TIM, PAT and ISAAC: synthetic speech on display at the National Museum of Scotland

Alison Taubman
National Museums Scotland

Telecomms Heritage Conference
Salford University
22 June 2019

In its communications gallery, staff at National Museums Scotland were keen to include a fundamental of human communication – speech. This paper will outline a display of speech mediated by machines, from the experiments first speaking clock to the now omnipresent synthetic voices of devices of satnavs and smoke alarms. It will explore how the Museum might approach collecting apps and devices to enhance communication for people with no voice and limited mobility.

I hope this presentation will form part of this conference's later discussions to explore not only what museums collect, but how that material can be interpreted for diverse audiences.

This paper attempts to set out how National Museums Scotland (NMS) tackled the issue of displaying speech in a gallery on communication. We chose to do so through speech mediated by machines. Today I will show you how we approached this and also include a few examples of speech related objects in other museums, as an indication of what can be, and what has been, collected, as well as examples of what is still to be collected.

The Communicate Gallery at the National Museum of Scotland, Edinburgh, opened July 2016

There is of course a question of definition of artificial and synthetic speech. The telephone can be viewed as a mediated speech device. For the purposes of this paper, I concentrate on devices or methods of communication beyond the person to person, in which speech is interpreted in some manner . My focus might be on machine generated speech, but, as I hope to show, in order to try to accomplish such speech required a profound understanding of the mechanics of human speech first and then required a method of visualising that understanding, to see the patterns of vocalisation, alongside an understanding of the mechanical, the electrical or the electronic.

Speaking machines

"From the time that the statues of Memnon emitted their mystical tones on the banks of the Nile, and the oracular responses were delivered in Delphi, through the period when a speaking head was exhibited by the Pope, towards the end of the tenth century, and others afterwards by Roger Bacon and Albertus Magnus, various surprising efforts have been made to produce a machine capable of articulating human words and sentences."

W & R Chambers, *Miscellany of Useful and Entertaining Tracts*, 1847

Chambers Quote

The National Museum of Scotland is located on Chambers Street, Edinburgh, named after William Chambers, one of the founding brothers of the famous educational publishers W&R Chambers. Among their many educational works was their Miscellany of Useful and Entertaining Tracts. In one of these, Volume 13, 1847, they enlighten their readership on 'Speaking machines'

"From the time that the statues of Memnon emitted their mystical tones on the banks of the Nile, and the oracular responses were delivered in Delphi, through the period when a speaking head was exhibited by the pope, towards the end of the tenth century, and others afterwards by Roger Bacon and Albertus Magnus, various surprising efforts have been made to produce a machine capable of articulating human words and sentences."
(Chambers' Miscellany of Useful and Entertaining Tracts, W&R Chambers, Edinburgh, Vol XIII, 1847, p 55)

Among those 19th century experimenters was Telegraph pioneer and Professor of Experimental Philosophy at King's College, London, Charles Wheatstone. By the 1820s Wheatstone was working on an improved version of the Baron De Kempelen's device for mechanically reproducing human speech, which he described in a paper to the British Association in 1835.
(Bowers, B., Sir Charles Wheatstone, London Her Majesty's Stationery Office, 1975, p33)

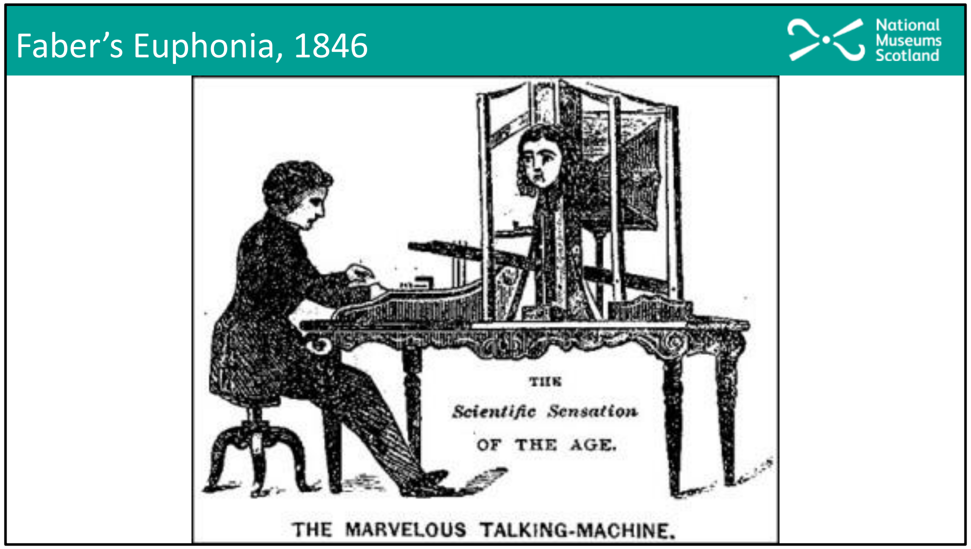
Von Kempelen speaking mechanism 1791

Von Kempelen speaking mechanism of c 1791 now in the Deutsches Museum

https://artsandculture.google.com/exhibit/4wIC7hLe64FKJA

Years later, in 1863, Alexander Melville Bell and his son, Alexander Graham Bell, visited Wheatstone to see this device demonstrated. Later in his life AG Bell recalled the visit;

'I saw Sir Charles manipulate the machine and heard it speak, and although the articulation was disappointingly crude, it made a great impression upon my mind.'
(Bowers, B., Sir Charles Wheatstone, IEE History of Technology series no 29, 2nd edition, London, 2001, p35)



Euphonia

The Miscellany of Useful and Entertaining Tracts also included a description of a device being demonstrated in London at the time by its inventor Joseph Faber, called Euphonia. Euphonia employed keys connected to a bellows system, which blew air across a passageway that resembles human vocal cords which created elementary sounds for words found in European languages. As far as I'm aware, no parts of Euphonia survive.

