MACREX – a history

Geraldine Beare interviews Drusilla Calvert

It is nearly 50 years since the Society of Indexers was founded by G Norman Knight and his fellow enthusiasts, and more than 100 years since the dissolution of the Index Society founded by Henry B Wheatley (see pp. 15–17). For most of that time, indexers produced indexes on 5-by-4-inch cards using pens and pencils. Twenty years ago form followed function when the first Apple Macintosh computer was unveiled (the company had been founded in 1976, producing only circuit boards for use by other companies), costing a hefty £1399 – well beyond most people's pockets – and a year later, in 1985, Microsoft Windows was launched. By the mid- to late 1990s, email and the wonders of the internet had joined the ranks of must-have technology. We take all of this for granted, and as indexers we could not survive without it. However, for computers to work, software is needed, and before both Steve Jobs and Bill Gates made their spectacular entrances, two people started up an enterprise which enabled the indexing profession to move to the forefront of the technology of the time. I am of course referring to Drusilla and Hilary Calvert.

There is a fine line between deliberately advertising a product and presenting it in such a way that it can stand alone as something to be admired and studied in a historical sense, and the pitfalls are enormous, yet MACREX is not something we can sweep under the carpet – for many of us, it is essential to our livelihoods. (I am aware of course that CINDEX too has a reputation and a history, but it is younger and needs a separate article altogether). In order to partially get round the problem of advertising, the following takes the form of a question and answer session which I hope will answer some of those frequently thought but never asked questions. This method may seem clumsy – but I hope it works.

GB: When did the idea for MACREX first come to you?

DC: In about 1981 Hilary wanted an excuse to buy a home computer because these were very expensive at the time. The deal was that he could do so provided he wrote a program to help with my indexing, and this is what happened. The first version was a RAM-based program which used about 8 kb of memory. The data was saved onto cassettes on an ancient cassette recorder. The computer, which cost £2500, was built from a kit. Later on a floppy disk was added for the cost of about £450. This led to our first program, which we called Micrex.

GB: How did you go about 'inventing' it?

DC: We were aware of the need for a program that could cope with large indexes and this was precipitated by a potential purchaser who needed to do an index of 80,000 entries. This led us to design a new program, MACREX, from scratch. The basic principles were that the capacity should be unlimited (as far as that was possible with the hardware available) and that the input of the index should be in a standard form, with the particulars of the layout etc. being determined at the time it was printed (or output into a word processor file). This was to accommodate the wide variety of styles that our users wanted, and to allow flexibility by separating the formatting functions from the others. This flexibility was even more necessary than we had originally imagined. Importantly, the limiting factor for the number of entries at this time was the amount of RAM available on common machines at the time we started, which was usually 64 kilobytes with about 56 k available for programs to use. (Compare this with 256 *mega*bytes [256,000 k] now standard on entry-level machines). By separating the index into batches, each containing a shorter alphabetical group, we could sort the index by taking it a section at a time into the RAM rather than all at once. We also devised a system where each batch could be automatically divided into temporary files which were sorted one at a time and then merged together again. This architecture proved very successful and permitted the 80,000 entries mentioned above to be sorted on a 64 k system. Although the use of batches is not usually necessary on modern machines, we have preserved it in the current program for the few people who have truly humungous indexes. This allowed MACREX to sort and print in a fairly complex layout indexes such as the index for the 25-volume *Letters to Delegates to Congress, 1774–1789* (cf the *General Index of Journals of Congress from the organization of the Government to the present time* mentioned by Wheatley in 1879!) published by the US Library of Congress which had, when it was imported into MACREX, over 300,000 entries. The final index was published in a separate 803-page volume.

Before we devised the program we had no idea how many different printing layouts would be needed, nor any idea of the range of reference styles used. We are still, more than 20 years after writing the first version of the program, finding the need to add printing options. One of the main precepts behind the development has always been that nothing is impossible; it would be good to think that there was a standard format used for all indexes, but this is unlikely to happen.

Our contact with Nancy Mulvany, our agent from 1985 until mid-1995, precipitated a raft of new developments for the US market. Most importantly, we needed to make adjustments to the program to make it compatible with the *Chicago Manual of Style* (we have to reassess this whenever a new edition appears). In addition, the Scandinavian market requires us to adapt the sort function to take account of the extra characters used in these languages.

Recently, several publishers have produced templates into which they require indexers to insert their indexes. This can be immensely time-consuming and we (or more accurately, our current US agent, Gale Rhoades) have been able

to produce specialized files to convert an index made in MACREX into these specialist formats.

GB: How did the name MACREX come about?

DC: A friend suggested that we call the original program Micrex (the small program) and the new one MACREX (the big program – Macintosh computers had not been invented at the time). Somehow the name has stuck. Gale Rhoades prefers the story that it stands for MACro indExing (macros being a collection of keystrokes executed by pressing a single key combination) because of the enormous advantage gained from the use of these and other keystrokes and time saving features built into MACREX – but this isn't in fact the case!

