Singing with Hatsune Miku: Vocaloids as a

Medium for Music Learning

Matthew D. Thibeault
The Education University of Hong Kong

Koji Matsunobu
The Education University of Hong Kong

ABSTRACT
In this short paper we explore the educational implications of Hatsune Miku. Miku is a Japanese vocaloid—software that can produce a sung performance. Following work in the field of sound studies, we characterize Miku as a medium—a contingent network of people, practices, institutions, and technologies that come to be recognized via patterns of engagement and recurring relations. We briefly consider three perspectives of Miku's network: first, Miku as "she" through her global pop stardom; second, Miku as "it" via the technological dimensions; and, third, Miku as "they" by considering the community of fandom and producers from which she emerges. Having established the basic outlines of the medium of Miku, the article presents engagement with Miku through the vignette of one of the author's first experiences singing through Miku. Finally, we outline some of the educational issues latent in Miku and other mediated music education.

Keywords: music education technology, media education, sound studies, Hatsune Miku, Vocaloid

The year 2007 could be when humanity first heard an important part of the future of music education. If so, it will be due to the birth that year of Hatsune Miku (初音ミク), whose name translates as "[初] *first* [音] *sound* from the [ミク] *future*." Miku is a Japanese *vocaloid*—software that produces sung performances. Listening to Miku, one hears the sound of a voice that is virtual and synthetic, yet also capable of great expressivity. It may well be a herald of things to come, a new creative vista which educators should explore. Miku, we contend, is a new *medium* for musical development—a network from which an extraordinary creative output has emerged, and with which some educators are already engaged.

Miku upends commonsense ways to think about singing and technology. Throughout the paper we welcome this instability. For instance, we are deliberate in our use of pronouns. We often use "her" to reference the persona at the center of Miku as a character or avatar—although one might use "he" when describing a performance that was entirely crafted by a male producer. Other times, we refer to Miku as "they" to refer to the larger community of vocaloid producers who operate as Miku. Since Miku is also software that runs on various hardware, and invoking this aspect leads us to also refer to Miku as "it." No one conception can dominate an honest discussion of Miku, and we are indebted here to queer theorists and technologists whose work sensitized us to these subtleties (Butler, 1990; Haraway, 1991).

The multitudes that make up Miku do cohere when she/he/they/it is regarded as *a medium*, a conceptualization we use throughout this paper. Drawn from the field of sound studies, we follow Sterne's (2003) argument that media are too-often misidentified only as gadgets—such as iPods or radios. For Sterne, these gadgets are merely "a placeholder for a recurring relation among people, practices, institutions, and machines" that, through repeated use, can be understood as constituting a medium (p. 226). A medium is, therefore, not the gadget but rather the larger network that emerges through recurrent use.

To regard Miku as a medium is to recognize the far-reaching aspects that together make it more than a piece of software: a contingent network that includes the musicians and visual artists who perform with and through her, the audiences who attend concerts and share her music, the technologists who develop and continue to innovate its software, and the various approaches to working with and enjoying Miku including CDs and video-streaming sites. This medium also evokes and works within a series of philosophical and ideological aspects that are of interest to those who wish to promote deeper educational engagement with and understanding of vocaloids.

Key for educators is the idea that Miku is contingent: she exists *only* in the recurring yet evolving connections of software, creators, distributors, and audience members who come together to enjoy and celebrate her work. Miku is a singer, then, whose existence depends on a robust ecology of people, technologies, and practices all working together to maintain her existence. Below, we first explore Miku as a medium, then present reflections from first-hand experience singing with Miku, before presenting preliminary ideas on the mediated learning Miku makes real.

Miku as a medium

We first illuminate three of Miku's network aspects for consideration: first, Miku as a persona; second, Miku as a technological network; and, third, Miku as a networked community of fandom and creators.