Faber’s Euphonia, 1846

THE EUPHONIA, OR SPEAKING MACHINE.

Euphonia

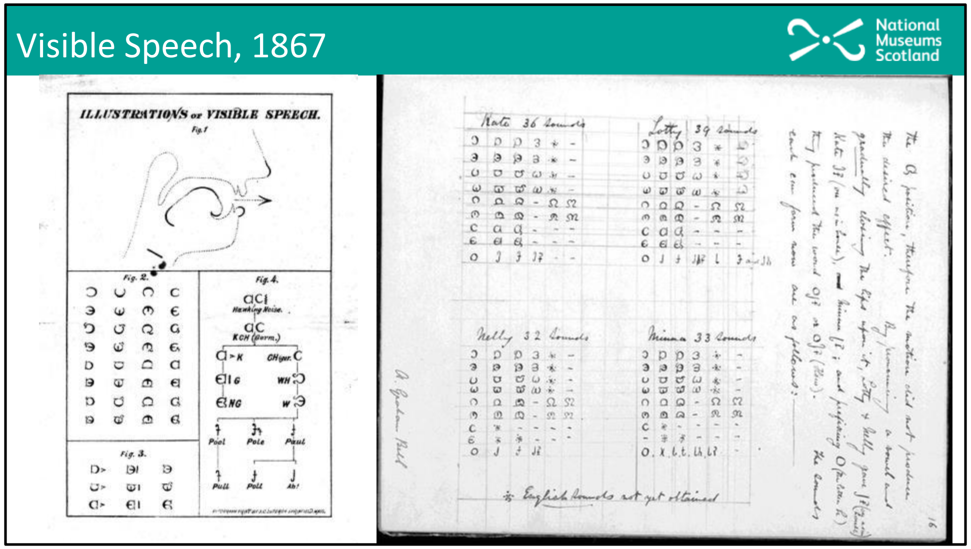
Already at this stage American electrical engineer and telegraph pioneer (and later Director of the Smithsonian Institution) Joseph Henry imagined connecting it to a telegraph system, converting the dots and dashes into a synthetic human voice- in effect text to speech.

(https://www.atlasobscura.com/articles/texttospeech-in-1846-involved-a-talking-robotic-head-with-ringlets accessed 20.5.19)

He wrote to a friend

“I have seen the speaking figure of Mr. Wheatstone of London but it cannot be compared with this which instead of uttering a few words is capable of speaking whole sentences composed of any words whatever.”

(Joseph Henry to Henry M Alexander, January 6 1846 in The Papers of Joseph Henry, Vol 6, Marc Rothenberg et al, eds (Washington 1992), p362)



Visible Speech

Alexander Graham Bell notebook, right, in the Library of Congress collections

Among the many visitors to Euphonia was Alexander Melville Bell. Melville Bell was an Edinburgh based teacher of elocution. In 1867 he published a phonetic transcription system he called Visible Speech: The Science of Universal Alphabetics. He used the term 'visible speech' because the symbols he developed specified the pronunciation of word sounds so accurately, that those learning each symbol's associated sound could repeat any given speech. Both Bells' chief interest in visible speech came from their deep commitment to facilitating speech for deaf people thereby mitigating the need for deaf people to communicate using sign language. This proved controversial at the time and remains so to this day, as Bell propounded a view that deaf people should integrate into a hearing world rather than be free to express themselves through signing; the promotion of oralism over manualism.
(https://rootedinrights.org/alexander-graham-bell-and-the-deaf-community-a-troubled-history/   accessed 31.5.19)

This extended to his advocating that deaf adults be prevented by law from marrying to avoid passing on a hereditary deafness gene and avoid what he referred to as 'a deaf variety of the human race'

(https://archive.org/details/cihm_08831   accessed 20.5.19)

On display at National Museum of Scotland

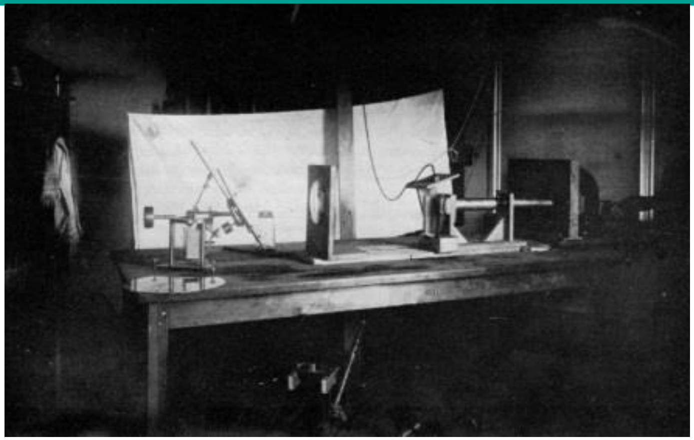
Teaching model of human tongue, larynx and vocal chords, made of papier mâché. This type of dismantleable model was used to teach anatomy students the complexities of human vocal physiognomy. By Dr Louis Thomas Jerome Auzoux, France, c.1860

Bell was known to have made a model of a larynx and vocal chords as a young man, in Edinburgh, as part of his research in to the mechanics of human speech. The juxtaposition of a teaching model of a human larynx, of a similar date – not the one made by Bell – next to a replica of Bell's liquid transmitter of 1875, encourages visitors to see the parallels between the construction of the two and understand how Bell's work on speech informed his early work on the telephone.