GB: Did you involve members of the Society of Indexers or publishers and printers from the start?

DC: SI showed interest in the program right from the start: I was summoned, knees shaking, to a Council meeting to describe the program, and I remember sending the original manual to Ken Bakewell and others for their comments. Although a large number of the features have been put in at the request of users, many of the time-saving and input features have been devised by us in response to a clear need, and were not requested by users. For instance it seems incomprehensible to indexers nowadays that a program should not be able to put entries in page number order, but that's because we introduced this feature very early on in development.

We had some interest from publishers, although they have tended to think that their IT departments, or even the printer, could produce an indexing program with no problems – something which has been proved wrong again and again. We also used *BS 3700* as the basis. However, we quickly found that the standard did not answer anything like all the decisions which we would have to make. For instance, there was no guidance on the order of bold, italic, annotated references to the same number, order of symbols, accented characters etc.

GB: How unstable was the program to begin with?

DC: MACREX has always been reasonably stable although, as with any other program, there have been (and probably still are) bugs which only emerge after years of use. Windows XP is probably the most stable operating system yet. The backup file that contains the whole index is a basic text file, which means that it can be edited in any text editor or in a word processor as a text file and then returned to the MACREX program. (Interestingly, it is possible to load backup files made with MACREX version 1 or even with Micrex into version 7 – something which, to our surprise, has proved very important).

GB: How difficult was it working with the variety of programs/computers around 20 years ago – and now?

DC: When we started there were a huge number of different microcomputers – Commodore Pet, Apple II, BBC Micro, etc. We only supported what was the then business standard operating system (WordStar ran on it) called CP/M. However, there were still problems because every CP/M machine used a different disk format, so we had to go to immense lengths – and expense – finding hardware to write all these different formats. In addition each machine would tend to have its own unique way of positioning the cursor on the screen, displaying bold and so on, so the user had to run a configuration program before MACREX would work, or we had to make a 'home visit'. Following the introduction of MS-DOS there were initially some differences between machines – for example Apricot controlled the screen in a different way from the IBM standard – but this has largely disappeared. With the introduction of Windows 2000 and XP we have had to make considerable changes to the workings of the program to accommodate the sophisticated features of these operating systems. MACREX version 7 and future versions of the program will not work on old MS-DOS systems, and MACREX version 6 (our last MS-DOS-compatible version) will not work reliably on XP systems.

GB: How about updates?

DC: Each new version contains new and/or upgraded features, and version 8, with built-in authority tables and autocompletion of known headings, with a redesign of the menu structure, is coming soon.

GB: Were you able to look at any commercial products before you started?

DC: There were really no equivalent commercial products – there was one in use in the USA written by an indexer for himself and his indexer wife which was sold for a time, but when MACREX came out he switched to using that. CINDEX and SKY came out later, but we have never sought to compete for sales. There is, in our experience, very little money to be made out of a specialist program for indexing.

GB: When did you start to give demonstrations of MACREX?

DC: At a SI meeting in London at the Library Association in 1981 at a session designed to assess the role of microcomputers in indexing, which attracted an unprecedented number of SI members and generated considerable interest. I think it was the only time that both A and B rooms in the suite at the LA had ever been open for an SI meeting.

We continue to give demonstrations and workshops, particularly at conferences.

GB: How much time was involved in getting the program up and running and how much time can you give it now?

DC: A lot of time – most evenings and weekends for years. Now, an hour or so a week with more time during holidays to add new features.

GB: Have you had many requests for one-off projects?

DC: Apart from *Punch* and *Nature*, for which we produced a system automatically generating the author index from the subject index, we have normally been able to make minor modifications to the program however complex the demand. Recently, we have also given advice to MACREX users, on SIdeline (the Society of Indexers' email discussion list) and individually, on how to deal with the CUP-XML indexes.

GB: How about the Macintosh system?

DC: We don't include a Macintosh system. We thought of it, but it was very expensive and time-consuming to develop a compatible program. However, a number of Mac users run MACREX successfully using add-in cards or software to emulate a PC or Windows environment.

GB: How difficult has it been keeping up to date with the new technology and for incorporating the new electronic indexing that seems to be coming in?

DC: We have a facility for indexing HTML (web) documents and for working with other types of embedded index. The techniques used for indexing websites can also be applied to other types of electronic copy. As far as new developments in electronic indexing are concerned it has been difficult to produce a general system because the publishers are tending to keep it in house. There is no generally accepted standard way to do it and it can be very difficult to find out what people actually want to do. If publishers could agree a common standard for electronic indexing (in the same way that HTML, and indeed the architecture of the PC itself have been agreed), then it would be more worthwhile for software developers to accommodate it.

Forecasting the future is foolish; the death of the book is as unlikely as Francis Fukuyama's 'end of history', and where information is needed, either in what we now call 'hard copy' or in electronic form, so too will indexes be needed. MACREX and CINDEX continue to lead the way, and we should not forget their pioneering spirits.