

Roman Kalinovski

The Grain of the Cyborg Voice

In November 2014, I attended Hatsune Miku's concert at the Hammerstein Ballroom. During the show's grand finale — the song *Starduster* — I was profoundly moved, emotionally and physically, to the point of tears. People cry at concerts all the time, so my experience isn't worth mentioning except for one detail: Hatsune Miku isn't human. "She" is a fictional character, the personification of a computer program that converts text to song. Her voice has human origins, sourced from the singer Saki Fujita, but the character created to represent this voice has taken on a life of her own in the popular imagination. This new voice is a hybrid of humanity and technology, but who is the cyborg, Miku or Fujita? Is Miku a prosthesis for Fujita, allowing her voice to transcend the limitations of singular embodiment, or does the human vocalist serve as raw material for the creation of a new kind of voice? Has any "grain" in Fujita's voice survived the violent digital transformation into Miku? Or does Miku's voice have a grain of its own, something not tied to a human body yet possessing the potential to be profoundly moving in its own way?

In his essay, *The Grain of the Voice*, Roland Barthes theorized that there is an aspect of the human voice — the titular "grain" — that is communicated non-linguistically: it isn't in the song's lyrics, the singer's emotive expressions, or even in the breaths between words. Rather, it is the bodily relationship that is felt between the listener and the singer, a sense of her individual being as expressed through the body. To Barthes, the moving power of song is to be found somewhere in this non-linguistic feeling of embodiment; this raises a question: can a voice not generated directly from a human body have this same power?

When Miku was created, micro-temporal samples of Saki Fujita's voice were recorded and saved as a database of waveforms. When the Vocaloid program is run, these samples are adjusted according to the user's input and stitched together to create Miku's voice. Is this output indexical to Fujita's actual voice, and is her individuality — Barthes's "grain" — retained in any of these sub-syllable samples? When listening to a song performed by Miku, the listener doesn't hear any of the actual waveforms as Fujita sang them. The Vocaloid program modifies the original samples based on particular parameters, numerical values that determine the pitch, timbre, and expressiveness of Miku's voice for each particular note. These parameters separate Miku's voice from Fujita's; a quantification of affect that gives the resulting voice its own sense of individuality.

Barthes wrote that the grain of the voice can be heard not merely in a singer's breathing — to him, breath was too simple a means to prove embodiment — but in the entirety of the anatomy: he described hearing a singer's tongue, teeth, and nose expressed solely through the voice. Miku doesn't have any of these features: her character dancing onstage may, but they are purely decorative and have no bearing on the sound of her voice. Instead of evidence of human anatomy, the astute listener can hear in Miku's voice the sounds of her digital anatomy. Rather than sound waves traveling from the lungs through the structures of the throat and mouth, the

digital signal of Miku's voice travels from a database of Fujita's samples through multiple layers of filters, pitch-shifters, and user-defined parameters that act similarly to a human singer's body, altering the output at each encounter.

The grain of the cyborg voice exists in the moments when the differences between two worlds, the material and the digital, are made undeniably apparent. *Starduster* — the song that brought tears to my eyes at the concert — has many such moments, particularly in the chorus, a litany of lines each beginning with “ai wo”. Miku's voice has a hollow ring as she sings these words; a charmingly clunky shift occurs between each A and I sound, evidence of the signal processing taking place in the space between notes. It sounds obviously mechanical and digital, yet is human enough to momentarily fool the ear, and that may be what is most moving about it. If Miku's voice was perfectly naturalistic, easily mistakable for that of a human singer, she wouldn't have nearly the popularity she enjoys today. It is the discernable evidence of her digitality that gives her voice personality, even if, as a fictional personification of a software program, any personality assigned to her is ultimately imaginary.