


RESEARCH ARTICLE

## Reconfiguring Voice in *The End*: Virtuosity, Technological Affordance and the Reversibility of Hatsune Miku in the Intermundane

Jessica Tsun Lem Hui 

University of Cambridge  
Email: [jtlh3@cam.ac.uk](mailto:jtlh3@cam.ac.uk)

### Abstract

This article explores the technological affordances of vocal production software in performance through a case study of Shibuya Keīchirō's *The End* (2012). In the performance of this 'humanless opera', desires for pliability and fantasies of control are realised through the affordances of a singing voice synthesis software known as Vocaloid. By reflecting on *The End*'s thematic focus on death and existentialism and on notions of vocal virtuosity, and by exploring the socio-technical processes by which the protagonist, virtual pop star Hatsune Miku, was constructed, the article provides an alternative narrative to vocal production and intermundane collaboration as it relates to the fluid and reversible configurations between voices, bodies and technologies in performance.

**Keywords:** Vocaloid; Hatsune Miku; Virtuosity; Singing Voice Synthesis; Music Technology

I can't recognize what is 'you' and what is 'me'  
I can't see a thing  
My lips, where am I?  
What is this voice of mine?<sup>1</sup>

So go the final lines of this 'humanless opera' in which a virtual singer reflects on the conditions of her own existence, and by extension, her voice. On stage, at the Théâtre du Châtelet in Paris in 2013, a sixteen-year old pop star with twin aqua ponytails is suspended in mid-air, her pixelated limbs dangling lifelessly by her side. The girl's mouth does not move as the song unfolds: there is no need, because this is a voice that arises from binary code, commanded not by any internal physiological apparatus, but by the stroke of a computer key. And yet, over the preceding eighty minutes, the audience has learned to pair this uncanny voice with the body of the virtual protagonist floating on the screens before them. As she vanishes behind the proscenium and the curtains close, however, the audience is left wondering: where does the voice come from?

For at least a century-and-a-half, audio technologies have served to disrupt the relationship between the singular voice and its body. Today, voice is entwined with the digital processes that have effectively transformed not only what it sounds like, but how it is

<sup>1</sup> Keīchirō Shibuya, *The End* (2012), DVD. Directed by YKBX (Tokyo: Sony Music Entertainment Japan, 2013). *The End* was performed in Japanese with English surtitles.

listened to and engaged with. Over the past thirty years, the subdiscipline of contemporary voice studies has emerged in response to the rapid development in voice-related technologies. It engages methodologies from a raft of disciplines, from vocal pedagogy to science and technology studies (STS), in order to destabilise and interrogate assumptions about the ontologies of voice.<sup>2</sup> In musicology, discussions about the fluidity of voices and bodies emerged in the late 1980s and early 1990s principally through feminist and queer theory.<sup>3</sup> If, as Martha Feldman put it, ‘in the 1970s and 1980s voice in musicology was typically as flat as a sheet of typing paper’, problematising the relationship between singers, vocal timbre and the performing body would seem a welcome development.<sup>4</sup> For all that it signifies and all the ways it operates, voice is fluid, multi-dimensional and confusing – and its technological and mediated versions even more so. With this in mind, it is unsurprising that posthumanism and voice studies have fused into a curious configuration, in the service of a common goal: voice, particularly the technological voice, can challenge the ideologies of a unified or enlightened body.<sup>5</sup>

This article begins from the premise that voices, bodies and technologies exist in malleable, inter-dependent and multifactorial configurations in performance. It illustrates the fluidity and fragility of these entangled structures through the case study of *The End*, and its protagonist, the virtual singer Hatsune Miku. By reflecting on *The End*’s narrative development and exploring the process by which Miku’s voice was produced using the Vocaloid singing voice synthesis software, it proposes a reading of technological voices and hologrammatic bodies in musical performance that troubles both the experience of the performing human voice and narratives of technological determinism relating to emerging voice technologies. The article suggests that synthesised voices disperse and reconfigure voice’s expressive dependence, casting its production in a variety of agential directions.

*The End*, dubbed the ‘world’s first humanless opera’,<sup>6</sup> centres on Miku, who ventures into the abstract world of posthuman existentialism, confronting a corrupted copy of herself. Look-a-Like, as this creepy simulacrum is known, compels Miku to obsess over and glorify humanity’s imperfect and mortal state – a state unattainable for an apparently infallible, digitopian idol. Miku is subjected to a series of tortuous experiments by an unknown force (she is drowned, poisoned with gas and stabbed), yet she remains bound to the world of digital immortality. Only by coming to terms with her undesirable state of perfection is Miku able to break through the screens that imprison her, crossing over to our realm, a place in which the audience imbues her with meaning. As Miku and her clone embrace, she is finally freed, passing away in peace. Ostensibly, then, *The End*

<sup>2</sup> Nina Sun Eidsheim and Katherine Meizel, eds., *The Oxford Handbook of Voice Studies* (New York, 2019), xiv.

<sup>3</sup> Carolyn Abbate, *Unsung Voices: Opera and Musical Narrative in the Nineteenth Century* (Princeton, 1996); Wayne Koestenbaum, *The Queen’s Throat: Opera, Homosexuality, and the Mystery of Desire* (London, 1993); Catherine Clément, *Opera, or The Undoing of Women* (London, 1997).

<sup>4</sup> Martha Feldman (convenor), Colloquy ‘Why Voice Now?’, *Journal of the American Musicological Society* 68/3 (2015), 653–8, at 655.

<sup>5</sup> Many scholars have explored the performance of the synthesised or digitally enhanced voice through a post-humanist lens with this purpose. See, for example, Joseph Auner, ‘“Sing It for Me”: Posthuman Ventriloquism in Recent Popular Music’, *Journal of the Royal Musical Association* 128/1 (2003), 98–122, Sarah A. Bell, ‘The dB in the .db: Vocaloid Software as Posthuman Instrument’, *Popular Music and Society* 39/2 (2016), 222–40, and Alexander G. Weheliye, ‘“Feenin”: Posthuman Voices in Contemporary Black Popular Music’, *Social Text* 20/2 (2002), 21–47.

<sup>6</sup> Leon TK, ‘Approaching *The End*: Face-to-Screen with the World’s First Humanless Opera’: [www.leontk.com/2014/01/09/approaching-the-end-face-to-screen-with-the-worlds-first-humanless-opera/](http://www.leontk.com/2014/01/09/approaching-the-end-face-to-screen-with-the-worlds-first-humanless-opera/) (accessed 17 October 2020). *The End*’s composer, Shibuya Keichirō, has repeatedly referred to the work as an ‘opera’, but the term is employed very loosely in this article as a means to refer to *The End* as a work which adheres to some structural conventions of opera (such as distinct recitative and aria sections) while rejecting other stylistic, aesthetic and performative conventions that would form what would traditionally comprise an opera performance.

can be read as a cautionary tale about the dangers of chasing a digital utopia, where the essence of humanity (and, by extension, the performing voice) is lost among immortal synthesians and their bootleg clones.

