked'" so that white audiences could discern them as Black (371). This continuous realignment undoubtedly informs contemporary ideas of labor, labor performance, and laboring bodies, further promoted by the sudden influx of Black voice assistants in 2019.

Specifically, these similarities demonstrate that contemporary IVAs are intrinsically haunted by Black women slaves and servants: built in accordance with and thus inevitably evoking these laborers in their positioning, programming, and task performance. Further facilitating this

alignment is the fact that advertisements for Black-voiced IVAs purposefully link well-known Black bodies in conjunction with their Black voices. Excepting Apple's Black-sounding voice options for Siri, all of the Black IVA voice options since 2019 have belonged to prominent Black American celebrities. Prior to Issa Rae, GA users could employ John Legend as their digital aide (April 2019 until March 2020). Samuel L. Jackson became the first celebrity voice option for Amazon's Alexa in December 2019, followed by Shaquille O'Neal in July 2021.



https://youtu.be/nstvu1duiHs

The ads for Black-voiced IVAs thus link these disembodied aides not just to Black bodies, but to specific Black bodies as a sales tactic—bodies which signify particular images and embodiments of Blackness. The Samuel L. Jackson Alexa ad utilizes close-ups of Jackson recording lines for the IVA and of Echo speakers with Jackson's voice emitting from them in response to users. John Legend is physically absent from the ad announcing him as the GA; however, his celebrity wife directs the GA to sing for her instead, after which she states that it is "just like the real John"—thus linking Legend's body to the GA even without his onscreen presence. Amazon has even explicitly explored the connection between the Black-voiced IVA and the Black body, releasing a 2021 commercial called "Alexa's Body" that saw Alexa voiced and physically embodied by Michael B. Jordan—with the main character in the commercial insinuating that he is the ideal vessel for Alexa.



https://youtu.be/Zd1HkPTZSqQ

By aligning these bodies with, and having them act as, labor-performing devices in service of consumers, these advertisements both re-align Blackness with labor and illuminate how these devices were always already haunted by laboring Black bodies—and especially, given the demographics of the bodies who most performed the types of labors IVAs now execute, laboring Black *women*'s bodies. That the majority of the Black celebrities employed as Black IVA voices are men suggests some awareness of and attempt to distance from this history and implicit haunting—an effort which itself exposes and illuminates the degree to which this haunting exists.

In some cases, the Black people lending their voices to these IVAs also speak in a way that sonically suggests Blackness: Issa Rae's "Now I'm just extra fly," for example, incorporates Black American slang through the use of the word "fly. As part of African American Vernacular English (AAVE), the term "fly" dates back to the 1970s and denotes coolness, attractiveness, and fashionableness. Because of its inclusion in Hip Hop, which has become the dominant music genre in the United States, the term, its meaning, and its racial origins are widely known amongst consumers. By using the word "fly," Rae nods not only at these qualities but also at her own Blackness in a manner that is recognizable to a mainstream American audience. Due in part to Hip Hop's popularity, U.S.-based media outlets, corporations, and individuals of varying races and ethnicities regularly appropriate AAVE and Black slang terms, often without regard for the culture that created them or the vernacular they stem from. The ad preceding Issa Rae's revelation as the GA specifically invited users to align the voice with a celebrity body, and users' predominant claims that the voice was a Black woman's suggest that something about the voice conjured Blackness and the Black female body.



"Alexa Voice" by Stock Catalog, (CC BY 2.0)

This racial marking was also likely facilitated by how people naturally listen and respond to voices. As Nina Sun Eidsheim notes in *The Race of Sound*, "voices heard are ultimately identified, recognized, and named by listeners at large. In hearing a voice, one also brings forth a series of assumptions about the nature of voice" (12). This series of assumptions, Eidsheim asserts in "The Voice as Action," is inflected by the "multisensory context" surrounding a given voice, i.e., "a composite of visual, textural, discursive, and other kinds of information" (9). While we imagine our impressions of voices as uniquely meaningful, "we cannot but perceive [them] through filters generated by our own preconceptions" (10). As a result, listening is never a neutral or truly objective practice.

