Published in: *Encyclopedia of Library and Information Science*, vol. 49 (New York: Dekker, 1992), pp. 268-78

Author's address: daniel.eisenberg@bigfoot.com

WORD PROCESSING (HISTORY OF).

Definition. The term and concept of "word processing" are by now so widely used that most readers will be already familiar with them. The term, created on the model of "data processing," is more vague than commonly believed. A human editor, for example, obviously processes words, but is not what is meant by a "word processor." A number of software programs process words in one way or another--a concordance or indexing program, for example--but are not understood to be word processing programs.'

The term "word processor" means a facility that records keystrokes from a typewriter-like keyboard, and prints the output onto paper in a separate operation. In the meantime the data is stored, usually in memory or magnetic media. A word processor also can make improvements in the stream of words before they are printed. At their most basic these include the ability to arrange words into lines. An "editor," such as the infamous EDLINE distributed with the Microsoft Disk Operating System (MS-DOS), lacks the ability to structure lines.

Commonly a word processor is understood to be a software program, and in the 1980s and 1990s it usually meant a program written for a microcomputer. However, preceding this period and continuing through it there have been hardware word processors. These are pieces of equipment sold for the sole purpose of word processing, containing in one package a keyboard, printer, recording and playback device, and in all recent examples a video or liquid crystal display screen. These machines include word processing software permanently stored in memory and processor chips, or in the earliest examples its equivalent in relays and wiring. A great disadvantage is that their storage media are usually incompatible with those of other machines, meaning words cannot be transferred from one system to another.'

Predecessors. The fundamental use of a word processor is to record, edit, and play back keystrokes. The printing press, whose storage medium was lead type that had been "input" by the human or mechanical typesetter, itself carried out this function. While the press itself need not be especially bulky or expensive, the cost, weight, and complexity of the equipment and procedures needed to prepare the type for the press were great. Thus printing has rarely been found in offices, and has been relegated to a separate business, the printing shop.'

Experiments in recording typewriter keystrokes, early in the twentieth century, used mechanical, air-driven mechanisms, borrowed from or inspired by the player piano. These were scarcely more than curiosities. The electric typewriter appeared in the 1930s, but despite its higher-quality output, it did not become dominant until twenty years later. It seemed to offer the potential for more economical recording and reproduction of keystrokes, but it was not until the 1950s that a machine carrying out these functions was manufactured and sold in quantity: the Friden FlexoWriter. The FlexoWriter used punched paper tape as a storage medium, thus providing some compatibility with other data processing equipment.'

Keystrokes could be deleted from the punched tape, which backed up as the backstroke key was pressed. The "delete" or non-printing character consisted of all possible holes punched; this was later implemented in the ASCII character set, in which the highest character, 127, with all bits set, was a "delete" character. However, insertions could only be made by splicing the paper tape, and there was no means to turn carriage returns into spaces or vice-versa. The difficulty of revision made the noisy FlexoWriter impractical for routine office correspondence. Its main use was for what soon became known as the "form letter," in today's terminology a merge operation: a standard text with different names, addresses, or other data inserted. These were used for political purposes, fundraising, and sales.'

The MT/ST. The first true word processor, and IBM's first entry into the field, was the Magnetic Tape Selectric Typewriter or MT/ST, introduced about 1969. Much larger and more expensive than the FlexoWriter, the MT/ST included a model of IBM's workhorse Selectric typewriter built into a small desk. Accompanying it was a console, about half as large as the desk, containing control relays, the tape recording and playback facility, and control buttons and dials.'

As typing took place, keystrokes were recorded on 16mm magnetic tape, using a head moved by solenoid and spring across the tape, one character per pass. The backspace key on the keyboard backed the tape up so the character could be overwritten with a new one. 100' cassettes of tape held approximately 25K of data. A 7-bit data set was used, providing different codes for upper and lower case letters, and an increased set of control codes. In a technique borrowed from mainframe computers, the operator could mark the tape with electronic divider codes, called "Search Codes." While the concept of named files was still in the future, the machine counted the number of Search Codes on the tape and could thus search for a file. In a crude way it could thus assemble standard chunks of text into a composite output document.'

