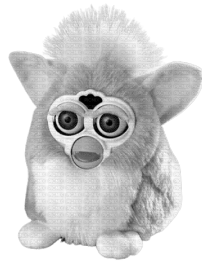


Talking toys: a ludic archaeology of the synthetic voices via automata, simulation, artificial intelligence, technological imaginaries, bodies, music, breath, and impairment

Seth Giddings, Winchester School of Art, University of Southampton

This is the transcript of a talk delivered at WSA on the 9th of October 2025. A video recording of the talk can be viewed here: <https://www.microethology.net/talking-toys/>



This project has its origins in a section I had to cut from my recent book *Toy Theory*. The book has a chapter on dolls and I had begun to trace a genealogy of the historical attempts to incorporate voice-generating devices into commercially produced dolls from the late nineteenth century. What I had imagined might be a couple of paragraphs quickly expanded, so I cut it and left it for later. And then more recently, due to some significant developments in my health and wellbeing, the topic suggested itself to me again, with new directions and emphases... as we will find out over the next forty minutes or so.

Here are three examples to set the scene:

The cinematography of Jean-Luc Godard's film *Alphaville* renders 60s Paris as a futuristic dystopia dominated by an authoritarian computer called Alpha 60, and dramatised through a film noir mise-en-scene - some 17 years before *Bladerunner*. The voice of the authoritarian computer was spoken through an electrolarynx, a prosthetic device used by people whose voices were affected by laryngeal cancer. I'll pick up on the salience of

disability and prosthesis later in the talk. After decades of popular TV and film depicting computers and robots with 'electronic' voices this rasping guttural sound seems scarier due to the traces of an actual human voice within the audio artefacts. This voice also predates another more famous departure from the cliches of technological speech: the creepily calm and composed Hal in *2001: A Space Odyssey* (just three years later).

My second example is also cinematic. Much of the humour and drama of Pixar's *Toy Story* films is driven by play with the characters' voices. On one level, all the main characters speak to each other as sentient subjects. But some have mechanical or electronic voices as operated by and listened to by the children who play with them, and often their subjective voice is quite different from the technological one in manner if not in sound. Both Woody's diegetic and pull-string voices are performed by Tom Hanks, but his everyday sensible and organizational manner sounds very different to the inspirational tone of his recorded phrases. For others, both voices and mannerisms are closer: Buzz's electronic and button activated voice is much closer to his conversational speech - as we'll hear later.

Wheezy then is an interesting case. His voice – and his name - are directly connected to his technical constitution. Squeaky toys could be considered as the closest to the physics and mechanics of actual speech in their sound is produced by the forcing of air through a constriction. The rubber toy's internal squeaker mechanism is broken and so he has a weak breathy voice. Interestingly, this technical breakdown is then doubled as a speech impairment resulting from an asthmatic or pulmonary condition. Which to me begs the question: was he called Wheezy before his squeaker broke?

My third example is ostensibly more playful, but perhaps just as creepy, as my first. Thomas Edison marketed his talking doll in 1890. The voice was produced from a phonograph inside the doll's torso. Each doll had a wax cylinder with a recording of a young girl reciting one of a small set of lullabies and bedtime prayers. Each recording is around 10 seconds long. They have recently been reanimated - the cylinders are too delicate to take a

needle so they have been photographed microscopically, producing a topographic map of each recording's groove, and a digital audio file generated from it.

Though these three examples set the scene and raise issues of technological processing and transmission of the voice, impairment and disability, and the relationships between the voice and subjectivity, they are not synthetic voices as such. The rasping voice in Alphaville is less a synthetic voice and more a prosthetic augmentation of an impaired human vocal tract. The doll voice is simply a recording of an actual human voice, distanced in time and space from its original utterance and speaking body. Novel at the time, but in a way that has become universal in communication and entertainment media since Edison's time. The unsettling effects of a speaking doll are no doubt amplified by its apparent animation, that is, its coming to life as a speaking figure. While the classical-realist narrative and comedy of *Toy Story* downplay the potential uncanniness of the living, speaking object, there are still mise-en-abyme moments or hints throughout of the horror of subject-object blurring.