For many consumers, these filters are informed by what Jennifer Lynn Stoever terms the sonic color line, "a socially constructed boundary that racially codes sonic phenomena such as vocal timbre, accents, and musical tones" (11). Where the racial color line allows white people to separate themselves from Black people on the basis of visual and behavioral differences, the sonic color line allows people "to construct and discern racial identities based on voices, sounds, and particular soundscapes" and to assign nonwhite voices with "differential cultural, social, and political value" (11). In the U.S., the sonic color line operates in tandem with the American listening ear, which "normalizes the aural tastes and standards of white elite masculinity as the singular way to interpret sonic information" (13) and therefore marks-as-Other not only the voices and bodies of Black people, but also those of non-males and the non-elite.



Voice bubble from 1940's print ad for Aunt Jemima Pancake mix: the sonic color line in sight and sound.

Ironically, the very listening practices which make consumers register particular voices and vocal qualities as Black also make Black voices inaccessible to Alexa and other IVAs. Scholarship on Automated Speech Recognition (ASR) systems and Speech AI observes that many Black users find it necessary to code-switch when speaking to IVAs, as the devices fail to comprehend their linguistic specificities. A study by Christina N. Harrington et al. in which Black elders used the Google Home to seek health information discovered that "participants felt that Google Home struggles to understand their communication style (e.g., diction or accent) and language (e.g., dialect) specifically due to the device being based on Standard English" (15). To address these struggles, participants switched to Standard American English (SAE), eliminating informal contractions and changing their tone and verbiage so that the GA would understand them. As one of the study's participants states,

> *You do have to change your words. Yes. You do have to change your diction and yes, you have to use... It cannot be an exotic name or a name that's out of the Caucasian round. ...You have to be very clear with the English language. No ebonic (15).*

This incomprehension extends to Black Americans of all ages, and to other IVAs. A study by Allison Koenecke et al. on ASR systems produced by Amazon, Google, IBM, Microsoft and Apple discovered that these entities had a harder time accurately transcribing Black speech than white speech, producing "an average word error rate (WER) of 0.35 for black speakers compared with 0.19 for white speakers." (7684). A study by Zion Mengesha et al. on the impact of these errors on Black Americans—which included participants from different regions with a range of ages, genders, socioeconomic backgrounds and education-levels—discovered that many felt frustrated and othered by these mistakes, and felt further pressure to code-switch so that they would not be misunderstood. Koenecke et al. concluded that ASR systems could not understand the "phonological, phonetic, or prosodic characteristics of" AAVE (7687), and that this ignorance would make the use of these technologies more difficult for Black users—a sentiment that was echoed by participants in the study conducted by Mengesha et al., most of whom marked the technology as working better for white and/or SAE speakers (5).



The speech recognition errors these technologies demonstrate—which also extend to speakers in other racial and ethnic groups—illuminate the reality that despite including Black voices as IVAs, these assistant technologies are not truly built for Black people, or for any person that does not speak Standard American English. And where AAVE is largely associated with Blackness, SAE is predominantly associated with whiteness: as a dialect widely perceived to be "lacking any distinctly regional, ethnic, or socioeconomic characteristics," it is recognized as being "spoken by the majority group or the socially advantaged group" in the United States—both groups which are solely or primarily composed of white people. SAE is so identified with whiteness that Black people who only speak Standard English are often told that they sound and/or "talk" white, and Black people who deliberately invoke SAE in professional and/or interracial settings (i.e., code switching) are described as "talking white" or using their "white voice" when doing so. That IVAs and other ASR systems have such trouble understanding AAVE and other non-standard English dialects suggests that these technologies were not designed to understand any dialect other than SAE—and thus, given SAE's strong identification with whiteness, were designed specifically to assist, understand, and speak to white users.

