

Note to Listeners/Readers: I ask that you please do not cite this in-progress project. I am planning to expand on this piece in various ways and would be very grateful for feedback. Feel free to email me with comments or questions: stephceraso@gmail.com.

What Is a Voice?

Steph Ceraso | University of Virginia

This essay's accompanying audio collage, which blurs the boundaries between human, animal, and machine voices, is a part of my ongoing exploration of what it means to have a voice. The collage starts out with a milestone of the human voice—a baby's first words (in this case my son saying "mama"). As the collage progresses, the concept of voice becomes increasingly complex through juxtaposed soundbites such as an interview with the human behind Amazon Alexa's voice (conducted in Alexa's voice), Stephen Hawking's iconic synthetic voice, the recomposed voice of a 3,000 year old mummy, and toads vocalizing in sync, among other examples.

The question I pose throughout the collage—"What is a voice?"—is first sounded as a straightforward, unmanipulated recording of my vocals; however, each repetition of this question is altered through digital audio effects to underscore the malleability of the recorded voice. Other soundbites reinforce this idea. For instance, the clip of the man stating "A favorite part of the holiday will always be getting together for a backyard barbeque with family and friends" is manipulated twice by changing its pitch via [VocaliD's AI voice synthesizer tool](#), a tool used to test drive purchasable custom voices for smart objects and products. While the manipulations of my own recorded voice and the demo of the AI tool amplify the changeability of voice in relation to digital technologies, the clip of the Brady Bunch song, featuring Peter Brady's puberty voice, suggests that even without technological intervention, voices are highly unstable and malleable.

In the final moments of the collage I layered the various soundbites used in this piece to prompt listeners to consider how these voices relate to one another—how common conceptions of voice as a uniquely human attribute (voice as the "grain" or essence of our humanity) might be retheorized to account for a broader ecology of vocal possibilities (Barthes).

* * *

There is a long history of scholarship on voice in writing and rhetoric studies. In *Voices on Voice* (1994), editor Kathleen Blake Yancey attempts to map out the scholarly landscape of voice in the field. In the introduction, Yancey gives an overview of the prominent trends in disciplinary scholarship on voice. This is my own succinct summary of the trends she identifies—many of which have overlapping or complimentary ideas:

- Voice in the Oral Context: refers to the physical voice in the rhetorical tradition, or “the medium that enables talk, through which we speak to each other and through which we learn about ourselves, language, and the world” (viii)
- Expressionism and Voice (associated scholars: Graves, Stewart, Murray, Elbow, Macrorie): voice as the self in writing; personal voice as marker of authenticity (ix)
- Voice vs. Voices (associated scholars: Harris, Tompkins): “a vehicle for expression of the self” but one that is plural rather than singular, or “a fluid composite of cultural voices and individual selves within the writer” (xi)
- Voices within Text (associated scholar: Elbow): “voice is created as much by the reader as by the writer and the text” (xii)
- Bakhtin, Voice, and Appropriation: voice is not singular, but multiple; voice is not necessarily confined to the self; we create voices out of the voices of others; language as heteroglossia (xiii)
- Non-Western Views of Voice (associated scholars: Gong, Powers, Carr); voice belongs to both the individual and culture/community; “the authority of voice derives from other and more than a self” (xvii)

While Yancey provides an in-depth and generous examination of the various areas of voice discussed in the field, the notion of *voice as self*—popularized by the expressionists in the 1970s and 80s—is an idea with staying power. The voice as self—as a marker of our humanity as writers and speakers—is still a powerful concept in writing and rhetoric studies, overshadowing some of the lesser known theories of voice within and beyond the field.

Since the publication of Yancey’s collection, there has been relatively few publications about voice writing and rhetoric. However, the materiality of voice, a topic that has been taken up in a recent cluster of scholarship (Alexander, Anderson), has the potential to push disciplinary work on voice in what I think are interesting and productive directions. More specifically, I want

to suggest that a consideration of contemporary vocal technologies and nonhuman voices—which call attention to the material nature of voice—might allow us to build upon, reconceptualize, and reenergize the concept of voice in writing and rhetoric, and particularly in the subfield of sonic rhetorics.