A couple of years ago, after persistent laryngitis, I had an endoscopy, and it showed that I had an ulcer on my larynx. The treatment included speech therapy and detailed education on the structure and working of the larynx and vocal cords. I found out about the complicated mechanical and pneumatic relationships between the mouth, tongue, throat, larynx and diaphragm in the production and nuance of the human voice. This reminded me of some material I had read for another chapter in *Toy Theory*, a chapter that tracked a genealogy of toy-like automata from classical antiquity to recent robots and AI via eighteenth century androids. The material in question was attempts to design automata that could produce human-like speech.

Most of the eighteenth-century automata looked like large dolls, and, with their ceramic faces and elaborate clothing, were made using materials, fabrication methods, and aesthetics similar to those of high-end toy manufacture at the time. Their imaginative impact was no doubt heightened by the simulacral sense of the doll as symbol or model of

the boundary between the animate and the inanimate, particularly with the verisimilitude these new manufacturing processes allowed. Though these devices were at the vanguard of technological innovation, they were not designed for instrumental or productive use – they were neither tools nor prototypes for tools. They were mechanical thought experiments, asking their audience to take the short imaginative step from watching a machine that is ingenious but explicable to one that can – preternaturally – think and reason for itself. The Writer captured this unsettling proposition as he scratched on his pad, in a play on Descartes, a question that might speak for future androids and simulacra: ‘I am not thinking, Do I therefore not exist?’

In particular historical moments, such self-moving machines personified the technological paradigm of their day. Fashioned from dominant technical processes, principles, and materials, they also conjured up imaginaries of these very technics. And very often these are imaginaries of the animate, of ‘living technology’ from hydraulics and molten metal in Ancient Greece and Egypt to the Newtonian cosmology of clockwork in the Enlightenment, to contemporary robotics and artificial intelligence. Writing at the beginning of the information age at the end of World War 2, the cyberneticist Norbert Wiener wrote: ‘At every stage of technique since Daedalus or Hero of Alexandria, the ability of the artificer to produce a working simulacrum of a living organism has always intrigued people. This desire to produce and to study automata has always been expressed in terms of the living technique of the age.’

These simulacra then are thought experiments realised in mechanical form, at once spectacular novelties for public entertainment and prompts to philosophical and scientific reflections on the nature of life, consciousness, and reason.

Recent critical reflection on robotics, virtual reality, AI and other cyborgian intimacies between the human body, mind, and technology often draws on this long genealogy of automata. It generally acknowledges their toylike appearance but downplays any

continuity with children's play objects. I argue that as well as looking like toys, these devices had fundamental toy-like characteristics: a performative mode of demonstration combining wonder, spectacle and amusement; and a strong and persistent relationship with play, games and the toylike. A toy is a concentration or condensation of materials, ideas, aesthetics, forces and mechanisms into a figure easily grasped the hand, eye, and imagination. And so the automaton's purchase on protoscientific and natural-philosophical enquiry was inseparable from its affinity with the child's plaything and its cheap distractions, simulacral trickery, and hands-on play.

There are precursors to these speaking machines in myth and legend. Ancient oracles are ambiguous in their being: disembodied voices in sacred grottoes, for example the head of Orpheus at Lesbos, or mortal priests channelling divine utterances. Various medieval speaking machines appear to be genuine attempts albeit with improbable claims made for them. For instance, the thirteenth century philosopher Albertus Magnus is reputed to have constructed an artificial man out of bronze with whom he would converse. His student, Thomas Aquinas, destroyed it, "tired of its great babbling and chattering".

The eighteenth century saw serious attempts to construct a working simulation of the human voice, including, in England, by Erasmus Darwin, and in France, a priest, Abbé Mical. The latter's pair of carved wooden talking heads were able to utter whole sentences and was presented to Louis XVI at Versailles in 1783. Wolfgang von Kempelen, a Hungarian engineer and contemporary of Abbé Mical, is a key figure in this movement to mechanically replicate speech. Better known now for his Chess Player, a technically ingenious and innovative but ultimately illusory assembly, his work on artificial speech however was driven by scientific enquiry rather than showmanship. In the development of his mechanical lungs, mouth and larynx Kempelen had to engage directly with the actual biological workings of the human body. Speech is simply too complex and multifactorial in its generation for any simple simulation.

