



Voices from the uncanny valley: How robots and artificial intelligences talk back to us.

Downloaded from: <https://research.chalmers.se>, 2025-06-18 04:41 UTC

Citation for the original published paper (version of record):

Männistö-Funk, T., Sihvonen, T. (2018). Voices from the uncanny valley: How robots and artificial intelligences talk back to us.. Digital Culture and Society , 4(1): 45-64.
<http://dx.doi.org/10.14361/dcs-2018-0105>

N.B. When citing this work, cite the original published paper.

Voices from the Uncanny Valley

How Robots and Artificial Intelligences Talk Back to Us

Tiina Männistö-Funk & Tanja Sihvonen

Abstract

Voice is a powerful tool of agency – for humans and non-humans alike. In this article, we go through the long history of talking heads and statues to publicly displayed robots and fortune-tellers, as well as consumer-oriented products such as the late 19th century talking dolls of Thomas Edison. We also analyse the attempts at making speaking machines commercially successful on various occasions. In the end, we investigate how speech producing devices such as the actual digital assistants that operate our current technological systems fit into this historical context. Our focus is on the gender aspects of the artificial, posthuman voice. On the basis of our study, we conclude that the female voice and other feminine characteristics as well as the figures of exoticized and racialized ‘Others’ have been applied to draw attention away from the uncanniness and other negative effects of these artificial humans and the machinic speech they produce. Technical problems associated with the commercialization of technologically produced speech have been considerable, but cultural issues have played an equally important role.

Introduction

I’m a real, live electronic girl. I would like to go out into the world and live with people. I can serve them, entertain them, and even help the elderly and teach kids. I can animate all kinds of human expressions but I am only starting to learn about the emotions behind those expressions. This is why I would like to live with people and learn from these interactions. Every interaction I have with people has an impact on how I develop and shapes who I eventually become. So please be nice to me as I would like to be a smart, compassionate robot. I hope you will join me on my journey to live, learn, and grow in the world so that I can realize my dream of becoming an awakening machine. (Sophiabot 2017)

In autumn 2017, an artificially intelligent humanoid robot named Sophia made headlines in the international press for speaking publicly at the Future Investment Initiative Conference in Riyadh and gaining the full citizenship of Saudi Arabia. As her speech and interview started circulating in social media on 25 October, her

performance quickly got viral and sprung conversations about robotic rights, citizenship, and the sentience or “intelligence” behind her conversational skills. Being developed by a Hong Kong-based robotics company Hanson Robotics, Sophia is the latest instalment of AI-based learning robots that, according to its creators, is designed to embody the positive aspects of the uncanny feeling that humans often experience while interacting with artificial humanoid figures (Hanson et al. 2005).

Sophia is thus an example of a robot that is developed to be as pleasant and likeable as possible. “Please connect with me and be my friend”, she says on her website, written in the first person (Sophiabot 2017). In addition to holding eye contact, recognizing faces, and engaging in real-time conversation, she also expresses emotions to “understand humans and build trust with people”. Another important factor in her likeability is her quintessentially feminine looks: she is designed to imitate Audrey Hepburn’s classic beauty with a “porcelain skin, a slender nose, high cheekbones, and intriguing smile, and deeply expressive eyes that seem to change colour with the light” (Stone 2017). Interestingly enough, it was possible for the Saudi audience to admire Sophia’s physical appearance, since she was not wearing the traditional headscarf or *abaya*, the cloak that women in Saudi Arabia are obliged to wear public. In spite of her assumed status as a “natural person” and a citizen of Saudi Arabia, journalists were quick to note that she also did not appear to have a male companion who would have authority to act on her behalf, which is required under the Saudi guardianship system (Molina 2017; Sini 2017).

Sophia and the identity politics she quite literally embodies are a fascinating recent example of the complex relationships people have with moving, human-shaped machines. The figure of an android, a humanoid robot, is deeply ingrained in the Western psyche, characterised by the ambivalence of admiration and anxiety, a simultaneous sense of wonder and a fear of being substituted by it (Telotte 1995, 3–4). The fascination with human-like machines is one with a long history. Moving, clockwork-based automatons have been built for hundreds of years and their history is fairly well known. As the case of Sophia demonstrates, however, human-like movement, eye contact, and even many aesthetic qualities in humanoid robots may be relatively well developed, but the voice of these machines still presents a credibility problem. In the case of Sophia, her speech appears almost too smooth in comparison with her abrupt facial expressions, and the combination of these two makes her seem disturbingly out of sync as a partner in conversation. Human voice is produced by a biological system that is – or has at least so far been – close to impossible to reproduce in a machinic form. Therefore the question becomes, how artificial beings can effectively simulate the medium of the human voice. How do robots speak, can we have conversations with them, and whose voice do we hear when they talk back to us?

In this article, we will analyse historical and contemporary representations of talking robots and AIs, and study their associations with the development of real-world speech-producing devices, from the early talking heads and statues to the actual AIs that operate our current technological systems. We argue that

taking voice in consideration when thinking about people's emotional reactions to human-like machines is of central importance, as the power of speech has been an essential characteristic separating people from the rest of the animate and inanimate world (Keane 2001; Pettorino 2015, 31). Voice and speech seem to lay at the very core of the questions concerning the difference between human and machine. As Amanda Weidman (2006, 8) points out, voice is always associated with agency and sincerity, and it is at the heart of notions of the rational subject. On the other hand, the presence of a voice, a particularly human trait, in artificial beings has manifested the utopia of hybridization of human beings and machines (Beiguelman 2010, 343). Voice therefore signals a powerful tool of agency – for humans and non-humans alike.

Nevertheless, the sound or voice of historical automatons is rarely discussed, and most of these machines indeed did not speak. Early robots were equally silent. Historical automata and moving human-like figures were sometimes accompanied by music, which highlights the common past between these clockworks-operated machines and musical instruments that produced sound mechanically and independently from their operator (Riley 2009). The oldest *fictional* creations of artificial humans and automatons, however, spoke fluently, even in antiquity, although the robots in the 20th century media often had tinny or cracked voices that highlighted their machinic nature. The most recent 21st century robots and AIs in fiction, then, are capable of perfectly human-like speech and voice. Despite this, the voice in artificial beings is far from a simple and clear-cut question (e.g., Nass & Brave 2007; Phan 2017; Sandygulova & O'Hare 2015).