* * *

My initial investigation of voice began after hearing Erin Anderson’s “Being Siri” (sampled in the audio collage). “Being Siri” is an experimental audio piece about Anderson donating her own voice to Boston-based voice donation organization, VocaliD. Much like a blood bank, VocaliD allows anyone with an internet connection to donate his/her/their voice via digital audio recordings. These recordings are then processed by technicians and used to create a “custom digital voice” for a voiceless individual. VocaliD provides an alternative to the mechanical-sounding assistive technologies used by people who cannot speak for themselves—the robotic sound of Stephen Hawking’s voice is a well-known example. VocaliD is able to manufacture voices that better match a person’s age, gender, race, ethnicity, and unique personality.

An estimated 2.5 million people in the US alone are severely speech-impaired, and there is very little variety in vocal assistive technologies (Romm). Given these facts, I don’t want to downplay the important service that VocaliD provides for people *who want* to have a new vocal identity that better matches their own ideas about who they are. However, I think there are problematic features of this organization that are worth examining further.

One of these problematic features is the way that VocaliD rhetorically frames their service by persistently linking voice to humanity. For instance, consider these quotes from Rupal Patel, the CEO and founder of VocaliD, in which she is emphasizing the need for voice donation technology:

“Here’s a way for us to acknowledge these individuals as unique human beings.” (Fastco Studios)

“I was talking to [a] girl we made a voice for. She told me that people are finally seeing her for who she really is.” (Medieros)

“Science is an incredible superpower when applied to improve the human condition.” (VocaliD website)

“This [generic assistive technology] is just continuing to dehumanize people who already don’t have a voice to talk.” (Kisner)

VocaliD’s claims that their unique vocal identities “humanize” individuals imply that one is not fully human unless one’s voice *sounds* human.

What I find especially disturbing about the rhetoric of humanity and individuality used by VocaliD is that the organization adopts the same type of rhetoric for a line of AI voices they sell for smart objects and products—as is evident in this passage from VocaliD’s website:

The universe of things-that-talk is growing, but the diversity of unique digital voice isn’t. This means lots of businesses don’t sound like themselves—they sound like each other. You’ve designed a device that solves modern problems for a diverse audience of consumers. Your robust operating system is cutting-edge and will grow with the demands of its users. Your team’s attention to detail has been meticulous, which is why you shouldn’t settle for giving your product a generic voice.

VocaliD’s discourse about voices for both humans and products emphasize uniqueness and individuality, attempting to “humanize” consumer goods in addition to voiceless people.

I could go on about the problematic nature of this rhetorical gesture, but what interests me here is that while VocaliD’s mission to make people *and* machines sound more human, what the organization is actually doing is exposing the fact that voices are not, in fact, unique markers of our humanity. The rhetoric of individuality and authenticity used by VocaliD evades how machine voices—and all voices, including our written and spoken voices—are composed of and grounded in an emergent ecology of material relations.

Put differently, voices are not a manifestation of our true or authentic selves. Rather, voices are a result of specific material and cultural conditions that involve much more than human-ness. Voices emerge and are developed through a blend of:

- unique biological makeup (or technological makeup in the case of machines with voices)
- specific environments and contexts (e.g. geography may determine the kind of accents humans have and the kinds of sounds animals are able to vocalize and hear)
- technologies (e.g. the phones, computers, digital recorders and editors, software, and assistive technologies that are used to preserve and circulate voices);

- others (e.g. humans often emulate the vocal patterns—or writing patterns—of the people they interact with most; many machine voices sound the same, which is why organizations like VocaliD exist; and even the toads featured in the audio collage sync their voices in chorus to mate and ward off predators)

Voices are not static; they are performed differently in different contexts and for different people/animals/machines, and they change as bodies/machines change, malfunction, and eventually break down. In short, voices are intentionally and unintentionally composed over time—shaped by ever-changing bodily states and engagements with the world. Voices are compositions by nature. They are alive and responsive and malleable because they are the result of a complex ecology that involves much more than a “unique” human being or animal or machine. Despite perpetuating narratives about voice and authenticity, then, VocaliD actually makes the composed nature of voices explicit through their technological processes.

