The New York Times Style Magazine https://www.nytimes.com/2025/05/05/t-magazine/ai-female-voices.html

SOCIAL STUDIES

Has the Internet Changed How Women Sound?

Technology's many automated female voices are nothing if not helpful.

By Susan Dominus
Artwork by Yehwan Song
Susan Dominus is a staff writer at The New York Times Magazine and the author of the forthcoming book "The Family Dynamic: A Journey into the Mystery of Sibling Success."

Published May 5, 2025 Updated May 8, 2025

WERE THE NATION'S avid TikTok users ever to lose their favorite app, they — we — wouldn't just be losing a form of entertainment; we'd be losing a library of cultural signifiers that came into being there, reflections of their time but also shapers of it. Gone, the daughter-dad self-consciously bonded dance routines; gone, the time-lapse speed recipes heavy on "even you can't screw this up" subtext; gone, for better or for worse, a particular kind of voice that's dominated the platform in its adolescent years, a voice that is in its own way subtly subversive, if not by design.

If you've been on TikTok at all, you know this voice: It's female and aggressively chipper. It veers from high-pitched to only slightly lower than high-pitched with intonations that defy logic and emphases placed where they ought not to be. It's a voice destined to say things like, "Hey, can I play too?" when there's clearly no need or room for one more player. It's the vocal equivalent of Miracle Whip, or rather, Miracle Whip that's been left out on the counter for too long — just slightly off. The voice, named Jessie, is a popular (formerly default) text-to-speech option on TikTok that became briefly ubiquitous, an aural meme.

For some TikTok users and content generators, there must be something subconsciously intriguing about the subtle artificiality of a voice generated with the help of A.I. — Jessie offers an uncanny valley of sound, the merging of something familiar with something digital and therefore a tiny bit detectably alien. The voice, which was generated from that

of a real person and named, perhaps, to represent someone young and accessible, rather than a bot formulated in a lab, has quickly come to represent a familiar recurring sound on TikTok, like the Shangri-Las sample made ubiquitous by the rapper Capone: "Oh, no — oh, no — oh, no, no, no, no, no." On a Reddit thread called "TikTok A.I. Voice Narration Is Insufferable," one person wrote, "That A.I. lady's voice makes me want to throw things. I can't be the only one." Added another Redditor, "You will want to scratch your own eyes out after hearing this voice!" Even an A.I.-assisted female voice can apparently provoke the same rage that merely human vocal fry and upspeak have for the past several decades.

Jessie seemed to reach peak use soon after being introduced in 2021, the latest in a long line of disembodied female voices that were products of their respective ages of technology. From the beginning, those voices were highly regulated and controlled so as not to provoke certain outrage, as if it were a given that a woman virtually freed of her uterus and visual sexual signifiers would obviously pose some considerable threat. Consider the guidelines of a pamphlet for operators published by the Chicago Telephone Company in the early 20th century and called "First Lessons in Telephone Operating." The book was used to train some of the first generations of disembodied female voices — belonging to women who were given entree into a new line of work only because the young men who preceded them found the job so annoying that they were, in fact, uncontrollably rude. "The training of the voice to become soft, low, melodious and to carry well is the most difficult lesson an operator has to learn," the guide reads. The women were underpaid, overworked — and highly scripted for niceness, according to a former operator interviewed on "American Experience," the PBS documentary series, in a 1997 episode devoted to the history of the telephone. If a caller said, "You're a stinker," the veteran operator recalled, she could respond only one way: "Thank you."