Bell's interest lay in the physiology of speech and in capturing speech as well as being able to reproduce it. Initially employing electric tuning forks to transmit speech like sounds, this ultimately informed his work on the telephone.

Visible Speech, 1884

National Museums Scotland



Smithsonian photo No. 44312-E

Volta labs Bell material preserved in the Smithsonian

Of particular interest to Bell, was the existence in vowel sounds of concentrations of acoustic energy, known as formants. These vital sound elements were very difficult to distinguish and isolate. Bell continued to develop his father's earlier developments in Visible Speech, funding work in his Volta Laboratory, on photographic techniques to capture visualisations of sound patterns, so that speech could been seen and read as well as heard. Photographs of this experimental process survive in the Library of Congress.

There was a notion such technology might enable deaf people to use the telephone, linked to a spectrograph to enable 'visible speech'.

The Voder, 1939

"PEDRO"

HE HASN'T any mouth, lungs, or larynx—but he talks a blue streak. His name is Pedro the Voder, and you may see him in action at the New York and San Francisco world's fairs. His creation from vacuum tubes and electrical circuits, by Bell Telephone Laboratories engineers, crowns centuries of effort to duplicate the human voice.

To manufacture Pedro's conversation, his operator employs a keyboard like that of an old-fashioned parlor organ. Thirteen black and white keys, fingered one or more at a time, produce all the vowels and consonants of speech. Another key regulates the loudness of the synthetic voice, which comes from a loudspeaker. A foot pedal varies the inflection meanwhile, so that the same sentence may state a fact or ask a question.

About a year's practice enables an operator to make Pedro talk glibly. He whispers or

How Vacuum Tubes and Filters Produce Speech

Voice-Operation Demonstrator Crowns Centuries of Effort by Scientists To Duplicate Human Speech Artificially

By ALDEN P. ARMAGNAC

POPULAR SCIENCE

The Voder

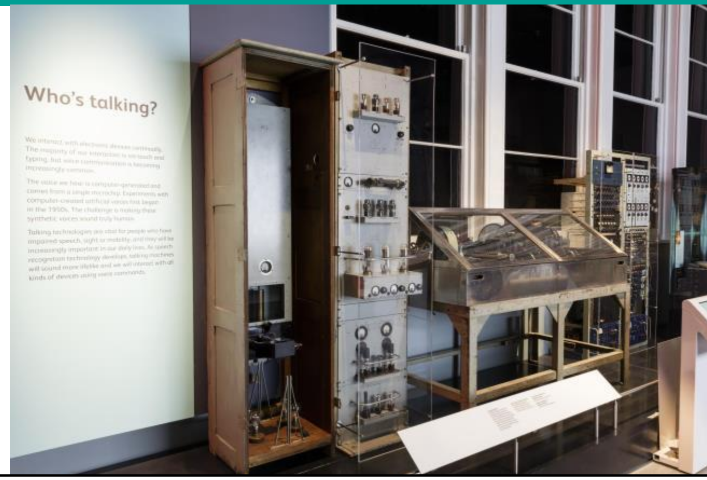
Building on Bells' visualisations, interest in the generation of synthetic speech was continued by scientists at Bell Laboratories, the descendant of the Volta laboratory, culminating in the "Voice-Operation Demonstrator," or Voder in 1939 (on the gallery as an AV clip). It recreated speech by imitating electronically the wide range of different sounds made by vocal chords. It was developed to research voice transmission along telephone lines but was a sensation at the New York World's Fair as an electronic talking machine, the twentieth century equivalent of Euphonia. The fundamental difference was that the Voder operator, by pressing keys, created electrical vibrations which were translated into sound waves rather than forcing air through vocal cord and mouth shapes.

(Life Magazine, 30 January 1939, p 24
https://books.google.ca/books?id=00wEAAAAMBAJ&lpg=PA24&ots=g3bryFQz2I&dq=%22franklin%20institute%22%20AND%20voder&pg=PA24#v=onepage&q=the%20voder%20does%20the%20same%20thing&f=false
accessed 22.5.19)

The device was also known as Pedro, after Brazilian Emperor Dom Pedro, who heard Bell's newly invented telephone at the Philadelphia World's Fair in 1876 and was said to exclaim "My God, it talks!"

TIM   The Speaking Clock, 1933

National Museums Scotland



TIM the Speaking Clock

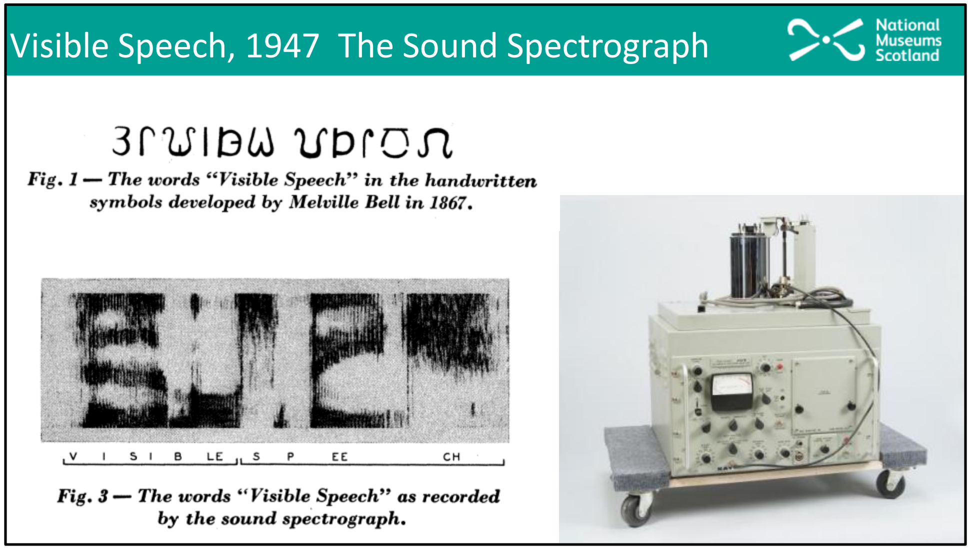
T.1964.33 TIM – pendulum clock, amplifier and optical reader

The Voder is present in the gallery as part of an AV. TIM the Speaking Clock is on display.

