

1 Introduction

A version of this paper was presented at the 1995 meeting of CAA under the light-hearted title *I just don't believe it (typesetting can seriously damage your health)*. At that time there was no intention to publish but comments from the (small) audience suggested that there was a need for the points made there to be more widely available and hence this paper. It has no particular originality, but aims to give advice to potential authors, editors and typesetters based on the experiences of one who has been involved in the process of computer-aided publication over the last six years. I also take the opportunity to air some thoughts concerning the future of publication in archaeology and academic publishing generally, especially in regard to electronic publication.

Before proceeding to the main aims of the paper, I wish to put the subject in context. Up until the end of the 1960s, the majority of publications in archaeology and other disciplines, was undertaken by professional publishing houses who either typeset the volumes in-house or sub-contracted to specialised firms. Generally, the result was high quality publications such as the *Society of Antiquaries Research Reports* (e.g., Wheeler/Richardson 1957). By the end of the 1960s the huge expansion of archaeology, both in the field and academia, created a need for a cheap and rapid form of publication, and it was this need which prompted the foundation of *British Archaeological Reports*, as well as other series. These reports were often little more than bound typescripts — either from the famous BAR typewriter (e.g., Casey/Reece 1974), or occasionally as submitted by the author. These publications filled a need. Another example of this form of publication is, of course, the early volumes of *Computer Applications in Archaeology* (e.g., Laffin 1986).

At the same time, computer technology was advancing rapidly. As cheaper PCs, functional word-processors and dot-matrix or daisy-wheel printers became more easily available, the roles of author and typist became conflated. All but a handful of technophobes now take for granted the process of writing one's own papers on a word-processor. This is a relatively new state of affairs.¹ Unfortunately, the wide variety of word-processing packages available often

resulted in the computer files not being re-usable and publishing houses having to re-key papers. This situation still exists — I know of at least one numismatic journal, and at least one recent book published by a major academic publisher, where the text was re-keyed. The majority of scientific journals re-key submissions. This is despite attempts at standardisation and development of exchange formats (e.g., Wilcock/Spicer 1986).

During the early to mid-eighties the increasing availability of reasonable quality laser printers and the popularity of typesetting packages such as TeX (Knuth 1984) and LaTeX (Lamport 1986), or DTP packages such as PageMaker, led to a further conflation of roles — this time the role of editor and typesetter. Rahtz (1986) summarises the situation in the mid-eighties. From 1987 onwards the CAA proceedings ceased to be produced as little typescript booklets and reasonably well-typeset volumes were published in a recognised series. These volumes, and all those subsequently, have been typeset by one of the editors using a variety of systems.

It is the conflation of the four roles into two: typist and author, editor and typesetter, along with the continuing pressure to publish quickly, which creates tensions in the whole process. In some cases, all four roles are combined into one: the most obvious example being the writing of dissertations and theses. Unfortunately, authors attempting to typeset texts which are for submission to a journal or edited volume often create extra work for the editor/typesetter. Hopefully, this paper will promote a little mutual understanding and thus help to ease these tensions.

A second facet of these developments is that the editor/typesetter is often now an individual who has no training, and possibly little interest, in typography or typographic design. Few of the available typesetting systems, or word-processors, actually offer any advice in these matters to the user, either in the manuals or on-line help, if available.² One is told *how* to do things, but not *when*, or *why*. This has resulted in the highly variable quality of publications. I do not claim any great merit for my own work; for example some pages in Lockyear and Rahtz (1991) are truly awful. Some self-typeset publications are of great merit; a good example is Dixon (1994) which was superbly

typeset by the author using PageMaker 5. The second aim of this paper is to provide potential editor/typesetters with a few basic typographic rules and tips.

2 Never mind the quality, feel the width...

The phrase which best illustrates why quality of production matters is 'don't worry about it, it's only a BAR...'. Whether we like it or not, first impressions do matter and the quality of production is often used as an indication of the quality and *credibility* of the contents. Of course, this equation is not necessarily true: some fundamental papers have been published in BAR volumes using the infamous wobbly typewriter, whereas the elegant and stylish *Cambridge New Directions* volumes include papers which are of far less value. To a lesser extent, the opposite can also be true: extremely well typeset and luxurious volumes with many colour photographs, such as *Symbols of Power at the Time of Stonehenge* (Clarke *et al.* 1985) can be dismissed as 'coffee table books' without the reader actually reading any of the text!³

Another aspect of design, which is important to publishing companies, is the creation and maintenance of a 'house style.' This allows readers to identify, perhaps subconsciously, a particular style with a particular publisher, and by association with certain topics, or even certain levels of quality and credibility. A good example of this is again the distinctive *Cambridge New Directions* volumes such as *Ranking, Resource and Exchange* (Renfrew/Shennan 1982). Given the title page of any article in this series most British archaeologists instantly recognise the source, and associate the paper with cloth-bound, thin volumes with a high theoretical content. CAA is trying to maintain a general house-style with volumes being in an A4 format, two columns, red cover and so on, although at a more detailed level there are many differences.

A more important aspect of quality and good typographic design is functional. Typography and book design is the result of a long process of development over many centuries. Many parts of a book allow it to be *used*: tables of contents, lists of figures and tables, page numbers, running headers and footers and the index can all be seen as navigation tools. Having found the piece of text one wishes to consult, the next stage is to *read* it. Typographic design, especially in academic publication, is primarily aimed at presenting the author's words in a manner which allows them to be easily read; just as grammar and punctuation is primarily aimed at helping one to understand the meaning. If the text is badly typeset, it becomes difficult to read: lines are lost, punctuation missed, emphasis is not apparent and the meaning becomes obscured. Tukey's superb *Exploratory Data Analysis* was typeset in an adventurous and unusual style (Tukey 1977). Unfortunately, the style

makes reading and using the book difficult which is ironic given that the techniques described therein are designed to make patterns in data clear.

