

Text-formatting concepts

David Keil

Common concepts of text formatting apply in different environments, including word processors in the office suites (MS Office, OpenOffice, Google Documents, Ajaxwrite), PDF-file editors, the Hypertext Markup Language (HTML), and professional formatting environments like LaTeX. Here are presented some of those common concepts. For help on how to use the corresponding features in particular environments and products, see documentation and help material for those products. Some tips for using features of Microsoft Word are included below.

1. Page headers and footers

Professionally formatted documents have the document, slide, or chapter name at the top or bottom of each page, with a page number. This is always done automatically by use of *page headers* or *page footers*. It is an example of the principle of *global control of formatting*.

2. Text formatting elements

Character formatting includes size, font, and style. The font of paragraphs in this document is Times Roman; that of subheadings is Arial Bold. Character styles include Roman/Italic, Bold/Medium, Underline, Superscript, and Subscript.

Paragraph formatting includes indenting, space-before, and space-after. A *hanging indent* indents all lines but the first. A paragraph indent is *first-line*.

In text editing, such as of email, the user must normally add space between paragraphs by pressing “Enter” twice. In word processing or text formatting, on the other hand, it is considered professional to space text vertically by using the space-after paragraph attribute. Vertical spacing, like font size, is measured in *points*. A point is about 1/72 of an inch.

3. Indents, bullets and numbering

Paragraph indenting is always done automatically in professionally formatted documents. Paragraph objects have an attribute called “first-line indent,” which may be zero for unindented paragraphs or a fraction of an inch for indented ones.

When a paragraph begins with a bullet or a number, it is standard to automate the bulleting or numbering and to apply a “hanging indent” to the paragraph. A hanging indent is an indent on all lines after the first, so that all text after the bullet or number is uniformly flush left.

4. Section headings

Subheads, or section headings, appear above paragraphs to divide a document in a logical way. It is normal to format them by adding space above and below, by using a different font from the paragraph text, by using boldface, and by constraining page and column breaks so that all headings are in the same column or page as the text below them.

5. Tables of contents and indexes

When named styles are used to format headings, many text formatters, such as MS Word, support automatic generation of tables of contents. The table will list the sections and the pages of the start of each section.

To generate an index, the user must mark one instance of each word in the document that is to be included in the index. In Microsoft Word, use *Reference / Table*.

6. Named styles

A *named style* (“Quick style”) is a package of formatting specifications, applicable to a paragraph, and *named by the user*. These are available in MS Word and HTML, among other environments. The user may create a style for a subhead, for example, or a paragraph, or a bulleted list, and may apply it to any text in a document. The named style becomes an attribute of that text. After applying a style, the user may modify it, and the change will be reflected throughout the document, to all text that has that style as an attribute. Any paragraph attributed may be used as part of a named style.

The use of named styles is characteristic of people who know how to use word processing efficiently and is required for generating tables of contents in Microsoft Word. It is the main case of the principle of *global control of formatting*.

7. Tables

Although tables of text may be formatted in a crude way using tab stops embedded in the ruler above the text in a text editor, it is common for text formatters and environments (including MS Word and HTML) to support tables as separately-formatted entities. This kind of table can have multi-line text within a single table cell, borders and common formatting of columns, among other features. Where a table is required, it is considered professional to use the table-formatting features of the text-formatting application to implement the table, because some future version of the document may require the more advanced features supported by the table concept but not supported by the tab-stop concept.

8. Bibliographies and footnotes

Academic papers are documented according to standard specifications for bibliographic entries, references, and footnotes, such as APA (American Psychological Association), and MLA (Modern Language Association). Some text-formatting applications support automated referencing and footnoting.

To insert a footnote into a Microsoft Word 2007 file, key Alt-S, F. A number will appear at the bottom of the page and you may type your reference or comment.

9. File encoding formats

Each environment for formatting text typically has an associated “file format” associated with it, by which is meant an encoding of the text-formatting commands. Each “file format” has an associated file-name extension. For example, an unformatted text file has the extension “.txt”, an MS Word file has the extension “.doc” or “.docx”, a WordPad file has “.rtf”, an HTML file has “.htm” or “.html”, a LaTeX file has “.tex”, an Adobe Acrobat file has “.pdf”.

Most text-formatting applications can export files to, and import files from, other file formats or encodings. For example OpenOffice reads and writes .doc files, and MS Word exports HTML and PDF files.

10. Formatted text versus text files

Text-file editors, such as email editors, Notepad, FrontPage, WinEdt, and Visual Studio, read and write unformatted files in which only visible and printable keyboard-generated characters occur. Some text files, such as HTML files, have formatting embedded in the text. HTML files may be edited textwise in Notepad or other text editors, and displayed with format showing in an Internet browser such as Internet Explorer or Firefox.