[W3C](http://www.w3.org/)

**Requirements for Tibetan Text Layout and Typography**

**W3C Working Draft 15 June 2015**

**This version:**

[TBD](http://www.w3.org/TR/2013/WD-klreq-20130514/)

**Latest published version:**

[TBD](http://www.w3.org/TR/klreq/)

**Latest editor's draft:**

[TBD](http://www.w3.org/International/docs/korean-layout/)

**Authors:**

TBD.

**Editor:**

TBD, W3C

This document is also available in this non-normative format: [Chinese version](http://www.w3.org/TR/2013/WD-klreq-20130514/korean/)

[Copyright](http://www.w3.org/Consortium/Legal/ipr-notice#Copyright) © 2015 [W3C](http://www.w3.org/)® ([MIT](http://www.csail.mit.edu/), [ERCIM](http://www.ercim.eu/), [Keio](http://www.keio.ac.jp/), [Beihang](http://ev.buaa.edu.cn/)), All Rights Reserved. W3C [liability](http://www.w3.org/Consortium/Legal/ipr-notice#Legal_Disclaimer), [trademark](http://www.w3.org/Consortium/Legal/ipr-notice#W3C_Trademarks) and [document use](http://www.w3.org/Consortium/Legal/copyright-documents) rules apply.

## Abstract

This document describes requirements for general Tibetan text layout and typography realized with technologies like CSS, SVG and XSL-FO. The document is part of Chinese Layout Requirement Task Force in W3C i18n Activity, which are focusing on language layout requirements for languages mainly used in Great China area.

## Status of This Document

This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the latest revision of this technical report can be found in the [*W3C technical reports index*](http://www.w3.org/TR/) at http://www.w3.org/TR/.

This document describes requirements for general Tibetan language text layout realized with technologies like CSS, SVG and XSL-FO. The document is part of Chinese Layout Requirement Task Force in W3C i18n Activity, which are focusing on language layout requirements for languages mainly used in Great China area. The document was originally developed by Tibetan language and typographic experts and standardization experts in China, then translated to English under the guidance of the authors. The [Chinese version](http://www.w3.org/TR/2013/WD-klreq-20130514/korean/) , and Tibetan version of this document is also available and is also the authoritative version. The Working Group expects this document to become a Working Group Note.

This document was published by the [Internationalization Working Group](http://www.w3.org/International/core/) as a First Public Working Draft. **We are looking for comments on the document before final publication. If you wish to make comments regarding this document, please send them to** [**public-i18n-clreq@w3.org**](mailto:public-i18n-clreq@w3.org) **(**[**subscribe**](mailto:public-i18n-cjk-request@w3.org?subject=subscribe)**,** [**archives**](http://lists.w3.org/Archives/Public/public-i18n-cjk/)**) before Q4 2015**. All comments are welcome.

Publication as a First Public Working Draft does not imply endorsement by the W3C Membership. This is a draft document and may be updated, replaced or obsoleted by other documents at any time. It is inappropriate to cite this document as other than work in progress.

This document was produced by a group operating under the [5 February 2004 W3C Patent Policy](http://www.w3.org/Consortium/Patent-Policy-20040205/)( ##need to verify and applied to new W3C Patent Policy##). The group does not expect this document to become a W3C Recommendation. W3C maintains a [public list of any patent disclosures](http://www.w3.org/2004/01/pp-impl/32113/status) made in connection with the deliverables of the group; that page also includes instructions for disclosing a patent. An individual who has actual knowledge of a patent which the individual believes contains [Essential Claim(s)](http://www.w3.org/Consortium/Patent-Policy-20040205/#def-essential) must disclose the information in accordance with [section 6 of the W3C Patent Policy](http://www.w3.org/Consortium/Patent-Policy-20040205/#sec-Disclosure).

## Table of Contents

TBD.

## 1. Introduction

### 1.1 Purpose of this Document

Every cultural group has its own language and writing system. Especially between East and West, the difference is great. Thus to digitize the writing system, a lot of data and technology is needed that accurately represents the language and writing system.

