Guidelines for Bibliographical Citations in LaTeX

Jean-Michel Hufflen

To cite this version:

Jean-Michel Hufflen. Guidelines for Bibliographical Citations in LaTeX. ArsTeXnica, 2007, pp.103–110. hal-00644464

HAL Id: hal-00644464
https://hal.science/hal-00644464
Submitted on 24 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Guidelines for Bibliographical Citations in \LaTeX

Jean-Michel HUFFLEN

Abstract

After a short overview of the schemes used for bibliographical citations, we give some guidelines to use some packages of \LaTeX\ 2\e and bibliography styles of \BibTeX\ in order to write adaptable citations, i.e., texts where switching a citation scheme to another is easy.

Keywords \LaTeX, \BibTeX, Con\TeX\t, methodology, \bbl\ module, bibliography citation schemes, natbib, jurabib, camel, opcit, amsrefs, M\LaTeX\b\.

Sommario

Dopo una breve panoramica degli schemi usati per le citazioni bibliografiche, l'articolo fornisce alcune linee guida per l'utilizzo dei principali pacchetti di \LaTeX\ 2\e e stili bibliografici di \BibTeX, allo scopo di ottenere citazioni flessibili, ovvero testi in cui risulti facile passare da uno schema di citazioni ad un altro.

Parole chiave \BibTeX, Con\TeX\t, metodologia, modulo \bbl, schemi per citazioni bibliografiche, natbib, jurabib, camel, opcit, amsrefs, M\LaTeX\b\.

1 Introduction

The \LaTeX\ typesetting engine [Knuth, 1984] has succeeded in implementing a clear separation between layout and structure and the word processors derived from it—in particular, \LaTeX\ [Lamport, 1994]—are unrivalled within this domain. However, if users wish to take as much advantage as possible of this feature, they should pay attention to write adaptable documents. For example, to put down a rule box between two paragraphs, as wide as possible, it is preferable to use:

\begin{verbatim}
\rule{\textwidth}{...}
\end{verbatim}

rather than a absolute length given in centimeters (or in another length unit) for the first argument. This allows such a rule to be adapted to the body’s width, and the \textwidth command is better than \columnwidth, because it prevents the possible use of a \textwidth option (see [Mittelbach and Goossens, 2004, § 2.1] about this option), as in the style used in the present proceedings. Other examples, more related to semantic markup, can be found in [Hufflen, 2006b].

This article also explores methodology, but about bibliographical citations. It does not replace the good descriptions of bibliographical citation schemes and their implementation by means

of \LaTeX\ 2\e packages given in [Mittelbach and Goossens, 2004, Ch. 12] and in more specialised reference manuals. It just aims to study how to write bibliographical citations according to a flexible way, especially when a word processor is used in conjunction with a bibliography processor. A good example of such a cooperation is given by \LaTeX\ and \BibTeX\ [Patashnik, 1988a], the bibliography processor most commonly used with that word processor. This subject is not artificial: there are several citation schemes used throughout written documents. Often an author is familiar with a particular citation scheme when writing an article or a book, and may be asked by a publisher to rewrite this work by using another scheme. This modus operandi may be more difficult than planned.

First, we give a short overview of the citation schemes used, then we show how some bibliography styles of \BibTeX\ and packages of \LaTeX\ 2\e can put them into action. In a fourth section, we give some ‘tricks’ in order for bibliographical citations to be easily adaptable to most schemes. Then we explain how new bibliography styles should be built.

Reading this article requires a good practice of \LaTeX\ and \BibTeX, but as end-users do. It is not needed to know how to write a new bibliography style for \BibTeX\ or new advanced commands for \LaTeX\. Let us only recall that when \LaTeX\ and \BibTeX\ are used in conjunction, bibliographical entries are given in bibliography database (.bib) files—an example is given in Figure 1—and can be referred by means of the \cite command of \LaTeX. This command’s argument must be a key used within such a .bib file, e.g., \cite{robeson1974c}. Then \BibTeX\ searches these .bib files and builds a .bbl file containing bibliographical references\footnote{1. Here we use a precise terminology, which is defined within M\LaTeX\b\ (‘Multi\LaTeX\b\’\), our reimplementation of \BibTeX\ [2003]: an entry (resp. reference) belongs to a .bib (resp. .bbl) file.} \LaTeX\ can process. This .bbl file, containing the ‘References’ section usually put at a document’s end, uses the \bibitem environment and the \cite command. See [Mittelbach and Goossens, 2004, § 12.1.2 & 12.1.3] for more details.