The UK's speaking clock service was launched on 24 July 1936 from Holborn Exchange in Central London (now the home of the BT Archive). A back up device was installed in Liverpool in 1942 and it is that apparatus which came to National Museums Scotland in 1964 shortly after the service was upgraded.

TIM is an early example of a general audience access to, if not synthetic speech, then certainly manipulated speech, created after sampling a human voice. In all, Jane Cain, the famous 'girl with the golden voice' recorded 79 individual short phrases and words, which were captured on four glass discs. The required time phrases were put together to create an accurate time message, by being read by photocell from the spinning glass discs. Her voice had to be manipulated to ensure each phrase was exactly the duration required to fit on the disc and to create a uniform message with the correct intonation.

Clearly, recorded speech was not a new phenomenon by the 1930s, but the technology to create meanings by sampling speech was new and the speaking clock would be an early example of hearing a machine 'speak'.



Sound Spectrograph

To truly produce synthetic speech, however, required a still greater understanding of how the individual components of speech come together to create an understandable voice. Fundamental to later 20th century developments in the field was the publication in 1947 of a second book with the title Visible Speech. The book is an account of work the authors carried out at the Bell Telephone Laboratories in the 1940s. The book's title is no co-incidence, acknowledging the pioneering work of Alexander Melville Bell. The work of the team at Bell labs was focussed on teaching deaf people to speak, thus linking back to both Bells' primary ambitions. The introduction to the 1965 edition states

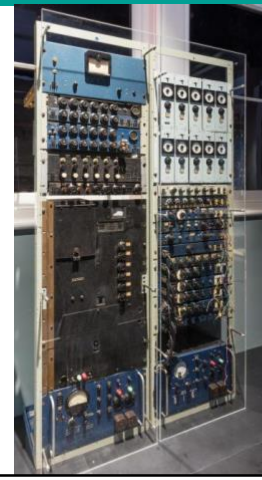
'In the field of speech itself, the work disclosed in Visible Speech has formed a foundation for work on speaker identification, automatic recognition of speech (alas not yet usefully accomplished), the improvement of electrical speech communication facilities and voice coding devices, the interpretation of important garbled or noisy speech messages, and the construction of speaking machines,'
(Visible Speech, 1966, iii)

One of their key developments was the production of a speech wave pattern recording device, known as a 'sound spectrograph'. National Museums Scotland currently doesn't have one of these but it would be an example of a device we would be interested in adding to the collections, if offered one with a relevant provenance of use (the sound spectrograph shown here is in a collection in Philadelphia).
Bell engineers initially proposed the sound spectrograph to improve telephone transmission as well as to support oral education and visual telephony for deaf people. It also appears this work being carried out during the war years was no co-incidence. The first spectrograph model was built as part of a project with the military to detect the cyphering of telephone communication. It was also used to monitor movements of ships and submarines, by analyzing their individuals engine sounds in water.
Now to an object we do have on display, and the second in the trilogy in this paper's title.

(https://griffonagedotcom.wordpress.com/2018/07/26/the-secret-military-origins-of-the-sound-spectrograph/ accessed 17.5.19)
(Mara Mills, "Deaf Jam: From Inscription to Reproduction to Information," p. 37. Accessed 17.5.19)

PAT The Parametric Artificial Talker, 1958

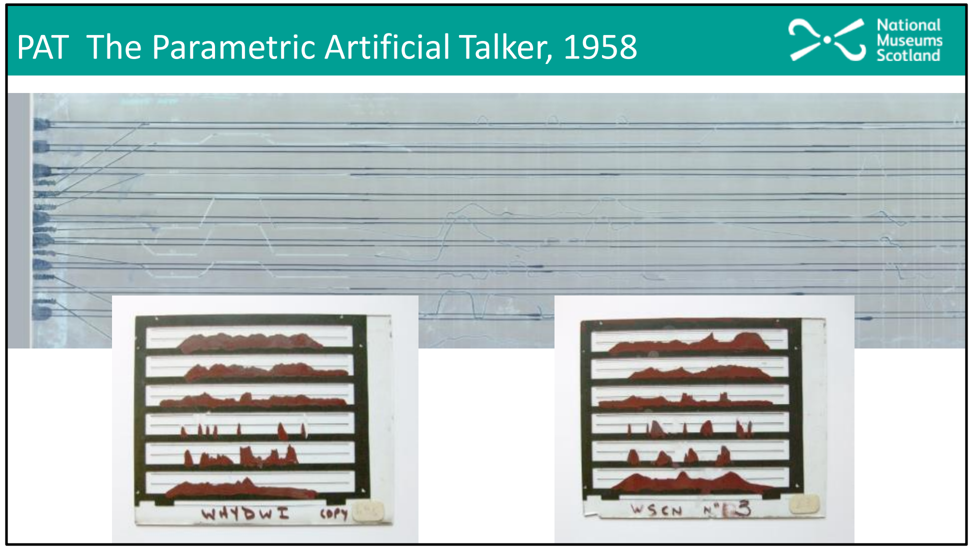
National Museums Scotland



PAT – Parametric artificial talker

The work begun during the war on speech patterns continued after it. Among them was OVE developed in the early 1950s at the Royal Institute of Technology in Stockholm, and, related to it, the University of Edinburgh’s Parametric Artificial Talker – better known as PAT ; one of the first computers to be programed to generate human sounds artificially, rather than use recorded speech. .