Further, synthetic voices make clear that voice is not the essence of our humanity, but rather a repeated practice that informs identity. Stephen Hawking is perhaps the best example of this. As Jordan Kisner writes,

Though the robotic quality of his [Hawking’s] digital voice (and the American accent) felt inappropriate at first, it came to be his trademark. Hawking reshaped himself around his new voice, and years later, when he was offered the opportunity to use a new voice that was smoother, more human-sounding, and English, he refused. This felt like ‘him’ now.

It seems that regardless of how “human” a voice sounds, through repeated use it can begin to feel like one’s identity. That is not to say that the voices we identify with can’t change—think of boys pre and post puberty or transgender people who undergo a vocal transformation in addition to other physical changes. Rather, repetition allows the voice we use or practice with most to feel like “our” voice.

The idea that voices are markers of our individuality or unique identity is complicated even more by the fact that voices are material actors that can operate in the world apart from us—even after we are dead. For instance, one of the clips in my audio collage featured a story about the London Underground agreeing to “reinstate the actor Oswald Laurence’s famous ‘Mind the gap’ announcement at Embankment station having learned that the widow of the actor had been regularly visiting this station since her husband’s death in order to hear his voice again” (McGettigan). Another clip features the story of Stephen Hawking’s voice getting beamed into space toward a black hole as a postmortem tribute. And perhaps the most absurd and ethically questionable clip from the collage features the recomposed voice of Nesyamun, an

Egyptian priest who lived more than 3,000 years ago. This mummy voice, which was made audible after scientists recreated Nesyamun's vocal tract, resulted in both praise and criticism. One vocal critic, John Kannenberg, director of the Museum of Portable Sound, expressed his outrage on Twitter: "This sound is playing a dead human being like a musical instrument – demonstrating a complete lack of respect (would you do this to your dead grandfather?), empathy and respect for the ancient culture and religion in which this body was buried" (Atkinson).

Voices have been able to operate independently of us since the dawn of recording technologies. However, newer technological developments—like the 3D printed vocal tract featured in the mummy story—raise even more ethical questions. It seems essential to ask: Just because we can share or recreate or repurpose a voice, does that mean we should? Is any voice recording or vocal composition fair game, even without the consent of the original source of the voice? This question is especially relevant as software like Adobe Voco, dubbed "Photoshop for Voices," is being tested and developed (Powers). With a very small sample of speech, this software can imitate an individual voice, making that voice say whatever the user wants it to say. While it is unclear whether or not Adobe Voco will ever be publicly available, the technology already exists. As Oliver Bendel writes, this raises a flurry of ethical questions:

What is a human, what belongs to him or her, what can be separated from him or her, and what about the identity, and how to secure it? Should people determine while alive what shall happen to their voice after their death? Should they give orders for the data and information they put out during their lifetimes, as already happening with regards to social media, and shall these invoice their actions of speaking and other actions?

I began this project by asking "What is a voice?" After considering some of the contemporary possibilities of vocal technologies and the ethics surrounding these possibilities, however, maybe a few more pressing questions for the study of voice—in the field of writing and rhetoric and beyond—should be posed: What can and *should* a voice do? What is the ecological makeup of a voice and why might that matter? What does it mean to (re)compose or (re)circulate voice ethically?

* * *

My brief exploration of vocal technologies and nonhuman voices has amplified several related ideas about voice more broadly. To conclude I want to offer these tentative ideas, which I believe have the potential to deepen and complicate how scholars of writing and rhetoric think about and compose with voice:

- 1) Voices are compositions
- 2) Voices emerge from a complex ecology (of bodies, environments, materials, technologies, etc.)
- 3) Voices are repeated practices that often (but not always) inform identity
- 4) Voices are material actors that can operate independently of us, and thus necessitate considerations

My hunch at this early stage in my research is that this compositional, ecological, practice-based, material approach to voice can offer a nuanced narrative about voice in the field while also reanimating and reenergizing previous disciplinary work that doesn't get as much attention as expressionist scholarship on voice. For example, writing and rhetoric has a rich body of work on non-Western ideas about voice. As Yancey explains in *Voices on Voice*, "Native Americans...work from a larger sense of context: voice belongs to the individual, but also to the choral, formed by the human and by the natural community. No writing can thus be only expression of the self; it inherently expresses others and nature, of whom the writer is a part" (xvii). Indigenous conceptions of voice seem key to developing an ecological theory of voice that could prove useful for theorizing the complexities of modern vocal technologies. This is one area of voice-related research I hope to explore further.

