

Artikel

ROBOT & THE END

ROBOT & THE END

A Comparative Critical Reading of the Staging of Synthesized Voices in Digital Media Performances

By Stina Hasse Jørgensen

Introduction

Today synthesized voices speak to people living in the rapidly developing smart cities in the Western world from technologies such as mobile phones, computers, and social robots, to name just a few. The increasing presence of synthesized voice technologies is also of interest for contemporary artists and the use of synthesized voices as an artistic medium is becoming more and more widespread in media performances, computer games, and new media artworks. It is therefore relevant to discuss how the representation of identities takes place through the use of synthesized voices in the arts. After all, the way we view and identify with “the bodies we see, whether in representation or in real time and ‘live’” informs the way we experience others and ourselves, as art historian Amelia Jones writes in her book *Seeing Differently* (Jones, p. xxi). Furthermore, the representation of bodies in media performances can be said to play a role in the ways we come to define ourselves and others. As media scholar Robin Sloan states: “the media stories that we experience can have significant impact on how we define ourselves as individuals and how we understand others” (Sloan, p. 53).

In this article, I will discuss how representations of gender identity are performed and staged using synthesized voices as an artistic medium in digital media performances. I will critically reflect upon this in my exploration of the 3D animated vocaloid performance *THE END* (2012) and the robotic multimedia dance performance *ROBOT* (2013). In my comparative critical reading of the two media performances I argue that they demonstrate two different ways of using synthesized voices as an artistic medium in digital media performances. More specifically, I will reflect upon the relationship between the sonic body and the visual body in the audio-visual staging and representation of gender identities in digital media performances.

Performance studies researcher Robin Nelson writes in *Mapping Intermediality in Performance* that “the digital doubling of bodies, virtual bodies, robots and cyborgs have entered the intermedial stage, if not to displace humans, then most assuredly to engage with them and question some of their most fundamental assumptions” (Nelson in Bay-Cheng et al., p. 23). Here the use and staging of synthesized voices could promote a proliferation and multiplication of categories, and expand “the range of alternatives, trading duality for multiplicity,” as feminist scholar Mimmi Marinucci writes (Marinucci, p. 47), complicating the fixed visible bodies, and troubling the idea of identity as something static that can be congealed into fixed binaries. In the following I will discuss the audio-visual staging in *THE END* and then in *ROBOT*, reflecting upon the different strategies the two media performances present. I will be concerned with questions inspired by the points made by Nelson and Marinucci, particularly the question of whether digital sonic and visual bodies in media performances have a potential to disrupt the notions of the heteronormative gender binary system.

My comparative critical reading of the media performances *THE END* and *ROBOT* is founded in an understanding of situated knowledge. In her essay “Situated Knowledges: The Science Question

in Feminism and the Privilege of Partial Perspective”, feminist scholar and science & technology theorist Donna Haraway describes situated knowledge as challenging the myths of traditional disembodied objectivity as something that is in fact always already partial, embodied, and specific. Haraway argues that one way of drawing attention to the politics of knowledge production as something that is not disembodied or universal is to make clear that vision is embodied and located. So-called disembodied objectivity, Haraway argues, is already partial, embodied, and specific. The discussion of the two media performances presented in the following is informed by my experiences of being brought up as a white person in the capital (Copenhagen) of a small country (Denmark) and as female assigned at birth. It is also grounded in the located vision defined, among other things, as coming from a situatedness in a Western academic environment where I, as a PhD student at a Department for Arts and Cultural Studies, have been trained mainly in Western art history, sound art, and new media performance with a special interest for artistic experiments with voice technologies.