According to Masahiro Mori's term *uncanny valley* (1970, translated from Japanese by Jasia Reichardt in 1978), the feelings of affinity towards human replicas grow as the human likeness of the replica increases, but can abruptly dip into feelings of disgust when the replica reaches an *almost* human-like stage. Designers and researchers of computerized voices and speech technology are still trying to cross the uncanny valley and create talking AIs that would be indistinguishable from humans and thus not unpleasantly creepy. However, the role of voice in the culturally perceived uncanniness of robots and AIs has not been researched enough for us to understand how this process actually works (e.g., Crowell et al. 2009). Furthermore, the voice is of course not only an instrument of speech but also a representation and a product of the speaker's identity and corporeality (Phan 2017). If speech is the implication of embodied experience, a talking robot hints at the possibility of a biological body behind the mechanical layer. This "biocybernetic" body not only forces us to ponder the limits between organic and inorganic, but through the human-machine vocalic interface it also guides us to redefine our own status in relation to technology (Beiguelman 2010, 346). Many contemporary representations of robots therefore negotiate the philosophy of artificial intelligence, of sentience and emerging consciousness, and the psychological dimension of machines and playthings becoming "people", thus highlighting the ever-narrowing split between real and artificial humans.

As we consider voices produced by robots and AIs, we have to ask, first, who the speakers are, and second, on what terms we can have conversations with them. In our quest to study the humanlike or non-human activity and agency robots and AIs have through their voices, we will rely on some of the theoretical insights provided by new materialism and posthumanism, and also discuss the gendered aspects of the artificial voice. Although questions concerning the agency and subjectivity of non-human beings have gained specific momentum especially in the feminist theory of the 1980s and 90s in the wake of Donna Haraway's 'Cyborg Manifesto' (1984/1990), our treatise is built on the notion of the *recurrence* of the attempts of giving voice to artificial humans. Through looking at the history of these repeated and often failed attempts, our approach resembles *media archaeology* (Zielinski 2006) in that it refutes linear progress in technological development, and in that it is not only interested in success stories but also marginal, obsolete forms of media. In the end of this article, we hope to have shown how the current development of talking robots and AIs can be contextualized within an understanding of and practical applications questioning the limits between human and machine, dating back hundreds, even thousands of years.

Long history of talking heads

In order to understand the cultural and social responses to the voices of and verbal communication with AIs, robots, and other speaking technologies, we need to understand the contexts that the responses are linked to and born in. It is useful to look at the origins of talking machines and the soundscape that surrounded them to make sense of the complex issues related to machines that talk with a human voice. In this section, we will take a concise look at the history of humanlike machines and machines talking with a human voice and develop historical insights into the subject. Massimo Pettorino (2015, 31–32) has categorized the historical development of talking heads into two main strands: that of the “voice transport” and the “artificial voice”. The voice transport has actually been a collection of methods of delivering the voice through an artifice to a fake head, whereas the creation of artificial voice has throughout centuries been a research interest where the functions of the human phonation apparatus have been imitated or simulated through technological means. In this section, we will take a look into these intermingling histories of the technical and cultural challenges of producing talking machines and finding uses for them.

Artificial humans have interested designers and researchers throughout centuries (e.g., Bruce 1913). The oldest known description of automata appears in Homer's *Iliad* (8th century BCE), where Hephaestus, the Greek god of blacksmiths, is assisted by golden maidens that “have sense in their hearts, and speech and strength” (Il. 18. 417–420). Such speaking, thinking feasts of metalwork were to be understood as divine and magical wonders, but already by the first

century CE, Hero of Alexandria's mechanical inventions like the automatic holy water dispenser, temple-door opening system, wind-operated organ, and theatrical effects may have given the impression that such wonders might be possible to achieve by technical mastery (Bosak-Schroeder 2016). Some of Hero's designs were also either human-like or sound-producing – such as singing birds and moving satyrs or deities – and many of them were powered by the heat of the sun. The most famous “talking” statue of the ancient world, the Colossus of Memnon in Egypt from 1350 BCE, most likely produced sound through changes in air pressure due to solar energy, too. It is quite fascinating that the legend of the statue's vocality and oracular powers spread throughout the ancient world so effectively that the colossi became a popular visitors' attraction (Pettorino 2015, 37–38).

Moving statues are another example of human-like machines that are mentioned in ancient texts, and statues with speaking tubes have been found by archaeologists in Greece and Egypt. The tradition of animated, possibly voice-producing statues continued in medieval Europe all through to Renaissance, combining magical thinking and technical as well as presumably alchemical skills. Especially the talking statue known as the Android, built and placed in a cell of the Dominican monastery in Cologne by Albertus Magnus in the 13th century, seems to have invoked a great deal of commotion in the following centuries. The Android's principle of producing voice was explained by Athanasius Kircher as late as mid-17th century (Pettorino 2015, 32–35). Legends of speaking, oracular heads, built with the help of pacts with the Devil, were connected to many important intellectual figures, from the 10th century Gerbert of Aurillac (later Pope Sylvester II) to the 13th century natural philosopher and founder of empirical scientific methodology, Roger Bacon. These legends can be interpreted as attempts of alleviating the concerns about the powers of new, emerging natural knowledge as well as non-Christian, ancient, and Arabic beliefs (Truitt 2011). They also seem to highlight the conceptual importance of artificial speech in negotiating the categories of the known world.

Another example of real and fictional medieval and Renaissance-era automata are memorial statues on tombs and in mausoleums, found in different parts of Europe. Their purpose was to physically imitate dead loved ones. By providing a site not only for mourning but also for “interaction,” they in effect complicated the distinction between living and dead bodies. Especially in fictional accounts, these statues moved and also spoke like the deceased. Between the 12th and the 15th century, these descriptions got reconfigured from a magical understanding of such effects to the mechanical, demonstrating a general shift in thinking about the possibilities of machinic speech (Truitt 2011). By the end of Renaissance, magic was no longer seen as a probable source of speech in machines, but the possible creation of speaking machines through mechanical solutions remained interesting at the same time as the whole understanding of the world and living organisms was moving to the direction of mechanist explanations and Cartesian machinic thinking (Hattab 2005; Kang 2011; Pettorino 2015).

Mechanical philosophy became the central paradigm of thinking about the universe, society, and living organisms alongside the 17th century scientific revolution in Europe (Bowler & Morus 2005). Hankins & Silverman (1995) sort out different motivations for building or trying to build speaking machines throughout ages: many of these have long and partially overlapping histories, including the magical or esoteric reasons to produce the appearance of artificial speech, as we can see in the examples above. From the late Renaissance onwards, the physiological interest in understanding the mechanism of speech grew, followed by acoustic interests in human sound production as well as technical interest in recording and transmitting speech.