PAT is a resonance analogue of the human vocal tract, in which electrical circuits resonate in a way similar to the cavities of the vocal tracts, when provided with information from the "larynx", or in PAT's case, from input tracings made on plastic sheets in conductive ink or produced on glass plates. The ones shown here say 'What have you done with it?' and 'What sound comes next?' and form part of the PAT collections at NMS.



PAT input tracings

The BBC broadcast a documentary series in 1958 entitled Eye on Research, which featured cutting edge university research projects across the UK. One of these programmes was called 'The Six Parameters of PAT'. It opens by saying:

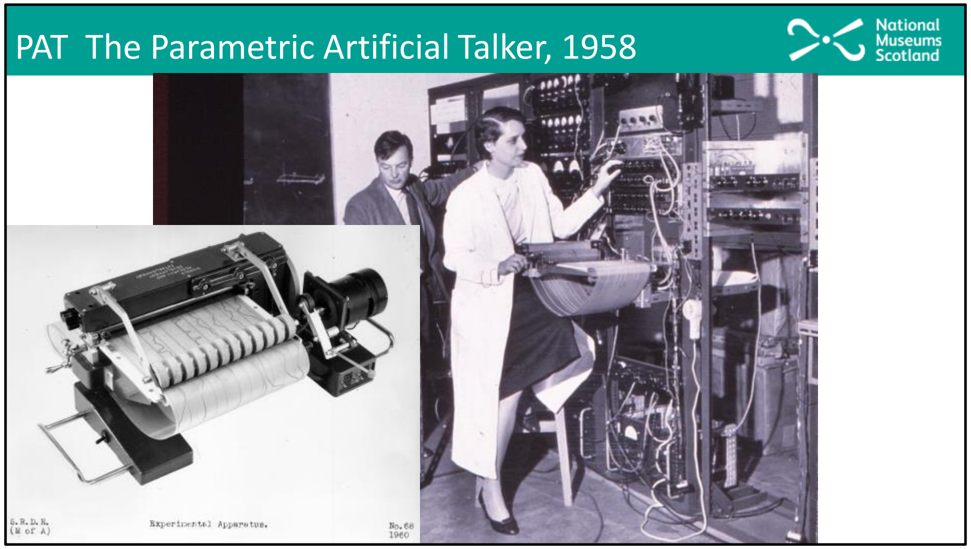
"Up in Edinburgh the scientists have, amongst their apparatus, what one can only describe as a talking machine….."
(Eye on Research Vimeo…..)

The Guardian's article on PAT in 1959 described the process.

"PAT creates speech by generating pulses corresponding to the sounds made by the vocal cords. These pulses are modified by electrical circuits to simulate the action of the human vocal apparatus."
(The Guardian, DATE, 1959, pXX)

The Daily Telegraph of November 30 1959 reported on "Teaching a Machine to Talk Back" where they said:

"It is not acting like a gramophone, a radio, a tape-recorder or a telephone. It is literally a talking machine, and it is producing, not reproducing, words."
(Daily Telegraph and Morning Post 30 November 1959, p xx)



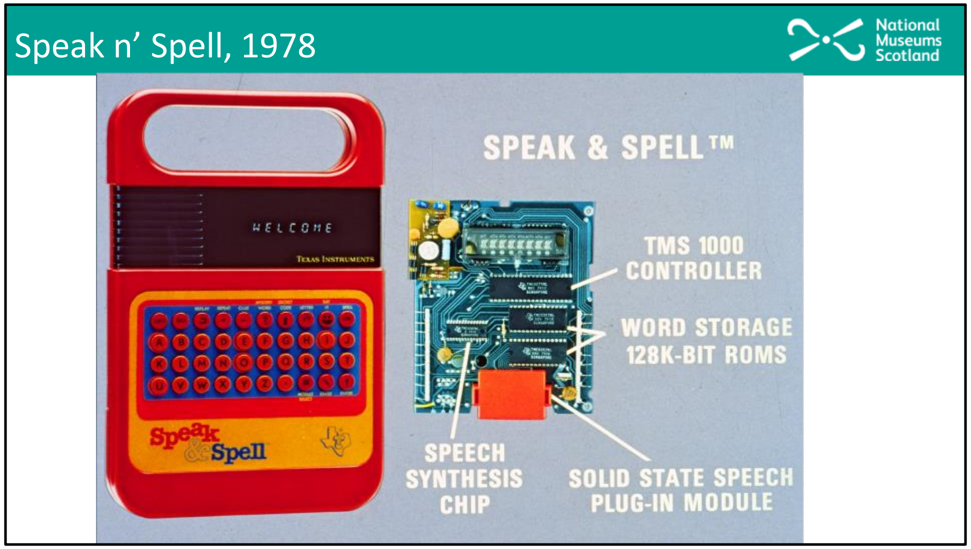
PAT input device

The first version of PAT was developed at the Signals Research and Development Establishment in 1952 to explore bandwidth reduction for use on submarine telephone cables, particularly the soon to be laid Transatlantic Telephone cable, TAT 1, in use from 1956. As with the sound spectrograph in the US, the British Military were also interested for communication cyphering so the original work on PAT was funded by the Ministry of Supply, though the work on telecommunications always seemed to be a key driver. As one of the scientists working on PAT said in the 1958 BBC documentary:

"It seemed to me we could make much better use of these cables if we abandoned sending signals which were faithful copies of the intricate soundwave patterns of speech and sent instead signals that described a simpler speech pattern but which could still be understood."
(BBC Eye on Research , broadcast 28 Oct 1958. Vimeo link https://vimeo.com/26005634)

