

Weaving a web of words

Melanie Butler learns from lexicographer Michael Rundell how new technology has fundamentally changed the nature of dictionaries

'If you ever own up at a dinner party to being a lexicographer, which I wouldn't recommend, you find people have the most extraordinary idea of what we do.' Michael Rundell winces. 'They think we spend hours in rooms discussing the most obscure words in the English language. I'm much more interested in simple words like take.'

For the editor in chief of the Macmillan Dictionaries, 'it is the core language that counts'. Rundell, who started his working life as a classics lecturer in Ghana – 'I don't know how useful that was to them, but I loved it' – began his dictionary career as a freelancer. He has worked for most of the big-name dictionary publishers but is best known for the Macmillan dictionary launched in 2002.

The most difficult thing about the Macmillan project, he admits, was launching a new product into a market that was so competitive. Its success, he believes, lay in the way that the Macmillan team, working with computational linguists such as Adam Kilgarriff, was able to harness technology to find new information about important features like collocation and new ways of presenting the information, such as the much-copied Macmillan collocation boxes.

'It was just the right time. The computers were more powerful, the corpora were bigger, there was cutting-edge analysis software.

'The corpus reveals things to you. It reveals, for example, how its collocates can help you find the meaning of a word.' He gives the example of the difference between 'economic' and 'economical' – hard to explain until you look at them with their collocates: 'economic' appears with 'policy', 'crisis' and 'growth'; 'economical' with 'price', 'car' and 'engine'.

For a corpus baby like Rundell who cut his lexicographic teeth in the golden age of corpus research, the internet offers infinite possibilities. It provides, for example, the largest corpus in the world. 'The first Cobuild Corpus had seven million words, the Macmillan corpus has two billion, but how many trillion words can we find on Google and Twitter?'

The internet, Rundell believes, has changed the very nature of dictionaries. 'Dictionaries are now part of search, and search is a very big enterprise. We all search every day. When we search on a word like oligarch we may be looking for information on dodgy Russian billionaires, ancient Greek politics or just the definition of the word – Google has no way of knowing.'



Courtesy of Chris Tribble

FUTURE PERFECT For Michael Rundell, Macmillan's decision to stop producing print dictionaries is a cause for celebration

More than 50 per cent of the users of Macmillan dictionaries online come through search engines, others are driven by news coverage. Some return

again and again and become contributors to the Macmillan open dictionary, the Wikipedia of ELT lexicography and a treasure trove of neologisms, street

languages and regional differences. Crowd-sourcing entry is only one of the things the online world also allows Macmillan to do: they can and have merged dictionary functions with a thesaurus – impossible in a book – to enable a switch between the two at the click of a button. They have words of the week, blog posts – 'I see the future as having all our knowledge, our corpus, our expertise in one big pot which users can access in many different ways.'

For Rundell the age of the dictionary as book is dead but the age of the dictionary as online resource has only just begun. For him Macmillan's decision to stop producing print dictionaries – the very last copies are now rolling off the press – is a cause for celebration. Isn't there anything he will miss about the old print dictionaries. 'Launch parties!' he laughs.

For the new world of lexicography the book had reached the end of its useful life – there were too many space constraints, they took too long to update, they were becoming too big and bulky. 'I looked at an old Oxford advanced dictionary the other day; it was 1,000

pages long, and in recent years they've come in at over 2,000 pages. Not something you can carry around in your pocket.' And that is what sets reference works like dictionaries apart from other books – you want to use them instantly when you need them, not read them from cover to cover or even browse them. 'As my friend Hilary Nesi says, dictionaries, encyclopedias and maps are things you need when you are busy doing something else.'

The world of lexicography has been transformed by computing. If there was one single electronic gizmo, one killer piece of software Michael Rundell would wish for, what would it be?

'That's easy, adaptive technology which gets to know you personally and makes suggestions about what you may want – a bit like the way Amazon offers you books. A dictionary which knows what level your English is, whether you like to hear the pronunciation of the word in British English, American or not at all.'

Some work has been done on 'adaptable' dictionaries – where users are asked to give their mother tongue, language level and personal interests – but so far it hasn't been outstandingly successful. 'What I want is a dictionary powered by a huge research database but that configures the content and functionality to suit me. A personalised dictionary. I think that's what everyone wants.' ■



Macmillan Dictionary Online: it's everything print dictionaries wish they were but never could be ... not in their wildest dreams.



It's your English!

www.macmillandictionary.com