Voice activated
Adrian Woolfson

Now You're Talking: the Story of Human Conversation from the Neanderthals to Artificial Intelligence
Trevor Cox

Bodley Head, 312pp, £20

Unlike the mineralised structures of bones and shells, human voices do not fossilise. As a result we may only imagine the "Half-penny half-pint!" cries of the milk-sellers of Victorian Billingsgate, the articulations of early hominids or the clamour of ancient Roman audiences in the Colosseum. Such ephemera form the intangible and unknowable dark matter of human cultural history.

It is perhaps unsurprising that having drawn our awareness to the richness of the unusual and esoteric sounds generated by the physical world in his book *Sonic Wonderland*, the author, broadcaster and professor of acoustic engineering at the University of Salford, Trevor Cox, should now turn his attention to the nature, history and future of human speech and conversation.

For as intriguing as the auditory aesthetics of the Whispering Gallery at St Paul's Cathedral, the sound of water lapping on an iceberg, or the Mojave Desert may be, the soundscape of the physical world lacks the complexity of human speech, which is uniquely imbued with conscious agency and defines individual identity.

Although we are unable to accurately track down the origins of human speech, Cox discusses how the use of CT scans to infer the dimensions of the ears of extinct hominids, coupled with modelling how sound waves oscillated within ancient ear bones, has enabled paleoanthropologists to establish that the basic auditory machinery necessary for speech receptivity was intact at least half a million years ago. A fossil from an ancient species of human, Homo heidelbergensis, meanwhile, indicates that the capacity for speech originated around the same time. The origins of protolanguages are harder to pin down. But if referenced to the emergence of symbolic art, comprehensive human language most likely emerged no more than 100,000 years ago.

Having established the origins of vocalisation, Cox sets out to explore its development, neural basis, genetics and other ramifications. He illustrates the deep imprint of evolutionary history on our auditory apparatus by revealing that up to the age of three months, babies prefer certain monkey calls

The Queen's voice has deepened each decade

to human speech. He also touches on the relationship between language and the brain. For example, feral children isolated from human contact in infancy struggle to acquire language competence, suggesting that language acquisition may be restricted to a critical developmental period. On the other hand, the relative preservation of language capabilities in a patient with a tumour in Broca's region of the brain, which facilitates language, indicates the organ's resilience to certain injuries and the anatomical fluidity of brain function. Inherited language defects suggest that some aspects of language have a genetic basis.

The voice, of course, is not constant throughout life. The lowering of a male's voice in adolescence is caused by testosterone. In a rare moment where sounds offend rather than delight Cox's heightened auditory sensibility, he describes listening

Newborn babies prefer certain monkey calls to human speech

to the scratchy wax 1902-04 recordings of the artificially high voice of the famous adult castrato Alessandro Moreschi. But it is not just the timbre of an individual's voice that changes with time. In a delightfully innovative use of archival recordings, Cox contrasts the earliest surviving episode of Alistair Cooke's *Letter from America* BBC radio series, from 1947, with his final broadcast in 2004. While the first broadcast contains three syllables per second, this drops to 2.4 in the second. The Queen's annual Christmas broadcasts provide another opportunity to trace the evolution of a voice. It turns out that hers has dropped by about one semitone each decade.

In a fascinating examination of how the nature of voices invoke prejudice and

preconceptions, Cox recounts the pioneering work of Tom Hatherley Pear, who in 1927 asked radio listeners to comment on their perceptions of nine individuals. He demonstrated how vocal stereotypes are used to build a detailed and vivid impression of the speaker's appearance, character and personality. Cox also explores the relationship between the voice, identity and sense of self, describing the distress of a brain-damaged patient with foreign accent syndrome, characterised by abnormalities in pronunciation and intonation.

In an age of spin, fake news, questionable facts and excessive social media, it is appropriate that Cox also addresses the issue of vocal charisma. One aspect of this is the association of the veracity of a statement with the accent of the speaker, with non-native accents being rated as less truthful. Lower pitched voices are, furthermore, associated with strength, integrity and competence. Candidates in the 2012 US House of Representatives elections took 4 per cent more votes and were 13 per cent more likely to win if they had lower voices. Faster speaking with varied pitch and changing volume is also perceived as more charismatic. Even frogs lower their croaks when attempting to increase their influence.

In the lively and quirky latter part of the book, Cox turns to the future of the voice and how it will be impacted by technology. The microphone was perhaps the first innovation to significantly change the voice. It freed up singers such as Bing Crosby from having to project their voices in performance theatres, enabling them to innovate, in his case by developing the conversational style known as crooning. More complex electronics enabled bands such as the Beatles and the Beach Boys, and more recently Kraftwerk, Cher, and Björk, to distort their voices in idiosyncratic ways.

Whereas the speaking machine constructed by the Hungarian Wolfgang von Kempelen in the 18th century produced just rudimentary vocalisations, Joseph Faber's Euphonia was in 1846 able to sing "God Save the Queen". But Cox makes it clear that the use of AI in the form of neural networks, and machine-learning are poised fundamentally to reconfigure the voices of the future.

Machines already behave as if their vocalisations are underpinned by a genuine mental process: whether or not they are actually able to become conscious may in the end turn out to be irrelevant. It appears as if humans have a tendency to attribute agency to anything that talks – such is our complex dependency on the voice. ●

Adrian Woolfson is the author of "Life Without Genes" (HarperCollins)