*The End* (2012) was composed by Shibuya Keichirō<sup>7</sup> and performed by Miku, a virtual performer with a digitally synthesised voice. She was ‘born’ in 2007, created by Itō Hiroyuki (the CEO of media company Crypton Future Media) in collaboration with Yamaha and their singing synthesis software, Vocaloid. Today Miku is an international pop-star, but back in 2007 she was merely a singer in a box, a digital voice instrument in a program. In 2012, Shibuya was commissioned by the Yamaguchi Center for Arts and Media to write an opera for Miku. YKBX was brought on as director and visual artist, alongside Vocaloid producer Pinocchio-P. *The End* premiered at the Yamaguchi Center for Arts and Media on 1 December 2012, making its European debut at the Théâtre du Châtelet in Paris almost a year later, on 12 November 2013.<sup>8</sup> *The End*’s official website describes the opera as a piece that ‘aims to dissect and transform opera with the goal of a radically new space/time creation that is neither traditional nor avant-garde’.<sup>9</sup> The opera is notable for its total absence of human actors; instead, the expressive nuances of every performance – voices, actors and *mise en scène* – are largely digitally reproduced.<sup>10</sup>

What is at stake in *The End* is the performance of the human. Living, breathing bodies are usurped by virtual avatars, and human voices – pitch perfect and digitally malleable in the Vocaloid software – are freed from their fleshy constraints. The increasingly common presence of the hologrammatic performer (notably, in operatic circles, BASE Hologram’s revival of Maria Callas in 2018<sup>11</sup>) has been a source of intrigue for media critics, understood as socio-technical phenomena in this context.<sup>12</sup> Yet, there are still areas to be explored relating to how these technologies play out in explorations of human empathy in the centring of virtual characters and synthesised voices, and in how Miku’s malleability as a cultural icon is simultaneously situated in a variety of narratives – even while she continues to be known predominantly as a popular virtual idol. Relocating the voice in this context also redefines the experience of performance and the performed: phenomena such as virtuosity or vocal failure are adapted from a different set of conditions that appear to turn focus away from the human body. Put another way, Miku and *The End* encourage us to reconfigure our expectations of performance by reflecting on the voice as a manifestation of the body in harmony with emerging digital technologies.

By understanding voices as ontologically separate from the technologies that enhance them, the concept of virtuosity reveals itself to be a paradox: it alludes to a transcendence of the biological body, but must also appear exclusively as its product. Technologies that ‘unnaturally’ enhance the voice (from amplification to Auto-Tune) are seen as an illegitimate augmentation of the body’s intrinsic capacities. Intense vocal training, the development of a rich timbral palate and the cultivation of higher formants and Wagnerian

<sup>7</sup> English translations of Japanese names in the body of the text follow the modified Hepburn system and are written in the customary order of surname, then first name. References follow the order of first name, then surname.

<sup>8</sup> ATAK, ‘Vocaloid Opera – THE END’: [http://atak.jp/en/theater/the\\_end/](http://atak.jp/en/theater/the_end/) (accessed 17 October 2020).

<sup>9</sup> ATAK, ‘Vocaloid Opera – THE END’.

<sup>10</sup> Technically, there is a human performer. During the opera Shibuya plays the keyboard, partially obscured behind a series of screens. He is, however, the only human on stage and he is not always visible to the audience.

<sup>11</sup> Tom Huizenga, ‘Raising The Dead – And a Few Questions – With Maria Callas’ Hologram’, NPR (6 November 2018): [www.npr.org/sections/deceptivecadence/2018/11/06/664653353/raising-the-dead-and-a-few-questions-with-maria-callas-hologram](http://www.npr.org/sections/deceptivecadence/2018/11/06/664653353/raising-the-dead-and-a-few-questions-with-maria-callas-hologram).

<sup>12</sup> David Rowell, ‘The Spectacular, Strange Rise of Music Holograms’. *The Washington Post* (30 October 2019): [www.washingtonpost.com/magazine/2019/10/30/dead-musicians-are-taking-stage-again-hologram-form-is-this-kind-encore-we-really-want/](http://www.washingtonpost.com/magazine/2019/10/30/dead-musicians-are-taking-stage-again-hologram-form-is-this-kind-encore-we-really-want/).

stamina all signify opera's dependence on innate athleticism. One might thus locate captivation with opera in the struggle of the performing body against its own limits. Empathy becomes wrapped up in spectacle, where the body may drastically fail, or it may succeed beyond expectation. The key point here is that virtuosity is predicated on the potential for failure, which is possible only because the body, however much it is trained or controlled, is always susceptible to malfunction within the confines of socio-cultural expectation.

While Miku is not in possession of a biological body and her voice is not physiologically produced, she is still limited, in a sense, by the humans and technologies that facilitate her. This thinking sits in line with the concept of 'the affordance of things'.<sup>13</sup> Influenced by Gibson's theories on the psychology of perception (that all organisms are oriented to objects in their environment relative to their affordances or 'the possibilities that they offer for action'), sociologist Ian Hutchby conceives technological affordance to mean both the 'functional [and] relational aspects of an object's material presence in the world'.<sup>14</sup> Affordances are 'functional in the sense that they are enabling, as well as constraining. ... Certain objects, environments or artefacts have affordances which enable the particular activity while others do not' – in other words, they are relational in the sense that the specific functions they have may not be immediately obvious, but are instead revealed or concealed in different environments.<sup>15</sup> As will be explored below, affordance thinking reveals Miku to be susceptible to failure and degradation. Moreover, it shatters the popular narratives of virtual performers as perfect digitopian replacements for human performance. While it is true that Miku has vocal capacities beyond human ability (singing at rapid speeds with perfect enunciation, for example),<sup>16</sup> she is also still prone to technical malfunctions – from code glitches to power outages – that disrupt the voice.

An international sensation in the popular music world, but a strange phenomenon on the operatic stage, Miku's existence seems to take on a new level of meaning in *The End*. Emerging technologies have often played a role in mediating and redefining opera: 'The advent of powerful new audio technologies', Linda and Michael Hutcheon observe, 'has distanced audiences and therefore has made the disembodiment and subsequent fetishizing of the operatic voice a particularly modern issue.'<sup>17</sup> Miku's operatic debut, I will suggest, reveals two seemingly opposing arguments: she consolidates opera's expressive dependence on the listening and performing body while also allowing us to conceive of the voice as a complex mashup of technology and human.

### Vocaloid and distributed subjectivity

Vocaloid, a singing voice synthesis (SVS) software, was developed by the Yamaha Corporation in 2004, to create speech or singing directly from a desktop computer. The program allows users to generate singing through a score editor interface: first, by allocating pitch and duration to create a melody, and second, by assigning phonemes to each syllable. Various vocal techniques can be simulated through the program, such as the level of vibrato, velocity, attack or dynamic range. More recent iterations of the software even facilitate a range of vocal timbres: 'growl', 'breathy', 'rich', 'bright' and 'ambient' are a few

<sup>13</sup> Ian Hutchby, 'Technologies, Texts and Affordances', *Sociology* 35/2 (2001), 441–56, at 447.

<sup>14</sup> Hutchby, 'Technologies, Texts and Affordances', 447.

<sup>15</sup> Hutchby, 'Technologies, Texts and Affordances', 448.

<sup>16</sup> This is evidenced by one of Miku's most famous songs, 'The Disappearance of Hatsune Miku', by Vocaloid producer cosMO, which hits 240 beats per minute. YouTube, '[60fps Full 風] The disappearance of Hatsune Miku -DEAD END-初音ミクの消失 DIVA Dreamy theater English Romaji': [www.youtube.com/watch?v=5qkTpjAhywg](https://www.youtube.com/watch?v=5qkTpjAhywg) (accessed 7 October 2021).

<sup>17</sup> Linda Hutcheon and Michael Hutcheon, *Bodily Charm: Living Opera* (Lincoln, NE, 2000), 10.

of the singing styles that a user can choose from for more expressive or natural-sounding voices. Vocaloid opens the possibility of synthesising and fine-tuning vocal performance outside of the biological body, literally transforming the voice, through hypermediation, into a virtual musical instrument. The consequences for cultural production are stark. As Daniel Black puts it, Vocaloid allows singing to be ‘mass-produced and used to create vocal performances that have never passed the lips of any living human being’.<sup>18</sup> This transforms the idea of a voice’s ‘body’ or producer: it becomes a strange coalition between user and interface, a dynamic of power that may be likened to a puppeteer and their marionette. As we will see, for all that Miku is an icon of an ostensibly democratised collaborative culture, she is also a symbol of a stereotyped femininity, a socio-technical construct that remains malleable under fantasies of control and commodification.