As I've tried to illustrate, voices are malleable compositions that are full of inventional and creative possibilities (Anderson, Alexander), but they can also be used in unethical and potentially harmful ways. Considering nonhuman voices and recent developments in vocal technologies in conceptions of voice can help scholars of writing and rhetoric broaden how voice is understood and used in a variety of contexts. My hope is that a robust retheorization of voice in the field would contribute to timely disciplinary and transdisciplinary conversations on sound, authenticity, embodiment, disability, identity, and posthumanism.

Works Cited

- Alexander, Jonathan. "Glenn Gould and the Rhetorics of Sound." *Computers and Composition*. 37 (2015): 73-89.
- Anderson, Erin. "Toward a Resonant Vocality for Digital Composition." *enculturation: a journal of rhetoric, writing, and culture*. 18 (2014) <http://enculturation.net/materialvocality>.
- . "Being Siri." *Unfictional*. KRCW. 3 Feb. 2017.
<https://www.kcrw.com/culture/shows/unfictional/you-want-a-piece-of-me/being-siri>.

- Atkinson, Rebecca. "New Research into Egyptian Mummies Leads to Calls for Major Ethical Review." *Museums Association*. 30 Jan. 2020.
<https://www.museumsassociation.org/museums-journal/analysis/2020/01/30012020-new-research-into-egyptian-mummies-leads-to-calls-for-major-ethical-review/#>.
- Barthes, Roland. "The Grain of the Voice." *Image/Music/Text*. Trans. Stephen Heath. New York: Hill & Wang, 1981: 179-89.
- Bendel, Oliver. "The synthetization of human voices." *AI & Society* 18 (July 2017): 83-89.
- Chandler, Michael Alison. "Volunteers Use New Technology to Donate Voices to Those with Speech Disorders." *The Washington Post*. 5 December 2016.
https://www.washingtonpost.com/local/social-issues/volunteers-use-new-technology-to-donate-voices-to-those-with-speech-disorders/2016/12/05/9442a0ee-b8a3-11e6-b994-f45a208f7a73_story.html.
- Fastco Studios. "How VocaliD Creates a Unique Custom Voice." [Video] *Fast Company*. 4 Nov. 2016. <https://www.fastcompany.com/3065364/how-vocalid-creates-a-unique-custom-voice>.
- Kisner, Jordan. "How a New Technology Is Changing the Lives of People Who Cannot Speak." *The Guardian*. 23 January 2018. <https://www.theguardian.com/news/2018/jan/23/voice-replacement-technology-adaptive-alternative-communication-vocalid>.
- McGettigan, Carolyn. "The social life of voices: studying the neural bases for the expression and perception of the self and others during spoken communication." *Frontiers of Human Neuroscience* 9:129 (2015): 1-4.
- Medeiros, João. "A Voicebox That Sounds Like a Human." *Wired*. 2 June 2014.
<https://www.wired.co.uk/article/voicebox>.
- Powers, Benjamin. "Adobe is Developing Photoshop for Your Voice." *Medium*. 27 Feb. 2018.
<https://medium.com/s/story/adobe-is-developing-photoshop-for-your-voice-f39f532bc75f>.
- Romm, Cari. "Donate Your Voice to Charity." *The Atlantic*. 18 December 2014.
<https://www.theatlantic.com/health/archive/2014/12/donate-your-voice-to-charity/383878/>.
- VocaliD. Accessed 16 Sept. 2020. <https://vocalid.ai/voicebank/>.
- Yancey, Kathleen Blake, Ed. *Voices on Voice: Perspectives, Definitions, Inquiry*. NCTE. 1994.

Audio Collage Transcript + Works Cited

Ceraso: [mid-range female voice] “What is a voice?”

Ceraso [rough phone recording]: “Can you say “mama”?”

Baby boy: [pause] “mama” “ma ma ma ma ma ma” [Ceraso giggles quietly in approval]

Ceraso: “Very good!”