The purpose of this document is to explain the important differences between the Tibetan writing system and others, in order to digitize it. However this document does not provide the solution for actual implementations or problems, but only describes the basic information that should be treated as important issues.

### 1.2 How this Document was Created

This document was created by discussing the Tibetan language related issues through China’s National Standards and surveying needs from actual users and technical experts.

The following types of experts were involved in the creation of this document:

1. Tibetan typography experts
2. International and Domestic (China) Standardization Experts
3. Academic and Industry Experts

As part of Chinese Layout Requirement Task Force of i18n activity, the contents of this document are related to the Tibetan writing system mainly in China and other related area in the world. Based on existing China national standards on Tibetan language character sets, encoding, typography and layout rules, all the discussions are took place in Chinese or Tibetan, and were then translated into English once the document had been finished.

The technical terms for discussing and describing the Tibetan writing system were carefully selected after considering potential differences in nuance even if a direct translation exists, and in some parts they were described in both languages so the discussion can be continued in the future. Also many figures are included to promote understanding for certain parts that are hard to describe in English.

### 1.3 Basic Principles for Document Development

This document describes the characteristics of the Tibetan language system along the lines of the following principles.

1. It does not cover every issue of Tibetan typography, but only the important differences from the Western language systems.
2. All text in the figures is in Tibetan, but the technical aspects of actual implementation are not covered by this document.
3. In order to help readers' understanding of how Tibetan is used, typical real life examples are provided to explain the core combination properties.
4. Text layout rules and recommendations for readable design are different matters, but it is hard to discuss these two aspects separately. In this document, these two issues are separated carefully.

### 1.4 The Structure of this Document

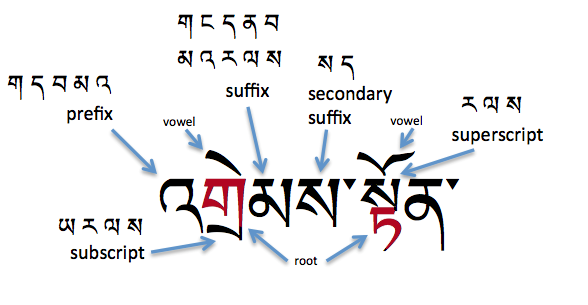
This document consists of six parts as below:

1. Introduction
2. Tibetan Character and fonts
3. Character typography
4. Paragraph typography
5. Components besides paragraphs
6. Page layout

## 2. Tibetan Language Overview

藏文是拼音字母文字，由辅音、元音、下标字、上标字等基本字母按照藏文拼写规律组合而成，形成藏文的字符和音节。

藏文的结构元素包括前缀（Prefix），根（root），下标（subscript），上标（superscript），后缀(suffix)，后后缀（secondary suffix），元音(vovel)及音节符（Intersyllabic mark）。



### 2.1 藏文字母（Tibetan Alphabet）

藏文字母（Tibetan alphabet）是用于拼写一个藏文字符或音节的辅音和元音所对应的图形符号。具体包括30个辅音字母及其对应的附加辅音字母、4个元音字母、4个下标字、3个上标字。此外，藏文也有自己的数字表示方法及对应的字母。

### 2.2 藏文字符（Tibetan Character）

藏文字符（Tibetan character）是由藏文字母纵向叠加而成的结构稳定的符号。一个藏文字符可能包括多个辅音，最上面的辅音称为主辅音，其他辅音称为附加辅音（Subjoined consonant）。

例如：རྐ་ སྤྱ 主辅音

附加辅音

### 2.3 藏文音节（Tibetan Syllable）

藏文是拼音文字，因此它的构字方式采用音节分写，不采用像汉语类词式分写。

#### 2.3.1 音节构成

一个藏文音节由一个或多个藏文字符组成，音节之间加音节间隔符（Intersyllabic mark），藏文排版不能对音节进行切分。

例如：གསལ་པོ་ /csa/ 清晰 其中有音节间隔符即隔音符“་”/tsek/ 作为区分两个音节。

#### 2.3.2 结构规则

要分析藏文结构必须先得找出根字母，然后其他的部分根据结构规则就能找到。根字母的判断方法如下：

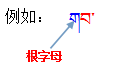
1. 一个辅音上有元音字母，那就是根字母，除非是如下上面是元音字母下面是辅音字母，在此中是根字母。