The most crucial component of Vocaloid’s architecture is the voicebank. This contains a selection of vocal fonts, essentially a collection of individual phonemes recorded from a human voice. This ensures that the voice will sound distinct (a vocal font will take on the likeness of its progenitor voice). Miku’s voicebank was generated using samples of the actress and singer Fujita Saki’s voice. While developments in the speech- and singing-synthesis world are closing the gap between the human voice and its digital sibling, they have not yet achieved the nuanced qualities of human singing. Vocaloid is not a purely synthesised voice, but recorded samples of a human voice transformed by varying levels of technological mediation – a hypermediated voice.

Appreciating the dynamics of Vocaloid production requires a thorough understanding of everything from the software’s architecture to the business tactics of Yamaha and Crypton Future Media, which cannot be covered within the scope of this article.<sup>19</sup> Instead, we will home in on the complex distributed subjectivities that Miku implies. I borrow the term ‘distributed subjectivity’ from Anahid Kassabian: ‘distributed subjectivity suggests a vast field, rather than a group of subjects or an individual subject, on which various connections agglomerate temporarily and then dissolve again. This field is significantly constructed through and with music.’<sup>20</sup> We can apply this theory to Vocaloid quite simply. Some Vocaloid fans have taken to transcribing Fujita’s original songs for Miku, and they provide insight into the approximation between the voice actor and the corresponding Vocaloid. For example, one can find many Miku covers of the song ‘Crystal Quartz’ on YouTube, where she sings in a variety of timbres and styles.<sup>21</sup> All of these covers highlight how the user’s variation in skill (musical proficiencies, software fluencies etc.) and the software’s affordances drastically impact the outcome of the song, despite all of them employing a voice from the same biological body. Furthermore, older versions of Miku’s software sound discernibly different from her newer voicebanks, particularly to those who are already well acquainted with the nuances of Vocaloid’s uncanny timbre.

<sup>18</sup> Daniel Black, ‘The Virtual Idol: Producing and Consuming Digital Femininity’, in *Idols and Celebrity in Japanese Media Culture*, ed. Patrick W. Galbraith and Jason G. Karlin (Basingstoke, 2012), 209–228, at 223.

<sup>19</sup> English-language scholarship on Vocaloid is limited, but the work of scholars such as Nick Prior (2018, 2021), Nina Sun Eidsheim (2009, 2019) and Daniel Black (2012) have brought the technology to the attention of musicology in recent years. Prior has argued that Vocaloid music and its fan culture developed out of a complex network of agencies and collective identities, from software engineers to DIY musicians, thereby confusing the typical notion of voice as a direct manifestation of individual agency or identity, what he terms ‘vocal assemblage’. Nick Prior, ‘On Vocal Assemblages: From Edison to Miku’, *Contemporary Music Review* 37/5–6 (2018), 488–506.

<sup>20</sup> Anahid Kassabian, *Ubiquitous Listening: Affect, Attention, and Distributed Subjectivity* (Berkeley, 2013), 26.

<sup>21</sup> Compare, for example, Fujita’s song ‘Crystal Quartz’ (YouTube, ‘Crystal Quartz – Fujita Saki’, <https://www.youtube.com/watch?v=ctlgHyqcgQo> (accessed 17 October 2020)) with a Miku cover uploaded by a user named DarkAngelAlhena ([Vocaloid] Hatsune Miku – Crystal Quartz + mp3 🎵: [www.youtube.com/watch?v=ycmewL1tLVg](http://www.youtube.com/watch?v=ycmewL1tLVg) (accessed 17 October 2020)).

Since her initial release in 2007, Miku has been updated through the release of several voicebank upgrades. Each serves to develop the capacities of Miku's performance, facilitating improved expressiveness and offering multiple languages and a wider variety of timbres.<sup>22</sup> Fujita is required to re-record each voice sample, meaning that every update of Miku's voice is an entirely new snapshot of Fujita's voice from a specific moment in time. While an apparent improvement on the last, each successive version is also susceptible to a different set of bugs and technical issues, as demonstrated by the wealth of forums dedicated to resolving users' problems.<sup>23</sup> Such affordances depend not on Fujita's skill as a voice actress or singer, but on the Vocaloid technology itself and the fluency of the software user. Thus the produced voice becomes an assemblage reliant on the distribution of co-dependent subjectivities. In other words, there is no single source for Miku's vocal production.

Although Miku is best understood as a complex, distributed entity, it is still possible to delineate her identity to some degree. Although she isn't human, Miku alludes informally to a kind of idiosyncratic existence. I could pick her out of a crowd. I can say that I've seen and heard her 'live' in concert. In this sense, she might be as 'authentic' as any other pop persona. The point is that she is not just an abstract set of concepts. You know her when you encounter her. She has an image that corresponds to a distinct voice and identity. And yet, when I speak of her I really evoke what facilitates her existence: the materialities of code and of Fujita's body, the skill of Vocaloid's programmers, the fluencies of the software user, and the support of the Miku fanbase, which all play into this distributed subjectivity. This concept allows Vocaloid to retain its position as complex and hypermediated, resisting any confinement to a fixed or unified body. Instead, the software, within its technological parameters, redefines bodies, agencies and information flow as fluid and open to reconfiguration through a series of ones and zeros. Vocaloid voices arise out of interactions with the software, creating a particular mode of vocal production that is distinct from straightforward bodily emission or even technologically mediated voice (which assumes that voices are physiologically produced and then subjected to technological effects).

Although a comparison of the voices of Fujita and Miku is appealing, we should not value Vocaloid primarily for its capacity for similitude. Far more pertinent are the differences that the software affords. Vocaloid allows the voice to be controlled to a microscopic degree of precision that is unavailable to the body. It is hardly surprising that a wealth of digital technologies in the past three decades have aimed to tame or control the voice – entrained to the theoretical norms of Western art music – when it breaks, or sings out of tune or out of time. In an article on the emergence of Auto-Tune, Catherine Provenzano points out that the voice is a particularly stubborn musical instrument and, given that 'the throat has no frets', a technology that can reliably tune the voice is both completely unsurprising and yet, to many, deeply unethical.<sup>24</sup> Even today, our digital technologies only 'fix' the voice post-production. Vocaloid is particular in

<sup>22</sup> After a period of time, each voicebank version will be retired and withdrawn from the market. The developer of Hatsune Miku's voicebank, Crypton Future Media, will terminate the generation of new serial codes for that product after the establishment of a newer version. Miku V2, for example, was released on 31 August 2007, and retired at the end of March 2016. VOCALOID, '「VOCALOID2 ライブラリインポート」サポート終了のお知らせ': [https://web.archive.org/web/20180307022738/https://www.vocaloid.com/articles/support\\_info\\_v2\\_import](https://web.archive.org/web/20180307022738/https://www.vocaloid.com/articles/support_info_v2_import) (accessed 17 October 2020).

<sup>23</sup> See, for example, the r/vocaloid forum on Reddit (Reddit, 'r/Vocaloid': [www.reddit.com/r/Vocaloid/](http://www.reddit.com/r/Vocaloid/) (accessed 18 October 2021)) and 'The Clinic' on the Vocaverse fan site, which allows users to 'Get help with any errors [they] experience with ... the Vocaloid editor, or other vocal synth software.' VocaVerse Network, 'The Clinic': <https://vocaverse.network/forums/the-clinic.54/> (accessed 18 October 2021).