It is also true that production of a well-typeset book is an aesthetic pursuit and the process of turning text into type is one which can afford great personal satisfaction. Some typographic rules are more concerned with appearance rather than function. Within the rules there is much scope for personal expression and experimentation, but the fundamental question should always be *is the book readable?* Many people regard books as objects both of value, and potentially of beauty. I will freely admit that seeing my work transformed from 300dpi laser print into 1300dpi type on high quality paper with a stitched cloth binding (e.g. Duncan 1993) is highly satisfying. Section 6 presents some tips for the novice typesetter.

3 The editor/typesetter's problems

What are the problems which face an editor who is also acting as the typesetter? These fall into two categories: those which arise from the authors themselves and problems with the software.

3.1 THE AUTHORS

Today, most books and papers are received on disk. The biggest task of the whole process is editing these files. Often, text is poorly input by the author. One of the main reasons is that usually the authors do not use the full functionality of their word-processors, or use ingenious but wrong methods to achieve the effects they require. A common error is to force a new page with hard-returns. For submission to a publication the authors should forget about pagination — the pagination during writing will have no relation to the pagination of the typeset article. If the text is to remain no more than a word-processed document, page-breaks should be forced using the correct command. The worst error, however, is the use of spaces to centre text, line-up columns and so on, rather than using the centre command or tabs. DOS word-processors, such as WordPerfect 5.1, have to use a fixed space font on screen, i.e. every character takes up the same amount of horizontal space. Some printer fonts, such as Courier, are also fixed space, but in a typeset publication these are only used in special circumstances such as quoting an e-mail address, as in the *Archaeological Computing Newsletter* (ACN). The main font in a typeset publication will be a proportionally spaced font where the letter i, for example, will take up much less horizontal space than the letter m. If columns are lined up on screen using spaces, they will be ragged when printed in a proportional font (fig. 1). In typesetting Duncan (1993) I had to remove the spaces from 113 tables and replace them with tabs and this task formed the major part

12	456	axe	LEC592
75	67	cup	LEC189
13	432	beaker	LEC592
101	2400	sword	LEC593
102	1200	dagger	LEQ67J
12	456	axe	LEC592
75	67	cup	LEC189
13	432	beaker	LEC592
101	2400	sword	LEC593
102	1200	dagger	LEQ67J

Figure 1. Lining columns with spaces on the left hand side (top) results in ragged columns when a proportionally spaced font is used (bottom).

of the job. That volume, input by a professional typing firm, also used an ingenious method of getting subscripts such as x_j . The ‘1’ was put on the next line, positioned using spaces, and then the line height set to 0.5!

Most journals and conference proceedings issue guidelines for authors. It is astonishing how often these are ignored. The instructions could be unclear but when I have issued these to contributors, I have never been directly asked for clarification. This is, however, no excuse for inconsistency — even if the authors have not understood the instructions given, it is easier to change the files using macros if they have been consistent.

Two problems which always occur are horrible graphics and references. I have aired my thoughts on graphics, and what is known in statistics as chart junk, elsewhere (Lockyear 1994). Graphics should be clear, uncluttered and simple. I have often returned graphics to the authors as unpublishable and the following three reactions are common:

1. the graphic is dropped (ideal — saves space)
2. the graphic is reduced on a photocopier (this does *not* improve the quality of the graphic)
3. the author pleads ‘this figure is really important and I haven’t got time to have it re-drawn as...
 - ... the drafts-person is away
 - ... I don’t have the data any more
 - ... my computer is broken... [insert as appropriate]
 - ... Could you include it anyway – I’ve seen worse published in...’ (I re-draw the figure myself).

Given that the graphics are the most obvious part of a paper, *why* does anyone want to ruin their otherwise carefully crafted *magnum opus* with a figure that looks like the scrawls of a five-year old with a felt pen?

References are a continuing problem not unique to the current situation. References are a fundamental part of an academic paper and distinguish them from popular publications. To check that everyone’s references are

correct is very time-consuming. As a minimum, the editor will check that all the parts of the reference are present: volume number, pages, journal title, etc., and in any case it is the author’s *duty* to ensure that the bibliography is complete. Despite this, approximately 60% of papers I have typeset have had incomplete references.⁴

A common failing with authors, and one which I am as guilty of as most, is late submission of texts. One advantage of computer-aided publication is that it is sometimes possible to include papers that have been submitted late, especially if the submitted text is well written and well input. It is when a late text contains the errors noted above that the editor/typesetter’s job becomes even more fraught.

3.2 SOFTWARE

I shall not waste space here with detailed criticisms of software but simply wish to warn budding typesetters that they will have to cope with:

1. incompatible formats and poor import filters
2. poor manuals
3. undocumented features (bugs).

The last problem will appear the day before you want to submit your camera ready copy and will never have happened to anyone else until you solve the problem, at which point everyone else will say it was obvious.

4 Authors’ problems

The authors’ biggest headache is, of course... the editors. All authors have a horror story of how the editor changed the meaning of a piece of text, left out a graphic, inserted something they don’t agree with or refused to publish a table of data crucial to the argument. My favourite is where the editors moved the data lines from one graph and combined them with the axes from another. Other cases include the editor changing the grand totals in a table by adding the phase 1 totals to the correct totals, or helpfully correcting the spelling of the name of a site, but in fact changing the name to that of an entirely different site. In defence, I would say that editor/typesetters are often under great pressure to complete the task quickly, and have to cope with the problems outlined above. One way to help prevent these problems is to follow Lockyear’s Golden Rules.