THE END

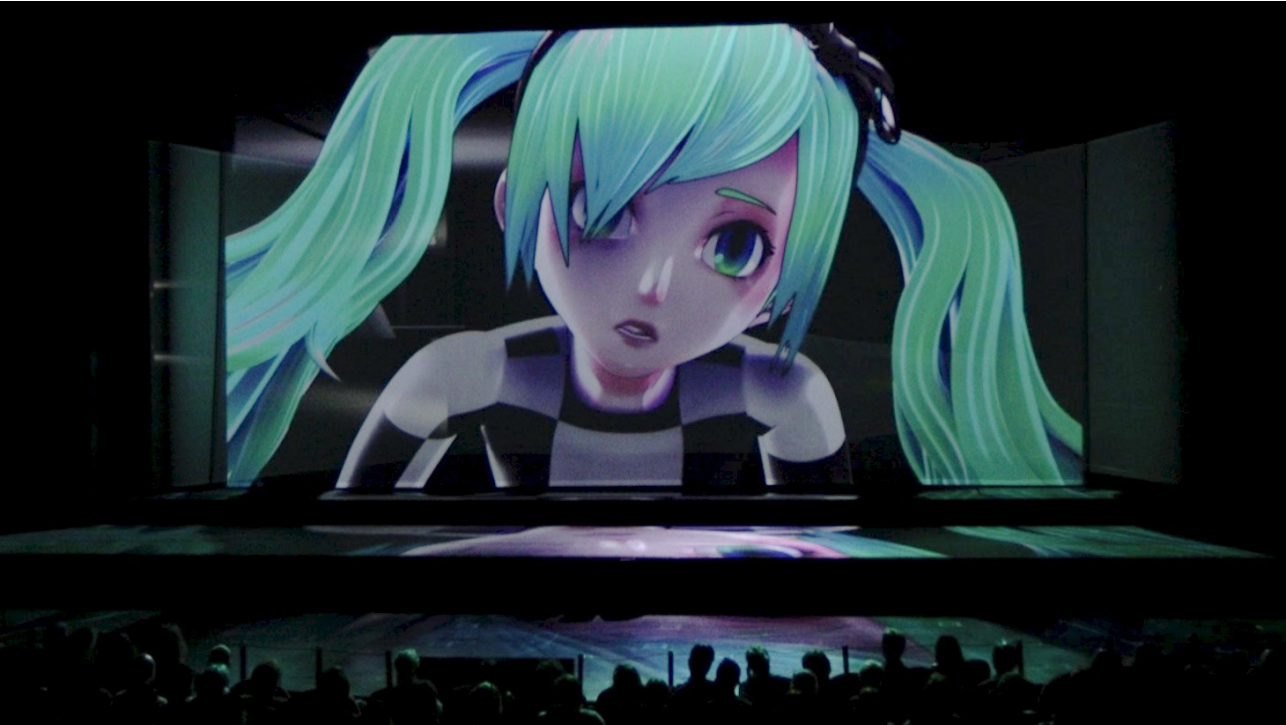
A helium sounding synthesized voice sings “ahhh” in a long tone, weaved together with a rhythmic pulsation of violins. The voice comes from the speakers of a 10.2 surround sound system. It oscillates around the performance space like an echo signal, swirling around me and the other audience members. “Maybe I was dreaming just now as if I was just breathing,” the female voice sings in a staccato rhythm. It comes from the pop star Hatsune Miku, who is the center of attention in the 3D animated vocaloid opera called *THE END*, composed by the acclaimed composer and DJ Keiichiro Shibuya, sound designed by sound artist evala, and 3D animated by art director YKBX. I experienced *THE END* on a summer night in August at Musikhuset Aarhus for the opening of Aarhus Festuge 2016.

THE END is a Japanese production, and its special relation to Japanese pop cultural icon Hatsune Miku, who is the lead character in the 3D animated manga-like story with electronic experimental music, attracts many different fans: cosplayers, cartoon fans, anime crowds as well as people interested in experimental music (Jørgensen, 2016a).¹

THE END is a sumptuous feast of technique and aural as well as visual effects. The speakers are placed around the audience in the concert hall and at the front, four large screens form a square around Shibuya, who stands in a yellow hoodie on the stage, surrounded by a smaller screen reminiscent of a DJ booth.

In the performance, a fragmented exploration of the possible disappearance of Hatsune Miku plays out. The narrative unfolds on the large screens on stage in the form of symbolic representations of the feelings the music in the concert hall creates. Throughout the performance, we follow Hatsune Miku’s mental journey and recognition process – from not knowing what it means to exist to understanding that she herself can disappear (like humans do when they die). The display of the end or death of Hatsune Miku has led to many protests from Hatsune Miku’s fans, but despite the protests *THE END* has been touring from Tokyo and Paris to Aarhus since 2012 (Jørgensen, 2016b).

1) There is much more to be written about the situated experience of a Japanese production in a Western context. The discussion of the cultural aspects of *THE END* and how gender performativity might be performed differently in a non-Western context is however not unfolded in this article and the argument presented here leaves room for critical readings by scholars working with intersectionality and postcolonial studies.



The 3D animated Hatsune Miku in THE END, data.tokyogirlsupdate (2015).

In Japanese, Hatsune Miku means “the first sound from the future” and first and foremost Hatsune Miku is a sound, a vocaloid. A vocaloid is a voice-based synthesizer technology developed by the music corporation Yamaha and used by media companies like Crypton Future Media, the company behind the release of the Hatsune Miku vocaloid in 2007. Hatsune Miku’s synthesized voice is made from large amounts of vocal recordings by the voice actor Fujita Saki. The digital voice trained on Saki’s voice can be manipulated in the vocaloid software, allowing musicians and producers to create vocal music digitally by typing in lyrics and drawing melodies. They can also automate pitch and singing rate and use effects such as gender factor, growl, and breathiness in the program.²

Hatsune Miku was at first codenamed CV01, but was, courtesy of the manga artist Kei, given a visual identity as a 16-year-old manga-like girl with long legs, big eyes, very long turquoise hair in pigtails, often dressed in a miniskirt. This visual depiction of the vocaloid has been further developed through manga drawings, music videos, and anime made by thousands of Hatsune Miku fans worldwide (Jørgensen, Vitting-Seerup, and Wallevik, 2017).