The MT/ST incorporated a number of other innovations. While it had no block movement capabilities, except by very cumbersome block export and import, it did have insertion and revision capabilities. At each carriage return, one space on the tape was reserved for insertions. More important, all except the very first models incorporated two tape stations, and the ability to play from one tape and while recording an updated tape on the other, incorporating insertions and deletions as wished. The two tape stations could also be played back simultaneously, typically merging a letter on one tape with a name and address list on the other.'

The MT/ST also introduced what may be called word processing concepts. Within a fixed zone (1") preceding the right margin, spaces were converted into carriage returns, and outside it returns were changed into spaces. The concept of a "soft hyphen" was used for the first time: a hyphen that would print only if it fell within the 1" hyphenation zone, and would otherwise be ignored. An overstrike code was also implemented; this permitted underscoring.'

Whereas the FlexoWriter only operated on a character basis, the MT/ST also operated on a word, line, and paragraph basis. The machine understood a word to be any string terminating with a space; a line was a string terminating with a carriage return. Two carriage returns constituted a paragraph boundary, and were not turned into spaces as single returns could be.'

However, the MT/ST had no concept of a page; the division of text into pages and the numbering of pages had to be done by a human operator. While possible hyphenation points in a word could be pre-marked, it had no hyphenation ability of its own. The machine did not permit operator hyphenation while playback was in process, and thus line endings were not always satisfactory. As it had no display screen, the only way to know what was on a tape was to print it out at 15 characters per second. No tape updating without printing was possible either. The operator needed to listen to the tape transport to know whether it was skipping text or awaiting keyboard input. The MT/ST required alertness and discipline from the operator.'

Other Dedicated Word Processors. IBM introduced a number of later word processors incorporating various improvements; other manufacturers followed IBM's lead. None of these achieved the success of the MT/ST. The Magnetic Card Selectric Typewriter (1969) incorporated the concept of the page into processing; one page was stored on each magnetic card. In 1971 the Lexitron (not an IBM product) added a display screen, using paper only for the final output. IBM's Magnetic Card Executive Typewriter (1972) had proportional spacing in a single type

face, sacrificing the interchangeable type elements of the Selectric. Magnetic Card Selectric Typewriter II (1973) added memory, allowing text to be stored and moved in memory and only stored on media when editing was completed. It also had the double pitch available on later Selectrics.'

IBM's DisplayWriter and word processors from Wang and Lanier moved closer to the word processor as we know it today. Floppy disks gradually replaced cassette tapes as storage media. Named files were introduced. On-screen menus of functions and labeled function keys made it easier to perform tasks: opening, saving, combining, printing, and deleting files; moving, centering, justifying, and underlining text; combining standard text (a letter or a financial document) with varying data; adding page numbers, headers and footers.'

Word Processing on Multipurpose Computers. Although not customarily thought of as such, the above pieces of hardware are computers, with input and output devices and data storage and manipulation ability. However, their programming was limited to that supplied by the manufacturer, and they could not be used for other tasks. Word processing has been done on both mainframe and minicomputers, in both of which the computer simultaneously served several users, but these never achieved acceptance outside of large businesses. Besides the cost of the computer, the printers were the primary limitation. Only daisy-wheel printers could produce office-quality output in the early 1980s, and they were slow in comparison with the processor speeds. Software for mainframe and minicomputer word processing soon trailed, in features and ease of use, that available on microcomputers. Mainframe administrators did not consider word processing a task particularly appropriate for their computers, as it required large blocks of memory but did not need mainframe processing speed.'

Early microcomputers. The microcomputer arrived on the scene in the mid-1970s. Early models used tape cassettes for data storage, with data encoded in an unreliable audio (analog) signal. There were two main families of central processing chips, the 8080 and the 6502 chip, which were later to evolve into the IBM-compatible and the Macintosh lines of computers respectively. The commands used by the chips were the same for all machines using that chip. However, the formats by which data was stored and displayed on a screen were not standardized, meaning that each model of computer required a different version of the same program, and computers could not share data. Early dot-matrix printers, typically using a 7-pin matrix, produced a legible but poor-looking product, unsuitable for business correspondence.

Very primitive software was hobbyist-written, or provided by hardware manufacturers.'

The addition of floppy disk drives to computers in the late 1970s marked a great advance in the reliability and speed of data storage. It then became practical for independent programmers to write and sell software. The only word processing program from this period to achieve recognition was Michael Shrayer's Electric Pencil, which introduced word wrap, in which it was not necessary to press a carriage return at the end of each line. This was a great innovation for its day.'