1. 一个辅音上有上标字或者下标字那么这个辅音字母也是根字母。



在一个具有两个辅音字母且没有元音字母的字中，第一个字母是根字母。



1. 如果一个字是由三个辅音字母构成且最后一个字母若不是 ****，则中间的那个字母即第二个最可能是根字母。

例如：  

但是当末尾出现时中间的字母不一定是根字母。例如： **ལགས་**/lag/ 从这个发现最末尾是**ས**所以根字母是第一个字母。

1. 在一个具有四个辅音字母的字中，第二个辅音字母是根字母。

* ****即前缀（prefix），放在第一个位置，即放在根字母前面，通常有五个字母，分别是ག ད བ མ འ

即根字母（root）一个音节的中心字母。

即后缀（suffix）放在根字母后，通常有10个后缀字母。如下十个： 

即后后缀（secondary suffix），放在后缀字母后的字母。通常只有两个 。即： 

即上标字（superscript），放在根字母上面，通常有三个字母 。即： 

即下标字（subscript），放在根字母下面。即右边两个： 

即元音字母（vowel），标准的藏文有五个元音，其中一个是发声固定的音 /a/，此音在通常带在各辅音字母后面，所以不需要加，除非文字结构需要加，其他的四个元音为非固定发声音，元音字母跟辅音字母通过上下形式相叠加，产生特定的音同时与其他相对结构较近似字区分。例如： /tji/(一般)

### 2.4 藏文的字体与字符的Unicode编码

藏文目前的字形字体为喜马拉雅字体。根据《GB？？？》，藏文字符编码覆盖了Unicode的U+0F00至U+0F0E、U+0F10至U+0F1E、U+0F20至U+OF2E U+0F30+范围。具体如下：



## 3. Typography for Tibetan Characters

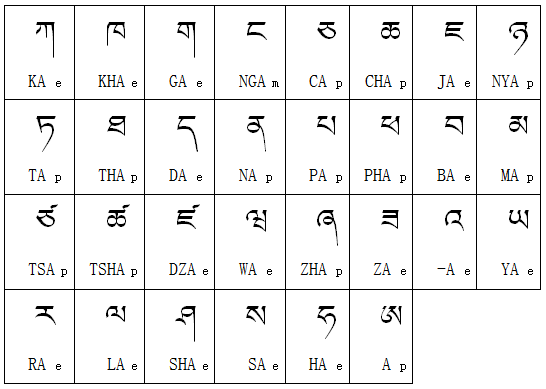
### 3.1 藏文的字符集分类（Character Category）

TBD

### 3.2 辅音与元音

#### 3.2.1 辅音字母（consonant）

藏文有三十个辅音字母，具体如下表所示：



#### 3.2.2 附加辅音字母（Subjoined consonant）

在现代信息处理过程中还有增加了附加辅音字母，辅音字母在字符的上下中位置不同，可能存在不同的编码，一般最上面的辅音称为主辅音，其他辅音称为附加辅音。在ISO/IEC 10646中，主辅音和附加辅音拥有不同的编码。

附加辅音在藏文的实际意义跟主辅音字母一致。

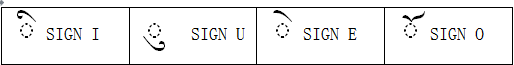
例如： ྐ加上其他字母后变形为叠加字如རྐ

每个辅音所对应的附加辅音的变形方式如下表：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ྐ  KA e | ྑ  KHA e | ྒ  GA e | ྔ  NGA e | ྕ  CA p | ྖ  CHA p | ྗ  JA e | ྙ  NYA p |
| ྟ  TA p | ྠ  THA p | ྡ  DA e | ྣ  NA p | ྤ  PA p | ྥ  PHA p | ྦ  BA e | ྨ  MA p |
| ྩ  TSA p | ྪ  TSHA p | ྫ  DZA e | ྺ  WA e | ྮ  ZHA p | ྯ  ZA e | ྰ  -A e | ྱ  YA e |
| ྲ  RA e | ླ  LA e | ྴ  SHA e | ྶ  SA e | ྷ  HAe e | ྸ  A e |