5 Lockyear’s Golden Rules

If you want a high quality publication, quickly produced, *take responsibility for your own work.*

Firstly:

- follow the instructions; if they are unclear ask for clarification;

- do not waste time typesetting the text yourself – it usually creates more work for the typesetter who has to undo your work first.

Then, submit:

- well input text using the full facilities of your word-processor;
- high quality, i.e. well drawn, graphics;
- complete bibliographies;
- files in a variety of formats. For the text submit the file from your word-processor and an ASCII text. For graphics, submit several popular graphics formats such as POSTSCRIPT, or DXF format, as well as good quality hard-copy so that the typesetter knows what *you* expect. If at all possible avoid bitmap formats as they often scale badly.

The logic is simple. If each paper in a 40 paper volume requires 5 hours work for the editor, this results in 200 hours work, roughly 5–6 weeks full-time. If each author takes more responsibility for their paper and reduces the editor's input to 1–2 hours work, this results in 40–80 hours work, or 1–2.5 weeks and given that most editor/typesetters have a real job as well, this is a significant time saving. The size of the figures quoted may seem surprisingly large but in reality they are quite conservative.

One final comment on style: try to use headings and subheadings sparingly. There is a trend to use endless headings — sometimes for every paragraph — and this is unnecessary: it breaks the text up too much making it more difficult to follow, it is very ugly and it uses excessive paper.

6 Typographic basics – some hints for the novice typesetter

This section contains a few basic guidelines for novice typesetters. For more detailed advice one should consult the large number of books on this subject; for example two old, but excellent texts are Oliver (1945) and Williamson (1956); more up-to-date books including advice on DTP methods include Felici and Nace (1987) and Sassoon (1993). *The comprehensive manual on style*, which also includes advice on all aspects of publication, is the *Chicago Manual of Style* (University of Chicago 1993). Zapf and Dreyfus (1991) contains some interesting essays on classical typography in a modern environment. (Hermann Zapf is a well-known typographer and font designer — for example *Zapf Chancery Italic* is one of his fonts.)

6.1 FONTS

'Typefaces can do for words, and through words for ideas and information, what clothes can do for people; they can attract or repel, enhance or detract, emphasise or neutralise;

making a piece of text memorable or forgettable.'

Grosvenor *et al.* 1992.

Two choices need to be made: the font(s) to be used and the size of the font.

As regards size, it has been clearly shown that the human eye can only follow lines with a limited number of characters per line; if the line is too long it is difficult for the eye to back-track to the start of the next line and the reader tends to either miss lines or to re-read the line twice. The usual rule of thumb for English is 66 characters, or about twelve words per line. Therefore, the best font size⁵ for the main body of the text depends on the size of the paper being used. A4 paper, the most common size in the UK due to its ubiquitous use with laser printers and photocopiers, is less than ideal. Either the text has to be in a large point size (at least 12 point) and possibly with large margins (e.g., the journal *Archaeologia*) or the text has to be in a two column format (e.g., *Journal of Roman Pottery Studies*). Fitzpatrick and Morris (1994) has been typeset in a relatively small point size in a single column in an A4 format with up to 97 characters a line, resulting in an almost unreadable text. The first few CAA volumes published with BAR had large margins to keep the line length down but this resulted in the expensive two-volume proceedings for 1988 (Rahtz 1988). From 1990 the proceedings have been in a two-column format in 10 point despite the attendant increase in typesetting problems. The frequency with which this basic and fundamental rule is broken is partly due to software; for example, the default text size in the Word for Windows 'normal' template is 10 point — far too small for A4 paper in a single column. I would strongly advise anyone undertaking the typesetting of a book to investigate the possibilities of *not* using A4 — just because your desk-top printer uses A4 does not mean you have to fill the page, or that your printing company will have to use that size. There are many other smaller standard paper sizes which are more convenient for typesetting and reading.

The choice of font is also extremely important. Originally, the limited range of fonts packaged with laser printers resulted in non-professional typeset publications using a small range of fonts. The advent of TrueType fonts for use with Windows applications has increased the number of easily available fonts for the casual user; these can be purchased in most software outlets, usually on a CD-ROM. Many of these fonts are, however, fancy display fonts, suitable for posters and fliers but not for body text in book length publications. There are also many very poor quality TrueType fonts on the market. Most professional typesetting houses use POSTSCRIPT fonts which are more expensive. Chivers (1994) discusses some aspects of

Courier — fixed spaced font	Hello World!
Helvetica — sans serif font	Hello World!
Times-Roman — common font, quite dark	Hello World!
<i>Times-Roman Italic</i>	<i>Hello World!</i>
Palatino — lighter than Times-Roman	Hello World!
<i>Palatino Italic</i>	<i>Hello World!</i>
New Century Schoolbook	Hello World!
<i>Zapf Chancery Italic</i> — beautiful Italic font	<i>Hello World!</i>
Avant Garde Gothic Book — avoid	Hello World!

Figure 2. Some common fonts.

TrueType fonts in relation to POSTSCRIPT. If funds are available, I recommend investigating fonts other than the standard set supplied with laser printers. See Lawson (1990) for a discussion of typefaces.

For body text a Roman font with serifs, such as Times-Roman, Palatino, Bembo, Lucida Bright or Garamond, is preferable. The serifs, the little feet or bars on the top and bottom of the verticals of the letters, reinforce the horizontal line and help the eye to follow it. These fonts have other characteristics of their own. Garamond is quite thin in places and should only be used if the final production is to be on a high-resolution device (300dpi is *not* high resolution!). Times-Roman is a quite compact font compared to, for example, Palatino. A document in Times-Roman will be physically smaller than one in Palatino, but will also be ‘darker.’ Italic fonts, used for emphasis, are not just slanted versions of the Roman font, but are a separate font designed to complement it. Sans-serif fonts such as Helvetica or Arial, are unattractive for body text and are often difficult to read when used in large blocks. They can be used to good effect in headings, helping to differentiate them from the main text (e.g., Beck/Shennan 1991). Again, if funds are available, purchase a sans-serif font designed to blend well with the main font. Helvetica, which is supplied with most POSTSCRIPT laser printers, does not blend well with the Roman fonts supplied and mixing them in the text is ugly (Duggan 1994). Figure 2 presents a few common fonts and illustrates their characteristics.

