

Requirements for Chinese Text Layout

《中文排版需求》的进展

—— 中文字体排印需求的现状 ——

新一代 Web 技术标准论坛

W3C 2022 年度全球技术大会中国会场

2022 年 9 月 6 日 · 杭州

钱争予

《中文排版需求》编辑

q@laerhsif.com

汉字为核心的
书写系统

字体排印

中文排版需求

字 · 体 · 排 · 印

Typography

Requirements for Chinese Text Layout

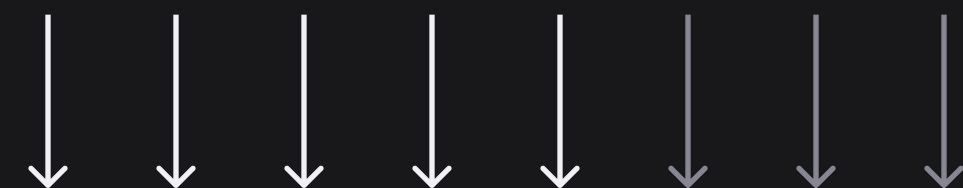
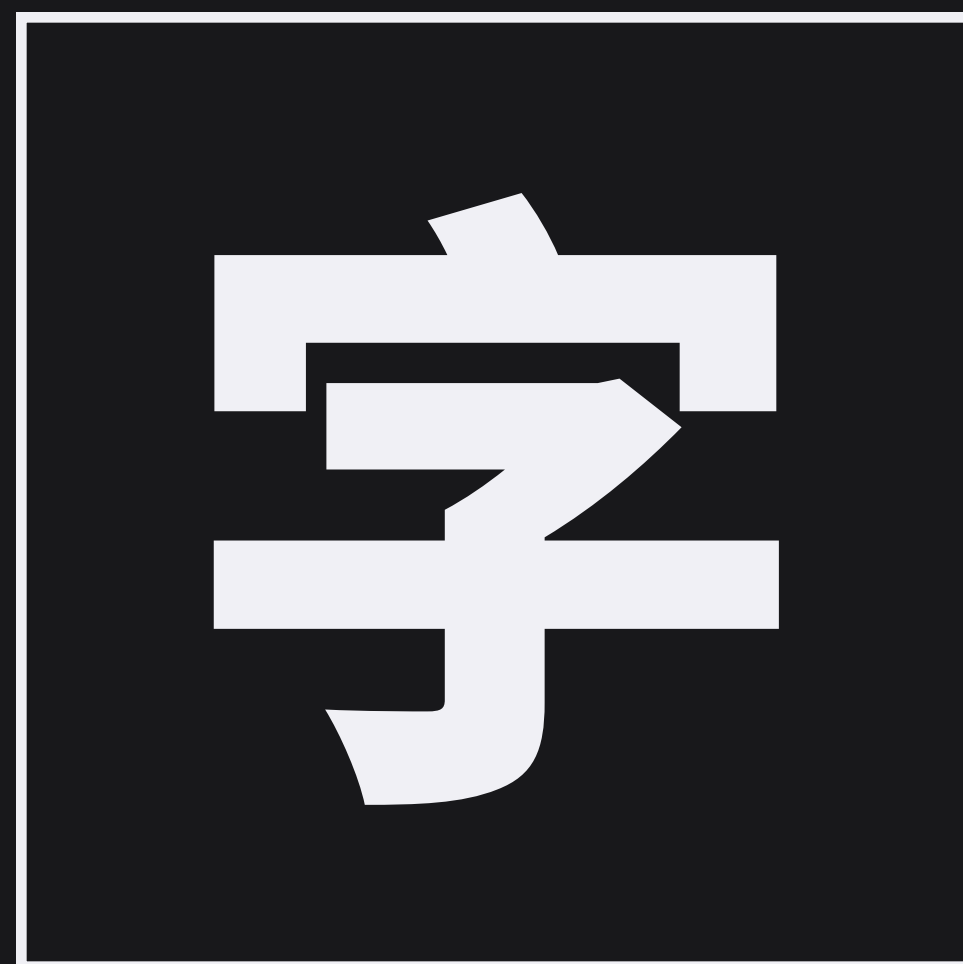
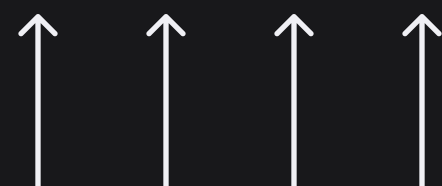
《中文排版需求》的进展

—— 中文字体排印需求的现状 ——

文字

字

宋 黑 仿 楷



陆 港 澳 台 星 日 韩 越

字形

Glyph

抽象字符

Abstract character



NOTE

Chapter 6.2 of v14 of the Unicode Standard suggests using U+22EF MIDLINE HORIZONTAL ELLIPSIS […] for ellipses.

在Unicode标准第14版的6.2章中，也推荐使用U+22EF MIDLINE HORIZONTAL ELLIPSIS […]作为省略号。

在Unicode標準第14版的6.2章中，亦推薦使用U+22EF MIDLINE HORIZONTAL ELLIPSIS […]作為刪節號。

This code point, with appropriate rotation and replacement mechanism, will be centered within its character frame regardless of whether it is in vertical or horizontal writing mode, which is more in line with the Chinese typesetting requirements.

