Liveness and Labor and Hologram Singers

Presented at *Bone Flute to Auto-Tune: A Conference on Music & Technology in History, Theory and Practice*, University of California, Berkeley, April 24-26, 2014.

In the past five years, a dynamic music industry has emerged around artificially-voiced singing humanoid effects in Japan. Hatsune Miku is one such pop star with an artificially simulated voice

and a 3-dimensional effect that draws on Japanese Manga illustration. Hatsune sings the songs of unsigned, independent songwriters but far from being part of a small musical underground, she has a large fan following. In fact, Vocaloid music was the 8th most popular genre in Japan last year, with 17.4% of the under-40 Japanese listening to Miku.[1] The aim of this paper is to investigate the aesthetic, social, and political stakes that come into play in the culture surrounding digitally-voiced humanoid effects. I nominate the study of affective attachments in mediatized musical practices in place of a preoccupation with musical Liveness, a concept problematized in music scholarship of the past two decades.[2] I propose that attention to Hatsune

[1] Tokyo Polytechnic University Department of Interactive Media, "[Vocaloid Survey]" a press release (February 26, 2013). Accessible online http://www.t-kougei.ac.jp/guide/2013/vocaloid.pdf

[2] Philip Auslander, *Liveness: Performance in a Mediatized Culture* Second Edition (New York: Routledge, 2008 [1999]). Jonathan Sterne, *The Audible Past: Cultural Origins of Sound Reproduction* (Durham, NC: Duke University Press, 2003). Jason Stanyek and Benjamin Piekut, "Deadness: Technologies of the Intermundane," *TDR: The Drama Review* 54/1 (Spring 2010). Paul Sanden, *Liveness in Modern Music:*

fandom's debates about the politics of language describing their musical culture can help musicologists conceptualize post-live musical practices outside of the binaries of live vs. mediatized, alive vs. dead, composer vs. listener, performance vs. recording etc.

Corporate Musicking

The culture industry surrounding Hatsune Miku could be characterized by the overlap of production, reproduction, and consumption. Christopher Small's term "musicking" lends itself to an analysis of this polynodal network. To Small, "the act of musicking establishes in the place where it is happening *a set of relationships,* and *it is in those relationships that the meaning of the act lies*."[3] Recognizing that many of these relationships are between human and non-human (corporate, technological) agents, and finding a language for describing these attachments is central to my analysis.

In 2007, Yamaha released Vocaloid software, which realistically simulates an assortment of human voices. The Japanese company Crypton Future Media developed a vocal effect of a character named Hatsune Miku, her name translatable as "The First Sound of the Future." Immediately upon its release, a large culture of independent songwriters started using this particular effect because it

Musicians, Technology, and the Perception of Performance (New York: Routledge, 2012). Justin Williams, Rhymin' and Stealin': Musical Borrowing in Hip Hop (University of Michigan Press, 2013).

[3] Emphasis mine. Small's concept spans not only de facto musical sounds, but chiefly all "people taking part, in whatever capacity, in [a] performance; [modeling relationships] between person and person, between individual and society, between humanity and the natural world and perhaps the supernatural world," Christopher Small, *Musicking: The Meanings of Performing and Listening* (Hanover: Unviersity Press of New England, 1998), 13.

delivered true-to-life results without the fee of a hired vocalist. Songwriters uploaded their songs online, gaining non-commercial, free exposure for their own creative work. One of the biggest websites housing this work has been the Japanese server Piapro (a portmonteau of 'peer production') run by Crypton.

Once Vocaloid users created a large online repertory for the Miku *voice*, the gaming company SEGA created an animated, three-dimensional visual effect of the singer, an image of a *body* that could appear in concerts. Immediately, users started producing multimedia variations of Miku's image and sharing their work on the Crypton-run Mikubook portal – a kind of Youtube/Instagram of Vocaloid culture curated by an algorithm gauging popularity votes by visitors. The works range from Manga illustrations, music videos, 3D imaging dance videos created with freeware called MikuMikuDance, and, of course, Vocaloid songs. With a boost in Miku's popularity, Crypton started its own music label called KARENT, featuring more than 3500 tracks written for Hatsune Miku and other Vocaloid characters. Highest rated fan content drives a whole network of promotional products and commercial events, and highest rated songwriters get signed by KARENT.