One aspect of fonts which is not often appreciated is that, with good quality fonts, a letter in 12 point is not simply a scaled version of the 5 point character or vice versa; SMALL CAPITALS fonts are not a simple mixture of point sizes and so on.⁶ Few computer-fonts attempt to implement these design characteristics although they do make a large difference to the appearance and legibility of text. The font creation program METAFONT takes a base font design, and then a series of parameters, to produce bitmapped fonts for

each point size. To design a complete font from scratch is not an easy task and Computer Modern Roman remains the most common MetaFont font. Adobe’s Multiple Master fonts also implements these design aspects, but using these fonts with many systems is, again, not a trivial task (Goossens/Rahtz 1995).⁷

A final problem is special characters. Few fonts have the wide variety of characters and symbols required to typeset mathematics such as often appears in the proceedings of CAA. Often the solution is to mix fonts but this is usually ugly. There are some fonts which contain a wide variety of maths symbols — Computer Modern Roman (mainly for use with [La]TeX) and Lucida Bright for example. Foreign accents are also a problem. Most systems, especially Windows applications using TrueType fonts, do not easily allow for many accented characters such as the Romanian character ‘ă’, although it is sometimes possible to create the required character by combining a floating accent with a letter. It is an anglophone conceit that accents are relatively unimportant.⁸ Part of the problem is that even the extended ASCII character set is simply not big enough to contain all characters. The advent of Unicode may help the problem (Fairbairns 1995). Meanwhile, the typesetter can either use systems which will create the accented characters on the fly such as WordPerfect or (La)TeX; or has to buy special fonts that contain the accented letters. Programs also exist which allow one to create accented or extra characters and encode them into a font but this, yet again, is not a trivial task.

6.2 WHITE SPACE

As important as the black ink on a page, is the white space. White space serves both functional and aesthetic purposes but I shall concentrate on the former. Within the body text, white space is obviously used to separate words — *thisdoesnotmakemuchsensedoesit?* It also helps to differentiate sentences; in English typography an inter-sentence space is larger than an inter-word space. When I was at school I remember the chant ‘full-stop space shift...’ echoing out of the business studies classroom. In the conflation of the role of typist and author, this rule has been generally lost. In a WYSIWYG system such as a word-processor, the inter-sentence space has to be inserted manually and the typesetter has to insert the extra space if the author has omitted it. Systems such as (La)TeX interpret the sequence ‘full-stop space(s) capital-letter’ as the command for an inter-sentence space. Spacing is also language specific. It is usual practice in English typography to have no space between a letter and an exclamation mark (e.g., hello!) but in French typography a space is inserted (e.g., Mon Dieu !). Lastly, a recent trend has been to omit full-stops and spaces in initials, e.g., SPQ Rahtz. With

recognised acronyms, especially those pronounced as one word (e.g. CAD), this is acceptable. With personal names, mainly in bibliographies, I prefer to stick to the old-fashioned but elegant form: S. P. Q. Rahtz.

White space around headings also serves a functional purpose. The amount of white space is style specific but it is usual to have more white space above a heading than below it. This reinforces the fact that the heading belongs to the following text. A problem occurs in many systems such as Word for Windows or PageMaker when a subheading follows a heading immediately, resulting in too much white space, which is ugly. In the case of those two packages the typesetter either has to define a style ‘subhead after heading’, or manually adjust the spacing.

Sufficient white space should also be left to differentiate tables, figures, headers and footers from the main text, and to separate multiple columns. The last is known as the *gutter space*. The space can be reduced if a thin line or *rule* is used to enforce the separation.

6.3 TYPOGRAPHIC SYMBOLS

Use of correct symbols also helps improve the quality, and the legibility of text. Modern computer keyboards have three keys for quotation marks: the single opening quote (top-left of the main key pad on an English keyboard), a single closing quote (the usual quote key used) and the double-quote (shift-2 on an English keyboard). The use of the double-quote results in a typewriter-style double quotation mark —". Avoid it as it is ugly. Use the correct opening and closing quotation marks, either single or double, from the base font, even if they have to be accessed in a special manner. In Word for Windows 2 they are not automatically used and have to be input using ALT-145, 146, 147 and 148 or the ‘insert symbol’ menu. Word for Windows 6 tries to use the correct quotation marks automatically, even if one only uses the closing or double-quote key. It usually succeeds but can be fooled. WordPerfect 5.1 will use the correct single quotes but not the correct double quotes. Do not confuse the single and double quotation marks with the symbols for feet and inches, or minutes and seconds which are somewhat different, i.e. 5' 7".

Use the correct types of dash:

- hyphens in words, e.g., co-ordinate
- en-dash in ranges e.g., AD 43–410
- em-dashes as phrase markers — like this

These dashes occur in most fonts. In WordPerfect 5.1 use control-V n- for an en-dash and control-V m- for an em-dash. In Windows systems use the ‘insert symbol’ menus. These dashes are important because they help to differentiate between different meanings. For example, if an

ordinary small dash ends a line, is it a hyphen, or a phrase marker, or a range? One should also note that the mathematical minus symbol is also a separate character.