此符号配合适当的转向与取代机制，在显示上无论直横排，省略点皆居中，更符合排版需求。

此符號配合適當的轉向與取代機制，在顯示上無論直橫排，刪節點皆居中，更符合排版需求。

5. Emphasis Dots.

着重号

着重號

Emphasis dots are symbols placed above or beneath characters to emphasize the text, strengthen the tone, or avoid ambiguity. For horizontal writing mode, emphasis dots are placed under the characters, whereas in vertical writing mode, they are usually placed to the right side of the characters. Both U+25CF BLACK CIRCLE [●]

NOTE

The *General Rules for Punctuation* (GB/T 15834—2011) does not state the corresponding Unicode code point for the three types of connector marks. However, we can make the deduction that the long connector mark [—] is U+2014 EM DASH [—] and the tilde [~] is U+FF5E FULLWIDTH TILDE [~]. Since the short connector mark should take half the width of the long connector mark, it should be U+2013 EN DASH [–]. The actual length of these connector marks may depend on the writing system as well as the typeface.

《标点符号用法》（GB/T 15834—2011）中没有指定这三个符号的码位，但是基本上可以推断一字线是U+2014 EM DASH [—]，浪纹线是U+FF5E FULLWIDTH TILDE [~]。但是对于“短横线”，该标准5.1.6节规定“短横线比汉字『一』略短，占半个字位置”，因此可以是 U+2013 EN DASH [–]。这些连接号的实际长短根据所用处理系统和使用字体会有区别。

《標點符號用法》（GB/T 15834—2011）中沒有指定這三個符號的碼位，但是基本上可以推斷一字線是U+2014 EM DASH [—]，浪紋線是U+FF5E FULLWIDTH TILDE [~]。但是對於「短橫線」，該標準5.1.6節規定「橫短線比漢字『一』略短，佔半個字位置」，因此可以是U+2013 EN DASH [–]。這些連接號的實際長短根據所用處理系統和使用字體會有區別。

7. Interpuncts.

间隔号

間隔號

NOTE

Due to the fact that Big5 code does not give a detailed definition of interpuncts, sometimes U+FF0E FULLWIDTH FULL STOP [·], U+2027 HYPHENATION POINT [·] or U+2022 BULLET [•] are used. It is recommended to use U+00B7 MIDDLE DOT [·] altogether. U+30FB KATAKANA MIDDLE DOT [·] is from JIS code, and is therefore not recommended.

过去因大五码未有详细的语意定义，所以有时混用U+FF0E FULLWIDTH FULL STOP [·]、U+2027 HYPHENATION POINT [·]、U+2022 BULLET [•]等字符作为间隔号的例子，建议使用U+00B7 MIDDLE DOT [·]。而U+30FB KATAKANA MIDDLE DOT [·]来自日文JIS编码，并非中文编码，不建议使用。

過去因大五碼未有詳細的語意定義，所以時有混用U+FF0E FULLWIDTH FULL STOP [·]、U+2027 HYPHENATION POINT [·]、U+2022 BULLET [•]等字元作為間隔號的例子，建議使用U+00B7 MIDDLE DOT [·]。而U+30FB KATAKANA MIDDLE DOT [·]來自日文JIS編碼，並非中文編碼，不建議使用。

8. Book title marks.

书名号

書名號

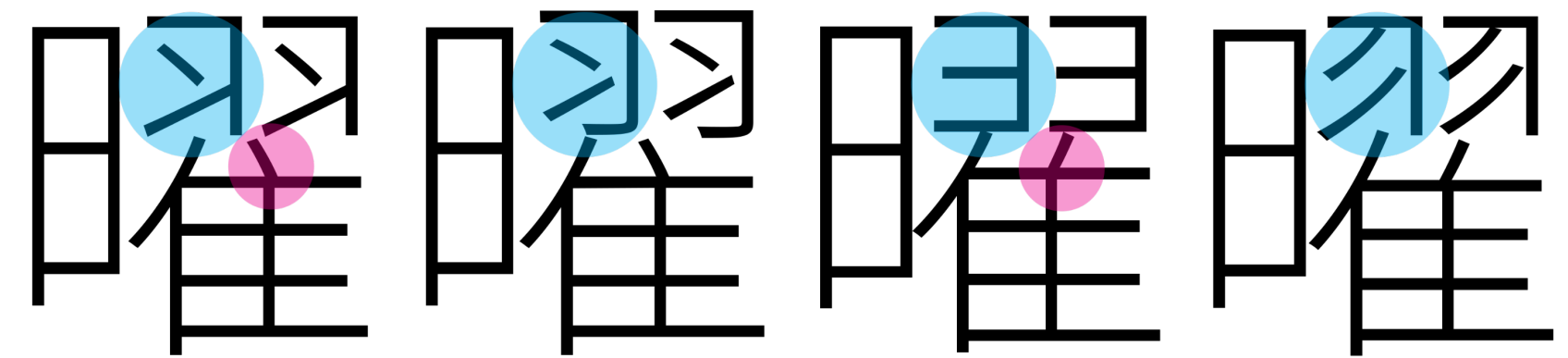
According to *Punctuation Guidance* (revised edition in 2008) issued by the Ministry of Education in Taiwan, book title marks are used to indicate the names of works which usually include books, articles, songs, movies, files, calligraphy and paintings. Generally, there are two types of book title marks: wavy low lines or angle brackets.