Back in Musikhuset Aarhus concert hall Hatsune Miku appears on the four large screens on stage. ‘She’ is lying down; her long blue hair floating above her head in what could seem like dark water, with two huge cartoon-drawn eyes looking at ‘her’ from behind. ‘She’ looks very

2) In *Synthesis of the Singing Voice by Performance Sampling and Spectral Models* (2007) Jordi Bonada and Xavier Serra write more about how to create singing voice synthesis using audio recordings from voice actors.

fragile as ‘she’ starts to sing with an airy, bright “little-girl voice” as Birgitte Rahbek writes in her review of the performance.³

Gender performativity

In the article “Performativity, Precarity and Sexual Politics” philosopher and queer theorist Judith Butler states that gender performativity is a practice through which gender is constructed through normative constraints and through relations of power: “that gender is performative is to say that it is a certain kind of enactment; the ‘appearance’ of gender is often mistaken as a sign of its internal or inherent truth” (Butler, 2009, p. i). Butler also writes about the notion of gender performativity in *Gender Trouble* (1990) and in *Bodies That Matter* (1993). Here Butler argues that the normative understanding of bodies prompts a performance of gender within a strict binary frame, as a gender binary. The gender binary is the classification of sex and gender into two distinct, opposite and disconnected forms of masculine and feminine. For Butler, the gender binary is part of a regulatory practice “whose regulatory force is made clear as a kind of productive power, the power to produce – demarcate, circulate, differentiate – the bodies it controls” (Butler, 1993, p. xii). Butler examines how the power of heterosexual hegemony shapes normative understandings of bodies, sex, and gender as essential and immutable. Her point is that the regulatory norm of the gender binary is not a given, but is rather constituted by gender performativity, through acts that reiterate the binary of



Hatsune Miku and catwalk models in clothes designed by the fashion brand Louis Vuitton (Next Nature).

male and female. She explains gender performativity as something that “must be understood not as a singular or deliberate ‘act’, but, rather, as the reiterative and citational practice by which discourse

3) My translation of “lillepigestemme” (Rahbek, 2016).

produces the effects that it names” (Butler, 1993, p. xii). Butler explains that these normative constraints produce a regulation of which bodies come to matter as socially intelligible and visible subjects, and which bodies are produced as unintelligible and invisible.

The gender performativity of Hatsune Miku

In *THE END*, Hatsune Miku is depicted with a miniskirt in accordance with previous widespread fan character productions, but here the luxury brand Louis Vuitton has designed the miniskirt. This visual design of Hatsune Miku stages the vocaloid as a digital femininity. This visual depiction of the vocaloid in a miniskirt, with long hair and small waist, supports the auditory experience of Hatsune Miku’s ‘little-girl voice’ as a female audio-visual body, a human sociality.⁴ The staging of Hatsune Miku’s sonic and visual body in *THE END* can be argued to play with the notion of presence and human sociality.

Media scholar Thao Phan argues that gendering of synthesized voices, for instance in smart assistants, within the gender binary makes the synthesized voices appear as a “believable performance of human sociality [...] diverting attention away from the act of mediation” in order to design “the illusion of immediacy by which the subject is perfectly seduced by the medium, and in this seduction, indulges in the fantasy that there is no medium at all” (Sloan, p. 28)). The gendering of technologies can create an illusion of immediacy and presence through the design of gendered voices that reiterate the Western binary system. In other words, through a gender performativity constituting the visible subjects of male and female. Maybe the staging of Hatsune Miku plays on the same performance of human sociality as experienced in voice assistants?

Media artist and theorist Aneta Stojnic writes that the increasing personalization of digital devices framed as “smart, intuitive, friendly, responsive, personal, sophisticated” makes them seem human-like, stressing that “all of these terms both attribute and delegate human characteristics to the technologies that become our anthropomorphic companions and our self-extensions” (Stojnic, p. 75). The digital devices are personalized and created as human sociality through the design of, among other things, synthesized voices with personality traits and human paralinguistic characteristics such as gender, age and accent (Nakano, et al.; Baird, et al.).⁵ Sound theorist and media philosopher Frances Dyson writes in her book *Tone of Our Times* about this development of synthesized voices: “The quantification of vocal tone, the synthesizing of the voice, and the development of human-computer interfaces based on human-machine conversation, create agents that provide the ‘who’ in this post-theological era” (Dyson, p. 69). The ‘who’ might be agent-based interaction devices such as smart assistants communicating with human users through synthesized voices, as designer Bert Brautigam writes: “You call it Siri, Alexa, or Cortana [...] It defines itself as female or male. The sound of a voice assistant imitates human sound and intonation” (Brautigam, u.p.). The development of synthesized voices to become increasingly human-like has been discussed in connection with what is known as the personification debate (Harris). This debate treats synthesized voices not just as neutral technologies that can be optimized in terms of engineering and programming, but as technologies that are designed within specific cultural understandings of identity construction (Benyon; Faber; Nass & Brave; Robertson).