Ligatures improve the appearance and readability of a text. Few systems allow for ligatures and many professional publishing houses do not now use them which is a pity. In words such as ‘official’ the fs and the i, especially the dot of the i, usually clash and look ugly. Ligatures are characters where the fs and the i are joined together, the i usually using the bulb on the descender of the f as its dot. Ligatures include the sequences fl, ff, fi and ffi. LaTeX will use any of these letter sequences as the command to insert the correct ligature. The German double-s character ‘ß’ is, in fact, a ligature of a sharp s⁹ and a regular s.

6.4 TABLES

Tables are a matter of design and personal preference. However, many non-professional typesetters over-use rules (i.e. lines) in tables. For example, the default table style in WordPerfect 5.1, as used in CAA91 (Lock/Moffett 1992), is extremely heavy and ugly and should be avoided at all costs. Originally, I used vertical rules in tables (e.g., Beck/Shennan 1991; Lockyear/Rahtz 1991) but quickly decided these were extremely ugly. I now avoid vertical rules if at all possible using a style derived from the Cambridge University Press (e.g., Duncan 1993). The columns in a table can be easily separated using white space, and groups of columns using varying sizes of white space. Kroonenberg (1994) and Allan Reese (1994) discuss various aspects of table design.

6.5 UNDERLINING

Don’t. Use *italics* for emphasis in Roman fonts like this, *or use Roman in pieces of italic text like this*. For headings, use combinations of font size, bold face and italics.

6.6 CONSISTENCY

Above all, whatever design choices are made, use them consistently. For example, use of AD, A.D., AD or A.D.¹⁰ is up to the designer, but they should be the same throughout the document.

7 Types of system

There are four main types of system for typesetting and all have their advantages and disadvantages. These systems are:

- generic markup (e.g., LaTeX)
- advanced use of word-processors (e.g., WordPerfect 5.1)
- desk-top publication systems (e.g., PageMaker)
- professional typesetting systems (XYvision, Miles 33)

I have *not* conducted a comprehensive survey of different packages and have no experience of the professional systems or Quark Xpress, Ventura Publisher, WordPerfect for Windows, troff, etc. I have, however, used one of each of the first three types of system in the production of book-length publications: Lockyear and Rahtz (1991) and Beck and Shennan (1991) were typeset using LaTeX; Duncan (1993), and others using WordPerfect 5.1; *Diatribes* volume 4 and Wilcock and Lockyear (1995) using PageMaker 5. Each has advantages and disadvantages. The output of these systems can be examined easily as the proceedings of CAA since 1987 have been created by all three types of system: CAA87–90 were typeset using LaTeX; CAA91 in WordPerfect; CAA92 and 93 in PageMaker 5 and CAA94 (Hugget/Ryan 1995) in Word for Windows 6. This volume of CAA will be the first to be professionally typeset.

PageMaker and LaTeX have vastly superior hyphenation, kerning¹¹ and justification routines compared to, e.g., Word for Windows 6. If one compares the text from CAA94 with CAA93 one can quickly see PageMaker's superiority. Conversely, editing text in PageMaker is painful and the process of adjusting the text after editing is quite slow. A good combination, if one wishes to follow the Windows/WYSIWYG route with all its attendant advantages and disadvantages, is to edit files in a word-processor such as Word for Windows, then complete the typesetting in a program like PageMaker. Unfortunately, PageMaker has some severe problems, the worst of which is its table editor. The table editor only allows one font in one size in a cell of a table and thus constructions such as σ^2 are impossible. It is also extremely memory-hungry and 8MB of RAM is barely sufficient and occasionally causes crashes. These were less frequent when a new hard disk allowed for a larger virtual memory file. Also, the program needs to know exactly which POSTSCRIPT printer you will be using, and changing printers can create some unpredictable results.

WordPerfect 5.1 as a word-processor has some excellent advantages. It is one of the few packages to be able to create many accents easily including a full range of Greek characters with breathings and accents. It uses a sensible default font size. Its 'reveal codes' facility, which enables one to explicitly examine the control codes in the text and thus edit them, is invaluable when 'typesetting' texts.¹² There are other features I have not found easy to copy in other word-processing packages. Many, however, hate the package because its opening screen, which is almost empty, is unfriendly; its use of function keys, possibly with shift, alt or control, seems difficult to learn; tables have to have a line spacing of 1.5 and the default rules are *extremely* ugly; the style-sheet system is not obvious and rather counter-intuitive. Many of these issues are solved by using WordPerfect for Windows 6.

Most modern word-processors use some system of style sheets. Style sheets define the visual characteristics of sections of text with a specific function. For example, a heading can be defined as being in 12 point bold Helvetica with a 1.5 line space before the heading and a 0.5 line space after the heading. Each heading is then tagged as such, and will inherit those visual characteristics. Changing the definition of the style, or using a different style sheet, will result in the visual format of all headings being altered without the user having to manually change the settings for every heading. These systems are not perfect but are improving rapidly. For example, it is usual in a numbered list for there to be some extra space above and below the list. To achieve this in many systems requires the definition of three styles: first numbered item, numbered item and last numbered item. In a complex document this can lead to a proliferation of styles. Authors are, however, strongly advised to make full use of the style sheet system of their word-processor.