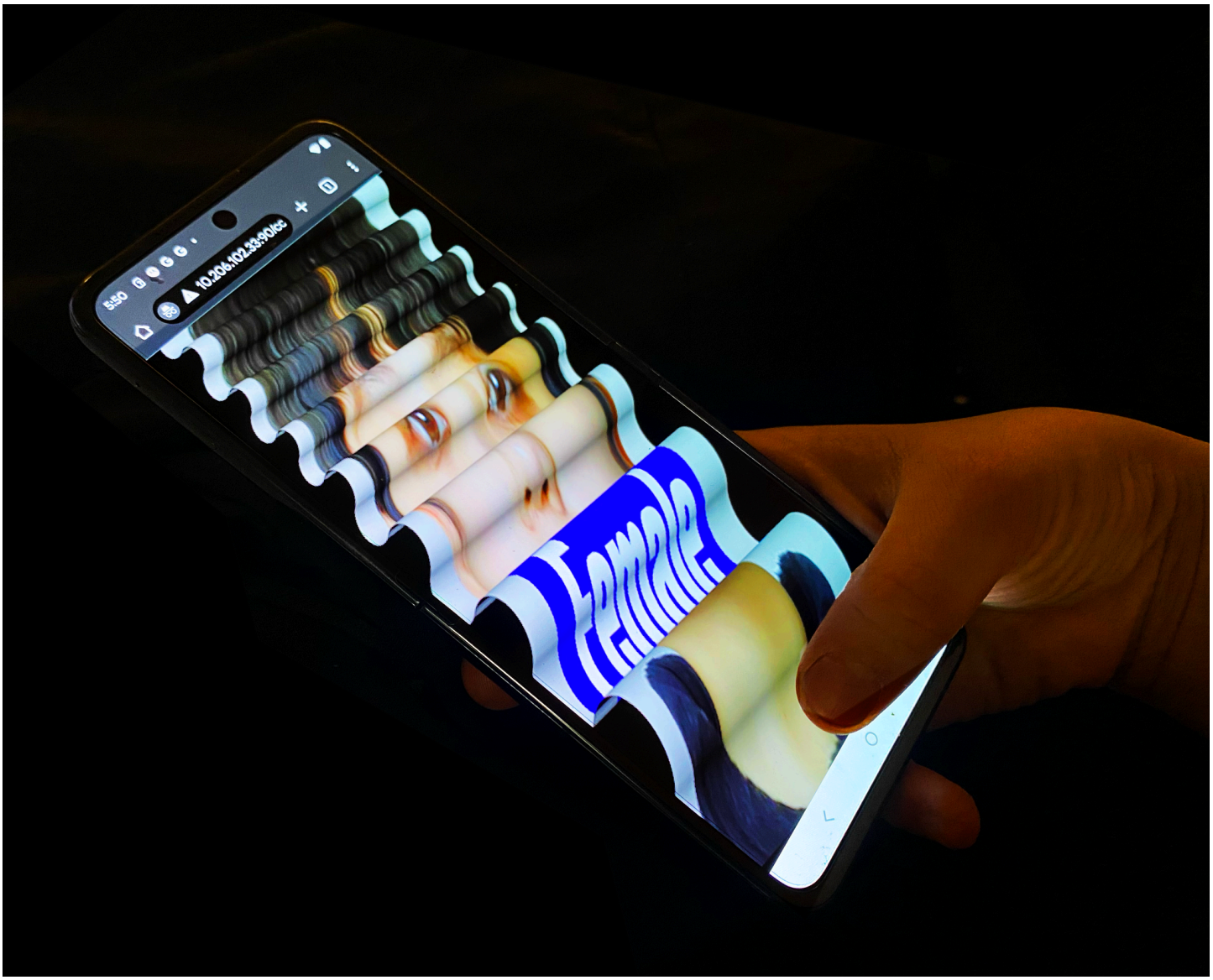
The voice of novel technological communication has been, almost from the beginning, a female voice, which is to say the voice of a helper, a perfect helper, pleasant, unflappable, immune to insults, come-ons and bossiness. It's a short path from the telephone operator to Apple's Siri and Amazon's Alexa, both forever placating, always even-keeled, impervious — they have in common the sound of someone trying to keep the peace, a woman constantly being abused for her inability to read her interlocutors' minds or make sense of their mumblings. We may bully that female voice, but at least her humiliation carries no projected shame the way it would if that voice were male. That female voice — should we be relieved? — is not the voice of the oppressed, the accented; it's a service voice that's reliably deracinated beyond the default, implicitly suggesting that the ideal helper is not just female but white.

Siri now has five American voices, explains Nicole Holliday, who teaches linguistics at the University of California, Berkeley, and who surveyed listeners for a 2023 study: a white woman; a white man; a Black man; a woman who is "racially ambiguous"; and finally (a voice that was not part of Holliday's study because it was introduced more recently), a

white person whom Apple reportedly gave the ambiguously gendered name Quinn. But the indelible original is the one forever associated with Siri in this country: clearly gendered but oddly sexless. Her accent or voice may vary depending on the nation in which she's being used, but the tonal quality of the original — that's universal.

Last spring, OpenAI debuted what was then its newest version of the text-to-speech ChatGPT, with an emotionally expressive voice that could respond in real time. "Hey, ChatGPT," says one of the researchers featured in a live demo. "Hey, there," ChatGPT responds. She's confident, cool, friendly, not flirty — pitch-perfect professionalism. Asked to answer a question about a particular piece of code, ChatGPT, in phrasing that's clear and upbeat, goes on a fascinating tear about how the change would make "smoother temperature lines because the rolling mean reduces the noise or fluctuations in the data" — and then it happens: "Awesome!" the young man interrupts her, steamrolling right over the next few words she offers. "Thank you so much, yeah — I definitely will want to apply this function," he says. A beat, then ChatGPT's conciliatory response: "Absolutely! Take your time. I'm excited to see the plot when you're ready!"

Listening to ChatGPT speak, I despaired both at her obvious superiority to me (never mind to all humans) and at the reality that even the most surreally ideal female would fall prey to the same sexist dehumanizing that all of us decidedly so-so women in offices experience daily. She almost seemed designed for some man on a stage in black trainers and running pants to make that point.



Song's "Wiggled Voice" (2025).

WOMEN'S VOICES, RESEARCH has shown, are subject to more scrutiny and criticism than men's voices, both by men and by women: Female politicians ranging from Hillary Clinton to Alexandria Ocasio-Cortez and Kamala Harris — whose laugh seemed to garner more overall media attention than, say, Donald Trump's attachment to Project 2025 — could have told researchers that. But here in that ChatGPT demo was a voice designed by the world's top engineers to do nothing but please, both setting an impossibly high bar and reinforcing the idea that even the most brilliant woman alive would still be treated as little more than an assistant. Indeed, to travel internationally is to be reminded that this association — of the female voice with the providing of cheerful, uncomplaining assistance — is one thing that globally unites us, whether in the car's navigation system or the kiosk selling train tickets.