Paul Théberge has written about the way different electronic music publications place value on very different and particular types of "meaningful musical behavior" – he contrasts illustrating articles with snippets of music notation vs. electronic circuit diagrams.[4] Vocaloid's interface doesn't use music notation, it translates text and diagrams of melodic contour straight into sound. A 2013 Vocaloid study by the Tokyo Polytechnic University celebrates the software for redefining musical literacy in fairly utopian language: "Although not everyone understands

[4] Paul Théberge, *Any Sound You Can Imagine: Making Music / Consuming Technology* (Hanover, NH: University Press of New England), 126.

music theory, everyone can create and nurture a melody within themselves, express this melody using this new musical language called Vocaloid, and create a song in the process."[5] These low barriers to participation – at least in terms of musical literacy – coupled with Crypton's encouragement of musical borrowing on Piapro and Mikubook, create a bouncy platform for jump-starting one's own composition. And using other composers' work to make one's own fosters a creative intimacy within the community of users, such that the common project of building Hatsune is as important as building a community.

Piapro and Mikubook users are not entitled to any compensation and only allowed to display their artwork non-commercially because Hatsune is protected by Crypton's copyright. All user-created artwork thus becomes property of Crypton.[6] This sounds harsh, but as a result, "unlike restrictive content tied up by [individual songwriters' copyrights], this environment is friendly to viewers and to creators who can freely participate in the creation of derivative works…"[7] Indeed, most of the work on Mikubook is collaborative, moving outside of modern copyright law and other restrictions governing the mainstream music industry.

---

[5] Professor Kuhara Yasuo as quoted in the press release by Tokyo Polytechnic University Department of Interactive Media. Translation by HIGHTRANCESEA, "Tokyo Polytechnic University Publishes Results of VOCALOID Study," *Vocaloidism.com* (March 7, 2013).

[6] The Terms of Service on Mikubook reads: "You grant us a perpetual, irrevocable, worldwide, non-exclusive, sublicenseable, transferable and royalty-free license to use, reproduce, prepare derivative works of, allow comments on, distribute, display, or otherwise exploit, in any way, your other Content or any portion thereof in any media now in existence or as hereafter may be developed." However, it also concedes: "You grant each other user of the Service a perpetual, non-exclusive, non-transferable royalty-free license to access, view, reproduce, prepare derivative works of, distribute, and publicly display your Content as and to the extent permitted by the Service."
http://www.mikubook.com/index.php/aboutus/termofService

[7] Tokyo Polytechnic University Department of Interactive Media.

Collaborative culture is very much at home in Japanese anime – a topic rendered book-length by MIT's Ian Condry in *The Soul of Anime* and Japanese cultural critic Hiroki Azuma's book on Japanese *otaku* (i.e. cute anime) culture.[8] In a recent critical edition, Azuma writes:

> It is quite ambiguous what the original is or who the original author is, and the consumers rarely become aware of the author or the original. For them, the distinction between the original and the spin-off products (as copies) does not exist, the only valid distinction for them is between the settings created anonymously (a database at a deep inner layer) and t=-he individual works that each artist has concretized from the information (a simulacrum on the surface outer layer).[9]

Interestingly, a Tokyo Polytechnic's survey shows that in spite of Crypton's clause about non-commercial creation, Vocaloid fans are very aware of the *value* of free content: more so than fans of other music genres, Vocaloid fans feel that the option of giving voluntary donations to providers of free content should exist.[10] Similarly, Japanese firms do not see open source as a giving up of control, but rather very much as a way of increasing their value, ultimately leading to profits, says U of Chicago's scholar of Japanese literature Michael K. Bourdaghs.[11] To Crypton,

---

[8] Ian Condry, *The Soul of Anime: Collaborative Creativity and Japan's Media Success Story* (Durham: Duke University Press, 2013). Hiroki Azuma, *Otaku: Japan's Database Animals*, Trans. Jonathan E. Abel and Shion Kono (Minneapolis, MN: University of Minnesota Press, 2009).

[9] Hiroki Azuma, "Database Animals," in *Fandom Unbound: Otaku Culture in a Connected World,* eds. Mizuko Ito, Daisuke Okabe, Izumi Tzuji (Yale University Press, 2012).

[10] Tokyo Polytechnic University Department of Interactive Media.