LaTeX does not meet with much enthusiasm from many people. This is because it has an old-fashioned interface: processing a LaTeX document is analogous to compiling a program rather than word-processing, and obviously the programming language has to be learned. The text input to LaTeX is in ASCII but they are full of commands such as:

```
\documentstyle[11pt,a4]{article}
```

which makes reading the text on-screen difficult. Figure and table placement in LaTeX is automatic which should be an advantage, but in a paper with large numbers of figures and tables in relation to the quantity of text, the results are poor and have to be manually adjusted. Installation of the system can be difficult with large numbers of font files, style files and so on. It is also a 'guru'-based system reflecting its origins on UNIX work-stations. The advice 'consult your local TeX guru' is not useful when working on a DOS machine from home, for example. However, many of these problems have solutions. The new LaTeX, LaTeX2 ϵ , addresses and solves many programming problems. Use of the Eddi4TeX editor on DOS machines¹³ helps reduce typing, and on-screen colour-coding of parts of the screen helps with writing. The GnuEmacs editor from the Free Software Foundation has excellent capabilities for editing TeX files and is available in Unix (and X-Windows) and DOS (and Windows) formats. Installation of TeX for PCs is automatic if one uses the 4AllTeX workbench, available on CD-ROM from the Dutch or UK TeX Users Groups, or less conveniently by down-loading the package from the *Comprehensive TeX Archive Network*.¹⁴ The various TeX User Groups and electronic discussion lists (such as TeXhax) can give you access to a world of TeX gurus, even if it does take time (most gurus seem to take TeXhax in digest form).

Despite its age, LaTeX remains an extremely powerful typesetting system. The best recommendation I can give it is that having started using LaTeX2.09 in 1988, I then tried various word-processors and typesetting systems, but returned to LaTeX (although I used the new LaTeX2 ϵ) to write Lockyear (1996). The main reasons for this choice were:

- it is free
- it uses the BibTeX bibliography package (also free)
- document portability (i.e. the ASCII text files can be processed by LaTeX on PCs, UNIX machines, VAXs, Macs and so on).
- there is easy access to all accents¹⁵
- input text as ASCII files (the whole of Lockyear 1996 fits on one floppy disk and thus it is portable in more than one sense)
- easy generation of typeset tables from a database
- poorly structured POSTSCRIPT graphics (generated from CANODRAW, or written by myself) can be included without problems

Other advantages (although not all are unique to LaTeX) are:

- consistency of style enforced by style files
- excellent hyphenation, kerning and line spacing/justification routines
- cross references automatically created
- extremely powerful mathematics typesetting capability
- conversion between HTML¹⁶ and LaTeX is fairly easy with latex2html and html2latex
- conversion to SGML is not too difficult (see below)
- it is possible to create PDF files with rich mark-up (see below).

In the hard sciences TeX is used widely. Compiling a volume of papers ought to be relatively easy. In archaeology, where most contributors use a word-processor, conversion to LaTeX is dull and time-consuming. Writing a large book (such as a thesis) from scratch is less painful and many LaTeX features make life relatively easy and I would recommend it highly. Conversely, few archaeological journals, or editors of proceedings, are able to process LaTeX documents, although the ACN is an exception. For papers to be submitted on disk, I either have to convert the paper to another format, or use a word-processor.

8 'Electronic publication' and archaeology

To conclude this paper I would like to make a few comments about 'electronic publication', especially publication over the World Wide Web (WWW). For other interesting comments see the recent debate in ACN

(Hollidge 1995; Kilbride 1995; Rahtz 1994, 1995). Despite the current burst of interest in the topic, and the immanent formation of electronic journals (see Heyworth/Ross/Richards, this volume), the concept has been around for some time (e.g., Rahtz 1986). However, the rapid expansion of the Net, and the recommendations of the Follett report (JFCLRG 1993), seem destined to make this style of publication a reality. Some predict that electronic publication will cause the death of traditional academic journals (Odlyzko 1995).

Current electronic journals take a variety of formats. For example, *Psychology* publishes as simple ASCII files; others maintain articles as POSTSCRIPT or TeX files (e.g., *The Journal of Artificial Intelligence Research*) and more recently journals are being published as HTML WWW publications, or using *Portable Document Files* (PDF). The latter is an extension of POSTSCRIPT with 'rich markup' which allows the document to contain hypertext style links and buttons. PDF files are read with Adobe Acrobat which is available free. The current debate around electronic journals tends to contrast *paper* publication with *electronic* publication. I feel this is an artificial and unhelpful division as each medium has inherent advantages. Firstly, it is an oft-quoted truism that most people cannot read large quantities of text off a computer screen. This is most clearly seen when people edit an article they are writing — most still print it out, mark corrections and then edit the file. Multimedia teaching packages also prove the point — if the screen contains too much text the student scrolls by until the next link or graphic is reached. Any electronic journal either has to:

1. create a new style of writing in word-bites
2. provide the journal in a form that allows large pieces of text to be printed out in a high quality format

The first option seems unlikely for academic work. If the second route is taken, the points made about typography in section 6 must be taken into account. Unfortunately, HTML is, as yet, not capable of producing well typeset text.

Conversely, paper publication is not feasible for large collections of plans, photographs, data etc., much of which ends up as an archive or 'grey literature' (i.e. limited production, privately printed material). Microfiche, silver-jacketed or not, is a data cemetery, not a data archive. As Rahtz (1986) notes, electronic publication is ideal for this sort of material. I recently conducted a straw-poll of non-computing archaeologists on the value of electronic publication which revealed that the greatest enthusiasm for electronic publication was the possibility of access to data archives that lie behind traditional publications. The much-vaunted non-linear nature of multimedia publications was

met with apathy, although hypertext-style links to other documents and data resources was seen as an advantage. The technology exists to produce electronic publications which include such links, high-quality typesetting and printing possibilities and thus to combine the advantages of both styles of publication. The largest hurdle, however, is the fact that most people *like* books and as stated above they are regarded as objects of value and beauty; many still prefer a well bound volume to stapled laser-print. This psychological hurdle is one which will have to be met by electronic publications, and is not a problem that I feel will go away quickly. After all, it is still more convenient to read a bound book on the train rather than a pile of loose sheets of paper.