4) When I write about Hatsune Miku as a ‘female’ character, I do not intend to imply that ‘female’ should be understood as a unified category. On the contrary it is a constructed and complex category highly contested by theorists working intersectionally (e.g. hooks, 1984).

5) For example in IBM- Bluemix- you alter such human-like features like “pitch, pitch range, glottal tension, breathiness, rate, and timbre of spoken text” to the synthetic voice (IBM Bluemix 2017).

Just like Siri, Alexa, and Cortana, the synthesized voice of Hatsune Miku is staged as a 'who' – a human sociality with a socially visible and audible body – a female within the gender binary system. The gender representations created in *THE END*, as well as in the smart assistants, demonstrate citational practices that reiterate the normative constraints of the binary system in order to make a “believable performance of human sociality” (Phan). The staging of Hatsune Miku’s gendered digital anime body and synthesized ‘little-girl voice’ constitutes the reiterative and citational practice Butler describes as gender performativity in the context of digital media performance. This citational practice created by the audio-visual presentation of Hatsune Miku can further be argued to fixate ‘her’ synthesized audio-visual body as a fetish object.

Fetishizing the audio-visual body of Hatsune Miku

In “Dis-Embodying the Female Voice” critical theorist Kaya Silverman writes about voice-over, arguing that the female voice is always brought back to the female body in cinema – as opposed to the male voice-over, which is often detached from the body, creating an experience of the omnipresent and powerful subject.⁶ In *THE END*, Hatsune Miku’s voice is never disembodied in the sense that it is always heard as connected to a visual depiction of a 3D animated female body. As such, the visual body and the sonic body of Hatsune Miku are staged in a way that emphasizes the embodiment of the voice in connection with a gender stereotype of the female in the Western gender binary. Following Amelia Jones’ reading of the female nude in Western art history, I will argue that Hatsune Miku’s audio-visual body is inscribed into the long tradition of the fetish of the ‘idealism’ of the female form – as it can be seen in e.g. iconic art historical pieces such as *Olympia* (1856) by Édouard Manet, or Alexandre Cabanel’s *Birth of Venus* (1863). Here the female nude “must not have any actual genitals: her sex must be erased in order for her body as a whole to function as fetish. The desired female body must, paradoxically, have no orifice, no actual sex” (Jones, p. 65). In *THE END*, Hatsune Miku is not presented with actual genitals or explicit sexual vocal utterings, but appears like an innocent young sexy girl with a miniskirt, big eyes and a ‘little-girl voice’. In this way, the presentation of Hatsune Miku can be understood in relation to the aesthetics of the female nude in Western art history and artworks such as Cabanel’s *Birth of Venus*. Jones writes that in this particular painting, “the woman, portrayed by the man, is ‘deceitful’ in her exquisite fleshy offering, and apparently inherently sexually available for heterosexual male gazing.” (Jones, p. 65). The vocaloid, Hatsune Miku, can be said to be a female sculpted, quite literally, by a man – namely Shibuya as composer, fan and DJ on stage. Here Hatsune Miku is presented as an aesthetic beauty just as the female (nude) body is presented as a trope of aesthetic beauty in the history of Western art. This is an aesthetic that, following Jones, can be understood as:

A container to enframe and control the threat of the unbridled female sexuality [...] as a strategic mode of discourse that operates to cohere the male subject, always anxious about the perceived power of female sexuality and social access [...] most often operated in the past through structures of fetishism (Jones, p. 65).

Just as the control of female sexuality might have been enacted through structures of fetishism throughout Western art history, I will argue that these structures of fetishism are still in play in *THE END*. The staging of Hatsune Miku’s voice as dependent on the male composer’s desire to

6) Silverman’s focus on the female body has been critiqued, arguing that this focus prevents Silverman from seeing a female subject (Fèvre-Berthelot; Sjogren).

listen and play with the voice creates Hatsune Miku as a fetish object. Hatsune Miku is ideal in her feminine desirability. Feminist film theorist Laura Mulvey writes: “Women are simply the scenery onto which men project their narcissistic fantasies” (Mulvey, p. 13). Fetishism functions through a system of binaries, Jones writes, referring here to identity-related discourses on fetishism and especially to psychoanalyst Sigmund Freud’s model of “how objectification occurs in the self/other relation,” where the other can be seen as a “projection of the desires of the empowered self” (Jones, p. 63). In *THE END*, the male composer, fan, and DJ Shibuya projects onto Hatsune Miku, as a representation of the female, his “narcissistic fantasies,” turning her body into a phallic substitute, wholly sexualized and available for his pleasure while not possessing her own sexual identity and body parts (Jones).