The arrival in the late 1970s of the standard operating system CP/M (Control Program/Micro) brought vastly increased standardization. Formats for file and memory structure were standard across all machines of that class. By 1980 IBM Selectrics and Diablo and Qume daisy-wheel printers were coupled to computers, giving a good-quality printed output and a variety of type styles, although at a high price. Changing type style meant stopping the printer and a manual change of printing element (in the case of the daisy-wheel printers this was a disk about three inches in diameter). The addition of memory to printers made bi-directional printing possible. This eliminated the delay caused by the return of the print head to the left margin position at the beginning of every line.'

While the logical disk structure was standard, the physical disk organization was unique to each computer. The standardized 8" disks were replaced with less standard 5.25" disks. The screen display codes were also not standardized. This meant that slightly different versions of a program were still required for each brand, and in some cases each model, of computer. It also meant that different brands of computers could not exchange disks, although conversion software and facilities appeared in the mid-1980s.

WordStar. Throughout the CP/M period the predominant word processing program was WordStar, although there were other minor competitors, such as Select and Perfect Writer. By recent standards CP/M WordStar was very primitive, lacking features now taken for granted, such as a spelling checker, footnotes, and proportional spacing. On-screen help was limited, and installation often required a technically knowledgeable operator. It could only support one printer at a time; multiple printers required that different versions of the program be prepared. Nevertheless, WordStar took full advantage of the features of the printers of the early 1980s. It supported super- and subscript, bold or shadow print, underscore, page numbers, and one-line headers and footers. It achieved a great success with its keyboard structure, in which cursor movement commands were arranged in a logical, one-hand design (Control-E for up, Control-S for left, Control-D for

right, and Control-X for down). The MicroPro Corporation, publisher of WordStar, successfully sold its product to hardware manufacturers, such as KayPro and Morrow. It was thus "bundled" with the computer: by purchasing that brand of computer one got WordStar included, at a price much less than an individual could have obtained it. "WordStar To Go," an abbreviated version, was stored on memory chips and sold with some of the first laptops.'

MS-DOS Word Processing. In the early 1980s the CP/M computers were replaced by more powerful machines using the Microsoft Disk Operating System (MS-DOS). This system was chosen by IBM for its microcomputers, and although IBM was never as successful with microcomputers as it wanted to be, it served to set a common technical core for a large family of "IBM-compatible" computers. Among other things, IBM's floppy disk format became standardized over a large number of brands of computers, permitting exchange of data with an ease that CP/M computers, with their manufacturer-specific formats, never achieved. A software producer needed only to prepare a single version of a program, which would then run on the machines of many manufacturers. The MS-DOS computers also vastly increased available memory, added subdirectories to their disk structures, added function keys and an "Alt" key to the keyboard, and were easily customizable without sacrificing compatibility.'

This new standardization, plus the added speed of the machines, caused a great hardware and software boom in the 1980s. The microcomputer software industry grew at a great pace; customers purchased the newly standardized machines for which much software was available. Prices of hardware continually dropped. Software gained dramatically in features and ease of use, while also declining in price.'

General characteristics of MS-DOS word processing programs. MS-DOS word processing programs offered more of everything. The increased memory and disk storage capability meant that more extensive on-screen help could be made available. Dot-matrix printers improved the quality of their letter shapes dramatically; the print quality was never as good as that of daisy-wheel and Selectric printers, but the dot-matrix printers were much faster and far more flexible. They offered proportional spacing (available on some daisy-wheel printers, but never successfully supported by CP/M programs), italic, a variety of type sizes, better support for foreign language characters, and the ability to print pictures. All of this kept the word processing programs scrambling, successfully, to keep up.'

In addition, new features were added to the programs yearly: footnotes; endnotes; the ability to edit more than

one document; descriptive, non-printing file headers; split screen editing; mathematical functions on data contained within the document. The features of third-party (supplementary) programs were gradually incorporated into the main programs: macros (master commands to automate repetitive operations); data exchange with other programs; spelling checkers; thesauruses; data encryption; outline processors; graphics (illustration) creation and editing; charting; file comparison; indexing; an equation editor. A word processing program actually became a package of programs, filling many disks. WordPerfect 5.0, for example, in addition to installation and setup programs came with PTR.EXE, a program to edit printer definition files; CONVERT.EXE, a file conversion program to import and export text and data; GRAPHCNV.EXE, a program to import illustrations; SPELL.EXE, to edit the spelling dictionaries; and WPINFO.EXE, a program to provide information about the user's configuration to assist in diagnosing installation problems. Available for a small additional cost was a macro editor program, later incorporated into the main program.'