2 Citation schemes

Roughly speaking, bibliographical citations serve two purposes:

1. This article also explores methodology, but about bibliographical citations. It does not replace the good descriptions of bibliographical citation schemes and their implementation by means
• they can refer to a work where more information could be found,
• or they can acknowledge anterior or comparable works.

In both cases, we require that references to citations are unambiguous, so looking for them within the ‘References’ section should be easy. Some examples could be:

[3] is known as one of the best adventures of Doc Savage. [...] Brian Herbert and Kevin J. Anderson continued the *Dune* saga [1].

where a ‘plain’ style is used, that is, references are labelled with numbers.

The style of the first example is controversial: some people think that a text should remain intelligible if references to bibliographical citations are removed. A good rewriting of this sentence would be:

*The Devil Genghis* [3] is known as one of the best adventures of Doc Savage.

Let us remark that the title has been typed ‘manually’ in this previous sentence. In fact, we can think that there is information redundancy between the title both typed and given in the ‘References’ section. If there is a mistake about this title, it must be fixed in several places: the bibliography database file and all the occurrences within the text’s body.

The second example does not meet this criticism, that sentence makes sense even if the citation is removed, but let us rewrite it using a ‘long’ style, that is, references are labelled with authors’ names, followed by the publication’s year:

Brian Herbert and Kevin J. Anderson continued the *Dune* saga [Herbert & Anderson 2000].

First, the same problem of information redundancy occurs, about the authors’ names. Second, the citation could omit them, since they are given earlier in the sentence:

Brian Herbert and Kevin J. Anderson continued the *Dune* saga [2000].

Now let us come to a short synthesis of the main citation schemes used throughout printed documents. As mentioned in [Butcher, 1992, § 10], there are four schemes:

**number-only** publications are sequentially numbered in the ‘References’ section, and citations refer to these numbers;

**author-date** citations use authors’ names, followed by the publication’s year; if an author published several works in a year, that year is suffixed with lowercase letters (e.g., ‘1974a’, ‘1974b’); in addition, references to citations can be abridged: only the year is given if the author’s name is known, and multiple citations concerning the same author supply this name only once (e.g., ‘[Robeson 1974a, 1974b]’ rather than ‘[Robeson 1974a, Robeson 1974b]’); more details about this scheme can be found in [Chicago, Ch. 16];

**author-number** this system is similar to the author-date system, except that each author’s publications are numbered and citations use authors’ names, followed by this number, e.g., ‘[Robeson (2)]’;

**short-title** references are given directly in the text at the first mention, the others may consist of an abridged form using the author’s name only, followed by an abbreviation of the title if there are several works of this author; a full bibliography at the document’s end may

---

2. According to Bt\TeX’s terminology.

3. Also known as ‘Harvard system’.

4. From our personal point of view, ‘short-title’ is not good terminology for this system, because titles may be omitted for repeated references, if there is no ambiguity. In other words, the short-title system does not use titles always.
Herbert, Brian and Anderson, Kevin J. continued the *Dune* saga.\footnote{1}


**Figure 2:** Citation as a footnote.

To sum up all the references, it can also be omitted, since a complete specification of each reference is provided either inside the text, or in footnotes.\footnote{5}

To go on with our previous examples, the first could be rewritten using the short-title system:

Robeson, *The Devil Genghis* is...\footnote{5} if this book has already been referred in a text where other books of Kenneth Robeson are cited. Let us remark that this sentence must be reformulated differently if references are given as footnotes. In this last case, our second example looks better, as shown in Figure 2.