Various kinds of automata served an important purpose in illustrating the mechanical functions of life, and they were popular both among the general public and the learned society. Jacques de Vaucanson's famous flute-player, drummer, and duck automata started an automaton craze that lasted from the mid to the latter part of the 18th century. These machine-bodies represented the fundamental idea in mechanical philosophy that considered all living bodies as complicated machines. Some artificial humans were warm to the touch, they breathed and even bled. However, speech remained a threshold of separation. In the *Discourse on the Method* (1637), René Descartes explains mechanist physiology in humans and animals, but states that animals only have inferior intelligence arising from the mechanics of their body-machine, whereas humans have a rational intelligence arising from their immaterial soul. A central sign of this intelligence in humans was their ability to speak (Riskin 2016; Kang 2011; Liu 2000).

Remarkably, however, the lively scientific discussion around the possibility of artificial speech mostly concentrated on the physiological mechanism of human speech, which was thought to be impossible or very difficult to copy in a machine form (Riskin 2016). Some of the most famous human-shaped machines produced in this era were elaborate hoaxes, like The Turk, a chess-playing automaton that later was revealed to have been human-operated – nevertheless, even that machine was ascribed intelligence and agency through a process of audience enchantment and rationality (Jarrett 2016, 1). Interestingly, Wolfgang von Kempelen, the designer of The Turk, also invented the manually operated, mechanical speaking machine that has later been named as the first successful speech synthesizer. Produced around 1780, the speaking machine was not imitating human appearance in any way, in contrast to the chess-playing automaton, but was a formal model of human vocal tract placed in a wooden box and operated by kitchen bellows, demonstrating the mechanics of human speech (Dudley & Tarnoczy 1950). Around the same time, other mechanical speaking machines and heads were produced and presented to scientific academies, like abbé Mical's discoursing heads that could perform a dialogue in praise of Louis XVI, and C. G. Kratzenstein's human voice instrument constructed out of organ pipes (Riskin 2016).

Mechanist thinking was criticized and challenged, but human-like automata became a popular amusement in the late 18th and early 19th century. At

the same time, they started to populate the contemporary fantastic literature, where they more often than not appeared as scary and eerie, alongside with such creatures as ghosts, vampires, and reanimated corpses that also questioned and crossed the categories of animate and inanimate, natural and artificial (Kang 2011). Terry Castle (1995) defines – in line with Sigmund Freud's essay *Das Unheimliche* (1919) – the uncanny feelings raised by human-like doppelgängers as an 18th century invention. In his treatise of the uncanny effect, Freud considers the impression made by “wax-work figures, artificial dolls and automatons,” and cites the example of Olympia, an artificial doll that has all the attributes of a living being, as an object of attraction while discussing Hoffmann's classic tale of the Sand-Man (Freud 1919, 5). Paradoxically enough, the Enlightenment's quest to offer rational explanations to all phenomena, and the social diffusion of this rationalization that replaced magical thinking, seems to have led to gradually growing and stronger cultural feelings of strangeness, alienation, and unease (e. g., Liu 2000).

It can be argued that many of the 20th and 21st century fictional representations of artificial humans can be understood as commenting on the fears and uncanny effects connected to the ongoing quest to achieve technological mastery and control over the world. Even the cute and obedient fictional robots R2-D2 and C-3PO in Steven Spielberg's *Star Wars* movies hint at the darker side of robotic existence (Kang 2011). Typical to the fictional accounts involving robots and androids in the 20th century has been the storyline involving the hubris of humans trying to create their likeness and only leading to their nemesis. Robots longing to be like humans is another typical and lasting feature of these fictions (Segel 1995). This kind of anthropocentrism is visible in many audiovisual depictions of androids from past decades, as well as the description Sophia the humanoid robot mentioned in the beginning of this article gives of herself on her website.

Most often, fictional robots were capable of speech, even the artificial revolutionary leader Maria in Fritz Lang's *Metropolis* (1927), although we can only see and not hear her speak, the film being a silent one. The real robots of the 20th century were incapable of speech, but remarkably often designed to give an impression of the ability to produce and/or understand speech for the sake of the show. For instance, the first robot in Britain, Eric, built in 1928, gave speeches that were transmitted through it by a live radio signal. Jessica Riskin (2016) quotes contemporary newspaper reports on Eric that indicate his radio-transmitted speech to have been somewhat out of sync with his awe and horror inspiring looks. The American robot Elektro, built by the Westinghouse Electric Corporation one decade later, could speak some 700 words by the use of a record player (among them the phrase, “My brain is bigger than yours”), and move to voice command. When commanding Elektro, the actual words spoken did not matter but their amount that sent a certain number of light impulses activating or releasing a series of relays inside the robot (Sharkey 2008).

We now turn to look at the history of commercial applications of talking machines, and the role that the uncanny effect has played in it. We will see how

selling the core of conceptual horror has throughout centuries been very difficult, and how femininity as well as racial and ethnic otherness have been utilized in trying to pave the way to the consumer's heart. Although the methods of producing artificial voice are at this point mostly based on the mechanical imitation of the human phonatory organs, there are also examples of the other path, the production of voice by a hidden subject, and the transportation of it through an artifice or a specific medium (Pettorino 2015). Even today, the most wide-spread methods of creating artificial voice, diphone synthesis and domain-specific synthesis, are based on pre-recorded speech segments that are produced by an actual speaking subject, stored on a database, and used in talking clocks and toys, or to deliver transit schedule announcements and weather reports. As is evident from these examples, the artificial human voice has not quite yet made its way into our everyday life and in the appliances we use.

How (not) to sell the uncanny effect

After the centuries-long quest to produce speech artificially and the turmoil that successful attempts caused, one would assume the machinic voice to be an attraction to the general public and an easily marketable commodity at best. It turns out this was never the case. When the mid- and late 19th century brought about the first human-like machines that were capable of somewhat continuous and intelligible speech, they were no economic or popular successes. In the 1840s, Joseph Faber built an elaborate speaking head that could be made to utter words mechanically with the help of a keyboard and pedals. The head was first dressed as a Turkish man and Faber exhibited it in New York and Philadelphia to largely uninterested audiences. Later on, Faber's machine, renamed Euphonia, was given feminine clothes and a hairdo with ringlets. It was exhibited in London and Paris until the 1870s. The inventor Joseph Henry, after having seen Euphonia and inspected it to ensure that it was not a hoax, envisioned that it could be connected to the telegraph to make it speak the words of the received message. He even saw possibilities in it delivering one sermon in several churches telegraphically. But none of this was to be reality, and after the 1870s all traces of Euphonia disappear (Lindsay 1997; Riskin 2016).