字体

体

思源字体项目
Source Han / Noto CJK

字形笔画规范的 地区差异处理



资料来源：Source Hans Sans ReadMe
github.com/adobe-fonts/source-han-sans

思源字体项目
Source Han / Noto CJK

标点符号横竖排
及地区差异处理

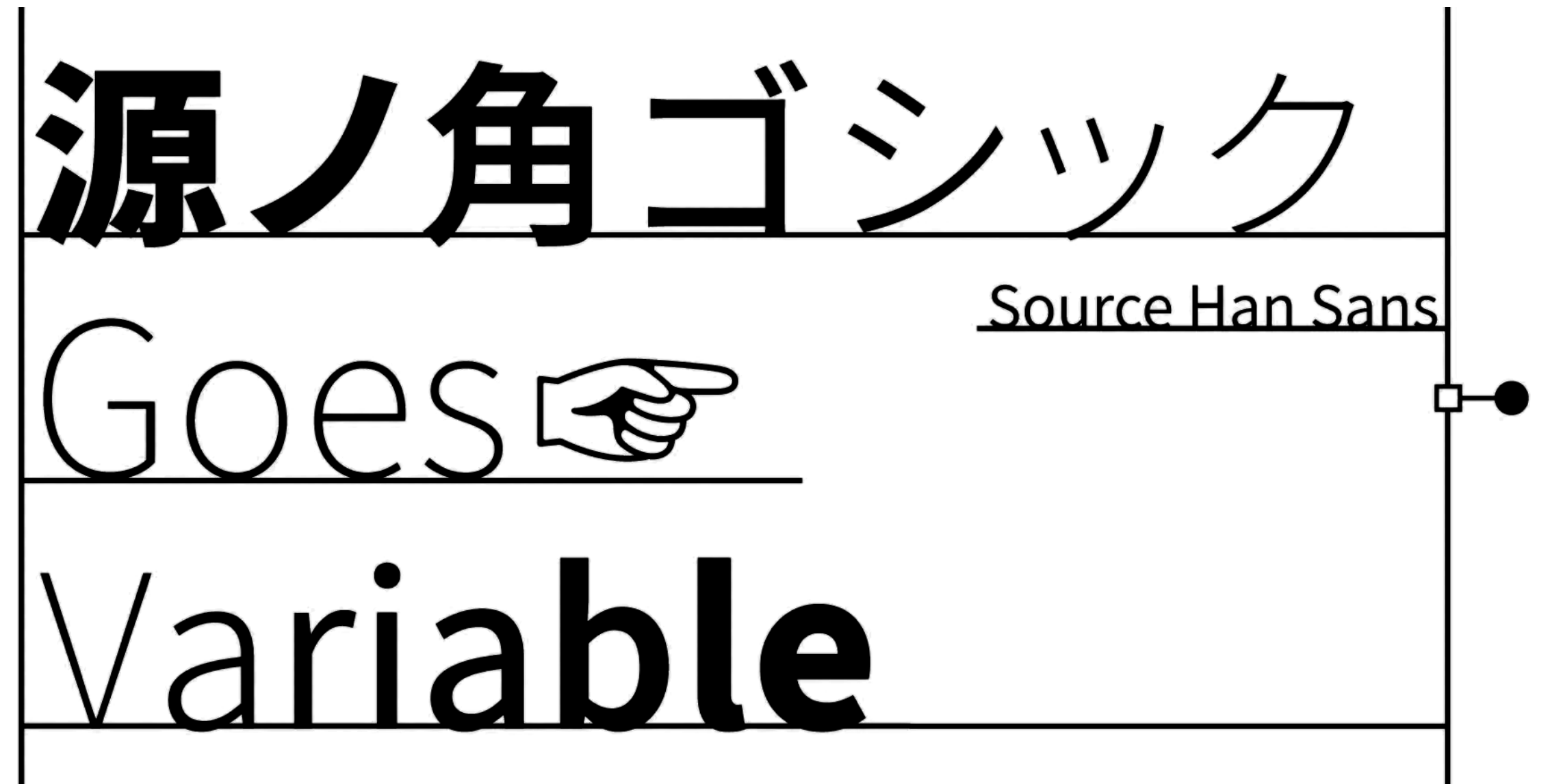
资料来源：Source Hans Sans ReadMe
github.com/adobe-fonts/source-han-sans

Unicode	Simplified Chinese	Traditional Chinese	Japanese	Korean
U+2018 ¹	‘ → ﹀	‘ → ’	‘	‘
U+2019 ¹	’ → ﹁	’ → ’	’	’
U+201C ¹	“ → ﹂	“ → ”	“	“
U+201D ¹	” → ﹃	” → ”	”	”
U+3001	、 → 、	、	、 → 、	、 → 、
U+3002	。 → 。	。	。 → 。	。 → 。
U+FF01	！ → ！	！	！	！ → ！
U+FF0C	， → ，	，	， → ，	， → ，
U+FF0E	． → ．	．	． → ．	． → ．
U+FF1A	： → ：	：	： → ．	： → ：
U+FF1B	； → ；	；	；	； → ；
U+FF1F	？ → ？	？	？	？ → ？

Unicode	VS1 (U+FE00)—Corner-Justified Form	VS2 (U+FE01)—Centered Form
U+3001	あ 漢、 漢あ	永、 永
U+3002	あ 漢。 漢あ	永。 永
U+FF01	漢！ 漢	あ 永！ 永あ
U+FF0C	あ 漢， 漢あ	永， 永
U+FF0E	あ 漢． 漢あ	永・ 永
U+FF1A	漢： 漢	あ 永： 永あ
U+FF1B	漢； 漢	あ 永； 永あ
U+FF1F	漢？ 漢	あ 永？ 永あ

思源字体项目
Source Han / Noto CJK

可变字体技术



资料来源: Adobe Blog
blog.adobe.com/en/publish/2021/04/08/source-han-sans-goes-variable

CSS Fonts Module
Level 4



serif

宋

sans-serif

黑

fangsong

仿

cursive?