In their eschewing of illusion and performativity, these devices are, then, on a trajectory away from the playful and toy like paradigms of Jaquet-Droz's performative automata. These practical challenges necessitated a departure from elegant natural philosophical models towards and more pragmatic engagement with mechanics and materials. This distinction reminded me of Baudrillard's historicizing of simulacra, from the courtly automata to the modern industrial robot:

A world separates these two artificial beings. One is a theatrical, mechanical and clockwork counterfeit of man where the technique is to submit everything to analogy and to the simulacrum-effect. The other is dominated by a technical principle where the machine has the upper hand, and where, with the machine, equivalence is established. The automaton plays the man of the court, the socialite, it takes part in the social and theatrical drama of pre-Revolutionary France. As for the robot, as its name implies, it works; end of the theatre, beginning of human mechanics. (Baudrillard 1983, 53).

Along with toys in general, this voice machine complicates any linear model of the history of simulacra.

Kempelen worked with wood and leather, and as well as drawing on recent anatomical studies of the mouth, neck, and lungs, explored musical instruments to see which most closely resembled the human voice. There was much discussion about whether the primary workings of the voice were closer to the flow of air through tubes like woodwind instruments or the bowing of the vocal cords as strings. Adapting existing mechanical and technical elements such as tubes and bellows, he approximated the workings of the diaphragm and lungs, the trachea, the vocal cords, and the shaping of sound by the mouth, lips, and nose. He realised that voice production is a complex phenomenon, relying on the inter-operation of all parts of the vocal tract, with speech sounds also modulated by voicing, aspiration, frication, and nasality. Additions to achieve these modulations

included nostril tubes and a supplementary bellows to rapidly inflate and deflate the artificial mouth to create plosives. Modern reconstructions of Kemepelen's apparatus have demonstrated that it could produce whole intelligible sentences.

However, the conjunction of various elements meant that the whole system had to be 'played' like a musical instrument, an operation that demanded expertise and practice. This music-like technics persists through subsequent speaking machines. The early nineteenth century "Euphonia" was a development of Kempelen's machine. Its inventor, Joseph Faber, replaced the simple mouth tube with a moving tongue, jaw, and lips – and this demanded a complex control mechanism, a keyboard with 14 keys and three pedals. The keys produced the basic sounds whereas the pedals controlled pitch, nasality, and voicing. A third pedal operated the bellows. Like the earlier musician automata, Faber's device was presented with a showman-like performativity. He would ask the audience for sentences, transcribe them phonetically in his head, then play them back. In London in 1846, it sang God Save the Queen, and so appears to be the first machine to ever sing a song. As we'll see the Euphonia has remarkable similarities to a key early attempt at electronic speech synthesis nearly 100 years later.

I'll give one example before we leave the courtly androids: Jacques Vaucanson's flute player from 1738. It used mechanisms and materials that would influence attempts to simulate speech, established a compelling interrelationship of breath and music that we'll see runs through this genealogy. It actually played the flute, with a repertoire of twelve tunes, with air blown from three sets of bellows through three windpipes, through flexible leather lips and tongue. As I have noted, recent critical work on 18th century automata has explored in depth the role of clockwork as a technological paradigm and imaginary. The techniques and technics of music, musical instruments, performance and dance are less evident. The two overlap in the practicalities of - and philosophical reflection on - *breath*. The Flautist and the voice machines both harnessed the flow of air as the fundamental motive force of much music-making, all vocalising – and as the essence of life and spirit. If

the voice exteriorizes intelligence, reason, and the soul, breath itself drives that expression and sustains the life of the expressive subject. As Jessica Riskin puts it: of all the various efforts to simulate bodily functions and organs, speech is particularly significant, as it crosses "the Cartesian boundary between mechanical body and rational soul."