Using synthetic speech

Research with synthetic speech at the time had unexpected influences. Daphne Oram is now perhaps best known for her foundation of the BBC Radiophonic Workshop in the 1950s and her development of the Oramics machine to produce electronic music. Oram visited PAT on several occasions at the University of Edinburgh, where she experimented with its inputting devices and parameter controls.
(Wilson D., The Woman from New Atlantis, in The Wire, vol 330, August 2011)



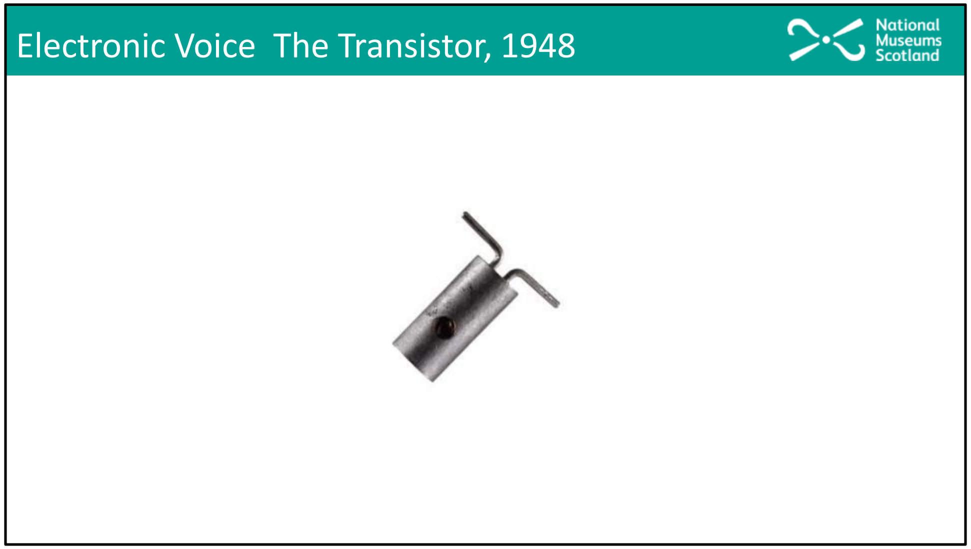
Speak n' Spell

For most museum visitors, their first encounter with electronic synthesised voices might have been in a generation of toys developed in the later 1970s. Companies such as Texas Instruments, better known at the time for making electronic calculators, had developed a voice on a chip. The quality was such that its uses were limited. It was good enough however to feature in one of the best-selling toys of 1978, the Speak N Spell.

Speak n’ Spell, 1980 ET phones home

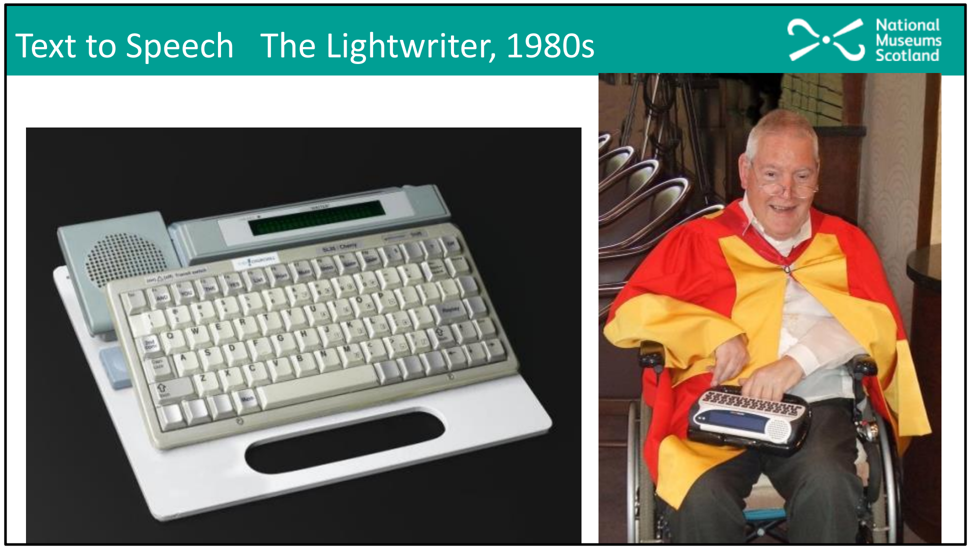
ET

And to feature in the film ET where the Extra Terrestrial famously uses one to 'phone home', sadly not in our collections



Type A transistor

The advances in electronic synthetic speech, enabled by the pioneering work carried out with PAT, and the miniaturisation made possible by the development of first transistors, then microchips (illustrated in our displays with a Type A transistor of 1948 and a range of microchips) had a more direct impact on the development of assistive speech devices for people who have lost the ability to speak or were born without speech.