All of these new features meant that the programs grew more complex, causing frustration for many users. A new industry started to provide training and learning aids (books and videos). Still, a significant body of users still do not make efficient use of their word processors, and find them intimidating and confusing. Even experienced users often find the complexities of the programs frustrating, and to master every feature of a major program is a full-time job.'

Specific programs. WordStar failed to make the transition from CP/M to MS-DOS machines. The greatest failure in the history of word processing was WordStar 2000, an improved version of the program with many new features such as footnotes, but with a completely different command structure and user interface. It was received indifferently by the existing user base, who had already put in a long time in learning the previous commands and were not eager to repeat the process. There were never enough first-time users of WordStar 2000 to mark the program as a success and force users of the older WordStar 3 to change to it. As a result, faced with user anger, the publisher retreated to its previous user interface, and released WordStar 4. In addition, WordStar was damaged by a group of dissident employees who founded a new company and offered a competing program, NewWord.'

WordStar 4 offered macros, improved installation, and better on-screen help, but was basically an adaptation of the older CP/M program. During the time all this had taken, WordStar was overtaken by Microsoft Word and WordPerfect, which were first in offering such new features as laser printer support and print preview. In one

of the most egregious examples of "version inflation," WordStar released versions 5, 5.5 and 6.0, while WordPerfect more honestly released version 5.0 and 5.1. The implication is of course that the program with the higher number is better; since this would only be the view of less knowledgeable users, it is tantamount to a declaration of defeat.

Microsoft Word. In the mid-1980s, Microsoft Word was the leading word processing program. Following it was WordPerfect, and very much in third place was XyWrite, a respected program for serious writers which never achieved the sales its partisans said it deserved. Word was written to take full advantage of the graphics capability of the MS-DOS computers. This allowed it to have a better display, including screen fonts for foreign language characters. It also offered superior formatting capabilities and style sheets; it was the first major program to incorporate an outline processor. Word was well-placed, by its graphics-based structure, to become the most successful program for the similarly graphics-based Macintosh, and to gain an early lead as the word processor for the IBM graphic interface Windows (also a Microsoft product). However, Word's lead, in the MS-DOS world, did not last long. Its menus were cryptic; function codes could not be searched for; and Microsoft, a controversial company which had many other software products, did not give Word as high a priority as the WordPerfect Corporation did its principal product.'

WordPerfect. During the late 1980s and into the 1990s the predominant word processing program was WordPerfect. This is only partly a result of the features of the program. WordPerfect is a character-based (as opposed to graphics-based) program, its function keys are arbitrary and hard to learn, and while it has been a strong competitor, it has never held the lead in flashy features.'

The triumph of WordPerfect as the predominant word processing program is due at least as much to good management and business decisions as it is to the program's features. It is, first, a program available for a greater variety of operating systems than any other, although the MS-DOS version is the most popular. It is today the only word processor that has a common file format for both MS-DOS and Macintosh versions. It has also been a leader in government and international sales, providing many more foreign language versions, and more support for foreign characters and alphabets, than any competitor. It has been more successful than any other company in providing a package of programs: a spreadsheet (PlanPerfect), a database (DataPerfect), a drawing program (DrawPerfect), and a shell providing mail, calendar, and file management functions (WordPerfect

Office). A number of other characteristics, while not as easy to proclaim in advertising, have endeared it to data processing professionals, whose recommendations have had a major impact in corporation decisions to adopt it: a macro programming language and editor, a merge programming language, soft keyboards, and the broadest printer support of any program. WordPerfect also has the best documentation in the industry, a significant advantage in a field long characterized by poorly-written or inaccurate documentation.'

Perhaps the biggest single factor in the success of WordPerfect is quite external to the program itself: the company's liberal policy for telephone help. It is WordPerfect's policy to provide purchasers of its microcomputer programs unlimited telephone assistance free of charge, and indeed it subsidizes the service within the U.S. by supplying "800" numbers. At first the company would verify that the caller had purchased its program; later it changed to providing service even to those with stolen copies.'