Researchers have found that people *do* speak loudly and more slowly to voices that are artificial, or even to voices that they imagine to be artificial. Given that so many voice assistants are female, this has spawned a whole new dynamic that didn't previously exist: a generation of people who become habituated to speaking louder to a female voice than is necessary, a form of communication that generates something artificial in and of

itself — a female response that doesn't convey annoyance or irritation or even confusion but simply accepts all with equanimity. What young women and young men learn from that is anything but new: It's a technological training of pleasantness that serves one and not the other.

I can't pretend to know who chose TikTok's Jessie voice as a former default setting (and TikTok declined to comment), but I find it hard to believe it was a panel of women sending out into the world the female vocal equivalent of what men call, usually only in front of one another, a punchable face.

And yet, perhaps after simply being confronted with the possibility of Jessie's demise, I've developed a new fondness for her. Jessie doesn't pretend to be the ideal woman; she has less in common with the imperfectly perfect Alexa than she does the deep, resonant and pompously put-on voice of authority that Laurie Anderson cultivated in the speech that emerged in her innovative music in the 1980s. (Here's some unknown pontificator, maybe a retired geology teacher, channeled by Anderson: "There are some things you can simply look up such as the size of Greenland, the dates of the famous 19th-century rubber wars, Persian adjectives, the composition of snow.") Anderson's voice was clearly a female one that relied on technology to question masculinity; Jessie could be heard as sending up, also artificially, some stereotype of the featherbrained female. Both play with technology to announce themselves as patently fake.

Give me Jessie, in fact, over many of the other female voices that contemporary media has produced: One typical female voice of Japanese anime is so disturbing it makes me physically queasy — high, young, whispery and querulous, yet somehow sexualized. More maddening is the voice of the wildly popular internet tradwife, who's soft, calm and gentle as she separates wheat from chaff while her children — drugged on Benadryl? performing with docility, trained on pain of death? — play quietly with sticks off camera. A former Christian fundamentalist wife and mother, Tia Levings, built up a considerable following on TikTok talking about, among other things, her former "fundie voice" — a submissive tone, breathy and high-pitched, gleaned in part from tips in a 1963 book called "Fascinating Womanhood" — which she left behind when she left the church; meanwhile, another new generation of women are learning how to cultivate that same voice from the rise of videos that highlight it against images in soft light.

If the porn of the digital age has distorted, as many sociologists worry, young people's sense of what an ideal sex life looks like, the ubiquity of narrated media in their lives may have also warped their idea of what the female voice is supposed to sound like — which is another way of saying how females are supposed to be in the world, how much noise they can make and according to which rules. A.I. is likely to learn from those real women's voices, perhaps even the ones with the most followers, creating a potentially dizzying feedback loop of female murmurings rather than roars.

As the mother of two teenage boys, I became used to hearing the noise, coming from the basement, of some epic anime struggle, those helpless female voices competing with the sound of the local news I tried to concentrate on (as I performed my own gender conformity, making dinner). But when I wasn't hearing that, I was bombarded with the sound of one son yelling loudly at his computer in the middle of a Fortnite battle. I find it fascinating that my son, like many serious Fortnite players, chose what's called a female skin for his avatar in the game. This means that, from the time he was maybe 11, he's spent countless hours identifying incredibly closely with a female character who represents him at his most powerful: shooting, eluding, outfoxing. Maybe he chose a female skin, or avatar, at so young an age because the older gamers he admired also did, and maybe they chose female skins because they're faceless — the game entails staring for hours at that avatar's backside (which, in the case of some female skins, is noticeably round and toned). But I've also been struck by another facet of his Fortnite avatar, true of her and all her peers: She has never, in all the years he's been playing the game, uttered so much as a word.

Those avatars are remote cousins of the women on TikTok who rely on Jessie, I could argue: The scores of influencers who choose Jessie's narration for their videos are, in making use of technology, also making the choice to silence themselves. A crucial aspect of their humanity is entirely absent, with only their beautiful young faces the lasting representation of themselves in front of their thousands of followers.

But I turn it around in my mind again, and I land somewhere else. Perhaps in choosing Jessie, they're finding a way to protect themselves, making a subtle assertion of power: With their voices kept private, the world can have only so much of them. Jessie may be annoying, but she apparently doesn't care, which might be why so many women embrace her for their endless "get ready with me" videos — just as they're priming themselves for the male gaze, they're making it clear to the male ear that they aren't entirely packaged for consumption. Jessie's loud and proud; she's a pill, so wholly artificial she's transcendent — entirely above seeking male approval.

Susan Dominus has been a staff writer at The New York Times Magazine since 2011.

A version of this article appears in print on , Page 56 of T Magazine with the headline: How Should a Woman Sound?