楷



语文教科书

宋

關關雎鳩在河之洲
窈窕淑女君子好逑

臺灣教育部標準宋體

关关雎鸠在河之洲
窈窕淑女君子好逑

方正规范书宋

關關雎鳩在河之洲
窈窕淑女君子好逑

Adobe 明體 Std Light

关关雎鸠在河之洲
窈窕淑女君子好逑

思源宋体（简体中文）Bold

關關雎鳩在河之洲
窈窕淑女君子好逑

思源黑體（繁體中文—臺灣）Regular

关关雎鸠在河之洲
窈窕淑女君子好逑

华文细黑

關關雎鳩在河之洲
窈窕淑女君子好逑

Hiragino Sans CNS W6

关关雎鸠在河之洲
窈窕淑女君子好逑

汉仪旗黑 80

黑

仿

關關雎鳩在河之洲
窈窕淑女君子好逑

華康仿宋體 W4

关关雎鸠在河之洲
窈窕淑女君子好逑

华文仿宋

關關雎鳩在河之洲
窈窕淑女君子好逑

文悅古體仿宋

关关雎鸠在河之洲
窈窕淑女君子好逑

方正聚珍新仿

關關雎鳩在河之洲
窈窕淑女君子好逑

臺灣教育部標準楷書

关关雎鸠在河之洲
窈窕淑女君子好逑

方正教材规范楷体

關關雎鳩在河之洲
窈窕淑女君子好逑

蒙納香港特粗楷二

关关雎鸠在河之洲
窈窕淑女君子好逑

华文楷体（简）黑

楷

国字标准字体

《通用规范汉字表》字形

關關雎鳩在河之洲
窈窕淑女君子好逑

臺灣教育部標準宋體

关关雎鸠在河之洲
窈窕淑女君子好逑

方正规范书宋

仿古（聚珍仿宋）正方骨架

新长仿宋骨架

關關雎鳩在河之洲
窈窕淑女君子好逑

Adobe 明體 Std Light

传承字形

喇叭口→

关关

W4

纯等线→

关关

河之洲
君子好逑

古體仿宋

关关雎鸠在河之洲
窈窕淑女君子好逑

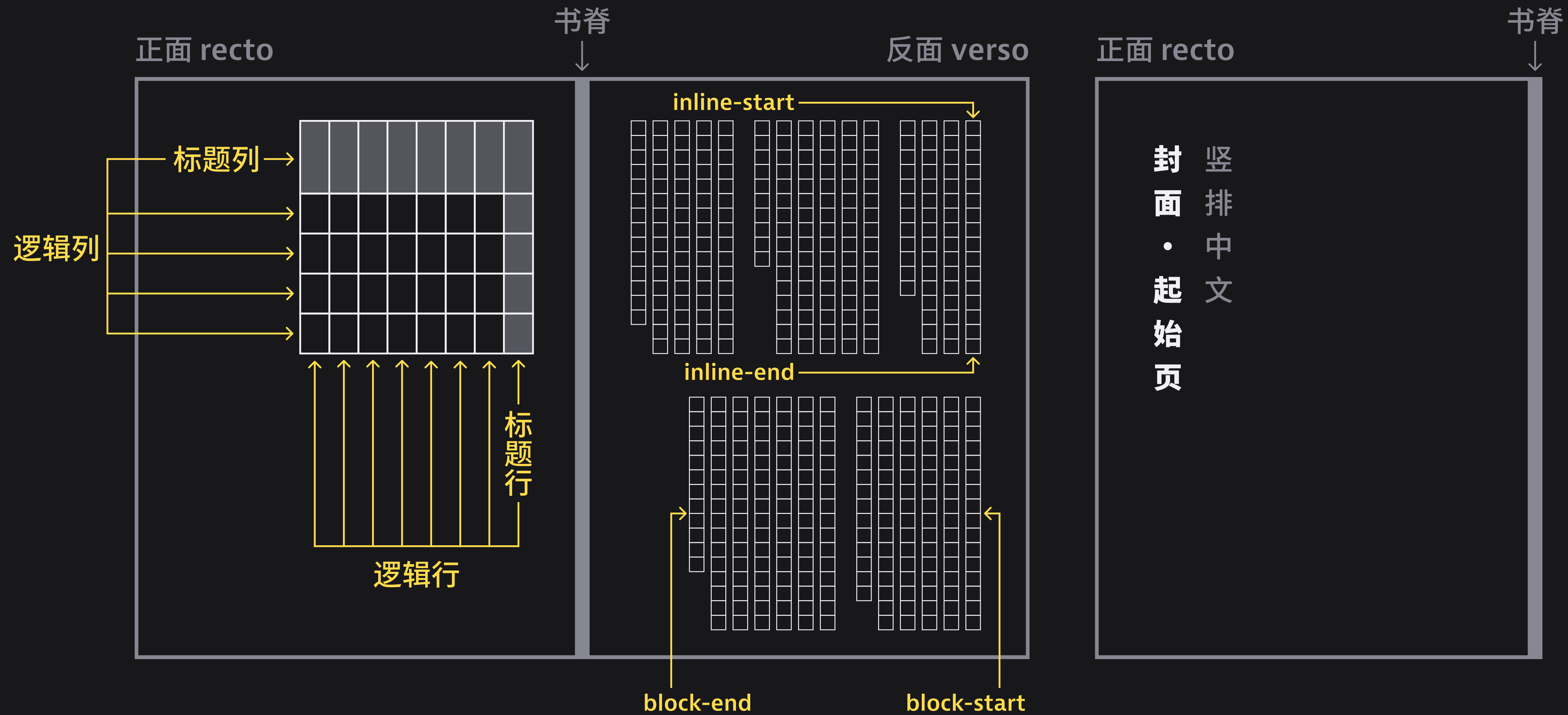
华文仿宋

关关雎鸠在河之洲
窈窕淑女君子好逑

方正聚珍新仿

排版

排



标题的地位