[11] Bourdaghs as quoted in Margaret Wappler, "I Sing the Body Electric: Teen Japanese-pop sensation Hatsune Miku proves corporeal reality isn't necessarily a prerequisite for fame," *LA Times Magazine* (June 2012). Accessible online. http://www.latimesmagazine.com/2012/06/i-sing-the-body-electric.html. Bourdaghs has also made the more thorough argument that Japanese musical commercialism, which is often represented as part of the 1990s boom, has a much older history going back to at least the 1930s and 40s, when the bourgeoning pop music industry at in significant part sustained Japan's military campaigns. Michael K. Bourdaghs, *Sayonara Amerika, Sayonara Nippon: A Geopolitical Prehistory of J-Pop* (New York: Columbia University Press, 2012), 4.

Hatsune was from the start designed as a promotional tool for the Vocaloid character suite, with users' free labor of voting on Hatsune media acting as built-in market research. Indeed, scrolling down the Mikubook, one finds clusters of promotional and commercial materials nestled comfortably among user content.

Vocaloid culture seems to represent a utopian-futuristic creative ethos uncharacteristic of the modern music industry, it even seems *countercultural* in its ostensible de-privileging of musical production. A number of journalistic articles refer to it as a *grassroots* culture. Conversely, very few articles on Vocaloid culture address the political tension between the intimacy of collaborative creation and free remix on the one hand, and the corporate weight of Yamaha, Crypton Future Media, KARENT, Piapro and SEGA on the other. The standout is a 2012 LA Times article by Margaret Wappler, who called Crypton's stance towards fan production "encouraging but watchful."[12] (This is mostly because there is a lot of *otaku* pornography, but also in the simple sense of Crypton protecting Hatsune's exact, copyrighted likeness.)

Jonathan Sterne and his colleagues at McGill wrote an excellent article critiquing the populist notion that the technology of the podcast (the term coming from the merging of iPod and broadcast) has allowed us to "escape from the vice-grip of commercial broadcasting." Sterne argues that podcasts simply advertise the idea that "one set of corporate products (a computer, iPod) might free us from all the others."[13] As much as I am a believer in open-source, I think the

[12] Wappler.

[13] In full, Sterne writes: "Independent podcasters share space with public and private media corporations looking to use podcasts as a profitable form of content delivery. Beyond its predictably 'Californian ideology', the podcasting-as-liberation argument also carries a little bit of corporate counter-culturalism with it, partaking of the knowing wink that one set of corporate products (a computer, an iPod) might set us free from all the others." Jonathan Sterne, Jeremy Morris, Michael Brendan Baker, Ariana Moscote

case of Vocaloid should be understood as similar, a commercial, $190-dollar commercial technology replacing the market technology of paying singers for labor, and the space on Mikubook shared by independent artists and corporations.

In Concert

The algorithmic curating of Piapro and Mikubook actually modulates how audiences *hear* Hatsune's stage concerts because it fosters the relationship between her likeness, way of moving, gesture, and the sound of her voice. Our perception of the Vocaloid voice as emerging from (or in any way linked to) the image of Miku's animated body is due to a famous auditory phenomenon called the McGurk effect, which describes the merging of auditory and visual information causing us to hear differently based on what we see.

In one of the only critical studies of Vocaloid to date, musicologist Nina Eidsheim excellently shows how Vocaloid essentializes race and gender.[14] She proposes that timbral hints at race, gender, age, or particular vocal personality should be considered nothing but **"**a set of inner *choreographies*.**"** Similarly, Suzanne Cusick has conceptualized the voice as a *transaction* **"**replicat[ing the] acceptance of patterns that are intelligible to one's cohort in a culture.**"**[15] Hatsune's cohort, so to speak, are the very timbrally specific female voices of *otaku* anime. In the 2010 Vocaloid suite, the tone of her voice runs the gamut of soft, sweet, dark, vivid, solid, and light, described by Crypton in even more gendered terms as 'gentle delicate voice', 'small voice',

---

Freire, "The Politics of Podcasting," *The Fibreculture Journal* 13 (2008). Accessible online http://thirteen.fibreculturejournal.org/fcj-087-the-politics-of-podcasting/

[14] Nina Eidsheim, "Synthesizing Race: Towards an Analysis of the Performativity of Vocal Timbre," *Revista Transcultural De Música* 13 (2009). Accessible online http://www.redalyc.org/pdf/822/82220946009.pdf

[15] Suzanne Cusick, "On Musical Performances of Gender and Sex," in *Audible Traces: Gender, Identity and Music,* eds. Elaine Barkin and Lydia Hamessley (Los Angeles: Carciofoli Verlagshaus, 1999), 32.