3 Bibliography styles and packages

The L\TeX{} distribution provides a rich set of bibliography styles usable with Bib\TeX{}, summarized in [Mittelbach and Goossens, 2004, Table 13.4]. Let us connect them to the schemes described in § 2. There is the ‘plain’ family, putting the number-only scheme, including styles such as plain, abbrv, and acm. We can consider that the ‘alpha’ family also puts this scheme into action. Labels are not numbers, but keys retaining some letters at the beginning of names, followed by the publication’s year. For example, the alpha style would produce the keys ‘[HA00]’, ‘[Rob74a]’, ‘[Rob74b]’ for the three entries of Figure 1. These keys are atomic, in the sense that it is impossible to split them into the information about author and year, so this style does not have the expressive power of an author-date system. The same for the styles belonging to the ‘long’ family\footnote{6} [see Hufflen et al., 1998, § 4.3.3]. Some bibliography styles of Bib\TeX{} were developed as early attempts of the author-date scheme: *authorodate*\footnote{5}, *chicago*, *harvard*, as mentioned in [Mittelbach and Goossens, 2004, § 12.3.1]. In addition to each bibliography style, a \LaTeX{} 2ε package provided more functionalities in order to cite only the authors of a reference—using an abbreviated form or in *extenso*—or the date only, or both.

Now, the best implementation of this scheme is the *natbib*\footnote{7} package [see Mittelbach and Goossens, 2004, § 12.3.2]. Let us consider again our second example:

```
citeauthor{herbert-anderson2000}
continued the *emphasis* (*Dune*) saga
\citeyearpar{herbert-anderson2000}.
```

the \texttt{citeyearpar} command putting the year between delimiters, whereas \texttt{citeyear} does not put them:

```
They’ve continued it since
\citeyear{herbert-anderson2000}.
```

We are close to the expected result, except that only last names are retained, first names are dropped out:

*Herbert and Anderson continued the* *Dune* saga [2000]. They’ve continued it since 2000.

The short-title scheme has been fully implemented by the *jurabib* package [see Mittelbach and Goossens, 2004, § 12.5.1]. This package is easily customisable by means of the \texttt{jurabibsetup} command, as shown in Figure 3. Another implementation of the short-title scheme is given by the *camel* package [see Mittelbach and Goossens, 2004, § 12.5.2], but it uses an enlarged version of \LaTeX{}. A package close to this scheme is *opcit* [Garcia, 2006], since it implements citations as footnotes, as often within this scheme. By default, this last package automatically uses its own bibliography style, *opcit.bst*.

4 Adaptable citations

4.1 Methodology

Let us recall that we aim to write adaptable citations in \LaTeX{} source file, in order to prevent a change to another scheme if an article has already been written. On another point, we aim to avoid...
information redundancy and attempt to extract as much information as possible from the metadata given in bibliography database files, and transcribed in the result files generated by BibTeX. Here we are interested only superficially in the look of the citation keys and the ‘References’ sections: besides, many packages allow end-users to enhance them (see [Mittelbach and Goossens, 2004], throughout Chapter 12).

Our first advice is to avoid the sentences that need the full mention of a title, unless such style is absolutely necessary. As far as we know, only one package provides a command returning the title only, analogously to the commands `citeauthor` and `citeyear` of the natbib package. If we look into the jurabib package, its `citetitle` command puts titles unconditionally, but after authors’ names\footnote{See Footnote 4, p. 2.}). This package also provides a `citetitleonly` command, but the title is given as a footnote if references are to be given in footnotes. In other words, this command would be unsuitable for our first example (concerning the Devil Genghis) within the source text given in Figure 3. The only way to get a title inside a text, regardless of the citation style—as footnotes or in-text—is `citefield{title}{...}`—see another example of using this command in Figure 3—provided by jurabib.

In fact, we think that authors should use the short-title scheme only if they are asked for that. Of course, this scheme offers undisputable advantages and is widely used in humanities. Moreover, the two implementations in \LaTeX~of this scheme are very powerful. But it is difficult to adapt texts written using this scheme if we have to switch it to another. It is even difficult to switch a package implementing this scheme to the other. First, they use specific bibliography styles for \LaTeX: e.g., jox, jurabib, jureco, jurunsrt for jurabib. Second, the camel engine uses its own commands for citations and bibliography management. In addition, jurabib and camel have been developed in order to process juridical works: they provide features unknown or rarely used in other fields. To sum up, the texts written with these packages may be very difficultly adaptable to other schemes, because of stylistic variants or because the commands used for citations are different.