Miku as performer

Miku is not universally known, but she is undeniably famous: a Google search conducted on June 6, 2017 found over 26 million entries for "初音ミク" and 7 million for "Hatsune Miku." A pop star-style biography hitting the high notes of this (perpetually) 16 year-old female android with long blue hair would note that over 22,000 compositions have been written for her, and that she has sung on over 100,000 recordings (Zaborowski, 2016, p. 122). While much of her work is stylistically oriented towards electronica and pop, her range includes nearly every genre, including heavy metal as well as her appearance in an opera written for her by the renown composer Kitaro (Conner, 2016, p. 122). In the year her software was released she was the number one selling software on Amazon.jp, and she also topped the music charts— *Exit Tunes Presents Vocalogenesis feat. Hatsune Miku* appeared in the number one spot for the Oricon charts in Japan ("初音ミク'ボーカロイドアルバム'が徳永を押さえ、初首位 | ORICON NEWS," 2010). Manga featuring her abound, as do countless visual depictions on the internet. A rhythm game by Sega that features Miku, *ProjectDiva*, exists in both arcade and console versions. She routinely performs in concert via holographic projection, including international tours that feature a human band and holographic Miku—the configuration that appeared in October 2014 performing on "The Late Show with David Letterman." High profile collaborations include the Japanese band "Bump of Chicken" and in the US with producer Pharrell Williams. As with other celebrities, Miku earns money by endorsing products, appearing in Toyota car commercials and even ads for Go Go Curry, a chain of Japanese restaurants, and she has sponsored the winter festival in Sapporo, Japan. Finally, in the wake of the 2011 Tohoku earthquake and Fukushima nuclear meltdown, Miku released an album to benefit affected survivors and families.

Miku as technology

Miku performances emerge through a rich network of technologies: the vocaloid software, the computers upon which that software resides, and a host of other digital music and video editing applications for realizing the final performance. Many typical recording technologies are frequently involved, for instance guitars with amplifiers the form the backing on Miku's tracks. And these tracks are shared via CD, TV, video games, and perhaps most prominently video sharing sites like YouTube and Niconico Douga. Miku is technologically contingent, as each technology in the network is routinely changing and updating. Consider the history of the core vocaloid software, which works through a process of concatenation— taking the sampled voice actress Saki Fujita, then filtering and synthesizing it to meet the particular musical demands of a given performance: changing volume, pitch, breath, and expression. This software has undergone profound changes since its creation. Miku's voice was released as part of Vocaloid 2 in 2007, the first in Crypton's "character" series, which introduced particular vocal banks. Version 3 added an optional English pack to allow singing in Fujita's Japanese-accented English. In 2010 an add-on provided access to different styles of singing: sweet, dark, soft, and vivid. Version 4 included several enhancements in the editing software as well as an enhanced way to work with sung vowel sounds (Wat, 2015).

A vocaloid producer thus is suspended between constantly changing technologies that all interoperate. To sing with Miku is to sing through this network, which involves understanding the technological and aesthetic aspect of singing *to* that network.

Miku as community

Just as a physical "radio" can distract from the larger network, so "Miku" as a persona can hide the much larger network that exist behind her every appearance. Not only do fans enjoy her videos, CDs, games, and plates of curry and rice—the larger network of Miku flows from the creative output of fans. In fact, there are no canonic or central or ur-text releases by Miku—Japanese CD stores typically have sections devoted to *vocalo* or *netto* (internet) music, but all of these contain music made by users of equal official stature. This is in part due to the intellectual decisions of Miku's creators—buying the software enters the user into a profit-sharing agreement with Crypton Media that allows one to commercially release performances by Miku while requiring revenue sharing with Crypton. This license agreement seems to have promoted the extensive recorded output of Miku, as well as the compositions for her.

Vocaloid not only sings for human fans; human fans sing Miku's music, too. Vocaloid is part of a larger practice of virtual reality, which is often called *Nijigen*, two-dimension, as opposed to the three-dimensional real world. *Karaoke* centers typically have a whole area devoted to vocaloid repertoire—people want to sing the music they love including that sung by androids. In Miku's case this presents some fascinating challenges because some of her music was programmed at tempi and with rhythm and range not normally possible for human singers. In many instances, human singers have pushed their technique to achieve the ability to sing what had been considered technically impossible. A significant subculture within vocaloid is devoted to the humans who specialize in vocaloid repertoire—such singers are known as *utaite* (歌い手). The attention that some *utaite* have attached has led to their becoming well-known in their own right, and some have used the attention they garnered to launch more traditional careers (Kimagure, 2013). And sometimes another synthetic voice covers Miku's repertoire. The fifth most watched video on YouTube in 2011, "Nyan Cat," features an animated cat with a Pop-Tart body flying through space to a cover of Miku's song "Nyanyanyanyanyanyanya!" (Netburn, 2011). The cover version is sung by Momone Momo—an Utau—a shareware version of a vocal synthesis software, resulting in one robot covering another.