'heartbroken voice', 'cheerful voice', 'loud voice', and 'innocent and heavenly voice' respectively. According to Crypton developer Wataru Sasaki, long before Miku *appeared* (so to speak), her musical, timbral embeddedness in anime directed the tone quality of her voice.[16] Research into the design of synthesized vocal interfaces shows that The McGurk effect thrives with synthesized voices completed *especially* by animated – and not altogether humanoid – avatars, because the various musical insufficiencies of certain types of vocal synthesis can be filled in and at the same time explained by the synthetic-looking visual information.[17]

The visual effect seen by audiences at concerts is not a hologram – it is a multi-source projection with multiple depths of field on an extremely thin, angled, transparent plastic screen suspended over the stage that creates the illusion of three-dimensionality when viewed from the front. This is to say, it does not *act* in real time, it is not *live*, and it is not a *performance*: it is just a complex video accompanied by Vocaloid playback and an all-male band in the background. It is tempting to stretch the concept of music's Liveness to such simulations – for instance, musicologist Paul Sanden, most of whose current work deals with categories of Liveness, established a category of "virtual liveness," which describes precisely such cases where "conditions for liveness (be they corporeal, interactive, etc.) do not *actually* exist. [What exists is only a] perception of liveness that is largely created *through* mediatization."[18] I am disinclined to

[16] "We didn't base our decision on voice quality alone, but also considered things like visual image and impact." Wataru Sasaki in Tara Knight, *Mikumentary* (2013). Accessible online http://taraknight.net
[17] Clifford Nass and Scott Brave, *Wired for Speech: How Voice Activates the Human-Computer Relationship* (Cambridge, MA: MIT Press, 2005), 65.
[18] Paul Sanden, *Liveness in Modern Music: Musicians, Technology, and the Perception of Performance* (New York: Routledge, 2013), 11. The full list of Sanden's types of Liveness comprises "Temporal Liveness" (the first utterance), "Spatial Liveness" (the space of this utterance), "Liveness of Spontaneity" (utterance demonstrating human spontaneity), "Corporeal Liveness" (evidence of a sounding body), and "Interactive Liveness" (interactivity between performer/audience, or performer/performer), and "Virtual Liveness."

such stretching of the term Liveness, because it stubbornly refuses to accept that non-Live events *can* and *do* produce the same affective attachments to sound, to celebrity, and to community. I recognize a related problem in Anglophone journalists' overwhelming use of the term hologram to describe the three-dimensional visual effect. This incorrect term conjures up a techology that has been established by Hollywood as a technology mediating presence (presence in another time and/or place) rather than a technology of virtuality. The uses of "Liveness" and "hologram" seem to reveal a certain anxiety about the real that is not shared by Vocaloid fans.

After Philip Auslander, I reject the often-invoked binary of Live versus mediatized performance, which is really just a discursive technology reinforcing the Romantic opposition of culture and technology**.** Understanding sound as a force and therefore as an intimate event of touching, and understanding musical culture as the product of labor, and therefore a network of social intimacies are two significant steps towards deflating the ideology of Liveness.

I find it more productive to understand Hatsune's concerts as a network of affective attachments. The most famous and proliferated music videos of Hatsune show the audience waving glowsticks and singing along. Many of her videos are actually produced by SEGA, filmed with roving light-sensitive cameras that make Hatsune appear larger and brighter than the effect seen by the audience – then again, most fans' experience of Hatsune is *precisely and only* through such videos. In waving glowsticks, which are typical in J-pop, the audience extends Hatsune's largeness and brightness, and in singing along, her voice.[19] Memorizing songs, writes intimacy scholar Lauren Berlant, creates "spaces with are produced relationally: people and institutions

[19] Fans' renditions of Vocaloid songs in karaoke bars and karaoke videos, often in costume – i.e. human vocalists' imitatation of synthesized voices and images – are one of the most fascinating culminations of Vocaloid culture.

can return repeatedly to them and produce *something*, though frequently not history in its ordinary, memorable, or valorized sense."[20]

In her book *Touching Feeling,* Eve Sedgwick draws on a Buddhist expression "pointing at the moon," where the act of pointing implicates both the finger and the moon.[21] This relational attachment plays out in a hyper-literal way with the audience's glowsticks, which look like they are made of Hatsune's flesh. Sedgwick collapses the distinction between somatic and psychic feeling into a notion she calls affective "besideness," which comprises a wide range of non-binary relations. (Berlant calls something similar "aesthetic of attachment.")[22] Interpretation at the level of affect allows us, as Sedgwick put it, "to enter a conceptual realm that is not shaped by the [dynamics of excess vs.] lack, subject vs. object, and means vs. ends."[23] Or, to name a few other dualisms that collapse in Hatsune's concerts, human vs. non-human, artist vs. audience, and live vs. mediated….