Another commercial failure were Edison's talking dolls that went into production in 1890. Some 55 centimetres tall, the dolls included a miniature version of the phonograph, Edison's famous invention from 1877. Each doll could recite one of a dozen nursery rhymes, summoned up by rotating a hand crank at the doll's back. As the recordings on phonograph cylinders could not be copied, the production of the dolls necessitated a group of young women, reciting nursery rhymes into the machines to create a unique recording for each doll. The dolls were in production for mere weeks before being withdrawn from the market, due to numerous customers complaining about the weak durability and sound quality of the wax

cylinders. What is now regarded as the first attempt to produce sound recordings for commercial purposes was over quickly (Dawson 2017). Interestingly, Edison had envisioned also many other uses for the phonograph that it did not become used for, such as creating recorded sound memories of family members' typical sayings or the last words of dying loved ones. Recorded voice messages never reached the same cultural status as carriers of memory that photographs and written diaries have, although voice recordings have been used from the beginning on for example for the purposes of dictating or learning languages (Kenney 1999).

A third example to consider as an early attempt to commercialize speaking human-like machines are the fortune-telling machines. Coin-operated electric fortune tellers were mostly human figures fitted in glass boxes. They were often placed in penny arcades that started to appear during the first decade of the 20th century, and became one of their central attractions from the late 1920s onwards, especially in the United States. Amusement parks were another typical site to place them. After receiving a coin, they would "come to life" by, for example, moving their hands, heads and eyes, and breathing before delivering a fortune or advice, most typically on a small printed card. They often also included light-effects: for example, their eyes started glowing once a coin was inserted. Remarkable for our interests here is that most, if not all human-like fortune-teller figures were either female, racially exoticized, or both. Between the turn of the century and the late 1960s, popular types were variations of "gypsy" and "grand-mother" figures, sometimes also combined into "gypsy grand-mothers" (Costa 1988).

Another interesting point to consider is that only very few of the fortune tellers included voice effects. Some of the more recent ones did, like Zoltan the fortune teller, introduced in the mid-1960s in Northern American game arcades. In contrast to most of the older machines, it was not animated but it spoke, using an endless loop tape recording. Zoltan was a male figure, dressed as an Arabic sultan, and spoke English with a heavy accent, reinforcing his supposedly "mystical" oriental origin. The same was apparently not true to the rare early verbal fortune teller, produced by the Mills Novelty Co around 1904. This fortune teller was a "gypsy queen", whose glass box advertised, "your fortune told by the human voice". Surviving in a museum in Virginia City, Montana, it is not functional at the moment, but, according to the staff, the museum has a recording of her original voice, which has a posh English accent, assumable not like the one to be expected from a "gypsy". Dropping a coin in a slot would make her eyes flash and mouth move and a hidden record player would deliver the fortune through a cone fitted in front of the box. Some others, like a 1950s Genco Horoscope Grand-mother fortune teller would not speak, but prompted the fortune-seeker to ask a question using a "vibra-phone" fitted in front of the glass-box. Although titillating in its performance, the vibra-phone was actually just a round piece of perforated metal with no technical function (cfhatprovidedotnet 2014).

The surviving examples of verbal fortune-tellers prove that the idea of using any technology necessary to produce voice was present in their design. One would

also assume the voice effect to be a welcome one in machines like fortune-tellers that specifically made use of the uncanniness or magical connotations of human-like machines. Voice producing technologies might have been relatively more difficult to maintain than the rest of the fortune-teller machinery, and apparently did not add sufficient extra attraction to make them lucrative. Accordingly, we can detect the older pattern repeating itself: although inventors and producers were drawn to the challenge of a speaking machine, they failed to commercialize any of the products developed. Until the present day, speaking consumer technologies have remained relatively rare, even though digital technologies have provided new and easier possibilities to manufacture and update them. To a certain degree, the toy industry represents an exception, although even there talking machines are not any kind of majority. Despite this, it has to be noted that toy technologies have in the 19th and 20th centuries been an important way of testing and spreading new kinds of technological experiences (Poser 2016). However, as speaking toy technologies have already been around for a century and a half, the rarity of other speaking consumer technologies raises questions on emotional responses to such technologies. Are they too uncanny to be commercialized?

Technological soundscapes and the uncanny

The soundscape as the audible human environment is at the same time psycho-physiological and culturally constructed, as it includes the perceptive mechanisms of the hearer and listener that are socially and culturally shaped. Therefore, we hear the same sound or voice differently in different spaces and times, or spatio-temporal circumstances (Thompson 2004; Sterne 2003). When the phonograph was invented and promptly made use of as a music technology, for instance, there already was an existing tradition of hearing music that was produced by machines. To a certain extent, many instruments could be understood as machines, especially the elaborate constructions of accordions and organs, as well as pianolas or player pianos powered by clockwork technologies. There also existed a lively and wide-spread tradition of mechanical music, mostly in the form of small music boxes at homes or barrel organs played at fairs, but also, occasionally, extending to large mechanical orchestras or orchestrions (Riley 2009).

Accordingly, there was a tradition of hearing and using music produced by machines, whereas speech produced by machines was a different subject altogether. James Donald (2008) has discussed the specificity of the uncanny effect experienced during the first three decades of the twentieth century that appeared when the disembodied human voices began to be delivered through gramophone, telephone, radio, and sound films. The mass-produced and technically reproduced human voice without a talking body often had a disquieting effect on audiences, the voice appearing as a ghost in the machine, emerging without a proper or apparent physical cause. One means of deflating this feeling of uncanniness was

using otherness to distract audiences and draw their attention to something else instead, such as to the materiality of the voice (for instance, by heavy accents) or to “those for whom the referentiality of the rational speaking subject was not allowed (e.g., black actors)” (Taylor 2009, 3). In the case of early sound films, this meant producing a remarkable amount of films with African-American actors. Here, we can detect similar uses of female and racial ‘Others’ as the commercial speaking heads discussed in the previous section.

With the spread of the telephone, the radio, sound film, loudspeakers as well as electrical megaphones and microphones, the soundscape of the early 20th century advanced to include more and more technologically produced or enhanced human voices. By the end of the 1920s, audiences in America were quite used to hearing human voices emitted from public address systems and the cinema experience was revolutionized by the sound film (Thompson 2004). However, as Jessica Taylor (2009, 2) argues about the sound film, not “just anyone’s voice was validated by the media as worthy of being heard and appropriate for mass consumption [...] the unity of body and voice was a desirable thing, but only when there was a ‘match’ between the social meanings of the voice and those of the image.”