For adaptable citations, the best solution is to use the natbib package as far as possible because it can produce author-date references as well as number-only ones. Using the numbers option:

\begin{verbatim}
\usepackage[numbers]{natbib}
\end{verbatim}

causes bibliographical references to be labelled with numbers, as in the number-only scheme, whereas the commands `citeauthor`, `citeyear`, and `citeyearpar` are still available. As shown in our previous examples, the `citeyear` command should be used when the year is to be put down in any case. When only the year information is given because the author information is already known, a ‘trick’ is to define a new command:

\begin{verbatim}
\newcommand{\citeend}[1]{\%\ifNAT@numbers\citep{#1}\else\citeyearpar{#1}\fi}
\end{verbatim}

(let us remark that the commands `makeatletter` and `makeatother` may be needed before and after this definition; see [Mittelbach and Goossens, 2004, § A.1.1] for more details). This `citeend` command, used within a sentence such as:

\begin{verbatim}
... continued the \emph{Dune} saga
\end{verbatim}

\citeend{herbert-anderson2000}.

will produce the citation of a numerical label if the numbers option has been selected, the year otherwise. In any cases, delimiters are put down.

For complete references, use the `\citet` command rather than `\cite` because the former always produces references between delimiters, as numbers if the numbers option is selected, as author-date information otherwise.

This \textit{modus operandi} is also suitable for \textit{unsorted} bibliographies, in which case the \textit{unsortnat} bibliography style should be used. Let us recall that in such styles, bibliographical items are not

\begin{verbatim}
\bibitem[Robeson(1974\textsuperscript{b})]{robeson1974c}
Kenneth Robeson.
\newblock \emph{The Devil Genghis}.
\newblock Bantam, November 1974\textsuperscript{b}.
\end{verbatim}

\begin{verbatim}
\bibitem[Herbert and Anderson(2000)]{herbert-anderson2000}
Brian Herbert and Kevin~J. Anderson.
\newblock \emph{House \textit{Atreides}}.
\newblock Bantam, February 2000.
\end{verbatim}

\begin{verbatim}
\makeatletter
\ifNAT@numbers
\citep{#1}
\else\citeyearpar{#1}\fi
\makeatother
\end{verbatim}
sorted with respect to the alphabetical order of authors or organisations, the order of such items in the ‘References’ section is the order of first citations of these items throughout the document. If references are labelled with numbers\footnote{In F rance, such a scheme is still used for Medicine Ph.D.s, but—strangely—not for the documents written by nurses, where the short-title scheme is usual. As far as we know, sorted bibliographies are commonly used within other scientific fields.}, finding an item within an unsorted bibliography is easy, but that becomes difficult if a large bibliography with unordered labels is to be searched. In addition, checking if a particular work is cited within a ‘References’ section may be tedious.

4.2 Towards more powerful bibliography styles

The citation scheme used by L\(\TeX\) defaults to number-only, in which case the `\texttt{\bibitem}` command has a mandatory argument:

\begin{verbatim}
\bibitem{robeson1974c}
\end{verbatim}

If we use a style belonging to the ‘alpha’ or ‘long’ family, an optional argument—built by Bib\(\TeX\)—specifies which label is to be put in front of the corresponding reference:

\begin{verbatim}
\bibitem[Rob74a]{robeson1974c}
\end{verbatim}

When a bibliography style suitable for the natbib package is used, this optional argument is used by the functions of the natbib package to retrieve the different parts of the author-date information: we show how this optional argument is built in Figure 4.