Throughout most of *THE END*, Shibuya can be seen standing on stage behind the DJ booth as if he were controlling the voice of Hatsune Miku and determining ‘her’ presence. In a way, Shibuya mimics the role of the Hatsune Miku fan, without which the vocaloid would not exist. Hatsune Miku is to a large extent a crowd-sourced phenomenon and there are more than 100,000 fan-made songs using the voice of the vocaloid to express their emotions and desires. The songs are shared and circulated by fans on social media platforms such as YouTube and Niconico (the Japanese equivalent to YouTube), and the number of songs is still growing. Fans, in this context, should be understood as anyone from bedroom musicians to professional producers and composers such as Shibuya. In *THE END*, Shibuya’s choice to leave the DJ booth could spell the end of Hatsune Miku’s singing. This is what *THE END* is about; the possible end of Hatsune Miku’s voice, abandoned by ‘her’ fans, who no longer desire ‘her’. Hatsune Miku’s audio-visual body is a desired object, a fetish object produced, consumed and shared by fans. In *THE END*, Hatsune Miku is presented as an erotic and commodified “sculpted object”, an object of exchange between fans (of all genders). Hatsune Miku’s audio-visual voice body is an object sculpted by fans, and a token of exchange that “circulates potentially endlessly across time and space, securing a network of future gazes” (Jones, p. 67). The future gaze, which one could argue is a heterosexual male gaze, also has the privilege of seeing without being seen. In *THE END*, this heterosexual male gaze is turned into a fetishizing or listening, where the man or fan, manifested by Shibuya as a male composer and fan, has the privilege of playing and listening to Hatsune Miku’s synthesized voice without his own voice being heard.⁷ In this staging of Shibuya as the male counterpart and orchestrator of Hatsune Miku’s feminine appearance and song, Hatsune Miku is inscribed in structures of fetishism as a gender stereotype and a fixed sexualized voice fully assimilated into the desire of others.

Staging a subversion of gender stereotypes

The staging of the synthesized voice technology that creates Hatsune Miku as a gender stereotype and fetish object can also be used to question gender binary stereotypes. Synthesized voices and digital bodies in media performances have the potential to disrupt the heteronormative gender binary system, complicating the understanding of immediacy as something that can only be achieved by a performance of human sociality through a citational practice reiterating the visible bodies in the gender binary system. Theatre and performance studies scholar Sarah Bay-Cheng

7) It is a difficult task to address the essentialist and dualistic gender politics in the experience of Hatsune Miku’s synthesized voice and visual staging in *THE END* without running the risk of undermining the critique by operating with other dualisms between ‘them’ (the male gaze) and ‘us’ (the feminist) and thereby reintroducing the essentialism being addressed. The notion of ‘us’ and ‘them’ are not global categories and can be parted up into other formations as it is argued for instance in intersectional thinking (hooks).

argues in her article “Virtual Realisms: Dramatic Forays into the Future” that there are many media performances and productions “that take up the questions of virtuality, technological dependence, and digitally transformed bodies” (Bay-Cheng, 2015, p. 689). Bay-Cheng writes that new media offer new possibilities for dramatic theatre to challenge the normative notions of e.g. gender identity (Bay-Cheng, 2015).

An example of this can be found in another media performance using synthesized voices, the multimedia dance performance *ROBOT*. Here seven 58 cm tall humanoid NAO robots are dancing side by side with human dancers.⁸ Choreographer Blanca Li, who has collaborated with Pedro Almodóvar, Beyoncé and Michel Gondry, is behind the production. *ROBOT* is a carnivalesque dystopian depiction of humans’ future life in coexistence with robots. The NAO robots have different personalities impersonating human characteristics and emotions, while the dancers mimic the somewhat rigid movements of the robots. In the show, the robots become more like humans. *ROBOT* brings to mind Czech author Karel Capek’s historical play *Rossum’s Universal Robots* from 1920, in which robots display human faults and vulnerabilities. Since *ROBOT* premiered in 2013 it has been performed in more than 100 theatres worldwide, from New York to Braga. I experienced the show in Teatro Circo in Braga, Portugal in the summer of 2015.

When I saw *ROBOT* there was one particular scene that caught my attention: the scene with the “clichéd flirtations” as Zoë Anderson writes in her review of the show. Here I heard the NAO robot’s default synthesized voice for the first and only time in the media performance. The default ‘boyish’ voice was used when the little robot, while performing calisthenics as a tiny version of a bodybuilder, asked the female dancer, lying on the floor next to it: “Don’t you think I’m strong?” making the audience in Teatro Circo laugh, including me. Nuance, the company behind the NAO robot’s default synthesized voice, calls it ‘Kenny’ and describes Kenny as a “he” that will match the NAO robot with a “custom text-to-speech voice specifically designed to match his personality” (Nuance, 2013).⁹ This framing of the voice fits with the depiction of the NAO robot as a ‘he’ on stage and in the press material for *ROBOT* (Li). The French male names given to the different NAO robot characters in the show (Sacha, Pierre, Jean, Alex, Lou, Dominique, and Ange) can also be said to promote an interpretation of the robots as masculine, reinforcing the normative constraints of the gender binary (Ellen Jacobs Associates, 2015).