Writing on this phenomena as a woman with a non-standard accent, Sinduja Rangarajan highlights in "Hey Siri—Why Don't You Understand More People Like Me?" that none of the IVAs currently on the market offer any American dialect that is not SAE. And while users can change their IVA's accents, they are limited to Standard American, British, Irish, Australian, Indian, and South African—which Rangarajan rightly highlights as revealing who the IVAs *think* they are talking to, rather than who their user actually is. That most of these accents belong to Western, predominantly white countries (or to countries once colonized by white imperialists) strongly

suggests that these devices are programmed to speak to—and perform labor for—white consumers specifically.



"Voice is Already Big": Adobe Sayspring Founder Mark Christopher Webster Presents At Entrepreneurs Roundtable Accelerator Demo Day in April 2017 (CC BY-SA 4.0)

When considering the primary imagined and target users of IVAs, the sudden influx of Black-voiced IVAs becomes particularly insidious. Though they may indeed make some Black users feel more represented, cultivating this representation is merely a byproduct of their actual purpose. Because these technologies are not built for Black consumers, Black-voiced IVAs are meant to appeal not to Black users, but to white ones. Rae, Jackson, and the other Black celebrity voices may provide a much-needed variety in the types of voices applied to IVAs, but they primarily operate as "further examples of technology companies using Black voices to entertain white consumers while ignoring Black consumers." Black-voiced assistants, after all, no better understand Black vernacular English than any of the other voice options for IVAs, a reality marking Black speech patterns as enjoyable but not legitimate.

By excluding Black consumers, the companies behind these IVAs insinuate that Blackness is only acceptable and worthy of consideration when operating in service of whiteness. Where Black people as consumers have been delegitimized and disregarded, Black voices as labor-saving assistants have been welcomed and deemed profitable—a reality which further emphasizes how historical constructions of Black people as labor-performing devices haunts these contemporary technologies. Tech companies reinforce historical positionings of white people as ideal consumers and Black people as consumable products—repeating historical demarcations of Blackness and whiteness in the present.

In imagining the futures of IVAs, the companies behind them would need to reconsider how they interact—or fail to interact—with Black users. Both Samuel L. Jackson and Shaquille O'Neal, the last of the Black-celebrity-voiced IVAs still currently available to users, will be removed as Alexa voice options by September 2023, presenting an opportunity for these companies to divest. Whether or not the brands behind these IVAs take this initiative, consumers themselves can be critical of how

AI technologies continue to reestablish hierarchical systems, of their own interactions with these devices, and of who these technologies are truly made for. In being critical, we can perhaps begin to envision alternative, reparative modes of AI technology—modes that serve and support more than one kind of user.

—

Featured Image: Issa Rae gif from the 2017 Golden Globes

—

Golden Marie Owens *is a PhD candidate in the Screen Cultures program at Northwestern University. Her research interests include representations of race and gender in American media and popular culture, artificial intelligence, and racialized sounds. Her doctoral dissertation, "Mechanical Maids: Digital Assistants, Domestic Spaces, and the Spectre(s) of Black Women's Labor," examines how intelligent virtual assistants such as Apple's Siri and Amazon's Alexa evoke and are haunted by Black women slaves, servants, and houseworkers in the United States. In her time at Northwestern, she has had internal fellowships through the Office of Fellowships and the Alice Kaplan Institute for the Humanities. She currently holds an MMUF Dissertation Grant through the Institute for Citizens and Scholars and Ford Dissertation Fellowship through the National Academy for Sciences, Engineering, and Medicine.*

—



REWIND! . . .If you liked this post, you may also dig:

Beyond the Every Day: Vocal Potential in AI Mediated Communication –Amina Abbas-Nazari

Voice as Ecology: Voice Donation, Materiality, Identity–Steph Ceraso

Mr. and Mrs. Talking Machine: The Euphonia, the Phonograph, and the Gendering of Nineteenth Century Mechanical Speech – J. Martin Vest

Echo and the Chorus of Female Machines—AO Roberts

Black Excellence on the Airwaves: Nora Holt and the American Negro Artist Program —Chelsea Daniel and Samantha Ege

Spaces of Sounds: The Peoples of the African Diaspora and Protest in the United States–Vanessa Valdes

On Whiteness and Sound Studies–Gus Stadler