Toby Churchill

One such is engineer Toby Churchill, who demonstrated his Lightwriter text to speech device on the BBC's Tomorrow's World programme in 1979.
(Tomorrow's World, BBC, 27 December 1979, http://www.bbc.co.uk/archive/tomorrowsworld/8019.shtml accessed 29.5.2019)

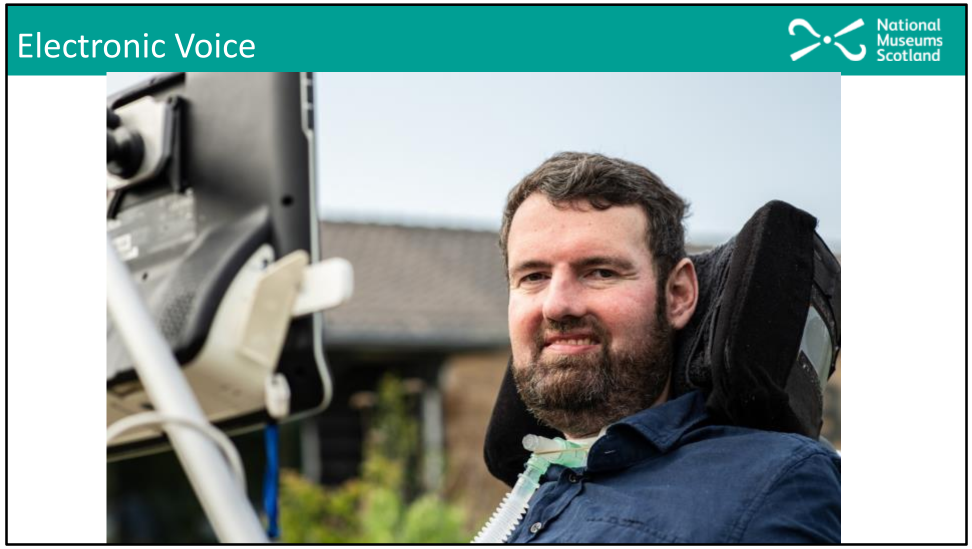
Engineering student Toby lost his speech and most of his mobility in 1968 aged 21. He set up a business making computer based text to screen communication devices, winning 3 Queen's Awards for Industry. He donated one of his prototypes from the mid-1990s to the Museum.

Electronic Voice in the home

National Museums Scotland

Claudius Converse and Nest smoke alarm

Synthetic speech has improved hugely since 1984 when the late Professor Stephen Hawking was first kitted out with his electronic voice, called Perfect Paul, which was developed as the default voice on an early telephone voice interaction device. Some of Professor Hawking's devices are in the Science Museum collections. The Museum illustrates a much more basic device - the Claudius Converse. Next to it is a NEST speaking smoke alarm of 2016, to represent the ubiquity of synthetic speech in the home.



Euan Macdonald

But those working with assistive electronic voices are still grappling with some of the issues experienced by the PAT scientists; how to create a truly human sounding voice with all the nuances that requires. In Edinburgh the Euan MacDonald Centre and the Anne Rowling Regenerative Neurology Clinic, working with the University of Edinburgh Centre for Speech Technology Research, are continuing the quest to make a machine that talks with their Speak: Unique research project. In the gallery we have a short film of Euan MacDonald's work with the Speak:Unique project.
(https://www.speakunique.org/ accessed 29.5.2019)

User experience and the born digital

We have come a long way since PAT began the 1958 BBC documentary by saying very indistinctly "Do you understand what I say to you?" ; the same words allegedly uttered by Bell to Emperor Dom Pedro in 1876.
(Eye on Research)

The irony is, of course, that all these physical devices are silent when on display in the Museum. They are interpreted through text and images and a few AV clips, such as the one on Euan MacDonald. There is one interactive text to speech generator on the gallery.

This brings me to the third in my trilogy - ISAAC

Augmentative and Alternative Communication

National Museums Scotland

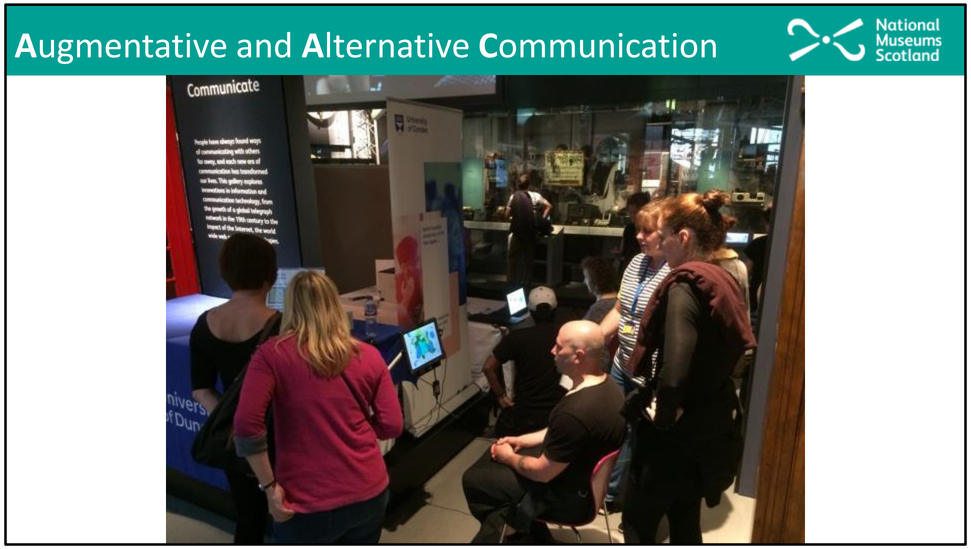
ISAAC

The **I**nternational **S**ociety for **A**ugmentative and **A**lternative **C**ommunication, founded in 1983

ISAAC

The International Society for Augmentative and Alternative Communication was founded in 1983 at a time in improvements in the technologies which made computer based synthetic speech accessible to those with no voice and limited mobility.