Another problem that these publications will have to meet is that they are now conflating a further two roles: that of publisher and that of library; they need to consider problems of the long-term archive. A recent debate on the mailing list ARCH-L centred around the viability of CD-ROM for archaeological publication, and concerns were raised about its long-term potential. This debate missed the point. Provided that the publisher/library is willing to move its publications from one machine/medium to the next at every upgrade, the problem is not the *medium* for data storage, but the *format*. For text, this must mean the use of *Standard Generalised Markup Language* (SGML; ISO 1986; Van Herwijnen 1994). It is to the shame of computing in archaeology, at least in the UK, that we have yet to develop an SGML document type description (DTD) for archaeology, specifically for excavation reports. In Norway, large quantities of text based information including that from archaeology, is being marked-up in SGML and the resultant files parsed to create relational databases of the document's contents (Holmen/Uleberg, this volume). The potential for searching documents is immense. The time has come when we should seriously examine the potential for the use of SGML in archaeology.¹⁷

9 Conclusion

This paper has quickly reviewed the past, present and future of computer aided and electronic publication in archaeology. It has shown that computers have conflated a number of roles: typist and author, editor and typesetter and now publisher and library. At each stage, past problems have been solved but new problems have arisen and some suggestions and advice have been offered. In particular, it was argued that centuries of development in typography, in terms of its aesthetic and functional role, should not be ignored, but on the contrary provides many valuable lessons for publication irrespective of the manner of that publication. The challenge for electronic publication is to learn the lessons of the past, and combine

them with the technologies of today, to provide us with a resource for the future.

Acknowledgements

I would like to thank Sebastian Rahtz, Dale Serjeantson and Tim Sly for their comments on this paper. Responsibility for the contents of this paper remains, of course, with the author.

notes

1 For example, in 1987 I was one of the first undergraduates in the archaeology department at Durham to word-process their dissertation (on a mainframe!). Most other students wrote their dissertations by hand and then had them typed, with consequent delays for corrections etc.

2 For example PageMaker 5. The LaTeX manuals (Lamport 1994; Goossens/Mittelbach/Samarin 1994) are an honourable exception.

3 I tested this theory during a class on typography, and this was precisely the result when they were asked to comment on this book.

4 It is often quicker to look them up in the library, or now, over the InterNet, than to ask the author.

5 Fonts for typesetting are measured in points: a point is 1/72nd of an inch. Some basic printer fonts are measured in cpi, *characters per inch*.

6 Goossens *et al.* (1994: Chapter 7) contains an excellent discussion about these matters.

7 I have been informed that Multiple Master fonts are easy to use with Macintosh computers.

8 In Romanian, the letter 'ă' is regarded as a separate letter, not just an accented 'a' and comes after the latter in the alphabet. This letter can change the meaning of a word, and thus the sentence. For example 'fată' is best translated as 'a girl' whereas 'fata' is 'the girl.'

9 A 'sharp s' is the older form of 's' that looks like an f with no cross-bar. It can be seen regularly in, for example, 18th century English texts such as the earliest volumes of *Archaeologia*.

10 Note that AD should always precede the date and BC should always follow it, i.e. Britain was invaded in AD 43; the battle of Actium was fought in 31 BC.

11 Kerning is the adjustment of space between letters. For example, the letter o can be placed closer to the letter K than to the letter H because of their shapes. If inter-letter spaces are of a fixed size they will appear variable, a kind of optical illusion.

12 Typesetting is here in quotation marks as purists would deny that a document could be typeset in a word-processor such as WordPerfect.

13 Contact Ulrich Jahnz, ulrich_jahnz@pe.maus.de, for details. A registered version of the Eddi4TeX editor is available to all UK TeX Users Group members.

14 FTP to ftp.tex.ac.uk in the UK, or to one of its mirror sites at ftp.dante.de or ftp.shsu.edu. The CTAN archives contain huge quantities of software and information relating to TeX and TeX related software including mailings to discussion lists etc. and complete TeX installations such as the emTeX package for MS-DOS machines. The archives run an enhanced version of FTP and have other useful site-defined commands. Down-load the files README.archive-features and README.archive-commands for an overview. For queries regarding the UKTeX Users Group contact them on uk-tug@tex.ac.uk.

15 I investigated the possibility of using WP5.1 with the Endnote bibliographic package. Although WP5.1 would create the accents I required, Endnote would not.

16 *Hypertext Markup Language*, used for documents that can be read over the World Wide Web using programs such as NetScape or Mosaic.

17 This argument is not new; Sebastian Rahtz has been arguing along these lines for many years. HTML can be seen as a SGML DTD, but it has a loose structure and recent developments, more concerned with document appearance rather than structure, are making it hard to justify HTML as a SGML DTD.