A contemporaneous review of Vaucanson's Flautist predicted that articulate artificial speech was impossible because of the complexity and mystery of its bodily processes. This prediction has been largely borne out: all subsequent attempts to mechanically reproduce speech have fallen short. Synthetic speech would only become viable with electronic processing, and convincing human-sounding voices are just appearing via artificial intelligence.

A little over a year ago, the materiality and physiology of my voice became an even more intense cause for concern. After experiencing symptoms such as cramp, twitching, muscle loss and weakness in my hands, I was diagnosed with motor neurone disease. The term motor neurone disease covers a range of connected conditions, the most common of which is amyotrophic lateral sclerosis, or ALS. Nerves, known as motor neurones, in the brain and spine stop sending messages to the muscles. This leads to muscles weakening, stiffening, and wasting, affecting the ability to walk, talk, eat, drink, and breathe. As speech is often seriously affected, people with the condition are encouraged to 'bank' their voice. This involves speaking a set of words and phrases into a software application that then generates a synthetic version of the voice that can be used in text to speech applications.

[in the talk the following passage was recited using a synthetic version of my voice generated on my iPhone]

I have tried two versions. One through a company called Acapela, and funded by the charity the Motor Neurone Disease Association, and one – this one you are listening to – is a facility available on all Apple Mac computers and iPhones. Though lacking

the cadence and intonation of natural speech, for many the banked voice is far preferable to earlier synthetic voices such as that made famous by Stephen Hawking.

Thankfully, I don't have bulbar onset MND, the variant in which the mouth and throat are affected, with marked difficulties with swallowing and speech from the outset. However, the strength of my voice has been affected by a weakening of my diaphragm. Thus, I am not dysarthriac but rather hypophonic – and the connection between speaking and breathing has become particularly evident.

For me, particularly during this research on the voice across the human and non-human, motor neurone disease has led me to alternative ways of thinking about the relationships between concepts of the body and technological paradigms and imaginaries. The condition draws attention to the bioelectric nature of the nervous system, and its sustenance and mobilisation of the skeleto-muscular mechanics of bodily movement. In all the discussion in recent decades over whether the human mind can be digitised, we forget that it is always already electronic and informatic. Indulge me with one last trip back to the 18th century, to Luigi Galvani's experiments with electricity and muscle movement, the electrical animation of both (and across) automata and organic bodies - both imagined and demonstrated experimentally at the same time Kempelen was developing his mechanical voice. Galvani's work demonstrated the animation of muscles via the bioelectric workings of the nervous system. His diagrams are, to me, a frank reminder of the fundamental relationships affected by motor neurone disease. And, along with Erasmus Darwin's speaking machine, Galvani's insights famously inspired Mary Shelley's *Frankenstein*.

I've mentioned Edison's doll and the Euphonia, but I'm going to skip through the rest of the 19th century with its advances in phonography and telephony. I will just note a shift in technological paradigms and imaginaries that have a bearing on twentieth century

developments. New electricity-based systems became the model for understanding the human body. As Sarah Bell puts it, telegraphy was regarded as the nervous system of the industrial nation, and concomitantly "the human nervous system [was understood as] a telegraph communicating sense perceptions through the body as electrical information."

This was a paradigm that would be further developed through cybernetics after the second world War. From the mid-20th century the convergence of cybernetics, information theory, and electronic media led to the theorisation of the human nervous system as an electronic communication system in itself – or as Marshall McLuhan famously asserted – a system extended out into the world through media networks. The advent of digital networks and media boosted interest in this connection through cyberculture and critical posthumanism.

Attempts at the *electronic* synthesis of the human voice were made at Bell Labs in the 1920s and 1930s, with the first public demonstration of an application at the 1939 World's Fair in New York: AT&T's The Voice Demonstrator, or 'Voder.' As we can see in these contemporaneous diagrams, the natural processes of speech were simulated informatically: simplified and modelled in electronic circuits and filters - but still imagined as analogues of the biophysical. Unlike recording-based toys and later entertainment-driven voice synthesis systems such as auto-tune and vocoders, The Voder did not process the human voice but generated its simulation from scratch.