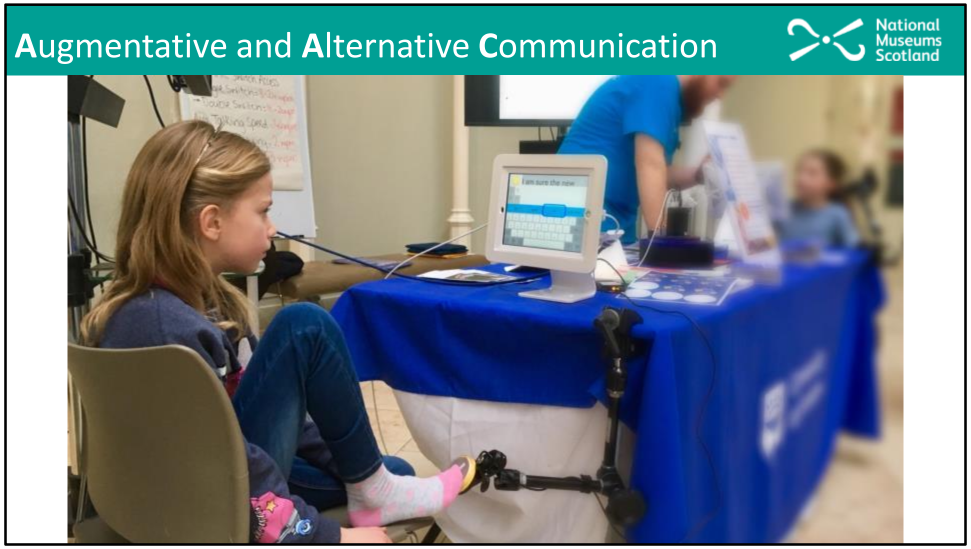
(https://www.isaac-online.org/english/home/ accessed 7.6.2019)



University of Dundee AAC Research Group Tobii Eye Gaze technology for Paint with your Eyes event, August 2017

The challenge for the Museum is that Augmentative and alternative communication is not so directly object-based but a grouping of methods of communication for those with speech and language impairment. These can range from manual sign languages (a link back to Bell and his antipathy to signing; an issue which remains fraught to this day) to computer programs, eye gaze technologies and synthesised voices.

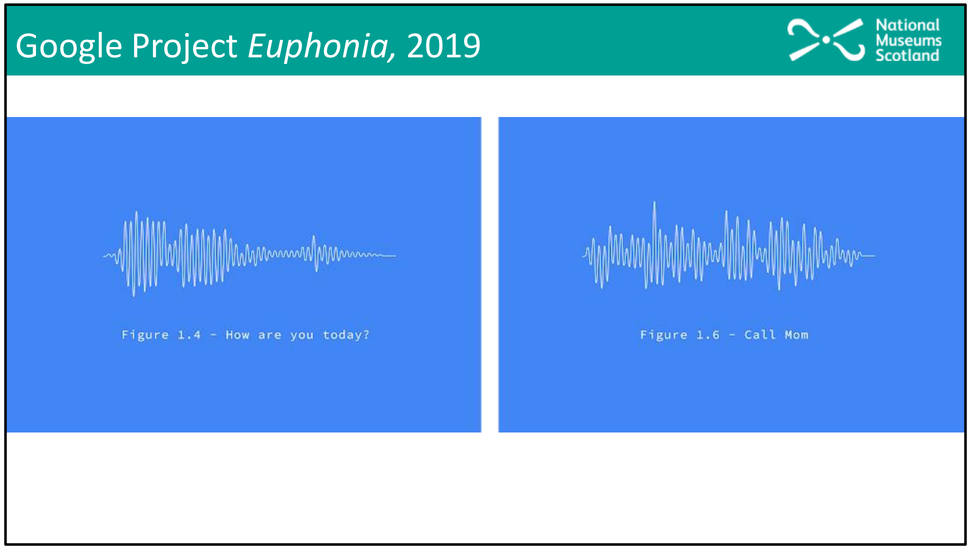
The question for us now is can we collect born digital examples of synthetic speech? At present, no, we do not have that ability. We are however making links with Scottish centres of research in the area, particularly the pioneering work in this field is being carried out on speech generating devices at the University of Dundee School of Computing.



AAC Science Saturday write with your Elbow

As part of their evaluation and outreach, they brought their research to some of the Museum's events to encourage visitors to engage with AAC and learn what it means to use assistive technology when accessing a computer for speech output and communication. The researchers encouraged the visitors to reflect on the situation of not being able to use natural speech for communication and engaged audiences with the realities of the difficulty in communicating for those with no speech and little or no mobility.

Younger visitors also learned about AAC in one of our Science Saturday events, where activities ranged from using any part of the body to step scan a keyboard for speech output, to the very origins of synthetic speech by building their own mechanical vowel resonator, an activity taking us right back to the earliest enquiries of De Kempelen and others.



SONOGRAMS FROM GOOGLE PROJECT EUPHONIA

Taking me back to the beginning of this presentation and earlier attempts at creating machines that speak, Google's Project Euphonia, launched in May 2019, uses Artificial Intelligence (AI) and algorithms to enable computers to recognise the speech of people with conditions such as Motor Nurone Disease (MND).
(https://techcrunch.com/2019/05/07/googles-project-euphonia-wants-to-make-voice-recognition-work-for-people-with-speech-impairments/ accessed 20.5.19)

The developers say

"We collaborated closely with groups to learn about the communication needs of people with MND, and worked toward optimizing AI based algorithms so that mobile phones and computers can more reliably transcribe words spoken by people with these kinds of speech difficulties….

To do this, Google software turns the recorded voice samples into a spectrogram, or a visual representation of the sound. The computer then uses common transcribed spectrograms to "train" the system to better recognize less common types of speech.
This illustrates the vital link between the work of both Melville and Alexander Graham Bell on the visualisation of speech and the ultimate aim to create devices that are not only "capable of articulating human words and sentences" as the 19th century Miscellany put it, but that do so with unique, individual and personalised synthetic voices.
(Chambers' Miscellany of Useful and Entertaining Tracts, W&R Chambers, Edinburgh, Vol XIII, 1847, p 55)