

## Editorial Principles

From the micro level of single characters or words, via sentences, passages and paragraphs, up to larger units such as chapters, works and composite works, text changes over time in various ways. A character, a passage or a chapter might be deleted or added, their order rearranged. And not only the text itself, its attributes too are unstable. A title, a creation date, or the authorship of a text might be forgotten, spuriously added or truthfully re-discovered. Changes can be intentional or unintentional, drastic or subtle, but they are evident for all forms and modes of text: as language, image, or music; oral, written, printed or digital text. Textual changes begin with the author and end only with complete oblivion: the state where we have forgotten that a text ever existed. It follows that every instance of a text that we perceive is only a snapshot, one stop in a succession of different instantiations of the process that is the text. This book – one of a series of Chinese Buddhist temple gazetteers – is no exception.

As a rule, all gazetteers in this series are composite texts, that is they consist of shorter texts from different authors, times and genres. They were originally printed – some with woodblocks, some with movable type – between the late 16<sup>th</sup> and the early 20<sup>th</sup> century. Between 2007 and 2011 they were scanned and digitized in XML/TEI format.<sup>1</sup>

Editors add information in various ways: obviously in the form of notes, apparatus, textual emendations, introductory material, revised punctuation and spelling, but also in less obvious ways such as through lay-out or paragraph breaks. In the digital medium these editorial interventions can be recorded in TEI. TEI is a markup standard that is expressed in XML and maintained and developed by the TEI community (<http://www.tei-c.org>) since 1991. This volume is

---

1 The base texts used for the digital editions are the gazetteer reprints found in Du Jiexiang 杜潔祥 (ed.). 1980-1994. *Zhongguo fosi shizhi huikan* 中國佛寺史志彙刊. Taipei: Mingwen shuju 明文書局.

printed from a PDF that in turn is a transformation of our TEI edition. Here, we present some of the markup we have added in a form that is suitable for print publication. The same TEI edition, however, can be transformed into an HTML view for an online interface, or be made available as an archive file. Because of the automated transformation of one format into another, our gazetteer editions appear somewhat different from traditional editions of classical Chinese texts. The following will describe in detail what kind of information has been added to the text.

## 1. Names

Person and place names pose one of the greatest obstacles to our understanding of the gazetteers as historical sources. Gazetteers obviously contain a large number of names: emperors and abbots, literati, monks and laypeople. In the digital medium names can be linked to authority databases containing background information on the person or place.

Building such a database is not trivial. In order to record someone's life dates one needs to model the fact that the exact birth or death date are not known exactly and must be expressed with a period rather than a simple date. Both dates also must be annotated as to their sources, because the information in the database must be traceable to be academically useful.

Historical place names have their own set of problems. An authority database must first of all try to identify the geographic location of a place, but in order to do so has to take into account the fact that places can move about (same name, different location) or change their names (different name, same location). Also, sources often refer to larger areas rather than single point locations and ideally an authority database would geo-reference these as polygons not as points. However, due to the lack of exact information regarding historical borders this cannot be done with any degree of certitude.

Since 2007 Dharma Drum Buddhist College has been constructing authority databases, where information about persons and places that

appear in various Buddhist sources is constantly improved and expanded. As of August 2012 these databases contain information on more than 22,000 persons and reference more than 55,000 places.<sup>2</sup> Online, these databases enable readers to e.g. click on a link that connects a name with its database entry, in print, however, such a function can only be approximated by the less convenient index.

Our printed gazetteer editions therefore contain indexes for people and places that provide readers with basic information. The indices in this series are generated automatically, they reflect the state of the databases in early 2013. In the digital medium, editions can be searched for character strings, in the print edition the page references in the index make up for this function at least for person and place names. The indexes are arranged alphabetically according to the Hanyu Pinyin romanization system mandated by the UN.<sup>3</sup> Compared to an index arranged by radical or stroke-order this has the advantage that entries can be found quicker. On the other hand any automatic arrangement based on pronunciation has to face the *pinyinzi* 破音字 problem. Some Chinese characters have more than one pronunciation, which is especially relevant in the case of names. The DDBC authority databases focus on referencing entities geographically and historically, they do not include information on pronunciation. In the case of the printed indexes for our gazetteer editions the pinyin arrangement is merely for look-up purposes, it should not be seen as a statement about the pronunciation of the whole name. Its order reflects merely the most common pronunciation the first character of each entry.

Every string that is marked as person name (straight underline) or place name (waved underline) in the text is findable in the index, except a very small number of names for which we were not able to find any information and which have no entry in the authority databases. Generally, we also mark honorific or other additions to names (dashed underline).

---

<sup>2</sup> The place name authority contains many entries provided by the Center for GIS, Academia Sinica, Taipei.