The Voder's developers cited Kempelen's voice apparatus as an inspiration, and the machine echoed the eighteenth century automata in its performative and non-instrumental presentation. And though its technical foundations were completely different, there was a remarkable similarity in the modes of operation necessary to generate a plausible voice. Highly trained operators, women dubbed 'Voderettes,' used keys and pedals to manipulate the signals in real time. The parallels with the musical instrument style of interface and performance of the Euphonia are marked.

Diagrams demonstrate the conceptualization of the machine with both the skeleton-muscular and nervous workings of the human body and prefigure cybernetics, the information-based technological paradigm shift that would emerge at the other end of the Second World War with electronic computers. The Voder also channeled the Enlightenment association of simulated speech with intelligence. Now, the emergence of computers was attended by a popular technological imaginary of "electronic brains," recognised as such by their actual or predicted ability to speak. Sarah Bell notes that before visual displays and screens became the primary computer interface in the 1970s, it was popularly imagined that human-computer interaction would, like SF robots and the computers of Star Trek and 2001: a Space Odyssey, conducted through spoken conversation.

Though vocoders *process* the human voice rather than generate a simulation from scratch like the Voder, they are a significant thread in this genealogy. They date back to the 1920s at AT&T's Bell Labs as part of attempts to code and compress the voice to make communication more efficient. Alan Turing developed the concept during World War II to disguise voice communication, and its principles contributed to both Claude Shannon's information theory and Norbert Wiener's cybernetics. Further developments from the 1960s to 1980s underpin the technology of cellular telephony and the compression techniques from CD-ROMs to the Web.

[video of Kraftwerk performing 'Robots']

As Mara Mills notes (and Kraftwerk demonstrate), the vocoder has been extolled as an archetypal posthuman technology, in its ability to render the human voice in a technological and robotic register, as an exemplar of virtuality, and as a challenge to humanist subjectivity.

In pursuing this research, I was struck by the persistent salience of disability and impairment throughout the genealogy of the simulated voice and related communications

technologies. Some innovations were responses to particular disabilities, whereas others seem to have been retroactively and even opportunistically applied to the needs of disabled people. Jacquet-Droz hoped that his mechanical hands might help amputees, Kempelen that his speaking machine might have the potential to help people with speech impediments or hearing impairments. Kittler notes that physical impairment was integral to the origins of mechanical sound recording, and that the first typewriters were designed by blind people. Edison's voice, captured by his phonograph, is heard as a scream, partly because amplifiers had yet to be invented and partly because Edison himself was hearing impaired. The new electronic media of the nineteenth century were at once analogies for, and prostheses of, human nerve and sense systems, "technological implementations of the central nervous system." The telegraph was an artificial mouth, the telephone an artificial ear, and their combination as the phonograph was an artificial brain - recording and replaying memory. One of the claims made for the Voder was its potential to "enable mutes to talk."

A significant source of funding for voice synthesis since the 1960s has been the development of text to voice applications for the visually and vocally impaired. The most famous vocal example of this post war commercialisation of assistive technologies and prostheses must be the synthetic voice used by perhaps the most famous person with motor neurone disease, Stephen Hawking. His first speech synthesiser was based on a system called Perfect Paul developed at MIT in the 1970s. It ran on an Apple II computer adapted for mobile use on a wheelchair. Hawking communicated at a rate of 15 words a minute using a hand clicker selecting words from a separate application. Though famously electronic- or robotic-sounding, over the years Hawking developed the voice's expressive - even performative - capacities, from rehearsing and editing lectures to cameo roles in television programmes including *The Simpsons* and *Big Bang Theory*. He refused to change the voice even as its technological platforms were upgraded over 32 years.