<sup>24</sup> Catherine Provenzano, 'Auto-Tune, Labor, and the Pop-Music Voice', in *The Relentless Pursuit of Tone: Timbre in Popular Music*, ed. Robert Fink, Melinda Latour and Zachary Wallmark (New York, 2018), 159–182, at 160.

the sense that correction, editing and synthesis happen within the same interface, suggesting for Provenzano a ‘fragile and fluid ontology of voice that demands constant re-enactment of its parameters and position – even more so in moments of acute confrontation between categories of human and machine labour’.<sup>25</sup> Vocaloid is just one of many digital technologies that encourage a reconsideration of the singing voice and its parameters. It plays into the development of so-called ‘intelligent’ voice editing and synthesis technologies,<sup>26</sup> as well as larger, ongoing investments in digital tools and their supposed democratising and streamlining of music production.<sup>27</sup> With Vocaloid, the pliable voice is no longer reliant on a physiological model but instead becomes a digital potentiality – a simulation.

### Vocal control and the contradictions of virtuosity

The postmodern voice is prized for its ability to be cloned, reconstituted, relocated, remediated and stored. James Q. Davies contends that today the voice ‘only apparently achieves optimal transcendence when it has been de-essentialized or bit-mapped’.<sup>28</sup> In *The End*, the recorded-synthesised voice – a ‘digitopian dream’, as Ken McLeod calls it – threatens with the excessive perfection and pliability of the digital.<sup>29</sup> When the labour of voice resides in the declarative clarity of code and key as opposed to the physiological strain of our speech organs, the basis of the virtuosic voice is called into question. The involuntary immediacy of the cough, hiccup, stutter or breaking voice exposes the very human, yet momentary, disconnection between physiology and technique.<sup>30</sup> And, as noted above, the spectacle of virtuosity also comes with the potential of failure in performance. As Emily Wilbourne puts it, in opera ‘we spectators seize upon and revel in the subtle symptoms of bodily betrayal as a guarantee of authenticity’.<sup>31</sup> In pursuing the lofty heights of vocal mastery, the nadir of failure is perpetually imminent. Consequently, Vocaloid, and all it stands for – ‘immortal’ voices, hypervocality, physiological transcendence, absence and simulation – appears to destroy the basis of this spectacle, and the prized labour of opera’s virtuosity.

There is, however, a potential reconciliation of the operatic voice with digital technology that counters this pessimistic outlook. As a number of scholars have observed, and has already been mentioned above, the scene of virtuosity is often paradoxical.<sup>32</sup> It must be, for even though it is transcendent, it can never be unattainable, but rather

<sup>25</sup> Provenzano, ‘Auto-Tune, Labor, and the Pop-Music Voice’, 172.

<sup>26</sup> Descript’s Lyrebird, DeepMind’s Wavenet and Google’s Tacotron are just a few of the ‘intelligent’ voice synthesis technologies that work under the basic premise of using neural networks to assume the qualities of a particular voice by ‘learning’ from short audio recordings.

<sup>27</sup> See, for example, Jonathan Sterne and Elena Razlogova, ‘Machine Learning in Context, or Learning from LANDR: Artificial Intelligence and the Platformization of Music Mastering’, *Social Media + Society* (2019), 1–19, and Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford and Ilya Sutskever, ‘Jukebox: A Generative Model for Music’, [arXiv:2005.00341](https://arxiv.org/abs/2005.00341) (April 2020). LANDR uses machine learning to automate part of a track’s production process. Jukebox is a generative model that its developers claim can ‘learn’ to produce entire songs in a variety of genres.

<sup>28</sup> James Q. Davies, ‘Voice Belongs’, *Colloquy ‘Why Voice Now?’*, *Journal of the American Musicological Society* 68/3 (2015), 677–80, at 678.

<sup>29</sup> Ken McLeod, ‘Living in the Immaterial World: Holograms and Spirituality in Recent Popular Music’, *Popular Music and Society* 39/5 (2016), 501–15, at 505.

<sup>30</sup> Norie Neumark, Ross Gibson and Theo van Leeuwen, eds., *Voice: Vocal Aesthetics in Digital Arts and Media* (Cambridge, MA, 2010), xxvi.

<sup>31</sup> Emily Wilbourne, ‘Demo’s Stutter, Subjectivity, and the Virtuosity of Vocal Failure’, *Colloquy ‘Why Voice Now?’*, *Journal of the American Musicological Society* 68/3 (2015), 659–62, at 660.

<sup>32</sup> Michal Grover-Friedlander, *Operatic Afterlives* (New York, 2011); Elisabeth Le Guin, *Boccherini’s Body: An Essay in Carnal Musicology* (Berkeley, 2006); Susan Bernstein, *Virtuosity of the Nineteenth Century: Performing Music and*

occupy a locale that is marginally above expectation. It is precisely this attainment of virtuosity at the interstices, this point between force and breaking, that opera's voices conjure. Elisabeth Le Guin's concept of the virtuoso sees the performer as epitomising the embodiment of technique, a reading consistent with her exploration of virtuosity as an extension of mechanical philosophy and a rejection of enlightened sensibility.<sup>33</sup> By way of a historical rationale, Le Guin cites Denis Diderot's ground-breaking text, *Paradox of the Actor* (1778/1830), in which the contradiction of virtuosity is located in the need for the virtuosic performer to simulate rather than embody emotion.<sup>34</sup>

An alternative definition of virtuosity is associated with the model of networked intentionality expressed by Richard Wagner in his 1840 essay 'Der Virtuos und der Künstler' (The Virtuoso and the Artist). He asserts,

the composer's intentions are to be conscientiously reproduced, so that the thoughts of his spirit may be transmitted unalloyed and undisfigured to the organs of perception. The highest merit of the executant artist, the Virtuoso, would accordingly consist in a pure and perfect reproduction of that thought of the composer's; a reproduction only to be ensured by genuine fathoming of his intentions, and consequently by total abstinence from all inventions of one's own.<sup>35</sup>

According to this definition, the virtuoso's merit is not located in their ability to interpret the work, but rather in their ability to stay 'true' to it, to channel the desires of the composer (as they were at the time).

We might begin to see how a composer's desire for proximity to the performance would be framed in the context of Vocaloid. The software, which collapses distinctions between score and performance, raises questions about the hermeneutics of Miku as a singer who can ostensibly reproduce her performances perfectly. Miku (at least in her current version) has no capacity to ad lib or perform of her own volition. She must be assigned vocal material by a user, and the words and melodies she is tasked with singing are just as much part of her identity as her voice. To put it another way, lyrics and melodies act as vehicles by which Miku's voice is brought into the world. She would remain silent if she were given nothing to sing, a fact that makes her an ideal technological instrument for fantasies of control and display concerning techno-orientalist and hyper-feminine stereotypes (a point we will return to later).

The affordance of vocal control and the collapsing of score/performance were utilised in the development of Miku's performance in *The End* by the well-established Vocaloid producer Pinocchio-P.<sup>36</sup> After Shibuya had composed and committed the vocal line to MIDI, Pinocchio-P was responsible for transferring this data directly to the Vocaloid software. Composer and programmer worked together remotely, screen- and audio-sharing to ensure Miku's performance aligned with Shibuya's vision.<sup>37</sup> The literal copying and pasting of MIDI data from the computer to the voice of Miku, combined with the minutiae of vocal manipulation, might be read as a Wagnerian dream come true – though of course all

---

*Language in Heine, Liszt, and Baudelaire* (Stanford, 1998); and Wilbourne, 'Demo's Stutter, Subjectivity, and the Virtuosity of Vocal Failure'.

<sup>33</sup> Le Guin, *Boccherini's Body*, 150.

<sup>34</sup> Le Guin, *Boccherini's Body*, 150.

<sup>35</sup> Richard Wagner, 'The Virtuoso and the Artist', in *Richard Wagner's Prose Works*, trans. William Ashton Ellis, vol. 7 (London, 1994), 108–22, at 111.

<sup>36</sup> Pinocchio-P is best known for the bubblegum J-pop song 'SLoWMoTIoN ft. Miku', which has accumulated over 12.5 million views on YouTube as of September 2020. YouTube, 'PinocchioP – SLoWMoTIoN feat. Hatsune Miku': [www.youtube.com/watch?v=ART2fVT33Lw](https://www.youtube.com/watch?v=ART2fVT33Lw) (accessed 17 October 2020).