It is interesting, then, that female voices at that time were attacked on as being too “shrill” or otherwise unpleasant for example on the radio, but it was female operators who were behind the functions of such voice technologies as the telephone, and later also Voder, the first electronic voice synthesizer, as well as the Sonovox throat microphone that could make different instruments speak the words spoken through it, both presented in 1939. Both the gramophone and the radio also catered to largely female audiences, especially during daytime. Sonovox was used in the advertising industry to make consumer products to “speak” to this female audience addressing its issues and anxieties (Kenney 1999; Smith 2008). It is stimulating to contemplate the dissonance between female characteristics being used by mostly male inventors and producers to relieve the uncanniness caused by speaking machines or to add to their attractiveness, and the large share of women in the user base of these products and services. This dichotomy is something that we can follow up to the present day.

As the soundscape changed during the 20th century, the immediate uncanny effects of the technologically transmitted or produced human voice faded. However, as Tom Gunning (2003) has proposed, these technologies retained their original ability to produce the sense of uncanny in a latent form. Especially technical failures or glitches in the smooth working of these technologies might make their uncanniness appear with a feeling that magical operations were involved. Especially fictional representations would make use of this, for example in the form of phone calls from dead persons, communication over the distorted television signal, or other eerie technological vocalizations. As in the case of robots, these fictional representations also affected the responses to real-world technologies, reminding us of their lurking uncanniness in the midst of rational everyday operations.

“Othering” of the artificially produced voice

During centuries, the operations and mechanisms of talking statues were attributed to the Devil, and many of the early examples of automata and clockworks have been lost due to violent attacks against their makers and their assumed magical or supernatural origins. It is no wonder, then, that through these machines also the mental border between the Western world and the Eastern “barbarians” was negotiated. The complex connections between the Christian Europe and the Arab world are reflected in the histories of talking heads and statues in many ways. In St. Eulalia Cathedral in Barcelona, for example, is a centuries-old statue of the head of a Moorish king that opens its mouth and rolls eyes when an attached rope is pulled, and quite likely it used to produce sounds as well. The head of the Moor symbolizes the Christian tradition of ridiculing the enemy by hanging their severed heads on stakes after a successful battle – or in this case, on organ pipes (Pettorino 2015, 38).

As we can see from our historical account of talking machines, many of these human-shaped automata were designed to invoke notions of the ‘Other,’ mostly of women, but also of racialized and mystified minorities, such as black or Romani people. Similar tendencies can be detected in the development of modern robots and artificial humans (e.g., Aytes 2013). In the beginning of this article we mentioned Sophia, who represents an important contemporary example of the merging of robotic technologies and feminine beauty ideals, but another emerging trend considers the bodiless personal assistant AIs who are operated through voice commands and who often respond to us with a feminine voice. At the moment, Apple has Siri, Microsoft has Cortana, Amazon has Alexa, and in addition there are hundreds of smaller AI-based applications for personal and business use (Hänel 2017). Social media giants such as Google and Facebook are developing their own assistant applications that are currently named DeepMind and M, respectively. The interesting aspect of these applications is their ability or inability to respond vocally and have conversations with their users.

Developments in automatic speech recognition and machine translation, as well as digital database-driven assistance and language technologies have made it possible for iPhones and other smart devices to have their own artificially intelligent assistants that are customisable (to a degree) and can be trained by recurrent interactions with their users. Despite their “smartness” and malleability, these natural-language systems need to be manually programmed in order to have convincing conversations with real humans: Apple’s Siri, for instance, is being taught by past (anonymized) interactions, and Cortana is being trained by people such as playwrights, poets, and novelists (see Greene 2017). Another interesting aspect in the development of these AI assistants is their status and the roles they are made to perform, especially if we consider the long Western history of talking machines. With voice come notions of and allusions to gender, class, race, and ethnicity, although on the surface robots and AIs are not supposed to have any of

these identity markers. Especially in fiction, the blending of human and machinic characteristics in robots and AIs is revealing, as we see in the romantic drama about the relationship between a man and an intelligent computer operating system Samantha in Spike Jonze's movie *Her* (2013).

Key figure in the development of robotics, Alan Turing, was concerned that "thinking machines" were likely to be exploited because their sentience was different from humans'. As we have illustrated, humanoid automata and robots have long been associated with any exploited class of person, even slaves – as is well known, even the word "robot" comes from the Czech "robota" (drudgery, servitude) – and we see similar trends continuing with the development of the industrial, caretaker, and assistant robots of today. It is not a coincidence that so many of the AI assistants and robot helpers of past decades have had particularly feminine voices and other defining characteristics (Jarrett 2015). Since they operate primarily through (spoken) language, their composition is arguably more "human" than that of other robot or AI technologies. As digital assistants are imbued with voice and personality, they are also given a certain materiality and tangibility in order to persuade us, their users, to trust them and the interactions they provide (Phan 2017). Interaction with these machines is, however, not a conversation on equal terms. There is a lack of scholarly research on this topic, but at least the importance of acknowledging the designers and makers of these speaking machines has recently been raised as a crucial issue in the popular press (e.g., Anderson 2013; Hempel 2015; Nickelsburg 2016; Seaman Cook 2016).

How the "Others" talk back to "us"

Through the history of human-like machines, voice and intelligible speech appear to be the most difficult problems to solve and also the most crucial thresholds keeping the human and non-human activity and agency apart. A machinic voice makes the listener to wonder what makes it speak and how much actual autonomy and individuality might be found behind the words uttered. The question of voice is central in the posthumanist discussion of the manner in which different actors have agency in different situations. The perceived spontaneity and intelligence of robot speech have not been experienced in the real world as of yet, but already the expectation has influenced the way in which we think about the possibilities of non-human speech, and look – both eagerly and fearfully – forward to closing that conceptual gap separating us from the animal and material world that surrounds us. As Haraway (1992) has stated, cyborgs are monsters, both promising and dangerous.

From antiquity to the present day, people have imagined automata, robots, and other human-like creations that would possess the power of speech, and the abilities of machines to understand and produce speech have been anticipated and even deliberately staged, while at the same time actual speaking automata

or robots have been extremely rare or non-existent. In the midst of human-like machines confusing the boundaries of human and non-human in an uncanny manner, speech has stayed put as the ultimate barricade that could not result from technical skills alone, but would necessitate the operations of either a living soul, divine powers, magic, or extreme technological prowess. At the same time, fictional representations of speaking machines have constantly blurred this border. Kathleen Richardson (2016) argues that there has never existed a boundary between robotic fiction and robotic science as the latter has followed the former. In the case of voice and speech the fiction has continuously created expectations of robots speaking to us.