I have found the term *impairment* more productive than the more commonly used *disability* in the context of this research, as it applies equally to the human body and technological systems. Jonathan Sterne, the media scholar with expertise in audio and music technology, explains impairment as working in "a shady place between function and non-function." He addresses impairment in communication systems as well as human bodies: videoconference glitches, radio waves with static and noise, faulty spam filters, and so on. Moreover, he observes that impairment is not always fully negative. It can have diagnostic, analytical, and even aesthetic applications. The observation that the workings of technical systems become most evident and available for study in moments of breakdown is familiar from science and technology studies and media archaeology. And in the case of motor neurone disease, the relationship of impairment and glitch in communication systems and the human body is neither metaphorical nor analogous, rather it is continuous, identical. The Armenian American inventor Emik Avakian, who was born with cerebral palsy, made this connection between the impaired nervous system with cybernetic paradigms in the 1960s. "A man's body is a vast intercom system serviced by a central transmitter, the brain. Palsy victims such as I suffer a breakdown in communications." Avakian developed a number of inventions that addressed disability, including a breath operated typewriter, and a system that synthesised recorded phonemes to generate real-time stock reports over the telephone.

I mentioned the aesthetic potential of impairment. For media archaeology, creative technology art practice, and cyberpunk popular culture, glitches in systems and devices not only reveal their inner circuits and mechanics. They also suggest alternatives, often creative and playful alternatives, to the system or device's intended uses. And, to return to my toy theory, the prevalence of electronic and digital toys in the artistic and maker practices of circuit bending is far from incidental. Keeping the focus on the voice here are examples from maker and popular culture [video clips of Max Headroom, and the circuit-bending of a Furby and a Speak n Spell].

The early 20th century saw many patents filed for talking toy mechanisms, but none seem to be innovative. Rather they were attempts to nuance the simple mechanics and pneumatics of airflow with bellows and reeds. Dolls were squeezed, inverted, or blown into to generate simple sounds such as crying or the word 'Mama.' Edison's phonographic approach returned in the 1920s with Dolly Rekord, who came with a set of interchangeable cylinders to recite nursery rhymes or bedtime prayers. By the mid-60s these tiny record players were robust enough to sustain successful toy lines such as Chatty Cathy.

Sterne traces a genealogy of the amplified and displaced voice from augmentation to the cyborg to communication breakdown. There are intriguing resonances with the diegesis of the *Toy Story* universe, where voices are indices of subjectivity mediated and disrupted by amplification, displacement, and doubling. For instance, Chatty Cathy is the inspiration for Gabby Gabby, a character in *Toy Story 4*. As a character in the world of living toys, Gabby can speak normally, but as a mechanical toy and commodity in the human world she is disabled by her broken voice mechanism. She demonstrates this: [clip in which Gabby removes and displays her broken voice player].

Gabby has a gang of ventriloquist dummies as henchmen. Their creepiness derives in part from their appearance as at once adult humanlike figures and obviously mechanical with their hinged jaws. But mainly of course because of their channelling of the human voice in a display of trickery and arts. These dolls are a persistent mainstay of comedy and horror, often with the control of the ventriloquist's voice appearing to transfer from the human to the toy in stories of the uncanniness of animation. In *Toy Story 4*, however their horror arises simply from their impairment, their spastic movements, and their muteness.

Speak & Spell, produced by Texas instruments from the late 70s, was, according to Sarah Bell, "the most famous voice synthesiser of the 20th century" (we might add, in the US at least). It was certainly one of the first consumer electronic devices to feature digital voice synthesis, and sustained, for a little while at least, the dream that everyday computer

interaction would be voice based. The device was brought to life as Mr Spell in the first *Toy Story* film and is an interesting character. As, I think, the only smart toy in the *Toy Story* films, his subjective voice and actual voice are the same. His status as an educational toy in the actual world is reflected back into the *Toy Story* world in his role as seminar leader and instructor to the toy community. I would argue that the *Toy Story* films among the most sustained explorations of the interplay of subjectivity, technology, embodiment, intelligence, and the voice since the Enlightenment.