<sup>37</sup> Pinocchio-P, email to author, 21 September 2020.



performances play out in unique instances that make it impossible to reproduce all the influences of a given musical experience. The division of labour in performance is spread throughout all the distributed subjectivities in the theatre, not only between the composer and the performer.

### Virtuosity and vocal malfunction

Voice's liminality is exposed in its malfunction. If the virtuosic voice 'breaks' during performance, it draws immediate attention to the labouring body through the breaking of the physiological instrument, and the virtuosic simulation is broken. (Much the same argument has been made about media signals, and TV viewers ensconced and forgetful of the artificiality of their entertainment until the signal is interrupted.) Conversely, a recovery from such a malfunction signifies the overcoming of trauma and the pre-eminence of the body's limits. The experience of hearing and seeing Ben Heppner's infamous voice crack and his 'heroic' recovery (what Carolyn Abbate describes as an act of 'extraordinary raw courage and sangfroid') during a performance of *Die Meistersinger* only makes sense in the context of affordance. That is, in the context of Heppner's vocal ability, the perceived technical difficulty of the performance – and his ability to regain control over his voice after it had momentarily 'failed'.<sup>38</sup> As Hutcheon and Hutcheon point out: 'Audiences also pay to experience the excitement and, frankly, the unpredictability of live opera: *the body and the voice may be sublime, or they may fail.*'<sup>39</sup> It is this friction between body and technique that contributes to the perceived intensity of performance.

Of course, the rules of vocal failure must be reconfigured for a virtual performer (with, it should be added, a capacity for pre-rendered performances). For all that Miku may be conceived as a 'digitopian dream' in *The End*, she is not impervious to technological breakdowns. Miku can malfunction (there are many documented examples of her singing out of sync with her hologrammatic body, or failing to sing at all).<sup>40</sup> Miku's glitches during live performance have sometimes led to audiences cheering her on, or singing the song back to her in encouragement until she found her voice again.<sup>41</sup> In fact, Miku is limited by a unique subset of technological affordances that are activated and enmeshed in everything from the 10.2 surround sound setup required for *The End*'s performance, right down to the Vocaloid software itself.<sup>42</sup> How, then, might we conceive of a digital being such as Miku who will not deteriorate with age, yet is open to vocal failure through her dependence on technologies?<sup>43</sup>

Vocaloid's extraction and codification of the singing voice as a media phenomenon encourages us to consider ways of thinking about the human voice itself as a kind of technology. Miriama Young hints of a return to Cartesian mechanical philosophy in writing

<sup>38</sup> Carolyn Abbate, 'Music: Drastic or Gnostic?', *Critical Inquiry* 30 (Spring 2004), 535.

<sup>39</sup> Hutcheon and Hutcheon, *Bodily Charm: Living Opera*, 118–19. Emphasis added.

<sup>40</sup> Miku's best-known malfunction occurred during a set of thirty-nine songs at MikuPa 2013 in Kansai, Japan. Hands framing either side of her mouth, she appeared to yell, psyching the audience up for her next big song. To the audience's surprise, no sound came out, only an odd series of electrical pops and crackling sounds. YouTube, 'Hatsune Miku Live Party at Kansai 2013 Po Pi Po': [https://www.youtube.com/watch?v=RXTsN2geBpw&feature=emb\\_logo](https://www.youtube.com/watch?v=RXTsN2geBpw&feature=emb_logo) (accessed 17 October 2020).

<sup>41</sup> YouTube, 'Hatsune Miku Live Party at Kansai 2013 Po Pi Po'. I would like to thank the reviewers for drawing my attention to this, and to the VocaVerse community for providing video samples of Miku's malfunctions.

<sup>42</sup> ATAK, 'Vocaloid Opera – THE END'.

<sup>43</sup> However, Vocaloid is prone to 'bit rot', by which the performances of older versions of the software gradually deteriorate and are replaced by newer versions. Nick Prior, 'STS Confronts the Vocaloid: Assemblage Thinking with Hatsune Miku', in *Rethinking Music through Science and Technology Studies*, ed. Antoine Hennion and Christophe Levaux (London, 2021), 213–26, at 220.

that ‘a reconciliation of body/voice/electronics views the human voice itself as a highly sophisticated piece of machinery – perhaps the most elaborate and altogether mysterious piece of technology yet invented’.<sup>44</sup> Even to this day the vocal apparatus in action remains relatively enigmatic, perhaps a reason for recent attempts to simulate the voice through vocal synthesis software and 3D-modelled vocal tracts.<sup>45</sup> Under these tenets, any continuation of the pro-technological voice becomes far easier if we see Vocaloid not as a counter to the sonorous voice, but an extension of it. As Young contends, perhaps the best way to think about voice and technology is ‘coexisting along a continuum – in which the “voice” is always technology, the critical variable being the extent to which an external machinery is evident or explicit in the human medium’.<sup>46</sup> The questions of perceived authenticity and human agency remain, and the degree to which we feel they are contingent on technological processes. In realising the virtuosic body as fully technological, Miku becomes an extension, a remediation of the performing body that makes explicit what was arguably already implicit in singers’ bodies: an amalgam of various skills, voices, bodies, technologies and creative interpretations.<sup>47</sup>

In performance, Miku is a holistic entity, who can be identified and delimited visually and sonically. She is an authentic being in the sense that she invokes real (human) reactions and emotions from the audience. As Hutcheon and Hutcheon write, ‘represented bodies are always given meaning by audiences, and those meanings will reflect or challenge the dominant cultural norms at the time of the experienced performance, for they will engage in complex ways the belief systems and values of real audience members’.<sup>48</sup> To her fans, it makes no difference whether she has a set of vocal cords that can actually produce her voice; all that matters is that she provokes empathy. In other words, if we are emotionally affected by a performance, does it matter whether the cause (i.e. the performer) is real or virtual, if the outcome, the symptom of connection between (the virtual) performer and (the real) spectator, remains intact?

A reconciliation of this divide between human and synthesised performance may be found with an alternate reading, however. As Wilbourne suggests, ‘Voice promises access to the interior experience of ourselves and others, a writing on the body that can represent both the material world and our embodied experience of materiality’.<sup>49</sup> Could it not be argued that Miku’s voice is a representation not only of Fujita’s body, but of the materiality of the software itself? Matthew Fuller writes, ‘a glitch is a mess that is a moment, a possibility to glance at software’s inner structure... . Whereas a glitch does not reveal the true functionality of the computer, it shows the ghostly conventionality of the forms by which digital spaces are organized’.<sup>50</sup> In other words, vocal failure in both its physiological and technological – hardware and software – formats, momentarily lays bare the particular affordances of what, or who, is singing.

While *The End* paints a different picture of this debate, the opera can also be read as a social critique of ‘digitopia’ itself, as the analysis will demonstrate. Miku is subjected to a range of violent experiments but remains unscathed, since in this world human suffering

<sup>44</sup> Miriama Young, *Singing the Body Electric: The Human Voice and Sound Technology* (London, 2015), 6.

<sup>45</sup> David M. Howard et al., ‘Synthesis of a Vocal Sound from the 3,000 Year Old Mummy, Nesyamun “True of Voice”’, *Scientific Reports* 10/1 (2020), 45000.

<sup>46</sup> Young, *Singing the Body Electric*, 6.

<sup>47</sup> In the context of this article, I draw on Jay Bolter and Richard Grusin’s definition of remediation. Remediation speaks to the constant repurposing of old media in new forms. Vocaloid is a remediating technology as it codifies the voice and then transforms it into a pliable digital instrument. Jay Bolter and Richard Grusin, *Remediation: Understanding New Media* (Cambridge, MA, 2000).

<sup>48</sup> Hutcheon and Hutcheon, *Bodily Charm*, 39.