<sup>3</sup> ISO 7098:1982 Revised 1991.

## 2. Dates

Gazetteers contain a large number of dates expressed in the traditional Chinese calendar notation using the sexagenary cycle for years and days, with years further dated by emperor and/or era names. The traditional notation, however, has a number of disadvantages. Few readers have memorized all era names and it is sometimes difficult to say which of two events recorded in different contexts is earlier without resorting to look-up tables. In critical editions of ancient texts it is good practice to map traditional calendar systems to an international standard. Dharma Drum Buddhist College has created a time authority database that maps East Asian calendars to a common standard, the Julian Day Number, that is widely used in astronomy.<sup>4</sup> Based on the Julian Day Number, dates can be reliably converted into other formats. In our print edition we annotate all dates and periods (except dynasties) with their equivalent in the proleptic Gregorian calendar. This helps to understand the relative dating of events and aligning them on a timeline.

In traditional notation it is often difficult to see how far apart in time two different events did occur.

One text, for instance, mentions an event for the 7th month in the 1st year of the Longxing era of emperor Xiaozong 孝宗隆興元年七月, then another event somewhen in the Jiading era of emperor Ningzong 寧宗嘉定間.<sup>5</sup> Because of the order in the text, or perhaps because one has memorized the order of Song dynasty era names, one can assume that the first happened before the second. However, most of us who are not experts in Song history would be hard pressed to say whether the second event occurred 10, 50 or 100 years after the first. Mapping the events to a common standard (孝宗隆興元年七月 = 1163.8.9-1163.9.6; 宋寧宗嘉定間 = 1208.1.26-1224.9.23) improves our understanding of how far apart events in the text have occurred relative to each other, and it makes it easier to compare dates recorded in different texts.

---

<sup>4</sup> See: The XXIIIrd International Astronomical Union General Assembly: Resolution B1 “On the Use of Julian Dates”.

<sup>5</sup> ZFSH 9, p.177.

Texts often date events imprecisely by referring to the ‘beginning’ or the ‘end’ of a period (e.g. “the beginning of the era Tianbao” 天寶初年 or “in the latter years of the Wanli era” 萬曆末年). In these cases we interpret this to mean the first or last 25% of the period in question.

The mapping to the proleptic Gregorian calendar is given in brackets for all dates except for dynasty periods. For these we ask the reader to refer to the overview table.

### 3. Additions

Editing is the process of adding information to a text in order to provide readers with opportunities to further their understanding of the work. Next to the information on names and dates mentioned above we have added four types of information:

**3.1. Section/text titles:** Some gazetteers included texts without titles. In order to reference these texts more effectively (e.g. in the table of contents) we have provided simple, descriptive titles for them. In the printed edition titles added by us are marked with angular brackets: [*Added section title*].

**3.2. Footnotes:** We have added footnotes pointing out and explaining problems in the text. We kept these notes to a minimum as the gazetteers contain a wide range of texts of different periods, and a comprehensive, detailed annotation of the texts was not possible within the scope of our project.

**3.3. Supplying missing characters or passages:** Working from copies of woodblocks or manuscripts single characters or passage were sometimes missing or illegible. Wherever possible we tried to supply these with the help of other editions. All these supplied passages are clearly referenced in footnotes, and marked with double angular brackets in the text: [[*supplied text*]].

**3.4. Punctuation:** The new punctuation we have provided for the gazetteers is likely going to be the most controversial aspect of the

edition. It was also turned out to be the most time- and labor-intensive effort of our project. Punctuation of classical Chinese is a difficult art which mercilessly reveals one's level of understanding. As punctuation projects go, ours is larger and more ambitious than most. The texts we try to punctuate are from different genres, by many different authors and periods. Though the classical tradition has a sense of punctuation,<sup>6</sup> there is no generally accepted punctuation standard for classical Chinese. The main reason for this is the lack of a comprehensive, generally accepted descriptive grammar of classical Chinese that would allow us to disambiguate syntactic components without fail. The morphology and syntax of Classical Chinese were never standardized by a set of rules in the way classical Sanskrit was described by Pāṇini and Patañjali.

In any event, even if we had such a grammar it must be remembered that punctuation, like spelling, is not part of a language itself, but a convention, which is usually standardized by political or cultural institutions. What punctuation marks to choose and to what degree they are compulsory is an arbitrary decision. The difference between a full stop “ 。” or a semi-colon “ ; ” is often a matter of interpretation. Modern Chinese punctuation generally differentiates between two types of comma, the *dunhao* 頓號, used for enumeration, and the *douhao* 逗號, used to separate clauses, a distinction that is not used in the punctuation of texts in European languages.