At first, I thought that the gender performativity by the little robot was an example of how digital technologies are staged in media performances as something reiterating the regulatory norms of the female-male dichotomy in a Western gender binary system. I thought that the NAO robot’s performance of a stereotypical ideal of a strong male body in the flirting scene with the female dancer was staging the robot’s male audio-visual body as a fetish object, equivalent to how I have argued that Hatsune Miku is staged as a stereotypical ideal of a female body, as a fetish object or an other desired by the subject in *THE END*. Indeed the clichéd flirtations between the NAO robot and the female dancer in *ROBOT* can be argued to rely on a system of binaries, the gender binary and the binary relation between self and other. However, after reflecting more on the staging of

8) The NAO robots are all-purpose social robots used in service industry, education as well as entertainment (Jørgensen and Tafdrup, 2017).

9) NAO’s plastic robot body first got the ‘Kenny’ voice by Nuance in 2011. In connection with the launch of the collaboration between the then Aldebaran Robotics and Nuance, Steve Chambers, executive vice president in Nuance, talks about the ‘Kenny’ synthesized voice as something that should be perceived of as human-like; “By working with Aldebaran, we’re creating unique and compelling possibilities in the space of robotics where people can connect with NAO as if they were connecting with another human being – and that’s simply powerful” (Rigg, 2013).

the flirtation scene, it occurred to me that the reason why the audience in Teatro Circo laughed was because the little robot's question to the female dancer, which might play with referenced stereotypes of the ideal male body as well as the staging of robots as strong, seemed to fail bravely. It is, after all, ironic and a little bit silly that a tiny (58 cm tall) robot with a very high-pitched voice (as a young boy) that looks more like a cute toy than anything else (with a plastic body and clumsy movements), would ask someone that appears both much larger and stronger to adore the strength of its little body.



The “clichéd flirtations” between robot and human, Vaison Danses (2014).

The absurdity of the stereotype

The staged absurdity of the situation lies in its questioning of the robot as something performing human sociality. Here the play between the sonic body and the visual body is crucial. The synthesized voice Kenny and NAO's plastic robot body subvert the experience of the audio-visual robot body as something that can perform human sociality as a male stereotype. This is the opposite of the experience created in the staging of Hatsune Miku in *THE END*. In *ROBOT* the little robot can be understood as an object of attention for the female gaze on stage as well as the audience's gaze. The robot as an object and as clumsy technology is emphasized in its absurd attempt to perform a masculine stereotype and to perform as a visible subject. In this way, the whole scene operates within dichotomies of the binary system, of male-female and subject-other, yet subverts the gender performativity within the frame of the binary system by questioning our, the audience's, notion of gender stereotypes and robot stereotypes.

The subversion of gender stereotypes is also explored at another point during the show. Here, the same NAO robot is dressed up in a purple sequined dress and a pink boa, dancing, and singing to the song *Bésame Mucho*. This time, however, the song is not performed by NAO's default synthesized voice Kenny, but is a playback recording sung by a female singer. In this way the robot's body is connected to two different sonic bodies, a male and a female, underlining the arbitrariness (and absurdity) of the gender performativity in humanoid robot technology (and technologies designed to be human-like).

The absurdity of gendering the robot technology within a Western binary system as a performance of human sociality is further stressed by the use of playback technology. If this scene is an attempt to create an experience of a performance of human sociality, it fails, since the act of mediation, the playback technology, is obvious to the audience. The expressional vocabulary and gestures of the little plastic robot seem disconnected from the playback voice and do not match the emotional intensity of the playback song. Again, the robot is staged as an object in its absurd attempt to perform a stereotype and to perform as a visible subject.

This subversive play with stereotypes is something that Blanca Li, the choreographer of *ROBOT*, is known for. Her play with gender identities can be experienced in the music video *Around the World* by the electronic music duo Daft Punk, where she humorously had dancers perform different kinds of stereotyped bodies: skeletons, zombies, and robots.

Conclusion

In this article, I have discussed how representations of gender identity are performed and staged using synthesized voices as an artistic medium. In a comparative critical reading of two digital media performances, *THE END* and *ROBOT*, I have pointed out two different strategies for using synthesized voices as artistic medium, especially in relation to the interconnectedness of the synthesized voices and their visual framing (the sonic body and the visual body) in the audio-visual staging and representation of gender identities in digital media performances.