标题摆放的位置及其所占空间，应以基本版式设定中行的位置为基准，以“占几行空间”方式来设计，这个方式称作“占行”。实际定版工作中为计算方便，通常设置为行的整数倍。定版要求通常写成“**单行标题占三行**”或“**双行标题占五行**”等方式，“单行”“双行”指标题文字本身的行数，而“三行”“五行”则指的是基本版式中行所占的空间。标题缩排时，也应基本版式所设定的文字尺寸为基准，决定要缩排的量。

CSS Rhythmic Sizing

看到了伊夫斯彼得斯在推特上询问图中尖刺的正式名称叫什么，我简短地回答了一下（剧透，这叫「光陷阱」）。我觉得在这个话题上还有更多值得探索，因此决定写点东西，不只关于光陷阱，还会包含类似的概念和造型。

起初 神创造天地。地是空虚混沌。渊面黑暗。神的灵运行在水面上。

起初上帝创造天地。地是空虚混沌。渊面黑暗。上帝的灵运行在水面上。

記第九屆國際印刷品展覽大會

佛拉色著
劉龍光譯

印刷品展覽大會，(The Ninth Printing, Stationery and Allied Exhibition) 於一九三六年十一月二日開幕，同年十二月五日閉幕。會之久，每日的觀眾達數萬人。這個價值，是可以想像的。本文是美國誌(The Inland Printer)主編佛氏(L. Frazier)所撰的。佛氏不惜長渡大西洋而到了倫敦，將考察所國印刷界。本文登載於佛氏主編印刷雜誌，原名為“A Jolly Good

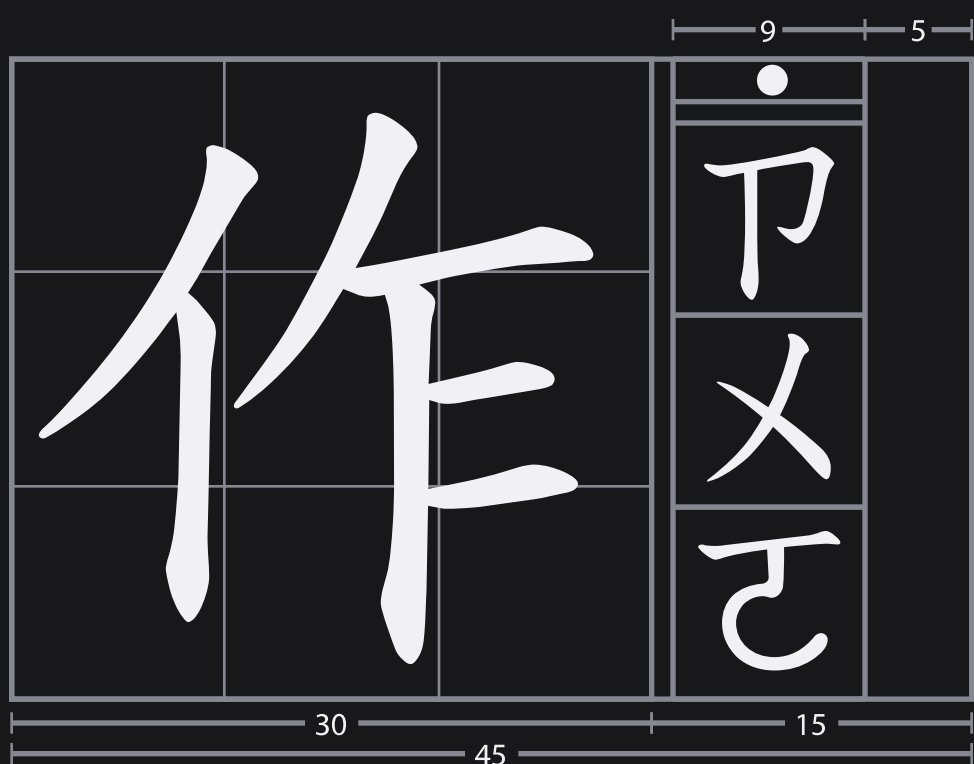
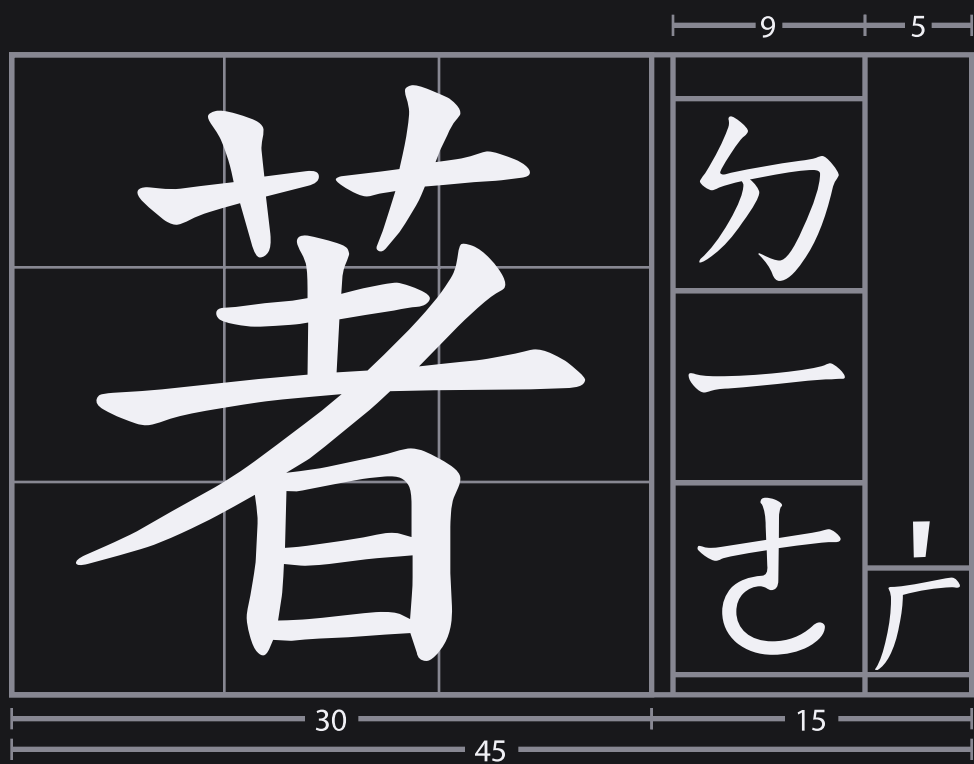
Show at Olympia”從一九三七年一月號起，陸續披露。譯者認為這個具有世界性的印刷品展覽大會，完全可以象徵現代印刷術進步的程度。在今日，歐洲各國的印刷術，究竟進步到一種什麼樣的程度，我想這是中國印刷界所十二分願意知道的。所以就佛氏的原文，譯在這裏。