I would note the significance of the fact that this pioneering technology and mode of interaction was realised not in a household or business application, but in a toy. The team developing it had to convince the company that it was possible to take the electronic and digital voice synthesis innovations from research labs in the 1960s and condense them into a single chip capable of real time speech synthesis.

Speak & Spell and Chatty Cathy's world of talkative toys blossomed in the 1980s, with products such as Mattel's Hug n' Talk doll. And in the late 90s and early 2000s new toy lines based on computer technology hit the market. The Talk With Me Barbie doll connected to a PC running a CD-ROM. As Lydia Plowman explains, 'Children were able to select their own name and a range of topics to discuss with Barbie [...] The choices were then beamed from the computer to the infrared receiver in the doll's necklace and stored in its memory. The Barbie Doll could then talk to the child without the need to be attached to a desktop computer.' The roboticist Rodney Brooks worked with Hasbro to launch My Real Baby in 2000, based on his research. The doll learnt vocabulary and phrases, and responded according to mood. Like Talk to Me Barbie and others such as My Dream Baby and Actimates, these toys were ultimately unsuccessful. They were expensive and limited by short battery life and unreliable interactivity. I would argue too that they profoundly misunderstood the nature of children's imaginative play with toys.

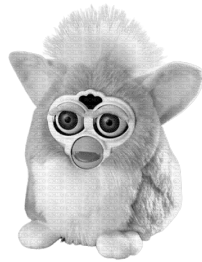
Such toys constitute an under-appreciated direct line of descent in the genealogy of speaking automata from the Middle Ages to chatbots. And attention to children's play with them suggest modes of material and imaginative engagement more nuanced and productive than most adult fantasies about artificial intelligence. As I argue in *Toy Theory*, children bring their toys to life in ways that are both immersive but also pragmatic. And, as Sherry Turkle observed in the early 1980s, children's play with smart toys offers a 'third way' of thinking about machine intelligence, between the animate and the inanimate. Her participants knew that Speak & Spell and Merlin weren't alive, but they were not inert as traditional toys and objects were, they were as one girl put it, "sort of alive." Turkle saw this as the emergence of a new social relationship with computational objects - I would be more specific, it was the beginning of a new *playful* relationship.

Yesterday, I tried a third version of voice banking, this time using artificial intelligence to add rhythm and cadence according to an assessment of sentence structure. This took only seconds to generate by uploading a short audio clip taken from a recording of an online lecture of mine. [the following section was spoken by a cloned version of my voice generated by Eleven Labs]

Donna Haraway's cyborg manifesto has been a key point of reference for me throughout my academic work. I have always been interested in how it sheds light on the everyday, intimate and mundane character of media and technoculture. Over the past year, it has unsurprisingly taken on new salience. The powered wheelchair is a vivid emblem of technological augmentation for impairment. However, for me, it doesn't feel that different to, say, one's everyday reliance on a car; both offer personally controlled mobility and navigation through, and nested within, technological systems of urban infrastructure. Similarly, most voice-based digital communication is pretty much a variation on the familiar lived environment of the smartphone and laptop. Where I sense a difference in kind rather than degree is when I hear synthetic versions of my voice. I'm not even sure if the possessive

determiner 'my' works here. I listen to a recording, and it is a recording of my voice. However, synthetic voices such as this are processed from recordings and sound somewhat like my voice, but are not my voice.

The Enlightenment investment in the voice as index of the humanist subject can be critiqued, but it's not easy to shed or unfeel. But on the other hand, like most concepts and objects it can be played with. These voices are toys.



Bak, Meredith. A., 2019. Between technology and toy: the talking doll as abject artifact, in: Hennefeld, M., Sammond, N. (Eds.), *Abjection Incorporated: Mediating the Politics of Pleasure and Violence*. Duke University Press.

Baudrillard, Jean, 1993. *Symbolic Exchange and Death*. London: Sage.

Bell, Sarah, 2024. *Vox Ex Machina: a cultural history of talking machines*. The MIT Press, Cambridge, MA.

Brandow-Faller, Megan (ed.), 2018. *Childhood by Design: toys and the material culture of childhood*. Bloomsbury, New York.