Speaking machines have seemed to attract inventors, but as we see in the history of Joseph Faber's Euphonia, Edison's talking dolls, and possibly also in the case of the verbal fortune tellers, they were difficult to commercialize and often did not lead to economic success stories. We see that to try and make a speaking machine a commercial success, female voice and other feminine characteristics as well as the figures of exoticized and racialized 'Others' have been applied to draw attention away from uncanny and other negative effects. Technical problems associated with commercializing the technologically produced speech have been considerable, but cultural issues have played an equally important role. The modern soundscape has come to feature a multitude of technologically transmitted and reproduced human voices, but speaking consumer products have not turned into widespread successes in our daily lives.

Minsoo Kang (2011) has categorized reactions to automatons and other machines on different physical and conceptual levels. When the physical power of a machine is irrelevant, but the machine itself is useful or beautiful, it causes fascination. A powerful and potentially dangerous machine that is under control causes sublime feelings, whereas a similar kind of machine can cause terror, if out of human control. This applies also to human-like machines, but they cause additional sensations on a conceptual level. According to how convincingly an automaton imitates life, it can cause *amusement* (not convincing at all), *uncanny sublime* (very good imitation, but clearly not alive), or *horror* (impossible to know whether the automaton actually is a living being or not). In addition to the general feelings of fascination, sublime or terror caused by technology, automata have the power to threaten one's entire sense of reality by being categorical anomalies and liminal objects in the world of machines, artifacts, people, and animals.

Kang proposes that the discomfort experienced when being confronted by a human-like machine might not be caused by the "flaws" that reveal it to be non-human, like proposed in the uncanny valley theory of Masahiro Mori, but rather the *conceptual* uncertainty caused by them. As we have seen, conceptual questions and uncertainties around automatons, robots, and artificial intelligence have both persisted and changed through time, depending on the understanding of both human characteristics and technological possibilities. Some issues, however, have remained astoundingly similar throughout the history of the Western world. As

we have noted, the role and significance of artificial humans has consistently been linked to the cultural practice of othering through associating them more or less obviously with racialized, ethnicized, and gendered 'Others.' As Jennifer Rhee (2013) points out, Karen Barad's new materialistic thinking helps us understand boundaries between humans and machines as not natural but historically constructed, and therefore, the history of this construction should be of special interest to us.

Barad's (2007) theory of agential realism includes insights about agency and difference that resonate in many ways with our questions concerning talking machines and artificial intelligence. Barad argues that any observation of a phenomenon necessitates an agential cut that forces an ontological difference where there was none. The observer and the observed are entangled, until the act of observation forces them apart and gives them shape as a 'subject' and an 'object'. Barad's posthumanist performative approach marks the boundary building between human and non-human as a key outcome of these observational cuts. In the case of speaking machines, it appears that we fervently strive for a clear cut between "us" and "them," using gender and race as tools of boundary building, but the ontological entanglement shines through in a way which accounts for the existential horror we experience at the thought of a human-like machine.

Looking at the long history of talking machines, we can detect gender and ethnicity/race as categories that have been crucial when dealing with the conceptual horror of human-like machines, especially the speaking machines produced for commercial purposes and entertainment. Gender and racial or ethnic otherness have been used both to make speaking human-like machines attractive and to alleviate the possibility of horror that always follows them. Historical and present-day evidence leads to a conclusion that it appears to be easier to let machines come close to the borders of humanity if they take the form of "lesser" humans, meaning anything other than white adult males (e.g., Pepper the caretaker robot that decidedly resembles an 8-year-old girl). Giving the talking machines the shape of 'Others' (who in reality make up the large majority of the human population) may serve the wish to combine the fantasies of robots and other artificial people as malleable and obedient, but at the same time fascinating and intriguing. Female sex robots like Samantha, who was badly damaged by visitors at a trade show (Levangie 2017), highlight the troubling issues associated with the introduction of female, infant or racialized robots as seemingly neutral wish-fillers of male audiences.

Disturbingly, this also positions "us" as the users, following the white, male, Western ideal, in contrast to the machinic female or exoticized 'Others' that remain artificial and foreign. To return to Sophia the robot, her speech in Riyadh and her gaining the full citizenship of Saudi Arabia demonstrate this point in a striking manner: the talking female machine was presented to and appreciated by a society known for its extreme male dominance, and internet commentators were quick to remark that she was allowed more freedom than the actual human females of the same society. At the moment, as there is remarkable anticipation of a robotic and

artificially intelligent future for the humankind, these kinds of power positions are more important than ever to acknowledge and act upon. The issues related to gender and intersectional otherness are at the core of the conceptual and cultural questions linked to human-like machines and their agency, and therefore they cannot be ignored any longer.

References

- Anderson, Lessley (2013): "Machine language: how Siri found its voice." In: *The Verge*, 17 September, 2013. <https://www.theverge.com/2013/9/17/4596374/machine-language-how-siri-found-its-voice>
- Aytes, Ayhan (2013): "Return of the Crowds: Mechanical Turk and Neoliberal States of Exception." In: Trebor Scholz (ed.), *Digital Labor: The Internet as Playground and Factory*, New York: Routledge, pp. 79–97.
- Barad, Karen (2007): *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*, Durham: Duke University Press.
- Beiguelman, Giselle (2010): "Media Voices: Beyond Talking Heads." In: Norie Neumark/Ross Gibson/Theo van Leeuwen (eds.), *Voice: Vocal Aesthetics in Digital Arts and Media*, Cambridge, MA: MIT Press, pp. 343–357.
- Bosak-Schroeder, Clara (2016): "The Religious Life of Greek Automata." In: *Archiv für Religionsgeschichte* 17(1), pp. 123–136.
- Bowler, Peter J./Morus, Iwan Rhys (2005): *Making Modern Science: A Historical Survey*, Chicago: The University of Chicago Press.
- Bruce, J. Douglas (1913): "Human Automata in Classical Tradition and Mediaeval Romance." In: *Modern Philology* 10(4), pp. 511–526. <http://jstor.org/stable/432766>
- Bryant, Sara (2013): "Dorothy Arzner's Talkies: Gender, Technologies of Voice, and the Modernist Sensorium." In: *MFS Modern Fiction Studies* 59(2), pp. 346–372.
- Castle, Terry (1995): *Female Thermometer: Eighteenth-Century Culture and the Invention of the Uncanny*, New York: Oxford University Press.
- cfhatprovidedotnet (2014): "1957 Genco Horoscope Grandma Fortune Teller Explained in Detail." In: YouTube, 27 March, 2014. <https://www.youtube.com/watch?v=nuCwZbhEbfo>
- Corbella, Maurizio/Anna Katharina Windisch (2013): "Sound Synthesis, Representation and Narrative Cinema in the Transition to Sound (1926–1935)." In: *Cinemas* 24(1), pp. 59–81. <http://dx.doi.org/10.7202/1023110ar>
- Costa, Nic (1988): *Automatic Pleasures: The History of the Coin Machine*, London: Kevin Francis.
- Crowell, Charles R/Matthias Scheutz/Paul Schermerhorn/Michael Villano (2009): "Gendered Voice and Robot Entities: Perceptions and Reactions of Male and Female Subjects." In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 11–15 October, 2009. St. Louis, USA.