Baby boy: “mmm mmm mmm...mama”

[a few seconds of upbeat percussive keyboard music play before Amazon’s Alexa begins talking; music continues throughout Alexa’s narrative; at times it sounds slightly glitchy, mimicking Alexa’s robotic speech pattern]

Amazon Alexa [friendly female AI voice]: “Hello. My name is Alexa. Actually, my name is Susan and I’m the voice of Alexa. People are always stopping me at the grocery store or at restaurants when they hear my voice and asking if I’m Alexa. Sometimes I say ‘yes.’ Sometimes I tell them ‘no.’ It just depends on how I’m feeling that day.”

[music fades; brief pause; the white noise associated with the beginning of a recording plays]

Stephen Hawking [iconic synthesized male voice with American accent]: “My theme tonight is science in the new millennium. Nearly all the visions of the future that we have been shown from H.G. Wells onwards have been essentially static.”

Ceraso [voice is manipulated by lowering pitch; sounds like a male voice]: “What is a voice?”

[moans of mummies and dramatic, suspenseful Hollywood movie soundtrack]

Female reporter: “With their moans [mummy moaning], groans [mummy groaning], and screams [mummy screaming], Hollywood’s mummies have always given us *chills*. [dramatic movie music continues for a few seconds and then fades out...minimal introspective keyboard music begins and fades out seconds later]. But this week researchers in England let us hear what a real mummy sounds like [a deep male voice in a tone of disapproval says: “mmmm....mmmm”]. That’s the voice of Nesyamun, an Egyptian priest who lived 3,000 years ago. [“mmmm....mmmm....mmmm”] Scientists were able to mimic Nesyamun’s voice by recreating his mouth and vocal chords with a 3-D printer. It allowed them to produce a single sound [mmmm....mmmm....mmmm].

[upbeat sixties guitar pop music]

Greg Brady: “Autumn turns to winter / and then winter turns to spring / it’s not just the seasons, [all of the Brady kids in chorus] ya know, it goes for everything”

Marsha Brady: “It’s even true for voices when boys begin to grow / ya gotta take a lesson from mother nature and if you do you’ll know”

All Brady kids in chorus: “When it’s time to change [other Brady kids’ voices fade and Peter Brady’s puberty voice awkwardly squeaks out] then it’s time to change [laugh track]”

All Brady kids in chorus: “Don’t fight the tide come along for the ride, don’t ya see? / When it’s time to change you’ve got to rearrange / who ya are and what you’re gonna be / Sha na na na na na na na na / Sha na na na na” [music fades]

Male voice [provided by VocaliD’s AI tool]: “A favorite part of the holiday will always be getting together for a backyard barbeque with family and friends.”

Same male voice but higher register [manipulated pitch with VocaliD’s AI tool]: “A favorite part of the holiday will always be getting together for a backyard barbeque with family and friends.”

Same male voice but lower, baritone register [manipulated pitch with VocaliD’s AI tool]: “A favorite part of the holiday will always be getting together for a backyard barbeque with family and friends.”

Ceraso [voice is manipulated by raising pitch; sounds cartoonish]: “What is a voice?”

Stephen Hawking: “Can you hear me?”

Bernie Krause [kind, authoritative male voice]: “At the end of its breeding cycle, the Great Basin spadefoot toad digs itself down about a meter under the hardpan desert soil of the American West, where it can stay for many seasons until conditions are just right for it to emerge again. And when there’s enough moisture in the soil in the spring the frogs will dig themselves to the surface and gather around these large, uh, vernal pools in great numbers. And they vocalize in a, in a chorus, that’s absolutely in sync with one another. And they do that for two reasons. The first is competitive because they’re looking for mates and the second is cooperative because if they’re all vocalizing in sync together it makes it really difficult for predators like coyotes, foxes, and owls to single out any individual for a meal. This is a spectrogram of what the frog chorusing looks like when it’s in a very healthy pattern.” [sounds of toad chorus]

Ceraso [voice is manipulated with reverberation effect]: “What is a voice?”