The structure passed from Bib\(\TeX\) to L\(\TeX\) is much more expressive when we consider bibliography styles suitable for the jurabib package. Figure 5 shows how the jurabib bibliography style process the entries of Figure 1, this .\texttt{bbl} file being usable by the source text given in Figure 3. This kind of structure would be very tedious to produce manually. But the jurabib package allows much customisation and such complexity is the price to pay. For example, `\texttt{\jbbtasep}' stands for ‘\texttt{Between Two Authors Separation}’, `\texttt{\jbdy}' for ‘\texttt{Dummy Year}’\footnote{By default, the \texttt{\jbdy} command skips the next token, and \texttt{\jbbtasep} inserts a `/\texttt{'} character. This last command can be redefined by means of the \texttt{\jurabibetup} command, as shown in Figure 3.}, the ‘book’ information is needed since jurabib allows a kind of entries to be processed specifically, …

However, transmitting a structured text to the packages dealing with bibliography citations is the only way towards more powerful commands for citations. In fact, a modern version of the `\texttt{\bibitem}' command should not only provide a label, but also

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{example}
\caption{References produced by the jurabib bibliography style.}
\end{figure}
encompass all the metadata for a bibliographic item. The first implementation of this idea is given by the amxport bibliography style [Downes, 2000], which generates output files like the example of Figure 6. Let us remark that this style’s main purpose is not the definition of commands accessing particular parts of a reference within a text’s body. It aims to allow the specification of formatting bibliographical references using \LaTeX-like syntax. That is done by means of the functionalities of the amsrefs package.

A more modern implementation of this modus operandi is the \texttt{bib} module of ConT\textsc{e}Xt [Hoekwater, 2006], the result of processing our bibliography sample being given in Figure 7. In ConT\textsc{e}Xt [see Hagen, 2001, § 2.2], commands like \texttt{\startpublication} and \texttt{\stoppublication} are analogous to environment delimiters in \LaTeX (\texttt{\begin{...}} and \texttt{\end{...}}). The structured information put between the commands \texttt{\startpublication} and \texttt{\stoppublication} may be customised as described in [Hoekwater, 2006, § 2.2] by means of the \texttt{\setuppublicationlist} command. In particular, this command allows authors’ names to be typeset in extenso or abbreviated, but such customisation exists for the ‘References’ section, not for citations of authors’ names inside a text. The \texttt{\cite} command also provides an advanced command for citations [see Hoekwater, 2006, § 3], e.g.:

\texttt{\cite[robeson1974b]{kenneth}}

Let us remark that only some options are available: author, authoryear, year, ... but this command cannot put only a title. The layout of the texts produced by this command may be customised by means of the \texttt{\setupcite} command.

11. Roughly speaking, a ConT\textsc{e}Xt module is analogous to a \LaTeX package.
\usepackage{bib}
\setupbibitem[database=...]
\starttext
\setupcite[author][left=,right=]
\setupcite[year][left=(),right=]
\cite[author][herbert-anderson2000] have continued the \{em Dune\} saga
\cite[year][herbert-anderson2000].
\setupcite[year][left=*,right=*
They've continued it since
\cite[year][herbert-anderson2000].
\placepublications
\stoptext

\section*{Conclusion}

We personally used the number-only scheme for a long time. Then we had to rewrite several of our texts using the author-date scheme. We also helped some people to prepare written documents about other fields than those we have usually dealt with. Some adaptations were difficult, especially about unsorted bibliography styles and texts using the short-title scheme. Now we are aware of the influence of citation schemes on the style of writing and began to experiment some solutions.

On another point, we are presently rewriting the bibliography styles usable with the packages \texttt{natbib} and \texttt{jurabib}, taking as much advantage as possible of the new functionalities provided by \texttt{MIBibTeX}, our reimplementation of \texttt{BibTeX} (see App. A). This work has definitely convinced us that the present form of the \texttt{\textbackslash bitem} command was no longer suitable for modern requirements. As a possible solution, we plan to investigate the use of \texttt{Lua}\TeX{} [Hagen, 2006], a new type-setting engine built out of \TeX{}, mixing the expressive power of \TeX{}’s features and service provided by a modern programming language. Within this framework, a bibliography processor could result in structures suitable for the \texttt{Lua} programming language [Ierusalimschy, 2006] when entries are processed. Then, when the word processor deals with a text, bibliographical references could be formatted by running fragments written in \texttt{Lua}.