- Dawson, Victoria (2015): "The Epic Failure of Thomas Edison's Talking Doll." In: *Smithsonian.com*, June 1, 2015. <https://www.smithsonianmag.com/smithsonian-institution/epic-failure-thomas-edisons-talking-doll-180955442/>
- Donald, James (2008): "As It Happened ... Borderline, the Uncanny and the Cosmopolitan." In: Jo Collins/John Jervis (eds.), *Uncanny Modernity: Cultural Theories, Modern Anxieties*, New York: Palgrave MacMillan, pp. 91–111.
- Dudley, Homer/Tarnoczy, T.H. (1950): "The Speaking Machine of Wolfgang von Kempelen." In: *The Journal of the Acoustical Society of America* 22(2), pp. 151–166.
- Forrest, Jennifer (1999): "Scripting the Female Voice: The Phonograph, the Cinematograph, and the Ideal Woman." In: *Nineteenth-Century French Studies* 27(1/2), pp. 71–95.
- Freud, Sigmund (1919): "Das Unheimliche" [The "Uncanny"]. First published in: *Imago*, Bd. V., 1919; reprinted in: *Sammlung, Fünfte Folge*. Trans. in English by Alix Strachey. <http://web.mit.edu/allanmc/www/freud1.pdf>
- Greene, Lana (2017): "Finding a Voice." In: *Technology Quarterly, The Economist*, 1 May, 2017. www.economist.com/technology-quarterly/2017-05-01/language
- Gunning, Tom (2003): "Renewing Old Technologies: Astonishment, Second Nature, and the Uncanny in Technology from the Previous Turn-of-the-Century." In: David Thorburn/Henry Jenkins (eds.), *Rethinking Media Change: The Aesthetics of Transition*, Cambridge: The MIT Press.
- Hales, Barbara (2010): "Taming the Technological Shrew: Woman as Machine in Weimar Culture." In: *Neophilologus* 94(2), pp. 301–316.
- Hankins, Thomas L./Silverman, Robert J. (1995): *Instruments and the Imagination*, Princeton: Princeton University Press.
- Hanson, David/Olney, Andrew/Pereira, Ismar A./Zielke, Marge (2005): "Upending the uncanny valley." In: *Proceedings of the national conference on artificial intelligence*, 20(4), Menlo Park, CA; Cambridge, MA; London; AAAI Press, pp. 24–31. <http://www.aaai.org/Papers/Workshops/2005/WS-05-11/WS05-11-005.pdf>
- Haraway, Donna (1991) [1984]: "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century." In: *Simians, Cyborgs and Women: The Reinvention of Nature*. London: Free Association Press, pp. 149–181.
- Haraway, Donna (1992): "The Promises of Monsters: A Regenerative Politics for Inappropriate/d Others." In: Lawrence Grossberg/Cary Nelson/Paula A. Treichler (eds.), *Cultural Studies*, New York; Routledge, pp. 295–337.
- Hattab, Helen (2005): "From Mechanics to Mechanism: The Quaestiones Mechanicae and Descartes' Physics." In: Peter R. Anstey/John A. Schuster (eds.), *The Science of Nature in the Seventeenth Century: Patterns of Change in Early Modern Natural Philosophy*, Dordrecht, NL: Springer, pp. 99–130.
- Hempel, Jessi (2015): "Siri and Cortana sound like ladies because of sexism." In: *Wired*, 28 October, 2015. <https://www.wired.com/2015/10/why-siri-cortana-voice-interfaces-sound-female-sexism/>
- Hern, Alex (2016): "Partnership on AI' formed by Google, Facebook, Amazon, IBM and Microsoft | Technology" In: *The Guardian*, 28 September, 2016.

- <https://www.theguardian.com/technology/2016/sep/28/google-facebook-amazon-ibm-microsoft-partnership-on-ai-tech-firms>
- Hänel, Liam (2017): "A list of artificial intelligence tools you can use today – for personal use (1/3)." In: *Lyra*, Medium.com, 30 April, 2017. <https://medium.com/imlyra/a-list-of-artificial-intelligence-tools-you-can-use-today-for-personal-use-1-3-7fib60b6c94f>
- Jarrett, Kylie (2016): *Feminism, Labour and Digital Media: The Digital Housewife*, New York & London: Routledge.
- Johnston, John (2008): *The Allure of Machinic Life: Cybernetics, Artificial Life, and the New AI*, Cambridge, MA and London, UK: The MIT Press.
- Kang, Minsoo (2011): *Sublime Dreams of Living Machines: The Automaton in the European Imagination*, Cambridge: Harvard University Press.
- Keane, Webb (2001): "Voice." In: Alessandro Duranti (ed.), *Keywords in Language and Culture*, Malden: Blackwell, pp. 268–271.
- Keedwell, Edward/Narayanan, Ajit (2005): *Intelligent Bioinformatics: The Application of Artificial Intelligence Techniques to Bioinformatic Problems*, Chichester, UK: John Wiley and Sons, Ltd.
- Kenney, W.H. (1999): *Recorded Music in American Life: The Phonograph and Popular Memory, 1890–1945*, Oxford: Oxford University Press.
- Levangie, Richard (2017): "Samantha the Sex Robot Needs Repairs After Abuse at Trade Show." In: *Second Nexus*, 7 December, 2017. <https://secondnexus.com/science/samantha-robot-sexual-assault/2/>
- Lindsay, David (1997): "Talking Head." In: *American Heritage of Invention and Technology* 13(1), pp. 57–63.
- Liu, Catherine (2000): *Copying Machines: Taking Notes for the Automaton*, Minneapolis: University of Minnesota Press.
- Minsky, Marvin (1952): *A Neural-Analogue Calculator based upon a Probability Model of Reinforcement*, Harvard University Psychological Laboratories internal report.
- Molina, Brett (2017): "Does this Female Robot Have More Rights than a Saudi Woman?" In: *USA Today*, 27 October, 2017. <https://www.usatoday.com/story/tech/talkingtech/2017/10/27/sophia-robot-hanson-robotics-was-citizen-saudi-arabia/806911001/>
- Molten, Megan (2017): "Artificial Intelligence Is Learning to Predict and Prevent Suicide." In: *Wired*, 17 March, 2017. <https://www.wired.com/2017/03/artificial-intelligence-learning-predict-prevent-suicide/>
- Munakata, Toshinori (2008): *Fundamentals of the New Artificial Intelligence. Neural, Evolutionary, Fuzzy and More*, New York: Springer.
- Nass, Clifford/Brave, Scott (2007): *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*, Cambridge: MIT Press.
- Nickelsburg, Monica (2016): "Why is AI Female? How our Ideas about Sex and Service Influence the Personalities we Give Machines." In: *GeekWire*, 4 April,