In short, there is no “natural way” of punctuating any written language, least of all classical Chinese. We have tried to provide a punctuation that will assist a Chinese reader in the early 21<sup>st</sup> century and have checked the new punctuation for each volume carefully. However, since the work was done by a large team of people, and the texts are sometimes difficult, results will vary slightly. Specialists in classical Chinese are bound to disagree with some decisions or find mistakes here and there. We hope that any mistakes will be offset by the value of the more than 99% of punctuation signs that correctly disambiguate the syntax.

---

6 Traditionally many Chinese readers punctuated their books on first reading, breaking the flow of characters with one single mark.

## 4. Treatment of variant characters

The graphic representation of Chinese characters varies just as the symbols in other writing systems. The variance is especially noticeable in manuscripts where the same symbols are almost never realized fully identical. Woodblock printing occupies a middle position between the manuscript and movable type printing. Like movable type it aims at uniformity and tends to regularize writing and spelling, while, due the production process, still preserving some of the flexibility of calligraphy. Sometimes calligraphic writing is preserved in a woodblock for honorific reasons. In the case of the Buddhist temple gazetteers, for instance, prefaces and postscripts are often printed in the author's original calligraphy. Even where not explicitly calligraphic, woodblock editions use many more character variants than movable type editions. In modern editions these are usually silently regularized. Gazetteers re-editions in China often even use simplified characters, thus moving even further away from the original appearance of the text. We have tried to walk a middle way, and, as in the case of punctuation, some readers might disagree with our strategy. Our main aim is to present the text as closely to the original as possible outside of a facsimile edition. On the other hand we want to aid modern readers by adopting the following procedure regarding variants.

For a digital edition the only relevant character set is the Unicode standard. Currently, every serious digital project that aims at long-term preservation and wide distribution has to work within the limitations that Unicode imposes on the range of characters that can be encoded.

The first question when encountering a character variant must therefore be:

Is this variant part of Unicode? If yes, our edition policy distinguishes between three possibilities:

a) The variant is clearly recognizable as such and is therefore used without further annotation. E.g. if the original has 德 it is used instead of the more common 德. However, as a rule, all CJK Unicode

compatibility variants (*jianrong yitizi* 兼容異體字) (Unicode F900-FAD9 and 2F800-2FA1D) are regularized.<sup>7</sup> Thus e.g. 益 (U+FA17) → 益 (U+76CA), 福 (U+FA1B) → 福 (U+798F), 僧 (U+FA31) → 僧 (U+50E7).<sup>8</sup>

b) There are cases where a modern reader might have difficulties in recognizing variant forms of the character: For example is not easily understood that the character 厶 corresponds to 去, or the character 灑 to 法. For these cases we have created a “Table of Difficult Variants” for each gazetteer, where we give the common reading for these characters. What counts as ‘difficult’ relies in practice on the judgment of the individual encoder and we do not pretend to follow a clear-cut standard as to which variants to include in these tables. Specialists in Classical Chinese might do without them, but we hope they might be of help for other readers.

c) A third difficulty involving variants in the wider sense can arise when the original uses synonyms that are not generally recognized as such anymore: For example 證 for 症, or 縣 for 懸. In these cases we have added footnotes to clarify the meaning.

If a particular variant has *not* been encoded in Unicode, it is usually possible to identify its common form.<sup>9</sup> In that case the common form is used in the text. In the very few cases where a common form cannot be determined beyond doubt, or the character belongs to the rare category of taboo characters which omit their final stroke, we have fonted the character in SVG.<sup>10</sup>

We hope that through the addition of information on person names, place names, and dates, the addition of notes and punctuation, as well as by the careful treatment of variants this edition of the text

7 Extensible Markup Language (XML) 1.0 (Fifth Edition) W3C Recommendation 26 November 2008 (<http://www.w3.org/TR/2008/REC-xml-20081126/>) Ch. 2.2 : “Document authors are encouraged to avoid “compatibility characters”, as defined in section 2.3 of [Unicode].”

8 For more details concerning see the project documentation at [http://wiki.ddbc.edu.tw/pages/佛寺志專案\\_特字處理](http://wiki.ddbc.edu.tw/pages/佛寺志專案_特字處理)

9 Authoritative for this is the Variant Character Dictionary 異體字字典 by the Taiwanese Ministry of Education 中華民國教育部國語推行委員會. (Online edition: <http://dict.variants.moe.edu.tw/> (Aug. 2012)).

10 See e.g. the taboo characters in the Hanshan Gazetteer 寒山寺志, p. 210.

## Editorial Principles

will be more reliable, complete and reader-friendly than previous print editions. The page numbers referenced in the margin point to the base text as reproduced in the *Zhongguo fosi shizhi huikan* to allow for easy comparison with that facsimile edition.