Sound is an affective medium *par excellance*: as a texture, we sense it as a force imprinted on our bodies; as a text, it means affectively. The intimacies activated in the concert preeminently include non-normative desire – an intimate attachment to an image and voice accessible through the fictional timbral cohort of *otaku* Anime. The largest consumer demographics of Vocaloid are 40% of Japanese teenage girls and 28% of women in their early

[20] Lauren Berlant, "Intimacy: A Special Issue," *Critical Inquiry* (Winter 1998), 285.
[21] Eve Kosofsky Sedgwick, *Touching Feeling: Affect, Pedagogy, Performativity (*Durham, NC: Duke University Press, 2003), 171.
[22] "Besideness," writes Sedgwick, "comprises a wide range of desiring, identifying, representing, repelling, paralleling, differentiating, rivalling, leaning, twisting, mimicking, withdrawing, attracting, aggressing, warping, and other relations." Sedgwick, 23.
[23] Sedgwick, 21.

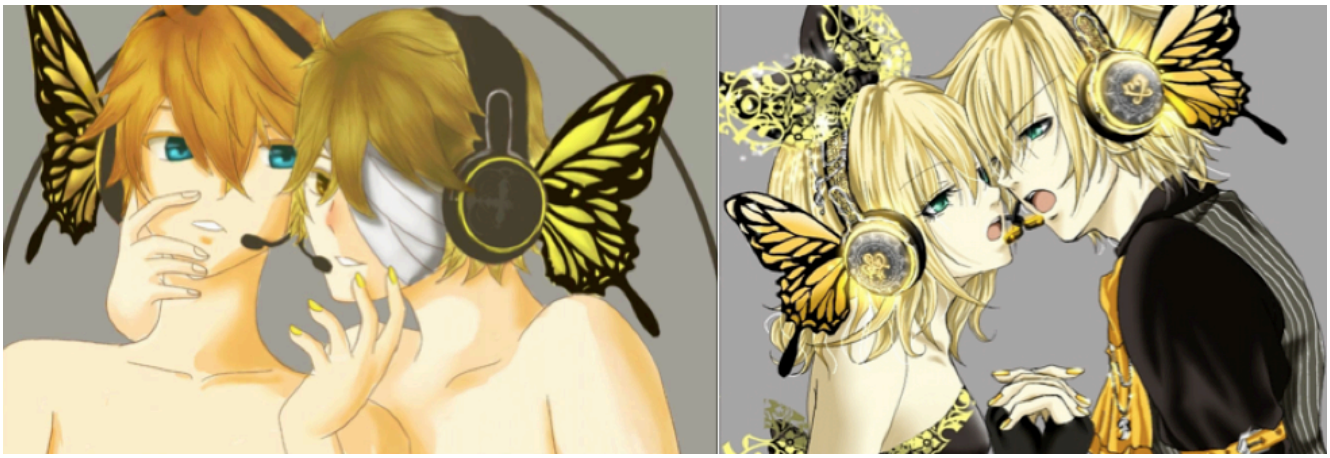
twenties.[24] To them, Hastune is not just an effect: she is both a projection of who Japanese teenage female fans want to be, an object of their desire, and a repository of their cultural affiliation. Miku culture gives teenage girls a sense of belonging not only in virtual *online* communities but also virtual *fictional* communities. Vocaloid fans commonly understand Hatsune as all – technological, virtual, and real (several fans articulate this dual understanding in a short, surreal documentary on Hatsune by UCSD's Tara Knight.[25]) and I think Sedgwick's articulation of "besideness" really captures this logic of AND that replaces the Cartesian logic of EITHER/OR.

Vocaloid composers often reflect their life experience and desires in a specific genre of autobiographical so-called Chara songs sung by Miku in first person. The most famous is "Melt" about a 16-yo girl in love with a boy, but if you dig around the Mikubook, *many* kinds of teenage experience and are represented: gay, goth, depressed, suicidal etc. On Yahoo Answers, an anonymous user asks: "Is Miku Hatsune a lesbian?," The best voted answer is telling: "Yes. Vocaloids doesn't have oficial defined personality or relationship. So, yes, if you want, she can be a lesbian," referring the asker to a famous song called "Magnet" about female-female romance between Hatsune and Luka. (The returning line of this song rings "Please make me believe that this is not a sin.")