譯者識

在一天的中間，能看見到五萬九千個印刷者嗎？能的，祇要你肯把一月二十八日(星期六)這一整天的功夫，消磨在英國倫敦亞令配亞(Olympia)展覽大會場的入口上。照實講，上面數碼字還祇包括着那

《国史大纲（修订本）》，商务印书馆 1996 年版

《艺文印刷月刊》影印资料
1937—1940 年间



资料来源：萌娘百科

zh.moegirl.org.cn/写作oo读作xx

GitHub w3c/clreq: #349 · #355 · #358 · #463

写作oo读作xx

萌娘百科，万物皆可萌的百科全书！转载请标注来源页面的网页链接，并声明引自萌娘百科。内容不可商用。

写作oo读作xx（{{ruby|oo|xx}}，效果^{xx}oo）指日文书写中旁注标记的用法，多指oo的正确含义为xx。

在漢字文化圈，古人讀經史時，碰到孔子的名字「丘」時，會將之讀作「某」，以示避諱，就連寫的時候也會缺筆，故這可說是「寫作『孔丘』，讀作『孔某』」是也。^{[1][2]}

旁注标记

日文中的旁注标记被广泛地用来标注表意文字（尤其是非常用字）的读音，其适用面比中文的汉语拼音广泛。因标记文字和被标记文字可以自由组合，因此可以作出读音和文字不同义的表达效果。

如右例，新闻的标题为「黒の騎士団の決起事件」（くろのきしだんのけっきじけん），但该处标注为「ブラックリベリオン」（Black Rebellion）。

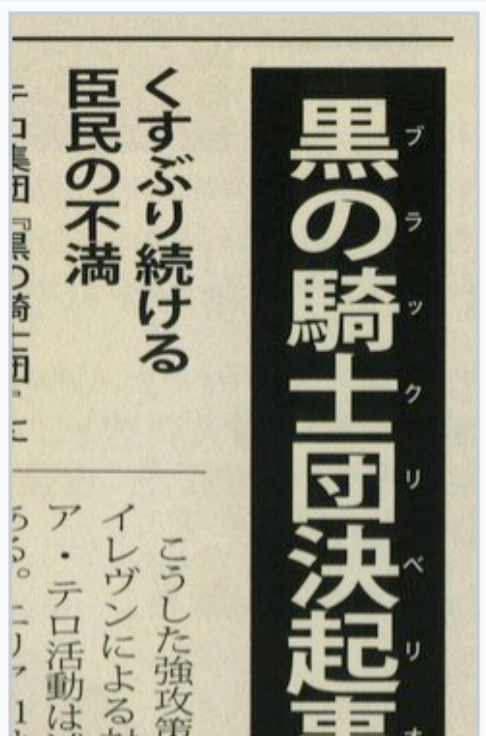
事实上，日语这种表述方法有着悠久的历史，在日语中被称作为**借字**，最典型的事例是：

- 平假名用作介词时：は^{wa}（原读ha）、へ^e（原读he）、を^o（原读wo）
- 在**游戏王**卡片中尤为常见（写作**青眼白龙**读作Blue-Eyes White Dragon）
- 《**Fate**系列》中大多数英灵的宝具也有使用这种表述方法（誓约胜利之剑^{Excalibur}、三千世界^{三段击}等）
- 《**约会大作战**》中，精灵的灵装以及天使/恶魔同理（如绝灭天使^{Metatron}、救世魔王^{Satan}、**鏖杀公**^{Sandalphon}等）
- 《**魔法禁书目录**》中的一些称谓以及剧集标题（如一方通行^{Accelerator}、超电磁炮^{Railgun}、魔女狩猎之王^{Innocentius}等）
- 《**JOJO的奇妙冒险**》中的一些替身名（如星の白金^{スタープラチナ}、**隠者の紫**^{ハーミットパープル}、**魔術師の赤**^{マジシャンズレッド}等）
- 姓氏：四月一日^{わたぬぎ}、五月七日^{つゆり}
- 作品名：《**天体のメソッド**》《**恋する小惑星**^{アステロイド}》