\section*{Using \texttt{MIBibTeX}}

\texttt{MIBibTeX} is a reimplementation of \texttt{BibTeX} with particular focus on multilingual features [Hufflen, 2003]. As we explain in [2006c], it uses XML\textsuperscript{12} as a central formalism and bibliography styles may be written using a variant of XSL\textsuperscript{13}. A compatibility mode allows bibliography styles of \texttt{BibTeX} to be run with \texttt{MIBibTeX} [2006a], although some change would be needed to run the bibliography styles associated with the \texttt{jurabib} package\textsuperscript{14}.

Concerning the problems raised by the different citation schemes, \texttt{MIBibTeX} is able to provide the following solutions.

- It eases the specification of styles according to the author-number scheme\textsuperscript{15}. These styles could be used with the \texttt{natbib} package\textsuperscript{16}.

- More generally, we think that the new language for bibliography styles eases the programming of labels for references. For example, we can associate the two citation keys ‘[Rob 74a]’ and ‘[Rob 74b]’ with the two entries of Figure 1 whose author is Kenneth Robeson, according to the \texttt{alpha} bibliography style. But a variant of this style could associate the two labels ‘[Rob 74]’ and ‘[Rob 74a]’ with these two items, that is, an additional letter is used only from the second work if an author published several documents in a year. Such variants are easy to program with \texttt{MIBibTeX}, in comparison to what should be done with the language of \texttt{BibTeX}’s bibliography styles.

- \texttt{MIBibTeX} provides additional markup when some information—person names and titles especially—belongs to a language different from the global language of an entry. An example is given in Figure 9: this bibliographical entry concerns a work written in Italian, so the author’s name and the title are supposed to be in Italian. This work is included in an international anthology whose two editors are American and French. In addition, the anthology’s title is given in French. These informations about languages are given in Figure 9, this entry being suitable for \texttt{MIBibTeX}. However, some rewriting of the \texttt{natbib} package would be needed in order for this package to take advantage of this feature.

\footnotesize
\textsuperscript{12} eXtensible Markup Language. Readers interested in an introduction to it can refer to [Bay, 2001].
\textsuperscript{13} eXtensible Language Style Transformations. This language is used to write transformations of xml texts [W3C, 1999].
\textsuperscript{14} That is because \texttt{MIBibTeX} and the styles suitable for \texttt{jurabib} both deal with a LANGUAGE field inside bibliographical entries, but this field obeys different conventions.
\textsuperscript{15} Although developing such styles using the language of bibliography styles of \texttt{BibTeX} [Patashnik, 1988b] should be possible. However, that has not been done yet, as mentioned in [Mittelbach and Goossens, 2004, § 12.4].
\textsuperscript{16} There is no current support of this scheme for \texttt{BibTeX}, unless we accept that references are numbered globally, in which case the \texttt{natbib} package can be used with the \texttt{numbers} option [see Mittelbach and Goossens, 2004, § 12.4.1].
Guidelines for Bibliographical Citations in \LaTeX

Jean-Michel Hufflen. \LaTeX{} architecture. \textit{Ars\TeX{}nica}, 2:54–59, October 2006c. In GUIT 2006 meeting.


Oren Patashnik. \BibTeX{}. Part of the \LaTeX{} distribution, February 1988a.

Oren Patashnik. \textit{Designing \BibTeX{} Styles}. Part of the \LaTeX{} distribution, February 1988b.


\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{figure9.jpg}
\caption{How to specify ‘foreign’ information in \MibTeX{}}
\end{figure}

Acknowledgements

I have learned very much whenever I helped some people to play with \BibTeX{} and the citation commands of \LaTeX{}. Initially, some of these people did not plan to use these two programs and had already begun to write important parts of their works. Thanks to them, because such cases gave me the idea of this paper. I am also grateful to Massimiliano Dominici, who has written the Italian translation of the abstract. Last, thanks to the anonymous referee, who suggested me significant improvement of the first version.

References


\begin{itemize}
\item Jean-Michel HUFFLEN
LIFC (EA CNRS 4157)
University of Franche-Comté
16, route de Gray
25030 BESANÇON CEDEX
FRANCE
hufflen@lifc.univ-fcomte.fr
\end{itemize}