[24] Tokyo Polytechnic University Department of Interactive Media.
[25] One fan says: "Yes, we all know that she is not an individual that exists, but seeing her disappearing from the stage like that makes us recognize that she has a virtual existence. She has a strong presence on the stage but we feel that she is also virtual." [Knight's translation.] Knight.

Manga and Anime

I find that the virtual genealogy of Vocaloid characters finds its most radical articulation in duets, where the characters' familial, generational, and romantic relationships change from song to song, outlining, again, a logic of nonexclusive multiplicity. One songwriter famously undertook the conceptual project of setting this love song "Magnet" for all possible vocal pairings of Vocaloid characters.



Bruno and Oliver, Spanish and English Vocaloid characters; Kagamine Lin and Kagamine Ren, Japanese Vocaloid 'twin' characters, who are however not necessarily siblings

The multiplicity of sexual and musical possibilities within every single character is very much enabled by their position within the Manga and Anime genres: scholars Tania Darling and Sara Cooper single out manga comics as a genre that "openly [deals] with a wide variety of gender expression as well as a range of manifestations of same-sex love."[26] In the musical practice of vocal remix, sexuality is remixed too.

Teenage girls' overwhelming participation in Miku culture situate the practice not only within Manga and Anime at large, but also specifically within the lineage of an older practice

[26] Tania Darling and Sara Cooper, "The Power of Truth: Gender and Sexuality in Manga," in *Manga: An Anthology of Global and Cultural Perspectives* (New York: Continuum, 2010), 157.

called *dojinshi*, where young Japanese girls exchange notebooks with variations on Manga drawings and they contribute to each other's notebooks. Scholar of free culture Lawrence Lessig writes about this practice having involved specifically "working-class girls," who often "have an incredibly active amateur life" making *dojinshi* next to their day job or student life.[27]

Conclusion

I continue doubting the extent to which Hatsune fandom is really teenage girls' exercise of cultural control, when its boundaries are to a large extent delimited by Crypton. However, the networks and options of non-normative intimacies between teenage girls and their virtual idol, teenage girls and Vocaloid technology, and teenage girls and other teenage girls, do defy the social configurations of both, teenage femininity and musical fandom.

I am also convinced by the *critical relationship, a critical stance* Hatsune fans have to their musical culture: with discussions about Hatsune's multiplicity of sexual identities, conceptualizations of Hatsune as BOTH a technology and a virtual subject, and even critical discussions about the language that should be used to refer to programming her voice. Some use the verb "to tame" (chokyo suru), some the expression "to tune a voice" ("chosei").

I got interested in this topic largely because of the technologies used but ended up developing my own multiple attachment to Hatsune as a virtual subject and especially as a gendered subject, which culminated in my academic leaning to refer to her as Hatsune (and a

[27] The breakdown of demographics participating in the various kinds of musicking that comprise Miku culture is more complicated and requires further study. Lessig quotes anthropologist Mimi Ito, who identifies "the AMV [Anime Music Video] scene [as] dominated by middle-class [W]hite men." Lawrence Lessig, *Remix: Making Art and Commerce Thrive in the Hybrid Economy* (New York: The Penguin Press, 2008). For a history of teenage girls' involvement in 1990s Manga *dojinshi*, see Sharon Kinsella, "Japanese Subculture in the 1990s: *Otaku* and the Amateur *Manga* Movement," *Journal of Japanese Studies* 24/2 (Summer 1998).

thought that I should upgrade *that* to Hatsune-san even though my knowledge of Japanese is null).

Singing hologram effects have started appearing across the United States.[28] Not only, but also because of this, music scholars should look to fans' debates about the politics of language describing mediatized music in order to articulate the plurality of our intimate relationships to these musical practices without resorting to structural binaries and vocabulary that exhibits our anxiety about the real.

[28] A 2Pac Shakur effect was used at Coachella music festival in 2012, an Ol' Dirty Bastard effect was used at Rock the Bells music festival in 2013. Most recently, Janelle Monae and M.I.A. effects appeared at one another's simultaneous performances on opposite coasts in April 2014. Hologram effects used in the U.S., which do not always employ Vocaloid voices, consistently refer to an imaginary presence in another time or place, riffing on the idea of a singer's resurrection or the 'live' videocast of a performance happening elsewhere. The anxiety about truly virtual subjecthood of hologram effects in the U.S. is a topic for another paper.