Erin Anderson [female voice that sounds slightly synthesized and robotic; the following statements overlap and blend into one another, making it increasingly hard to understand more than one sentence at a time]: “She hid his key. So strange was the dream. Tickle the frog with the feather. The internet of things. It’s always cold in San Francisco. You must earn it. Where’s Lord

Alfred? Please be patient. The ripe taste of cheese improves with age. Buy Bobby a puppy. Nod if you are alright. She needs a heart. You should remember that you are a young person. Gloves are more important than anything else. Can I have ten bucks? When can I go to sleep? No one else will help me. Your hands are bigger than mine. Brothers and sisters don't always get along. Pizza is my favorite food. I always spoil everything. We are cut off from all the rest of the world. It doesn't matter I guess. Hello. Welcome. This is my voice vault. It is still incomplete. I am a level 7 speech donor. I am donating my voice to someone I've never met, someone I may never meet at all. Every word I'm speaking will some day be hers. [a few xylophone notes ring out and fade]"

Oswald Laurence: [male British voice; aged but clear recording]: "Mind the Gap."

Male British Reporter voice: The 40-year-old recording of the "Mind the Gap" message played on the underground is to be used once more so the widow of the actor behind the warning can hear his voice. Oswald Laurence was heard for years on the northern line and later at Embankment Station, which his wife would visit just to hear him. The transport for London replaced the message until they heard about the widow's story."

Female reporter's voice: "On Friday Hawking's ashes were buried between the remains of two other giants of science [clips of Hawking's funeral mass can be heard in the background of the report], Isaac Newton and Charles Darwin. [Hawking's recorded voice plays quietly in the background] As he was laid to rest, the European space agency beamed his voice from a satellite dish in Spain toward the nearest known black hole more than 3,000 light years away."

Hawking [excerpt of voice recording beamed into space; mysterious/inspirational synthesized music plays as he speaks]: "We are here together and we need to live together with tolerance and respect."

Ceraso [voice manipulated by slowing speed substantially]: "What is a voice?"

Final Audio Collage: layered sounds of baby saying "mama"; Peter Brady's puberty voice squeaking out "time to change," mummy voice sounds ["mmm mmm"], toads vocalizing in sync, Laurence's "Mind the Gap" message, and Hawking saying "Can you hear me?"

Ceraso [clip is edited so that words are choppy and repeated for effect]: "What what is a voice, voice, voice, voice?"

Works Cited

- Anderson, Erin. "Being Siri." [Audio File] *Unfictional*. KCRW. 3 Feb. 2017.
<https://www.kcrw.com/culture/shows/unfictional/you-want-a-piece-of-me/being-siri>.
- "Custom Digital Voices for All." VocaliD. [Demo Tool] Accessed 14 Sept. 2020.
<https://vocalid.ai/>.

Krause, Bernie. "The Voice of the Natural World." *TED: Ideas Worth Spreading*. TEDGlobal June 2013. https://www.ted.com/talks/bernie_krause_the_voice_of_the_natural_world?language=en.

"London: Wife's Search for Husband's 'Mind the Gap' Announcement." [Audio File]. *YouTube*. Uploaded by BhamUrbanNewsUK. 29 March 2013. <https://www.youtube.com/watch?v=ey96gF9LXqU>.

"Researchers Recreate What Mummy's Voice Would Have Sounded Like." [Audio File]. *YouTube*. Uploaded by CBS This Morning. 25 Jan. 2020. <https://www.youtube.com/watch?v=Xy36R3vmE2A>.

"Stephen Hawking Honored, His Voice Beamed into Space." [Audio File] *YouTube*. Uploaded by CBS This Morning. 16 June 2018. <https://www.youtube.com/watch?v=7Feetc73e14>.

"Stephen Hawking Speech | From 1983 to Last Speech." [Audio File] *YouTube*. Uploaded by Top Famous Tube. 14 Mar. 2018. <https://www.youtube.com/watch?v=T7YrNJ7Np38>.

"Time to Change." [Audio File] *YouTube*. Uploaded by phoomd. 20 Sept. 2007. https://www.youtube.com/watch?time_continue=1&v=DyooALwfxO8&feature=emb_log_o.

"Who Is Alexa? The Voice Actress Behind Amazon's AI." [Audio File] *YouTube*. Uploaded by Mike Wiest. 16 May 2019. <https://www.youtube.com/watch?v=HkanSVgHeYU>.

Note: "What is a Voice?" narration and "Can you say mama?" audio clips are Ceraso's personal recordings.