I have argued that the staging of Hatsune Miku's synthesized voice in the vocaloid performance *THE END* can be said to perform and reiterate a normative binary system, creating the audio-visual body as a fetish object assimilated to the desire of others. In this way, *THE END* can be argued to display the design and performance of a synthesized voice operating within binaries such as male-female and subject-other. As such, it somehow seems to ignore the critique of the binary system made by feminists, queer theorists, and poststructuralists throughout the last 35 years.¹⁰

The staging of the synthesized voice Kenny in the media dance performance *ROBOT* also arguably displays a design and performance of the synthesized voice as reiterating the regulatory norms of the Western gender binary and the binary of the subject-object. Contrary to the staging of Hatsune Miku in *THE END*, however, the staging of Kenny and the NAO robot in *ROBOT* seems to disrupt the notions of the heteronormative gender binary system. By staging the irony of the gender performativity of a robot with a little plastic body, and by further complicating the understanding of the interconnectedness between the sonic and the visual body, giving the same plastic robot a male and a female voice, the citational practice of reiterating the visible bodies in the binary system is challenged. *ROBOT* is a media performance that troubles the idea of identities

10) The gender stereotypes performed with the use of synthesized voices in both *THE END* and *ROBOT* can also be experienced in technologies with synthesized voices, most widely known in from the smart assistants Siri by Apple, Alexa by Amazon, Cortana by IBM, that all operate with a categorized male or female voice. Another example is the IBM Bluemix Catalogue with thirteen voices categorized as either male or female (Baird 2018; IBM Bluemix 2017).

and gender representations as fixed through a subversion of the staging of gender binaries and stereotypes.

My comparative critical reading of the media performances presented here has demonstrated how the discussion of synthesized voices as an artistic medium should take into account the interplay between the auditive performance, the visual framing, and the situated context. As an artistic medium, the synthesized voice can be said to operate with a gender performativity that creates stereotypical gender representations. However, the staging of synthesized voices in media performances matters, since it is possible to use the connection between the synthesized voices and their visual framing (the sonic body and the visual body) subversively. Future media performances with synthesized voices might even challenge the normative notions of gender identity and promote a multiplication of representational categories and bodies of difference.

Stina Hasse Jørgensen

is a PhD student at the the Department of Arts and Cultural Studies at the University of Copenhagen. In her research, Jørgensen has focused on topics such as sound art, media art, feminism, gender theory, vocal theory, interaction design, performativity and performance art. She has written articles and contributions to anthologies, online magazines like *Kunsten.nu* and *Seismograf* as well as for journals such as *Digital Creativity: Transformations; Body, Space & Technology Journal*; and *Cultural Analysis Journal*.

References

- Anderson, Z., 2017. *Robot, Barbican Theatre, London, review: The audience gasps in dismay when it falls over, and laughs fondly when it wiggles its fingers, asking to be lifted.* [online] Available at: <http://www.independent.co.uk/news/robot-barbican-theatre-london-review-blanca-li-dance-a7601601.html> [Accessed: April 19th 2018].
- Baird, A., Jørgensen, S.H., Parada-Cabaleiro, E., Cummins, N., Hantke, S., Schuller, B., 2018. The Perception of Vocal Traits in Synthesized Voices: Age, Gender, and Human-Likeness. *JAES Journal of the Audio Engineering Society*, vol. 66(4), pp. 277-285.
- Bay-Cheng, S., Kattenbelt, C., Lavender, A., and Nelson, R., 2010. *Mapping Intermediality in Performance*. Amsterdam: Amsterdam University Press.
- Bay-Cheng, S., 2015. Virtual Realisms: Dramatic Forays into the Future. *Theatre Journal*, 67, pp. 687–698.
- Benyon, D., 2014. *Designing Interactive Systems: A comprehensive guide to HCI, UX and interaction design, 3/E*. Harlow: Pearson.
- Bonada, J., and Serra, X., 2007. Synthesis of the Singing Voice by Performance Sampling and Spectral Models. *IEEE Signal Processing Magazine*, 24(2), pp. 67 – 79.
- Brautigam, B., 2017. *The New Skeuomorphism is in Your Voice Assistant.* [online] Available at: <https://uxdesign.cc/the-new-skeuomorphism-is-in-your-voice-assistant-3b14a6553a0e> [Accessed: March 16th 2018].
- Bremers, A., 2015. “The End’: a Virtual Opera Singer, Millions of Fans and the Meaning of Death.” *Next Nature Network*, [image online] Available at: <http://nextnature.net/2015/09/end-opera-players-virtual/> [Accessed: April 28th 2018].
- Butler, J., 1993. *Bodies That Matter*. New York: Routledge.