<sup>49</sup> Wilbourne, ‘Demo’s Stutter, Subjectivity, and the Virtuosity of Vocal Failure’, 660.

<sup>50</sup> Matthew Fuller, *Software Studies: A Lexicon* (Cambridge, MA, 2008), 114.

cannot be uploaded onto a virtual persona. In this diegetic universe, it appears we cannot simulate the effects of bodily labour or its suggestions of human trauma without the body itself. The chasm that emerges within the performer through vocal failure grants access to an interior space of lived experience. A Freudian reading of vocal failure insists that it is a sonic exteriorisation of trauma, and that trauma is a consequence of lived experience. Miku sits precariously on the edge of this claim, for while she retains some semblance of Fujita's voice (and therefore, one might argue, Fujita's lived experience), Miku is a hypermediated entity and, unlike Fujita, she sings with no physiological consequence. Human fallibility, in *The End*, is precisely what Miku cannot afford. By embodying the perils of technological overdevelopment, permanence and vocal excess, Miku is animated as the antithesis of the human: the antiheroine on stage.

### **The End**

The analysis that follows is centred around *The End*'s European premiere on 12 November 2013 at the Théâtre du Châtelet in Paris.<sup>51</sup> The opera lasts approximately one and a half hours and is separated into twelve numbers: an instrumental overture, four arias ('Aria for Death', 'Aria for Time and Space', 'Aria for Voices and Words' and 'Aria for the End') interspersed with recitative-like sections (e.g. 'The Gas Mask and the Gas' [00:35:27–00:38:13]). The opera is performed in Japanese, with English and French surtitles, while the music is characterised by synthesised voices and an eclectic mix of J-pop, 'minimal techno ... EDM, modern and contemporary classical music and sound art'.<sup>52</sup> Layers of synthesised sounds – from strings to pad tones – contribute to the opera's ambient and glitchy sonic world, giving way to electric crackling and hissing noises that underscore the chaotic nature of the diegetic realm. Aria sections offer tonal respite from dissonant (often non-tonal) recitative sections, structured in more conventional pop forms comprising distinct verse, bridge and chorus sections that are harmonically and rhythmically stabilised by drum and bass loops.

Each performance of *The End* employs an intricate technological setup: '10.2 multichannel sound, through dual-layered 5.1 channels, as well as a cleverly constructed space formed from seven high-luminosity, high-resolution projectors and four giant screens, creates unique 3D acoustic and visual effects with the theatre space'.<sup>53</sup> The presence of this technological setup even plays a metadiegetic role in the opera: Miku is shown, virtually, to crash through the screens at the end of the performance, revealing not merely her projection onto, but her imprisonment within, the wall of screens. The experience is symbolic of the diegetic leakiness between *The End*'s world (the virtual) and the audience's reality.

Beyond technological innovation, the musical style of *The End* evades categorisation. The work's composer, Shibuya, claims that *The End* 'adopts [an] operatic tragedy and upholds operatic styles of aria and recitative' while borrowing from a motley range of musical styles including contemporary Western art music, electronica and dubstep.<sup>54</sup> The identification of the work as an opera has been a subject of interest for many critics. Leon TK recalls post-performance exchanges amongst audience members in the theatre,

<sup>51</sup> A recording of the English subtitled and partially dubbed performance is available at YouTube, '[VOCALOID Opera] THE END [English Subtitles]': [www.youtube.com/watch?v=Ey8oj8S-j3U](http://www.youtube.com/watch?v=Ey8oj8S-j3U) (accessed 17 October 2020). Subsequent timings refer to this recording.

<sup>52</sup> ATAK, 'Vocaloid Opera "THE END": Holland Festival, Amsterdam': <http://atak.jp/en/news/2015-06-04/> (accessed 13 November 2022).

<sup>53</sup> ATAK, 'Vocaloid Opera – THE END'.

<sup>54</sup> YouTube, '【VOCALOID OPERA】 "THE END" Artist Interview 【渋谷慶一郎・初音ミク】': <https://www.youtube.com/watch?v=Z1-YbcbAQ84&lc=UgjeI2kjj2zbYXgCoAEC> (accessed 17 October 2020).

questioning whether what they had just witnessed ‘should really have been billed as an opera’.<sup>55</sup> On the other hand, Stephen Whittington considers whether ‘*The End* might be a beginning of a revolution in opera’,<sup>56</sup> while Gordon Forester rationalises that ‘pre-programmed music and pre-produced visuals together with the absence of an orchestra and human performers, might lead some to question the label “opera”, but in the brave new world of Hatsune Miku, semantics seem superfluous’.<sup>57</sup> Indeed, if the contention about *The End* indicates anything, it is that musical genres are often expected to sit within a specific set of aesthetic, musical and performative conventions, and are thus defined by the audience’s expectations.

According to the production company’s own description of *The End*, the opera strives to challenge the boundaries of Western art music and notions of an enlightened liberal human subject:

A new world emerges from ‘THE END’, one that escapes from the European anthropocentrism that was conventionally bound to civilization and art, a world that dissolves the boundaries between life and death, public and private, parts and the whole, layer and delineated, human and animal, existence and production. In this world, Miku, who has had a presentiment of her fate, talks with animal characters and degraded copies of herself to ask the age-old questions, ‘what are endings?’ and ‘what is death?’.<sup>58</sup>

This simultaneous adherence to and defiance of the philosophy of opera is acknowledged by critics. Murray Bramwell notes that the tragic suicide of Shibuya’s wife is imprinted on Miku as a surrogate of opera’s foundational myth: ‘Shibuya has imbued the work with tenderness, and an innocence, which in the final stages of the opera, signals a recognition – with its insistent melodic repetitions and the composer’s simple, anguished ... libretto – that, like Orpheus separated from Eurydice, his beloved is lost for all time.’<sup>59</sup> From a musical perspective, Leon TK notes that ‘Shibuya’s turbulent music sends the listener off-balance as it consistently evades any sense of conventional structure; computer-generated pixel flurries [visual effects] add to the chaos. Beneath relentless waves of sensory overload, the resultant cognitive dissonance feels quite appropriate given the work’s central themes: death, uncertainty, and fear.’<sup>60</sup> Yet it bypasses any substantial critique of the performance, instead questioning the categories of musical tradition within which one might begin to analyse *The End* in the first place. Such reactions indicate that, as far as musical genre is concerned, *The End* has no fixed home.

If opera is so centred on the perceived presence of the labouring body, then a key to unlocking this ambivalence of genre in *The End* may be Jason Stanyek and Benjamin Piekut’s concept of ‘deadness’. If *The End* thematises the meaning of death in a

<sup>55</sup> Leon TK, ‘Approaching The End: Face-to-Screen with the World’s First Humanless Opera’.

<sup>56</sup> Stephen Whittington, ‘The End, by Japanese composer Keiichiro Shibuya with anime star Hatsune Miku, is opera – but not as you know it’: [www.adelaidenow.com.au/entertainment/arts/the-end-by-japanese-composer-keiichiro-shibuya-with-anime-star-hatsune-miku-is-opera-but-not-as-you-know-it/news-story/852a2a179a2719bb0bebad8256e36dcc](http://www.adelaidenow.com.au/entertainment/arts/the-end-by-japanese-composer-keiichiro-shibuya-with-anime-star-hatsune-miku-is-opera-but-not-as-you-know-it/news-story/852a2a179a2719bb0bebad8256e36dcc) (last updated 5 October 2017).

<sup>57</sup> Gordon Forester, ‘The End (Keiichiro Shibuya & Hatsune Miku, OzAsia)’: [www.limelightmagazine.com.au/reviews/review-the-end-keiichiro-shibuya-hatsune-miku-ozasia/](http://www.limelightmagazine.com.au/reviews/review-the-end-keiichiro-shibuya-hatsune-miku-ozasia/) (last updated 4 October 2017).