Cowan, Ron, 2015. Ghostly voices from Thomas Edison's dolls can now be heard. *New York Times*. <https://www.nytimes.com/2015/05/05/science/thomas-edison-talking-dolls-recordings.html>

Faber, Liz, 2020. *The Computer's Voice: from Star Trek to Siri*. University of Minnesota Press, Minneapolis.

Giddings, Seth, 2019. Toying with the singularity: AI, automata and imagination in play with robots and virtual animals, in: Mascheroni, G., Holloway, D. (Eds.), *The Internet of Toys: Practices, Affordances and the Political Economy of Children's Smart Play*. Palgrave Macmillan.

Kittler, Friedrich, 1999. *Gramophone, Film, Typewriter*. Stanford University Press, Stanford CA.

Lister, Martin, Dovey, Jon, Giddings, Seth, Grant, Iain, Kelly, Kieran, 2009. *New Media: a critical introduction* (2nd ed.). Routledge, London.

Mills, Mara, 2012. Media and prosthesis: the vocoder, the artificial larynx, and the history of signal processing. *Media, Culture & Communication* 21, 107–149.

Park, Julie, 2020. Making the automaton speak: hearing artificial voices in the eighteenth century, in: Cave, S., Dihan, K., Dillon, S. (Eds.), *AI Narratives: A History of Imaginative Thinking about Intelligent Machines*. Oxford University Press, Oxford.

Plowman, Lydia, 2004. “Hey, hey, hey! It’s time to play”: children’s interactions with smart toys, in: Buckingham, David & Brougere, Gilles, (Eds.), *Toys, Games, and Media*. Lawrence Erlbaum, Mahwah NJ, pp. 207–224.

Ramsay, Gordon J., 2019. Mechanical speech synthesis in early talking automata. *Acoustics Today* Summer, 11–19.

Rhodes, Alice 2024. *British Romanticism and the Matter of Voice*. Cambridge: Cambridge University Press.

Riskin, Jessica, 2024. “You are my friend” Early androids and artificial speech. *The Public Domain Review*.

Riskin, Jessica, 2016a. *The Restless Clock: a history of the centuries-long argument over what makes living things tick*. University of Chicago Press, Chicago IL.

Riskin, Jessica, 2016b. Frolicsome engines: the long prehistory of artificial intelligence. *The Public Domain Review*.

Riskin, Jessica, 2003. The Defecating Duck, or the ambiguous origins of artificial life. *Critical Inquiry* 29, 599–633.

Siebers, Tobin, 2006, Disability in theory: from social constructionism to the new realism of the body,” in Lennard J. Davis (ed.) *The Disability Studies Reader*, ed. New York: Routledge, 173–83.

Smith, Marquard, 2013. *The Erotic Doll: a modern fetish*. Yale University Press, New Haven, CT.

Sobchack, Vivian, 2006. A leg to stand on: prosthetics, metaphor, and materiality, in Marquard Smith and Joanne Morra (eds) *The Prosthetic Impulse: From a Posthuman Present to a Biocultural Future*. Cambridge: MIT Press, 17–41.

Stafford, Barbara, 1999. *Artful Science: enlightenment entertainment and the eclipse of visual education*. MIT Press, Cambridge MA.

Stafford, Barbara, Terpak, Frances, 2001. *Devices of Wonder: from the world in a box to images on a screen*. Getty Research Institute, Los Angeles CA.

Star, Susan Leigh, 1999. The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), pp.377–391.

Sterne, Jonathan, 2022. *Diminished Faculties: a political phenomenology of impairment*. Duke University Press.

Tiffany, Daniel, 2000. *Toy Medium: materialism and modern lyric*. University of California Press, Berkeley CA.

Truitt, E.R., 2015. *Medieval Robots: magic, mechanism, nature and art*. University of Pennsylvania.

Turkle, Sherry, 1984. *The Second Self: computers and the human spirit*. MIT Press.

Wood, Gaby, 2002. *Living Dolls: a magical history of the quest for mechanical life*. Faber & Faber, London.