#### 3.2.3 元音字母

藏文中，共有4个元音字母，如下表。



元音字母禁止单独使用，单个也不能表达意思。跟辅音字母使用后产生特定的音。它们需要和辅音（或附加辅音）联合起来表达音节和含义。

例如：上述四个元音字母与辅音ྐ联合使用：ཀི ཀུ ཀེ ཀོ

### 3.3 下标字与上标字

#### 3.3.1 下标字

藏文中，有4个下标字，如下表：



这些下标字加在辅音下方，辅助产生音。例如：



#### 3.3.2 上标字

藏文中，有3个上标字，分别为ར ལ ས。这三个字母也可作为上标字跟辅音字母叠加产生特定形式的字以及音。例如：



### 3.4 后缀与后后缀

TBD

### 3.5 藏文的数字

藏文有自己的数字表示。藏文数字跟阿拉伯数字本质上没有不同，只是符号表示形式不同而已，但是在藏文中增加了半数字的表示形式。

藏文的半数字是指1到9数字一半的表达方式。例如：1的一半是0.5，2的一半是1，3的一半是1.5等等。在藏文中，可以通过半数字直接表达而不用书写小数树。例如，43.5在藏文中可以直接书写为 ༤༬。

下表是藏文数字（及半数字）与阿拉伯数字表示方法的对应：



藏文的文章或文字在表示数量大小的时候，通常可以使用阿拉伯数字进行表示。一般来说，藏文的半数字较少使用。但在大多数藏文书籍页码页脚等场合中，也可以使用阿拉伯数字代替藏文数字。

### 3.6 藏文的标点符号

藏文标点符号，主要为“分界线”或“分界符号”。藏文标点符号形体简单、种类极少，而且，其使用规则也与其他文字的标点符号有别。

藏文标点符号共有6种形式：云头符、隔音符号、云头符、单垂符、双垂符、聚宝垂符等。其中：音节之间的隔音符号使用频率最高。此外，还有云头符，用于书题或篇首；蛇形垂符，用于文章开头处；单垂符，用于短语或句终；双垂符，用于章节末尾；四垂符，用于卷次末尾；聚宝垂符等。

#### 3.6.1 云头符

云头符“”必须出现在每篇文章题目的起始或者每本书名的起始位置。表示文章及作品从这里开始，也起一个修饰作用（预示写作能顺利完成）。如果不加云头符，我们就不知道文章是从何开始。

#### 3.6.2 音节符

音节符 ，加在每个音节之后，起划分音节的作用，让读者正确读出文字。如果缺少音节符文字混论无法读出正确的字，也不构成字。

藏文音节符必须出现在每个音节之后，才能构成一个完整的音节。在排版时，每行都是由完整的音节组成，行末音节不能拆分。藏文的行首不能出现音节符。当最末尾的一个音节不能完整排下时，要把该音节移到下一行，同时该行末用音节符补齐，如果需要补齐的音节点较多而影响排版美观时，则把该行段句符距离拉大进行调节。

#### 3.6.3 单垂符

单垂符通常出现在一段文字的末尾表示该文字段是一个句子，如果该句子恰在一段的末尾，表示该段落。表示一个句子或者段落的结束。单垂符出现在词语、短语或者句子末尾，相当于汉语的顿号、逗号、句号、叹号等。