《搞笑漫画日和》熊吉（左）

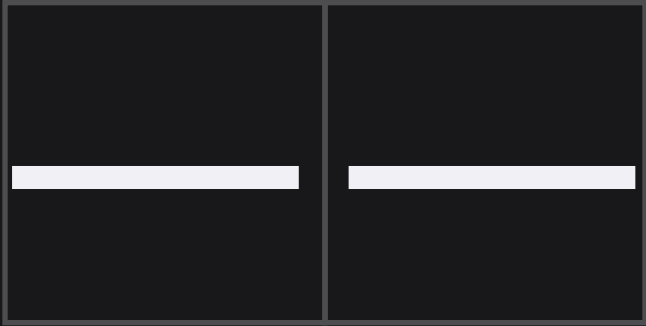
基本资料	
用语名称	写作oo读作xx
用语出处	《xxxholic》中主角的名字。写作「四月一日」，却读作「わたぬぎ」（WA TA NU KI）
相关条目	^{biàn tài} 绅士、濑户灿、xxxholic、借字、CLANNAD是人生



印刷 呈现 渲染

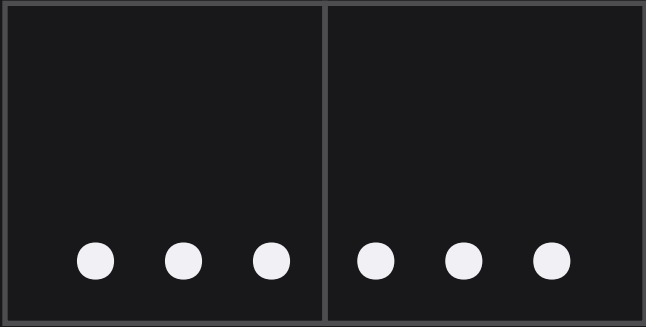
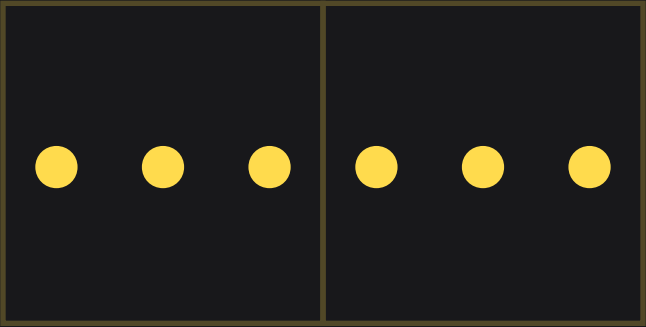
印

破折号



Em dash

省略号



Ellipsis

引号



Quotes

短横线

U+002D
HYPHEN-MINUS

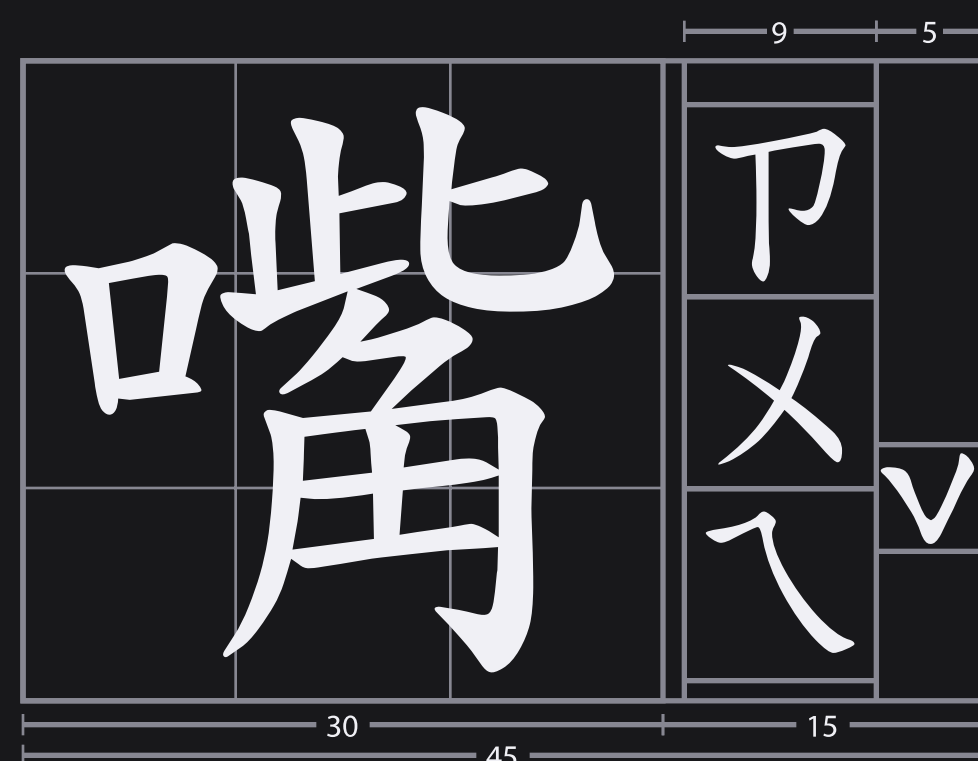
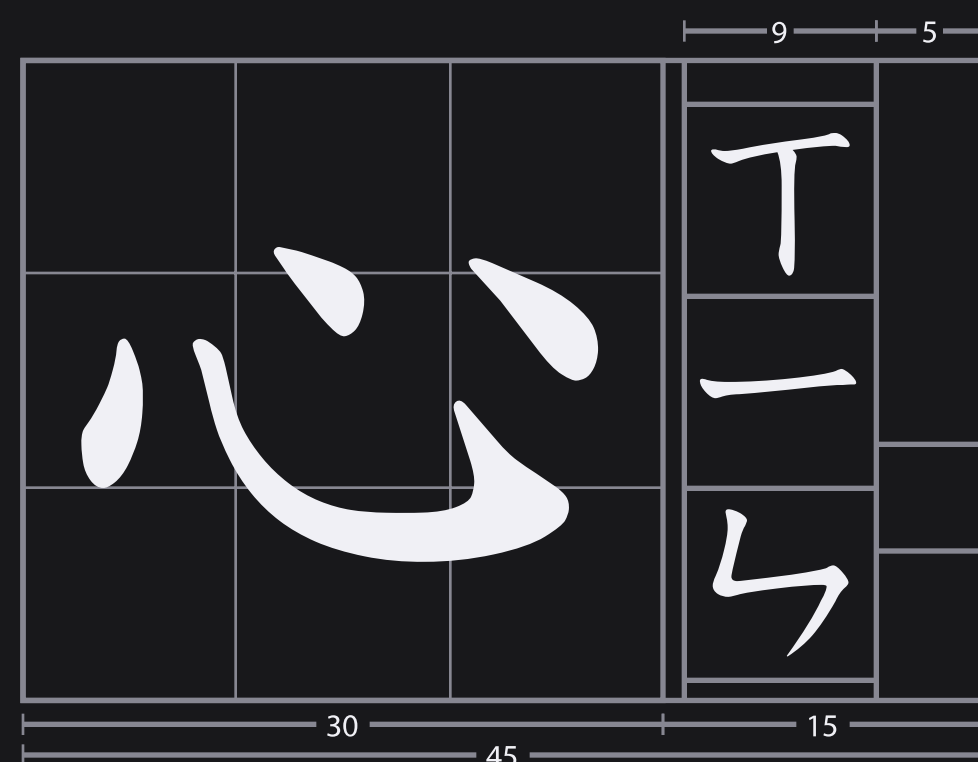
U+2013
EN DASH