The contribution of this decision to WordPerfect's success cannot be exaggerated. For new users, confused or faced with installation problems, it offered reassurance and help; for the intermediate it offered help with advanced features; for advanced users faced with program bugs it offered confirmation that bugs exist and indirect access to expert assistance. The telephone help served to reduce user frustration and greatly increase loyalty to the product. A feature much commented on was the addition of "hold jockeys," similar to radio disk jockeys, to calm users on hold during busy periods by playing music, giving progress reports on their calls, and offering news about WordPerfect products. The telephone help service also provided WordPerfect with invaluable feedback on user perceptions of its programs, and the company has shown an admirable ability to assemble and use the data so generated. However, the costs of providing this service have been heavy. It remains to be seen whether WordPerfect will be able to maintain this free service indefinitely.'

Simpler word processors. As the major word processors grew so complex, a number of smaller, cheaper, and simpler alternatives emerged in the mid-1980s. These included PC Write, Bank Street Writer, First Choice, and a number of others. While each has a number of users, none of them has achieved predominance over the others. In 1990 WordPerfect released LetterPerfect, a simpler and smaller version of its program that retained file compatibility.'

Macintosh computers. The CP/M and IBM-compatible microcomputers all use members of a single family of processing chips: the 8080 and its descendents the Z-80,

8088, 8086, 80286, 80386, and 80486. The other main family of computer processing chip, the 6502 and its descendents, gave birth to the Apple and then to the Macintosh computer, with the short-lived Lisa as an intermediary. The Apple computer, whose color, joysticks, and oversized letters made it very successful in schools, was never used for serious word processing. Its video display was too poor to have an 80 column by 24 line screen; thus it was not possible to see at all what things would look like on the screen.'

The Macintosh, introduced about 1984, quickly gained a loyal and vociferous group of followers. Its operating system, self-installation, and graphical user interface made it far easier to install and use than IBM-compatible computers, which the company emphasized in its advertising. Documents could be seen on the screen just as they would be printed, and the writer had precise control over fonts and font sizes. In contrast with the IBM microcomputers, patents and copyrights protected the Macintosh from competitors, and to date there are no Macintosh-compatible computers. This meant, however, that prices remained much higher than in the more competitive IBM-compatible environment.

The Macintosh was never much used in an office environment or for word processing. The dot-matrix printer sold with it produced an output only marginally suitable for correspondence. There was for some time no intermediate step between that printer and an expensive PostScript laser printer. The mouse was awkward for good typists, as it required that a hand be removed from the keyboard, and the control over typeface and type size was not only unneeded but even a distraction in writing.[1] The Macintosh's greatest success was in publishing and advertising, for which its graphics screen and precise preview of type, illustrations, and layout were invaluable.

Word Processing and Typesetting. Reducing the cost of typesetting so that it could be incorporated in an office or small business budget has been a long-standing goal of office equipment. Traditional lead type required expensive equipment and highly trained operators; it was slow to produce and correct. Adaptations of the electric typewriter provided proportionally spaced output of typeset quality, but justification—making an even right margin—meant that each line needed to be typed twice. The first keyboarding mechanically counted the length of the line, and the repetition inserted fractional spaces between words so as to align the right margin. IBM's entry in this field was the Selectric Composer, a piece of hardware vaguely resembling but much more complex than the Selectric typewriter, but with proportional spacing and with three pitches (15, 17.4, and 19.8 characters per inch). A special ribbon and paper were used to obtain a very black and

crisp output, suitable for offset printing.'

A model of the Selectric Composer was soon fitted to read tapes created and corrected on the MT/ST. The machine calculated interword spaces, and it was thus necessary to type only once to achieve a justified output. Hyphenation during playback, though operator-assisted, was implemented for the first time. While the Magnetic Tape Selectric Composer was still very expensive for an office, and was limited to a maximum type size of 12 points, its input could be prepared by a secretary without the special training of a typesetter. It marked an important first step in the marriage of word processing and typesetting.'

From early in the computer age, typesetters developed the hardware needed to use computer files as input. However, this was limited to letters, numbers, and punctuation; formatting, control of type face or size, and use of special characters were only possible if complex codes were added to the computer files. The results could not be previewed on the computer, nor could the transferred files, once processed on the target hardware, be returned for further editing on the computer. The output devices were far too expensive for office or small business use.'