- Butler, J., 2009. Performativity, Precarity and Sexual Politics. *AIBR. Revista de Antropologia Iberoamericana*, 4(3), pp. i-xiii.
- Data.tokyogirl., 2015. *Miku Hatsune's Vocaloid Opera "THE END" to Be Performed in Shanghai Next!* [image online] Available at: <https://tokyogirlsupdate.com/miku-the-end-shanghai-20150751579.html> [Accessed: May 26th 2018].
- Dyson, F., 2014. *The Tone of Our Times: Sound, Sense, Economy, and Ecology*. Massachusetts: The MIT Press.
- Ellen Jacobs Associates., 2015. *From Paris with Robots*. [image online] Available at: http://www.ejassociates.org/press_releases/from-paris-with-robots/ [Accessed: May 18th 2018].
- Faber, L. W., 2013. *From Star Trek to Siri: (Dis)Embodied Gender and the Acousmatic Computer in Science Fiction Film and Television*. Ph.D. Southern Illinois University Carbondale.
- Fèvre-Berthelot, A. L., 2013. Audio-Visual: Disembodied Voices in Theory. *InMedia*, [online] Available at: <http://journals.openedition.org/inmedia/697> [Accessed: May 20th 2018].
- Haraway, D., 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, 14 (3), pp. 575-599.
- Harris, R. A., 2005. *Voice Interaction Design: Crafting the New Conversational Speech System*. San Francisco: Morgan Kaufmann Publishers/Elsevier.
- hooks, bell. 1984. *Feminist Theory: from margin to center*. New York: Routledge.
- IBMBluemix, 2017. *Using SSML*. [online] Available at: <https://console.bluemix.net/docs/services/text-to-speech/SSML.html#ssml> [Accessed: April 25th 2018].
- Jones, A., 2012. *Seeing Differently*. New York: Routledge.
- Jørgensen, S. H., 2016a. *Popstjerne af lys, lyd og software*. [online] Available at: <https://kunsten.nu/journal/popstjerne-lys-lyd-software/> [Accessed: April 15th 2018].
- Jørgensen, S. H., 2016b. *THE END: En Teknisk Operaoplevelse*. [online] Available at: <http://seismograf.org/the-end-en-teknisk-operaoplevelse> [Accessed: May 26th 2018].
- Jørgensen, S. H., Tafdrup, O., 2017. Technological Fantasies of Nao - Remarks about Alterity Relations. *Transformations*, 29, pp. 88-103.
- Jørgensen, S. H., Vitting-Seerup, S., and Wallevik, K., 2017. Hatsune Miku : An uncertain image. *Digital Creativity*, 28(4), pp. 318 – 331.
- Li, B., 2015. *Robotic pop ballet for all future generations*. [online] Available at: blancali.com/en/download/Robot2015-english.pdf/pdf_120_file_fr.pdf [Accessed: March 18th 2018].
- Marinucci, M., 2016. *Feminism is Queer: The intimate connection between queer and feminist theory*. London and New York: Zed Books.
- Mulvey, L., 1989. *Visual and Other Pleasures*. New York: Palgrave Macmillan.
- Nakano, Y., Neff, M., Paiva, A., Walker, M. ed., 2012. *Intelligent Virtual Assistants*. Berlin: Springer-Verlag.
- Nass, C. and Brave, S., 2005. *Wired for Speech: How Voice Activated and Advances the Human- Computer Relationship*. Massachusetts: The MIT Press.
- Nuance, 2013. *Aldebaran robotics and Nuance revolutionize human-machine interaction*. [online] Available at: <http://whatsnext.nuance.com/connected-living/aldebaran-robots-nuance-revolutionize-human-machine-interaction/> [Accessed: May 17th 2018].

- Phan, T., 2017. The Materiality of the Digital and the Gendered Voice of Siri. *Transformations*, 29, pp. 23-33.
- Rahbek, B., 2016. *Hun er 16 år, har millioner af fans - og findes ikke*. [online] Available at: <https://www.b.dk/nationalt/hun-er-16-aar-har-millioner-af-fans-og-findes-ikke> [Accessed: February 18th 2018].
- Rigg, J., 2013. *Nao robot to become even more of a chatterbox with new software (video)*. [online] Available at: <https://www.engadget.com/2013/10/30/nao-robot-new-nuance-voice-software/> [Accessed: May 26th 2018].
- Robertson, A., 2016. *Google's DeepMind AI Fakes Some of the Most Realistic Human Voices Yet*. [online] Available at: <https://www.theverge.com/2016/9/9/12860866/google-deepmind-wavenet-ai-text-to-speech-synthesis> [Accessed: May 25th 2018].
- Silverman, K., 1984. Dis-Embodying the Female Voice. In: Doane, M. A., Mellencamp, P., Williams, L., eds. 1984. *Re-vision: Essays in Feminist Film Criticism*. Los Angeles: University Publications of America, pp. 131-149.
- Sjogren, B., 2006. *Into the Vortex – Female Voice and Paradox in Film*. Chicago: University of Illinois Press.
- Sloan, Robin J.S., 2015. *Virtual Character Design for Games and Interactive Media*. Boca Raton, FL: CRC Press.
- Stojnic, A., 2015. Digital Anthropomorphism. *Performance Research: A Journal of The Performing Arts*, 20(2), pp. 70-77.
- Vaison Danses., 2014. *Blanca Li - Robot ! - Festival Vaison Danses 2014*. [image online] Available at: <https://www.youtube.com/watch?v=VHCA1eEW7jI> [Accessed: February 26th 2018].