references

- Allan Reese, R. 1994 LaTeX and Tables, *Baskerville* 4(5), 2-6.
- Beck, C.W.
S.J. Shennan 1991 *Amber in Prehistoric Britain*. Oxford: Oxbow Monographs No. 8.
- Casey, P.J.
R. Reece 1974 *Coins and the Archaeologist*. BAR British Series 4. Oxford: British Archaeological Reports.
- Chivers, I.D. 1994 A tutorial on TrueType, *Baskerville* 4(3), 14-16.
- Clarke, D.V.
T.G. Cowie
A. Foxon 1985 *Symbols of Power at the time of Stonehenge*. Edinburgh: National Museum of Antiquities of Scotland.
- Duggan, A. 1994 Mixing and matching POSTSCRIPT fonts, *Baskerville* 4(1), 2-5.
- Duncan, G.L. 1993 *Coin circulation in the Danubian and Balkan Provinces of the Roman Empire AD 294-578*. London: Royal Numismatic Society Special Publication No. 26.
- Dixon, P. 1994 *Crickley Hill. Part I: The Hillfort Defences*. Nottingham: Crickley Hill Trust and Department of Archaeology, Nottingham University.
- Fairbairns, R. 1995 Report of visit to CERN to attend presentation of Ω , *Baskerville* 5(3), 14-17.
- Felici, J.
T. Nace 1987 *Desktop publishing skills: a primer for typesetting with computers and laser printers*. Reading, Massachusetts: Addison-Wesley.
- Fitzpatrick, A.P.
E.L. Morris 1994 *The Iron Age in Wessex: Recent Work*. Salisbury: Trust for Wessex Archaeology.
- Goossens, M.
F. Mittelbach
A. Samarin 1994 *The LaTeX Companion*. Reading, Massachusetts: Addison-Wesley.
- Goossens, M.
S.P.Q. Rahtz 1995 Using Adobe Type 1 Multiple Master fonts with TeX, *Baskerville* 5(3), 4-8.
- Grosvenor, J.
K. Morrison
A. Pim 1992 *The PostScript Font Handbook. A Directory of Type 1 Fonts*. Revised edition. Reading, Massachusetts: Addison-Wesley.
- Holledge, S. 1995 Ranter's Corner I: Sebastian Rahtz/World Wide What? (ACN40): A Confessed Cyberphile Replies, *Archaeological Computing Newsletter* 42, 1-2.
- Huggett, J.
N.S. Ryan (eds) 1995 *Computer Applications and Quantitative Methods in Archaeology 1994*. BAR International Series 600. Oxford: Tempus Reparatum.
- ISO 1986 *Standard Generalised Markup Language (SGML)*. Geneva: International Standards Organization Technical Report ISO 8879.
- JFCLRG 1993 *The Joint Funding Councils' Libraries Review Group: Report*. Bristol: Higher Education Funding Council for England.

- Kilbride, W. 1995 Ranter's Corner II: Mary Queen of Scots, Thomas Edison and the Problems with Electronic Publishing. *Archaeological Computing Newsletter* 42, 2-4.
- Knuth, D.E. 1984 *The TeXbook*. Reading, Massachusetts: Addison-Wesley.
- Kroonenberg, S. 1994 Table design, *Baskerville* 4(4), 2-4.
- Laflin, S. (ed.) 1986 *Computer Applications in Archaeology 1986*. Birmingham: University of Birmingham, Centre for Computing and Computer Science.
- Lamport, L. 1986 *LaTeX. A document preparation system*. Reading, Massachusetts: Addison-Wesley. First edition.
- 1994 *LaTeX. A document preparation system*. Reading, Massachusetts: Addison-Wesley. Second edition updated for LaTeX2ε.
- Lawson, A. 1990 *Anatomy of a Typeface*. London: Hamish Hamilton.
- Lock, G.
J. Moffett (eds) 1992 *Computer Applications and Quantitative Methods in Archaeology 1991*. BAR International Series 577. Oxford: Tempus Reparatum.
- Lockyear, K. 1994 Ranter's corner – PUPs, *Archaeological Computing Newsletter* 41, 1-2.
- 1996 Forthcoming. *Multivariate Money. A statistical analysis of Roman Republican coin hoards with special reference to material from Romania*. PhD thesis to be submitted to University College London.
- Lockyear, K.
S.P.Q. Rahtz (eds) 1991 *Computer Applications and Quantitative Methods in Archaeology 1990*. BAR International Series 565. Oxford: Tempus Reparatum.
- Odlyzko, A.M. 1995 Tragic loss or good riddance? The impending demise of traditional scholarly journals. In: R.P. Peek/G.B. Newby (eds), *Electronic Publishing Confronts Academia: the Agenda for the year 2000*. Michigan: MIT Press. Preprint available from the AT&T preprint server in POSTSCRIPT format.
- Oliver, S. 1945 *Introduction to Typography*. London: Faber and Faber.
- Rahtz, S.P.Q. 1986 Possible directions in electronic publication in archaeology. In: S. Laflin (ed.), *Computer Applications in Archaeology 1986*, 3-13, Birmingham: University of Birmingham, Centre for Computing and Computer Science.
- 1988 *Computer and Quantitative Methods in Archaeology 1988*. BAR International Series 446. Oxford: British Archaeological Reports.
- 1994 Ranters Corner – World Wide What?, *Archaeological Computing Newsletter* 40, 1-2.
- 1995 Ranter Redux, *Archaeological Computing Newsletter* 43, 1.
- Renfrew, A.C.
S.J. Shennan 1982 *Ranking, Resource and Exchange*. Cambridge: Cambridge University Press.
- Sassoon, R. (ed.) 1993 *Computers and Typography*. Oxford: Intellect Books.
- Tukey, J.W. 1977 *Exploratory Data Analysis*. Reading, Massachusetts: Addison-Wesley.
- University of Chicago 1993 *Chicago Manual of Style*. Chicago: University of Chicago. Fourteenth edition.

- van Herwijnen, E. 1994 *Practical SGML*. Boston: Wolters-Kluwer. Second edition.
- Wheeler, R.E.M.
K.M. Richardson 1957 *Hill-forts of Northern France*. London: Society of Antiquaries Research Reports No. 19.
- Wilcock, J.D.
K. Lockyear (eds) 1995 *Computer Applications and Quantitative Methods in Archaeology 1993*. BAR International Series 598. Oxford: Tempus Reparatum.
- Wilcock, J.D.
R.D. Spicer 1986 Standardisation of word processing for publication: the application of ASPIC in common word processing packages. In: S. Laflin (ed.), *Computer Applications in Archaeology 1986*, 21-28, Birmingham: University of Birmingham, Centre for Computing and Computer Science.
- Williamson, H. 1956 *Methods of Book Design*. London: Oxford University Press.
- Zapf, H.
J. Dreyfus (eds) 1991 *Classical Typography in the Computer Age*. Los Angeles: University of California.

Kris Lockyear
88, Kent Road
St. Denys
Southampton SO17 2LH
United Kingdom
e-mail: K.Lockyear@soton.ac.uk