Becoming Miku

In search of answers, Thibeault decided to purchase and work with Miku, to actually become part of the world of Miku to see what it was like to participate in that creative medium. This section presents his preliminary reflections, drawn from extensive note-taking over a two-week period in which he first worked with vocaloid, to help frame and sharpen the educational issues that arise from Miku and other potentially similar technological tools that may become commonplace in education. Here is his account:

> I bought Miku. She cost about USD$170 for the version with English and Japanese language abilities. Actually, because she is dependent on other software, before my purchase I first joined an online community for advice as to which version would work given my laptop and the software I already owned. And, to be completely honest, before I even sought help for the version to purchase, I had been daydreaming about singing with/through/as Miku. Before I put down money, I wanted to make sure I had some projects that might be enjoyable or interesting or challenging. I decided that there were some Okinawan sanshin tunes I was trying to learn, and that having her sing them would help me work on the lyrics and practice—in essence, I thought Miku could be a tutor for Okinawan dialect Japanese as I learned to play the sanshin.

> When Miku arrived from Amazon.jp I set it aside for a few days until a Japanese friend came over with her children for a playdate with my son. She helped me to read through some of the installation instructions, and together we installed the software. I then began to work through a getting started tutorial, as Miku's interface is not completely intuitive. I was working with the PiaPro editing studio with Logic Pro. I input a few notes, and then played them back. The six or so people in the room grew silent, listening as I played them back again. I asked what lyrics I should have her sing, and my son suggested "あ―い―う―え―お" the basic vowels of Japanese that he had been practicing at his school. Everyone was slightly amazed when Miku voice sang it back. I then put in the opening phrase of "O Kina Kurino," Japanese lyrics to an English folk song that all children learn in preschool. As Miku finished singing, the visiting mother clapped. Miku and I had pleased our first audience.

> As with most software, there were many moments of friction: for some reason the first song I entered was always missing its first note when I pressed play, a failure that a comical number of workarounds failed to solve. I failed to enter Miku's notes through my MIDI piano keyboard, instead relying on a mouse for very slow individual note entry. Vocaloid producers

refer to "tuning" a performance, by which they mean listening and then fine-tuning parameters such as the volume or brilliance or "gender" to get the most compelling performance from Miku. I was still working mostly with rhythmic placement and rudimentary dynamics—I knew that I had barely started to understand how Miku sings, how her voice is shaped to make a more beautiful performance.

Over the next several days I noted that I often found myself thinking of Miku's voice. What would we sing? What would I have her sing? The medium was in my imagination. I mostly play strings in an informal manner: banjo, guitar, ukulele, and the sanshin of Okinawa. I daydreamed about the songs I could play and record. There is a song "Juku no Haru" that is a duet between a man and a woman from Okinawa, and I wondered what it would be like to create a recording of that song—an impulse to sing through Miku and also sing with her. Miku had made her way into my imagination. I also spent more time watching and listening to Vocaloid songs, with renewed appreciation for the nuances that can be heard from skilled Vocaloid producers. Some people can really make Miku sing beautifully.

Vocaloid use also made me reflective of my own voice, which is not typically a focus of my attention. I noticed myself singing much more throughout the day. Miku inspired me to try things out with my voice, singing along as I entered her notes, testing things and hearing how they sound in my voice. And this singing lead me to appreciate how much I can do with my voice that I cannot do with Miku's voice. Just as I think about what I can musically express with an ukulele, I begin to think about sounds that Miku could bring to life. I end up recording a bunch of banjo and sanshin tracks to test Miku with, a collection of short pieces so that I can find out how we'll sound together.

I am already part of Miku, just as she is part of my musical self. I have not finished tuning a song to the level that I shared it yet, but I expect I will. At that point I will be adding to the public superstar Miku, another voice and mind that collectively instantiate Miku. Even with a limited exploration, Miku and I have learned music together.