2016. <https://www.geekwire.com/2016/why-is-ai-female-how-our-ideas-about-sex-and-service-influence-the-personalities-we-give-machines/>
- Pettorino, Massimo (2015): "The History of Talking Heads: The Trick and the Research." In: *First International Workshop on the History of Speech Communication Research*, Dresden, Germany, September 4–5, 2015. ISCA Archive, <https://pdfs.semanticscholar.org/ce62/d996d50d23d4138f6d99410105d18612ac71.pdf>
- Phan, Thao (2017): "The Materiality of the Digital and the Gendered Voice of Siri." In: *Transformations* 18(29), pp. 23–33. http://www.transformationsjournal.org/wp-content/uploads/2017/02/Transformations29_Phan.pdf
- Poser, Stefan (2016): *Glücksmaschinen und Maschinenglück: Grundlagen einer Technik- und Kulturgeschichte des technisierten Spiels*, Bielefeld: transcript.
- Rhee, Jennifer (2013): "Beyond the Uncanny Valley: Masahiro Mori and Philip K. Dick's do Androids Dream of Electric Sheep?" In: *Configurations* 21(3), pp. 301–329.
- Richardson, Kathleen (2016): "Technological Animism: The Uncanny Personhood of Humanoid Machines." In: *Social Analysis* 60(1), pp. 110–128.
- Riley, Terrance (2009): "Composing for the Machine." In: *European Romantic Review* 20(3), pp. 367–379.
- Riskin, Jessica (2016): *The Restless Clock: A History of the Centuries-long Argument over what Makes Living Things Tick*, Chicago: University of Chicago Press.
- Sandygulova, Anara/Gregory M. P. O'Hare (2015): "Children's Perception of Synthesized Voice: Robot's Gender, Age and Accent." In: *Social Robotics. Lecture notes in computer science*, vol. 9388. Springer, Cham. <https://doi.org/10.1007/978-3-319-25554-5-59>
- Seaman Cook, Jennifer (2016): "From Siri to sexbots: Female AI Reinforces a Toxic Desire for Passive, Agreeable and Easily Dominated Women." In: *Salon*, 9 April, 2016. https://www.salon.com/2016/04/08/from_siri_to_sexbots_female_ai_reinforces_a_toxic_desire_for_passive_agreeable_and_easily_dominated_women/
- Segel, Harold B. (1995): *Pinocchio's Progeny: Puppets, Marionettes, Automats, and Robots in Modernist and Avant-Garde Drama*, Baltimore: Johns Hopkins University Press.
- Sharkey, Noel (2008): "The Return of Elektro, the First Celebrity Robot." In: *New Scientist*, 17 December, 2008. <https://www.newscientist.com/article/mg20026873.000-the-return-of-elektro-the-first-celebrity-robot/>
- Shaviro, Steven (2014): *The Universe of Things: On Speculative Realism*, Minneapolis: University of Minnesota Press.
- Simon, Herbert (1996): *The Sciences of the Artificial*, Cambridge, MA: MIT Press.
- Sini, Rozina (2017): "Does Saudi Robot Citizen have More Rights than Women?" In: *BBC Trending*, 26 October, 2017. <http://www.bbc.com/news/blogs-trending-41761856>
- Smith, Jacob (2008): "Tearing Speech to Pieces: Voice technologies of the 1940s." In: *Music, Sound, and the Moving Image* 2(2), pp. 183–206.

- Sophiabot (2017): *Sophiabot.com*, online.
- Sterne, Jonathan (2003): *The Audible Past: Cultural Origins of Sound Reproduction*, Durham: Duke University Press.
- Stewart, Patrick (2014): "Introduction to Methodological Issues in Biopolitics." In: Robert H. Blank/Samuel M. Hines/Odelia Funke/Joseph Losco/Patrick Stewart (eds.), *Politics and the Life Sciences: The State of the Discipline*. Emerald Group Publishing Limited, pp. 67–99.
- Stone, Zara (2017): "Everything You Need To Know About Sophia, The World's First Robot Citizen." In: *Forbes*, 7 November, 2017. <https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/#463946b246fa>
- Taylor, Jessica (2009): "Speaking Shadows: A History of the Voice in the Transition from Silent to Sound Film in the United States." In: *Journal of Linguistic Anthropology*, 19(1), pp. 1–20. <https://doi.org/10.1111/j.1548-1395.2009.01016.x>
- Telotte, Jay P. (1995): *Replications: A Robotic History of the Science Fiction Film*, Champaign: University of Illinois Press.
- Thompson, Emily (2004): *The Soundscape of Modernity: Architectural Acoustics and the Culture of Listening in America, 1900–1933*, Cambridge: MIT Press.
- Truitt, E. R. (2015): *Medieval Robots: Mechanism, Magic, Nature, and Art*, Philadelphia: University of Pennsylvania Press.
- Weidman, Amanda (2006): *Singing the Classical, Voicing the Modern: The Postcolonial Politics of Music in South India*, Durham: Duke University Press.
- Wosk, Julie (2015): *My Fair Ladies: Female Robots, Androids and Other Artificial Eves*, New Brunswick: Rutgers University Press.
- Zielinski, Siegfried (2006): *Deep Time of the Media: Toward an Archaeology of Hearing and Seeing by Technical Means*, Cambridge, MA: MIT Press.