<sup>58</sup> ATAK, ‘Vocaloid Opera – THE END’. Similar questions of mortality are also raised by animatronic ‘operabots’ in Tod Machover’s *Death and the Powers*.

<sup>59</sup> Murray Bramwell, ‘Meeting Points: Scary Beauty and The End: Vocaloid Opera (OzAsia, Adelaide)’: <https://dailyreview.com.au/meeting-points-scary-beauty-keiichiro-shibuya-australian-art-orchestra-skeleton-space-theatre-adelaide-festival-centre/> (last updated 8 October 2017).

<sup>60</sup> Leon TK, ‘Approaching The End: Face-to-Screen with the World’s First Humanless Opera’.

postmodern digital world, Stanyek and Piekut contend, then music and sound recording technologies have not only been associated with the dead and the preservation of the body through sound and voice, but now facilitate the rearticulating and splicing of the body 'into networks that extend beyond self-contained limits'.<sup>61</sup> In other words, it is not merely that voices from the past are recorded (preserved, stored) but that they are routinely upcycled (spliced) to produce new material, sometimes in configurations that appear to find the dead singing themselves 'back to life', as is the case with Vocaloid:AI.<sup>62</sup> Conceptually, this maps onto the process by which Miku's voice is generated by deconstructing and recombining samples of Fujita's voice, and the notion of 'intermundane' collaboration that underscores virtual performances (between synthespians and humans, or, in this case, 'lifeless' Miku, and the Miku that becomes human through death).

I read *The End* as a commentary on the paradox of opera: Miku fetishises mortality and its unattainability for her as a virtual body. As the show unfolds, it becomes clear that Miku's desire for death is synonymous with her desire to be 'real', to exist beyond the prison of the four screens, and to instead diffuse into our world. *The End* suggests that in 'dying', Miku will at last embody the living operatic diva. The work's central irony is played out through the character of Look-a-Like – a corrupted copy of Miku – when Miku's main affordance is her pliable, replicable, digital state. As described on the production company's website:

Opera always dealt with human death, creating a situation where the abnormal exertion of life's greatest energy by a person about to die was essentially linked to sound and acoustics. *The End* takes note of this habitual format and treats opera as an anthropological mechanism of critique where the distortion of life and death unfolds.<sup>63</sup>

How, then, can Vocaloid's uniquely remediated voice be read both as a complete undermining of bodily labour and an extension of the already machinic/virtuosic performer? Fujita is an *anime* voice actress and singer. This style of voice acting calls for a light and breathier style of singing, in contrast to what would typically be expected of a sonorous operatic voice, yet the digital faculty of Miku's voice allows her (and the other characters) to whisper-sing throughout a (theoretically) limitless vocal range. The possibilities of her hybrid, hyperactive vocality are even more pronounced in the opera's recitative-style sections. (At one point, Miku exclaims 'If you gave me some words I'd pronounce them perfectly, no matter how fast' [00:35:41–00:35:46].) Perfectly executed vocal leaps that do not sit within the *passaggio* [00:18:06–00:18:17] point to an uncanny sound that appears to bypass the sonic signifiers of a biological voice. Furthermore, Animal's voice is distinguished by an eerie, pitch-bended self-harmonisation [00:05:25–00:13:00]. These physiologically impossible voices, in which multiple characters draw – differently –

<sup>61</sup> Jason Stanyek and Benjamin Piekut 'Deadness: Technologies of the Intermundane', *TDR: The Drama Review* 54/1 (2010), 14–38, at 17.

<sup>62</sup> The emergence of Vocaloid:AI, a version of the software that utilises deep neural networks, is able to simulate the grain of any singing voice based on limited audio sample data, and is further confusing dynamics of body, voice and agency. In October 2019, this technology was used to reproduce posthumously the voice of beloved Japanese *Enka* singer, Misora Hibari (1937–89) using pre-existing recordings. 'Hibariloid' and a hologram in her likeness performed 'live' in September 2019 for the thirtieth anniversary of her death. Yamaha, 'Yamaha Vocaloid:AI™ Faithfully Reproduces Singing of Legendary Japanese Vocalist Hibari Misora': [www.yamaha.com/en/news\\_release/2019/19100801/](http://www.yamaha.com/en/news_release/2019/19100801/) (accessed 14 November 2022). A recording of this performance is available at NHK, 'Hibari Misora Revived with AI': [www.nhk.or.jp/special/plus/videos/20191007/index.html](http://www.nhk.or.jp/special/plus/videos/20191007/index.html) (last modified 7 October 2019).

<sup>63</sup> ATAK, 'Vocaloid Opera – THE END'.

upon Vocaloid voicebanks,<sup>64</sup> is a very odd experience to the ear: its weightlessness points uncannily to the lack of exertion upon the body. Even within the opera's humanly singable ranges, the characters sing too smoothly and too easily: it is strange to listen to a voice without stress, when traditional opera demands that its bodies labour intensively.<sup>65</sup>

On the other hand, in order for Miku's voicebank to sound distinctly Miku, she relies on the input of Fujita's voice. While Fujita cannot have any direct claim to agency on *The End*'s stage, the breathy and girlish timbre of Miku's voice, as she effortlessly plunges into the demands of a humanly impossible operatic vocality, cannot be entirely separated from Fujita's physiological body either. Mona Lalwani points out that, in detecting Miku's voice, the audience are really 'hearing modulations of [Fujita's] vocals'.<sup>66</sup> But in hearing Miku's voice, we experience the human and the technological as one, since the artificial processes that make Miku's voice distinct from Fujita's play a salient role in the simulacrum's unique timbre. Miku's voice is a combination of many agencies and forces, and this assemblage theory 'refuses to privilege human over non-human agency, instead seeing how they enmesh and activate one another', as Nick Prior puts it.<sup>67</sup> In this context, Miku and assemblage theory may remedy, or at least expose, ideologies of a unified subjectivity, revealing how bodies, voices and technologies are reconfigured amongst each other.

*The End* invites the audience to reconfigure the voice's position in relation to the operatic diva. For Sterne, there can be 'no diva without the countless techniques and technologies that make her audible, visible and sensible. Mediatic technologies form the diva's conditions of possibility'.<sup>68</sup> Clearly, there is no Miku without Vocaloid, or the countless other technologies that make her available as an idol, a performer and a voice synthesis software. Both the diva and the idol are, as Sterne alludes to, social and technical fabrications of femininity, sexuality and race: identities manufactured for the purpose of performance and spectacle.<sup>69</sup> Clearly embedded in this construction is the notion of control and power: from the user's capacity to control another voice with the Vocaloid software, to the media logics of the *Jimusho* system (the monopoly of performer management companies that are responsible for cultivating Japan's biggest idols).<sup>70</sup> Since the

<sup>64</sup> Pinocchio-P utilised two voicebanks for the opera, Miku, and Kagamine Rin, another well-known Vocaloid pop star whose image was not featured in *The End*. Pinocchio-P, email to author, 21 September 2020.

<sup>65</sup> Miku employs a limited vocal range in her arias (the vocal line in 'Aria for Death' sits between  $D\flat_3$  and  $A\flat_3$ ). Elsewhere, she peaks at  $B\flat_5$  [00:18:16].