Educational issues resonant with Miku

Music education is often mediated—such as the rich networks of pedagogic practice around sheet music or recorded media—but the media of music learning are only gradually coming to be more fully appreciated (Thibeault, in press, 2016). At this early stage it is impossible to be definitive about Miku as a medium for music learning, but several issues and questions are salient enough for initial consideration and further discussion.

Perhaps the most prominent issue rests with how educators understand Miku: as an extension of the voice, or as a replacement for the human voice. If one views Miku as an extension, perhaps she will be more warmly welcomed into learning situations, and used in combination with other forms of musicking. If, however, one worries that Miku will replace and perhaps discourage students from filling their lungs to sing out, then educators will likely be wary to include her. This year the makers of vocaloid introduced a simplified version of Miku, one that runs on Windows tablets, along with a curriculum designed to help students work with her.

A secondary consideration has to do with whether we try to enjoy Miku within traditional contexts for making music, rather than follow her to new contexts. The kinds of multimedia music making Miku invites is not yet common in schools that typically segregate the visual arts from music, but the creative work around Miku might entice more collaborative work among arts educators.

One additional consideration concerns community and its cultivation among educators. The rich online community surrounding Miku, full of fan contributions and comments, is one that educators might encourage students to enjoy. Such a community might allow students to find like-minded artists with which to collaborate. And, yet, on the other hand at present working with Miku would seem to often be a solitary affair, leaving the student disconnected from others. In other words, it is not enough to use the tool with Miku, a full exploration should include engaging with the larger mediated network that comprises her.

These questions have not been considered from an empirical standpoint, but they could be researched. As more music is mediated, and as more students take up tools through mobile phones, tablets, and other digital tools for making music—or should the early promise of artificial intelligence and virtual reality bear fruit in musical realms—we might find ourselves more often in the kinds of digital spaces Miku

already inhabits. It would be good if, as a profession, we have considered the kinds of musical experiences we could promote in these contexts.

References

Butler, J. (1990). *Gender trouble: feminism and the subversion of identity*. New York, N.Y: Routledge.

Conner, T. (2016). Hatsune Miku, 2.0Pac and Beyond: Rewinding and Fast-Forwarding the Virtual Pop

Star. In S. Whiteley & S. Rambarran (Eds.), *The Oxford handbook of music and virtuality* (pp.

129–147). New York, NY: Oxford University Press.

Haraway, D. J. (1991). *Simians, cyborgs, and women: the reinvention of nature*. New York: Routledge.

Kimagure. (2013). カラオケに異変!? ボカロを歌う若者急増. Retrieved from

https://www.youtube.com/watch?v=RE1tzklbKnQ

Netburn, D. (2011, December 20). Talking Twin Babies, Nyan Cat among YouTube's top videos of 2011.

Retrieved from http://latimesblogs.latimes.com/technology/2011/12/talking-twin-babies-nyan-cat-

and-friday-top-youtubes-most-watched-videos-of-2011.html

Sterne, J. (2003). *The audible past: Cultural origins of sound reproduction*. Durham, NC: Duke University

Press.

Thibeault, M. D. (in press). Learning with sound recordings: A history of Suzuki's mediated pedagogy.

Journal of Research in Music Education.

Thibeault, M. D. (2016). Understanding sheet music as a medium to expand pedagogic practice. *Journal of*

Music, Technology and Education, *9*(2), 209–222. https://doi.org/10.1386/jmte.9.2.209_1

Wat. (2015, April 15). 初音ミク V4 TEST：開発状況レポート＠2015.04.24 – SONICWIRE ブログ.

Retrieved June 15, 2017, from http://blog.sonicwire.com/2015/04/v4-test-2015-04-24.html

Zaborowski, R. (2016). Hatsune Miku and Japanese Virtual Idols. In S. Whiteley & S. Rambarran (Eds.),

The Oxford handbook of music and virtuality (pp. 111–128). New York, NY: Oxford University

Press.

初音ミク"ボーカロイドアルバム"が徳永を押さえ、初首位 | ORICON NEWS. (2010, May 25).

Retrieved June 14, 2017, from http://www.oricon.co.jp/news/76554/full/