一字线

U+2014
EM DASH

间隔号

U+00B7
MIDDLE DOT



Text to Speech of Electronic Documents Containing Ruby: User Requirements

W3C Editor's Draft 03 March 2022
w3c.github.io/ruby-t2s-req

One more thing

术语 · 翻译

There is another method of specifying the type area that uses **line height** rather than line gaps. Line height is the distance between two adjacent lines measured from their reference points. The reference point differs from implementation to implementation. For example, in vertical writing mode, the horizontal center of the character frame can be used, and in horizontal writing mode, the vertical center of the character frame can be used. When the character size is the same for every character, the following calculation is used:

指定版心的方法中，有着不以行距，而是依**行高**来设定的方法。行高就是彼此邻接的两行之基准点间的距离。基准点依照处理方式而会有所不同。比如，在直排时可以行左右的中央线为基准点，横排时可以行上下的中央线为基准点。当配置的文字全部尺寸相同时，有着以下关系：

指定版心的方法中，有著不以行距，而是依**行高**來設定的方法。行高就是彼此鄰接的兩行之基準點間的距離。基準點依照處理方式而會有所不同。比如，在直排時可以行左右的中央線為基準點，橫排時可以行上下的中央線為基準點。當配置的文字全部尺寸相同時，有著以下關係：

- $\text{line height} = \text{character size} / 2 + \text{line gap} + \text{character size} / 2 = \text{character size} + \text{line gap}$

行高=文字尺寸÷2+行距+文字尺寸÷2=文字尺寸+行距

行高=文字尺寸÷2+行距+文字尺寸÷2=文字尺寸+行距

- **line gap** = line height - character size

行距=行高-文字尺寸

行距=行高-文字尺寸

统一文档中「em」「character width」「字宽」等表述 #465

🔒 Closed realfish opened this issue on 18 May · 0 comments · Fixed by #466

Assignees

Labels

內容修正/fix

realfish commented on 18 May · edited

Member

根据 5 月 17 日例会讨论，文档中的「em」「character width」「字宽」这三个术语的使用，有待重新检查并进行调整。

暂定的改进方向为：

1. 中文尽量避免使用 em 单位，改用「字宽」；
2. 中文「字宽」作为长度单位时，英语大多可考虑翻译作「em」；
3. 按语境重新检查所有「character width」的表述，并确认是否可替换成「em」。

<> Code Issues 113 Pull requests 3 Actions Projects Wiki ...

工作組編輯體例之簡繁體詞彙表

[Jump to bottom](#)

Fuqiao Xue edited this page on 24 Jan · 7 revisions

與排版相關之專業用語，編輯於CLREQ中的[詞彙表](#)。

此詞彙表主要供工作組在進行內容簡繁體轉換時，作為作業用參考，以確保用語符合地區需求，提升文件整體品質。



LATEST PUBLISHED

[w3.org/TR/clreq](https://www.w3.org/TR/clreq)

MAILING LIST

public-i18n-chinese@w3.org



EDITOR'S DRAFT

w3c.github.io/clreq

GITHUB

github.com/w3c/clreq