<sup>66</sup> Mona Lalwani, 'It Takes a Village: The Rise of Virtual Pop Star Hatsune Miku': [www.engadget.com/2016-02-02-hatsune-miku.html?guccounter=1&guce\\_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLnNvbS8&guce\\_referrer\\_sig=AQAAACZxyN8Pr77\\_pPDaYWls7jP9oQj5NEGGNvNP70Zhv97N6vCM\\_h7lb4HL9dTRQl9vMK75-8BYr3bc52TJ4PRkk3mDrZ-6U3Zult-s9dlez9FEo4dN6Fxx9jxBoz1DAEuij8RXZCYhEaZ0hzBrNqEyAzUWS88CQ\\_XwgXlcrvJVRQyEI](http://www.engadget.com/2016-02-02-hatsune-miku.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLnNvbS8&guce_referrer_sig=AQAAACZxyN8Pr77_pPDaYWls7jP9oQj5NEGGNvNP70Zhv97N6vCM_h7lb4HL9dTRQl9vMK75-8BYr3bc52TJ4PRkk3mDrZ-6U3Zult-s9dlez9FEo4dN6Fxx9jxBoz1DAEuij8RXZCYhEaZ0hzBrNqEyAzUWS88CQ_XwgXlcrvJVRQyEI) (last updated 2 February 2016).

<sup>67</sup> Prior, 'On Vocal Assemblages: From Edison to Miku', 492.

<sup>68</sup> Jonathan Sterne, 'Afterword: Opera, Media, Technicity', in *Technology and the Diva*, ed. Karen Henson (New York, 2016), 159–64, at 159.

<sup>69</sup> Sterne, 'Afterword: Opera, Media, Technicity'. Additionally, Patrick Galbraith and Jason Karlin's edited collection explores the industry of idol and virtual idol production in the context of Japanese media systems. While this goes beyond the scope of this article, it is important to place Miku within this broader context, particularly the *Jimusho* system of media production in which star personas are carefully crafted by a monopoly of talent companies. See *Idols and Celebrity in Japanese Media Culture* (Basingstoke, 2012). For a critical analysis of Vocaloid's earliest voicebanks and constructions of race through voice, see Nina Sun Eidsheim, 'Race as Zeros and Ones: Vocaloid Refused, Reimagined, and Repurposed', in *The Race of Sound: Listening, Timbre and Vocality in African American Music* (Durham, NC, 2019), 115–50, and 'Synthesizing Race: Towards an Analysis of the Performativity of Vocal Timbre', *Trans. Revista Transcultural de Música* 13 (2009), 1–9.

<sup>70</sup> David W. Marx, 'The *Jimusho* System: Understanding the Production Logic of the Japanese Entertainment Industry', in *Idols and Celebrity in Japanese Media Culture*, ed. Patrick W. Galbraith and Jason G. Karlin (Basingstoke, 2012), 35–55.

advent of virtual idols in the late 1990s, fantasies of control play out through fantasy bodies and are often concerned with producing images of ‘compliant femininity’ (Miku, for example, can only sing back to you if you program her).<sup>71</sup> The opera can be comprehended this way: through her voice and body she is constructed as the perfect, digitopian performer, which ironically becomes the very hellscape from which she breaks free.

There is, however, another perspective from which Miku can be read in *The End*. If, as Laura Miller and Rebecca Copeland argue, the diva ‘systematically [draws] our attention to the performative nature of identity, to gender, and to battles over control of female bodies and female sexuality’,<sup>72</sup> pushing ‘the boundaries of expression, asking us to question what is natural, what is normal, what is culturally appropriate’,<sup>73</sup> then Miku lays bare the very nature of how femininity is constructed and expressed when the diva in question holds no agency. In fact, in the opera Miku develops awareness about her own artificiality only when she is confronted by her corrupted clone, an experience of self-knowledge through alterity. The performing voice (in this case without its agential ‘voice’) thus becomes not only a pure instrument in the social and technical construction of hyperfeminine identity, but also a cause for reflection on unequal systems of power and fantasies of control that are often associated with the (human) talent industries and the performance of the virtuoso.

The simulation of trauma in *The End*, then, is also a fantasy of control. Miku becomes the perfect test subject for the work’s mysterious, diegetic powers because she remains unaffected by the experiments – and they are thus infinitely replicable. Similarly, opera’s compulsion with death is possible precisely because it is a simulation, a performance which is replicated night after night. Perhaps in this light, the status of performer remains undecided, a recalcitrant midpoint within the semiotics of body and digital reincarnation – virtuosity and technology.

### Final remarks

We build worlds around voices, worlds at once cultural-technological and natural-biological.

James Q. Davies<sup>74</sup>

The worlds of opera and technology have always intersected, and media has both soothed and complicated this relationship. *The End* offers a site where these worlds collide, fragmented in binary code. Miku stretches the capacities of the voice, revealing Vocaloid’s affordances within the context of opera. She is not relegated to some other world, however. She exists here, coded and switched on and off by a human. She is engendered with social and cultural meanings and returns as a mirage that facilitates a redefinition of the physical, conceptual and technological limits of the body in performance.

Beyond *The End*, Miku has made her name as an internationally renowned pop icon. As we imbue Miku with techniques associated with the body, we also reappropriate her performance within the human body. ‘The human voice returns as a simulation of the perceived authenticity not of humanity but of the digital machine’, writes Prior: ‘In

<sup>71</sup> Black, ‘The Virtual Idol: Producing and Consuming Digital Femininity’, 209. Black also notes elsewhere that the virtual idol phenomenon should not merely be ‘dismissed as a bizarre marketing gimmick’ as seen through the gaze of Western media. To Black, the socio-technical construction of femininity is just as embedded in Western culture industries. Daniel Black, ‘The Virtual Ideal: Virtual Idols, Cute Technology and Unclean Biology’, *Continuum* 22/1 (2008), 37–50.

<sup>72</sup> Laura Miller and Rebecca Copeland, eds., *Diva Nation: Female Icons from Japanese Cultural History* (Oakland, 2018), xi.

<sup>73</sup> Millet and Copeland, *Diva Nation*, 5.

<sup>74</sup> Davies, ‘Voice Belongs’, 681.

other words, with phenomena like beatboxing, the voice becomes a simulation of a simulation.<sup>75</sup> This bodily and stylistic reclamation of technique belongs not only to the voices of popular music, however. The technique and technologies of voice are intrinsically linked through opera's virtuosity too. A posthumanist might argue that it has in fact been our proximity to technology that has allowed us to move beyond the limits of performance. Miku and Vocaloid then become the foundation for a new tradition of virtuosity. Based on technological affordance, such remediated voices become embedded in performative culture, and are in turn consumed and reappropriated by the human body. As Prior witnessed during his fieldwork in Tokyo, Miku fans perform karaoke and attempt to copy her impossibly rapid vocalisations (not unlike human pianists copying Conlon Nancarrow's *Studies for Player Piano*).<sup>76</sup> So long as we continue to listen and perform, these techniques will always find meaning within the cultured body (that is, the physiological body culturally mediated by its technological dependencies). In coding machines to sing like us, we in turn may begin to sing like them.

In this sense, virtuosic singing renders the body a simulation, and is therefore an interface for controlling the voice. Vocal failure takes on new meanings with technological voices, as does virtuosity itself. By asking the question 'what is this voice of mine', Miku in the end reveals the agential complexity of virtuosic voices, encouraging us to wonder how such configurations of power and control may be appropriated, broken and redefined.

**Acknowledgements.** I am most grateful to my reviewers for their helpful comments on previous iterations of this article. I would also like to thank the following (alongside many others) for their guidance: Annette Davison, Fatima Lahham, David Trippett and the Sound, Voice & Music working group at the Theatre and Performance Research Association 2021.

**Jessica Tsun Lem Hui** is a PhD researcher in Music at the University of Cambridge, whose research interests include the histories of singing and speaking machines, the development of visual voice-identification practices, and the use of AI software in the recreation of singing voices.

<sup>75</sup> Prior, *Popular Music, Digital Technology and Society* (Los Angeles, 2018), 119.

<sup>76</sup> Prior, 'On Vocal Assemblages: From Edison to Miku', 502.

**Cite this article:** Hui JTL (2022). Reconfiguring Voice in *The End: Virtuosity, Technological Affordance and the Reversibility of Hatsune Miku in the Intermundane*. *Cambridge Opera Journal* 34, 364–379. <https://doi.org/10.1017/S0954586722000301>