Production of typeset output on the microcomputer itself required an affordable output device with good resolution, speed, and print quality, the ability to handle a variety of type sizes and special characters, and the processing power to calculate microspacing needed to justify lines. The latter is the simplest: the program uses a table containing the widths of each character, calculates the total length of the letters in a line by adding the width of each, subtracting this from the desired line length (measure), and dividing the remainder by the number of interword spaces. The result is then sent to the printer as fractional spacing instructions. In addition, good typesetting requires kerning: removing space between selected pairs of letters to achieve a satisfying visual effect. This added to the previous requirement the need to look up pairs of characters in a table and send fractional spacing adjustments to the printer between letters.

Proportional spacing was partially implemented on CP/M machines, and fully adapted on both MS-DOS and Macintosh machines in the mid-1980s. Kerning was implemented with WordPerfect 5.0, in 1988. The output device of choice proved to be the laser printer. Daisy-wheel printers could handle proportional spacing, but not varying type sizes. Changes from roman to italic required that the machine be stopped twice for manual change of print wheels. Dot-matrix printers could change to different type sizes and styles without operator intervention, but were unacceptably slow when operated at their finest resolution. The blackness of their output was also unreliable.

These problems were resolved with the laser printer,

which appeared on the scene in the mid-1980s. By the end of the 1980s it had so declined in price that it was available to almost anyone needing to produce typeset material. Using a print engine adapted from the photocopy machine, the laser printer offered high speed and reliability together with good print quality.

All laser printers came with some typefaces built in. All but the earliest accepted and housed in the memory electronic type stored in the memory chips of plug-in cartridges, or sent by the computer. The electronic type consisted of patterns of dots that would be printed on the page when the appropriate letter was sent from the computer. Any size or style of type could thus be accommodated. At the very end of the 1980s and early 1990s the production of electronic type grew into a new industry. The industry offered the microcomputer-based typesetter virtually all the choices in type style previously available on dedicated typesetting equipment costing many times the price. Many previously unavailable fonts appeared, both ornamental (based on handwriting, for example), and practical (exotic foreign languages; dead languages needed in Biblical and historical studies).

Page design programs, sometimes described with the misleading label "Desktop Publishing," allowed the microcomputer user easily and precisely to arrange headlines, text, and illustrations on pages. The leading programs were Aldus PageMaker and Ventura Publisher, and several simpler and less expensive alternatives.[2]

The future of word processing. Hardware has always set the boundaries for word processing. Hard disks and larger computer memories meant larger and more powerful programs could be stored and run. Better graphics displays meant more characters could be shown. Better printers allowed growth from letter writing to typesetting.

The huge storage capacity of the optical disk means that word processing will be linked even further with a multimedia environment including pictures and sound. At the same time, access to huge bodies of electronic text, locally or remotely stored on optical media, is rapidly becoming a reality. Reading will eventually take place primarily on a screen, and word processing software will be used for reading as well as a writing. The software offers the possibility of reading texts in new, non-linear ways, creating links between blocks of text or between files ("hypertext") which have no parallel in the conventional, printed world.'

As written communication shifts to electronic rather than paper transmission, word processors will become more closely linked with electronic mail systems. If past trends continue, functions presently carried out by separate programs from different manufacturers—grammar checking, concordance producing—will also be incorporated

into the concept of word processing. Typesetting will become even less distinguishable from word processing. Users will have the option of having display screens emulate the type styles of printed books. As operating systems become more standardized—Microsoft Windows, the very successful new user interface for the IBM-compatible, resembles the Macintosh in many ways—word processors will also come to resemble each other more. They will thus become easier to learn. Voice input and output is also in the near future, as is hardware designed to place less stress on the human body (arms, fingers, back, and eyes).

Daniel Eisenberg

[1] Marcia Peoples Halio, "Student Writing: Can the Machine Maim the Message," *Academic Computing*, January 1990, 16-19 and 45; Robin Nelson, "Wurd Processing," *Personal Computing*, August 1990, pp. 49-50.

[2] See Daniel Eisenberg, "In-House Typesetting on a Tight Budget," *Scholarly Publishing*, 21 (1990), 205-20.