在一个词语、短语或者句子最后一个音节的末尾出现 时，不写单垂符。在一个词语、短语或者句子最后一个音节的末尾出现“”时，必须在单垂符前加上一个音节符。排版的时候需要注意“”的位置，如果“”出现在其所在的音节中间，虽在音节末尾，但在句子中间时，不存在上述规则。

#### 3.6.4 双垂符

双垂符 跟单垂符用途基本相同，但是双垂符只用在段落的末尾或者文章的最末尾处。一般不用在句子的末尾。

双垂符“”出现在一个自然段的末尾和诗歌中每行的结尾处，表示段或诗歌行的结束。在自然段的末尾和诗歌中每行的结尾处出现“”时，不能使用双垂符号，而改用单垂符；在自然段的末尾和诗歌中每行的结尾处出现“”时，必须在双垂符号前加上一个音节符。

#### 3.6.5 四垂符

四垂符“”出现在长篇文章的结尾，表示文章的结束。在长篇文章的结尾处出现“”时，只需要出现“”。

#### 3.6.6 聚宝垂符

聚宝垂符“”主要用来断句，藏文印刷体书籍的一句尾，恰在一行的右端起首，不满三个字母者均该用聚宝垂符作句符。其最末一字为语终词，以及为颂歇句末尾者，则均用两个聚宝垂符。

#### 3.6.7 蛇行垂符

蛇行垂符“”用于段落层次之首，但用此符号后不能另行。

#### 3.6.8 其他符号

1. 增亮符

增量符，表示强调，增亮。

2. 

…

## 4. Typography for Tibetan Paragraphs

TBD。

### 4.1 藏文的书写方式

藏文文字为“从左至右”书写、“从上向下”显示，也简称为横排从左至右排版。

### 4.2 段落调整（Paragraph Adjustment）

### 4.3 标点符号的书写规则（Writing Rules of Punctuation Marks）

### 4.4 行与分行规则（Line and Line Breaking Rules）

### 4.5 多文种混排中的行布局规则

## 5. Components Besides Paragraphs

段间组件的布局规则。

### 5.1 Notes (Footnote, Endnote)

TBD.

### 5.2 Tables

TBD.

### 5.3 Images

TBD.

### 5.4 Lists and Numbering

TBD.

### 5.5 Other Components

科学公式?

## 6. Page Layout

### 6.1 Designing the Tibetan Text Layout System

The Tibetan text layout system is approached from two perspectives: overall page layout design and page body design.

### 6.2 Elements of Designing Page Format

TBD.

### 6.3 Designing the Page Body

TBD.

### 6.4 Running Heads and Page Numbers

Running heads and page numbers are placed outside of the page body, and typical positions are as shown below.

TBD.

### 6.5 Pagination and Multi-Column Text

分页与多栏文本.

## 7. Conclusion

TBD.

**Appendix A Terminology**

报告中用到的专用词汇及其英文、汉字、藏文对照，及含义。TBD。

| **Terminology** | **Chinese  Tibetan** | **Romanized transliteration** | **Definition** |
| --- | --- | --- | --- |
| back matter | 後付 | atozuke(拼音或发音) | Appendices, supplements, glossary of terms, index and/or bibliography, and so on, appended at the end of a book. |
|  |  |  |  |

**Appendix B References**

报告中引用到的中国国家标准，或其他标准文本

* China National Standards. 《信息技术信息交换用藏文编码字符集基本集》
* China National Standards. 《信息技术藏文编码字符集点阵字形-第一部分：白体》
* China National Standards. 《信息技术- 藏文编码字符集扩充A》
* China National Standards. 《藏文编码字符集键盘字母数字区的规划》
* TBD

**Appendix C Revision Log**

This is the first public working draft of the W3C Working Group Note. It is public to open as a W3C Working Group Note on XX. 20XX. The major modifications and additions to version 1 are as follows:

* TBD.

**Appendix D Acknowledgements**

This document has been developed with contributions from participants of the [Chinese Layout Requirement Task Force](http://www.w3.org/2007/02/japanese-layout/), with kind help from experts from 信标委中文信息处理分技术委员会及藏文信息处理工作组.