中国語組版

JEPAセミナー・2025/6/16

エリック・リュウ
Eric Liu (eric@thetype.com)

簡体字でしょ？

- 簡体字・繁体字という二分化の考え方はあまり良くない
- 地域化（L10N）の構造では、地域-言語-用字系の三段構造。

source: Wikimedia Commons

source: Wikimedia Commons

9 600 000 km²

10 180 000 km²

56 130 30

民族	言語	文字
漢民族 92%	漢語方言：10言語 (?) 標準語人口：70% (2010) ピンイン人口：68%	

source: www.gov.cn

བོད་སྐད་ Vahcuengh ئۇيغۇرچە

scolarnk.edb.hkedcity.net

Standing Committee on Language Education and Research (SCOLAR)
語文教育及研究常務委員會（語常會）

輸入查詢字串 | 文字大小 | Eng | 簡

學術界 | 小學 / 中學 | 幼稚園 | 社區

主頁 | 關於語常會 | 人工智能高峰會 (AIinLT) | 兩文三語運動 | 資助計劃 | 推廣語文 | 研究與發展 | 贊助項目 | 進修課程 | 資源

兩文三語運動

- 2024
- **2025**
  - 政府電視宣傳短片
  - 一分鐘短片製作比賽
  - 合作夥伴活動
  - 語常會支持的活動
  - 宣傳短片
  - 學生語文大使委任典禮

香港是一個多元及中西文化薈萃的國際化大都市，也是大灣區城市之一，讓港人認識兩文三語的重要

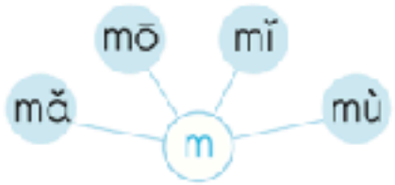
香港：兩文三語

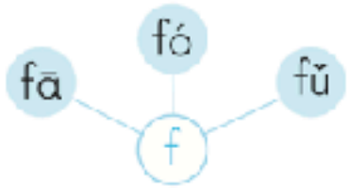
両文：中文・英文

三語：いわゆる北京語、
広東語と英語

中国人にとって
アルファベットは馴染みがある？

pīnyīn
拼音





b p m f

b b p p m m f f

bà mā
爸 妈

bà ba
爸 爸

mā ma
妈 妈

4

d t n l

dā dé dī dú

tā tè tī tú

ná né nǐ nù nǚ

lā lè lǐ lù lú

c d t t n n l l

d d t t n n l l

第五单元·阅读
1 秋天
天气凉了，树叶黄了，一片片
子从树上落下来。
天那么蓝，那么高。一群大雁往南
儿排成个“人”字，一会儿排

W3C Group Draft Note

w3.org

2. 罗马拼音

使用拉丁字母的汉语拼音是中国大陆推行的现代标准汉语（普通话）标音与拉丁转写方案，汉语拼音的全文标音也是中国大陆现代最常见的行间注用例。在中国大陆与台湾，台湾闽南语罗马字拼音等各种汉语方言的拉丁字母标音方案也有与汉语拼音情况类似的行间注用例。

因拉丁字母的特质，此类标音文本仅横排。针对幼儿母语者的文本以单字作为基文进行标音。针对双语学习者的文本除偶有与前者一致的格式外，大多分词连写，以词作为基文进行标音，基文之间如西文书写一般用空格隔开，并且注文有句首大写、专名首字母大写等格式以及与正文对应的标点符号，自成一体。汉语拼音早期印刷品的格式较为多变，一致性不强。大体上，单字标音与分词连写标音都常见，不作详细述。

ā Q fú xiàcù shìjìn le píngshēng de lìqi
阿Q伏下去，使尽了平生的力气
huà yuánquān tā shēngpà bèi rén xiàohuà lìzhì
画圆圈。他生怕被人笑话，立志
yào huà de yuán
要画得圆。

图 罗马拼音行间注的排版示例。

繁體中文
简体中文
English
AI

zh.wikipedia.org

维基百科
自由的百科全书

搜索维基百科 搜索

资助维基百科 创建账号 登录

汉字 [编辑]

137种语言

条目 讨论 汉 漢 大陆简体 阅读 编辑 查看历史 工具

- 大陆简体
- 香港繁體
- 澳門繁體
- 大马简体
- 新加坡简体
- 臺灣正體

目录 隐藏

- 序言
- “汉字”语源
- 汉字的历史
- 汉字的特点
- 汉字的语文知识
- 汉字文化
- 汉字变化
- 计算机处理
- 繁简体汉字的比较
- 参见
- 注释
- 参考文献
- 外部链接

维基百科，自由的百科

以下问题，请协助改善本条目或在讨论页针对议题发表看法。 [展开]

汉字，又称**汉文、**字、唐话字、唐人字、中国字，是书写汉语所使用的文字系统[世界上唯一仍广泛使用的语素文字[15]，主要流行于汉字文化圈区），也是联合国四种官方文字之一（另外三种皆为拼音文字，语与西班牙语使用的拉丁文字、俄语使用的西里尔文字和阿拉文字）。

汉字是世界历史上唯一被高度发展的语素文字[c]，从基础的象形字和指事字发展出了现今以形声字和会意字为主流的六书系统，加上各类异体字和生僻字在所有字典中总共收录超过十万字。每个汉字在结构上由最基础的笔画进行二维排列组成，每个笔画不包含任何语素，相当于表音文字的字母；由笔画组成的偏旁部首是汉字内部结构上的最小语素单位，相当于一些表音文字的词根。除形声字外的汉字通常不具备音素，而形声字的声旁则带有一定表音功能，但因为其大多基于上古汉语的发音所以与现代标准汉语有一些差异。现今汉字通常使用额外独立的标音帮助阅读，在中华民国时期主要使用注音符号，但在台湾以外的地区（特别是中国大陆）主要使用拉丁化的汉语拼音。虽然绝大多数汉字单字都拥有自身的独立含义，在语法上相当于英文的单词，但汉语词汇通常都是由两个以上单字组成的合成词。

汉字

类型	语素文字
使用时期	约公元前2500年至今
书写方向	横排从左到右，从上到下（现代） 竖排从上到下，从右到左（传统）
用于官方文字	汉语

- 中华人民共和国
  - 香港特别行政区
  - 澳门特别行政区
- 中华民国
- 新加坡
- 联合国
- 缅甸（仅部分地区）
  - 佤邦
  - 缅甸联邦第一特区
  - 缅甸掸邦东部第四

簡体字でしょ？

- 簡体字・繁体字という二分化の考え方はあまり良くない
- 言語は中国語だけではない。例えば、香港は基本「両文三語」。
- 文字は漢字だけではない。ピン音・ボポモフォなども。
- 簡体字≠中国本土。シンガポールの公式文字でもある＋α
  （例：zh-Hans-CN）
- 繁体字≠台湾。＋香港、マカオ。＋α　（例：zh-Hant-TW）
- 地域化（L10N）の構造では、地域-言語-用字系の三段構造。

中国本土の話し

- 簡体字は、日本の当用漢字（1946）も参考
- 1956 年「漢字簡化方案」→４度改訂
- 1964 年「簡化字総表」
- 1977 年「第二次漢字簡化方案草案」→即廃止
- 1986 年「簡化字総表」再発表
- 2013 年「通用規範漢字表」

通用规范汉字表

- 2013 年 6 月 5 日発表
- 8105 字
  - 一级字 3500 字　常用字集：
    99.226%のカバー率
  - 二级字 3000 字
    - 出版印刷・辞書編纂・情報処理
  - 三级字 1605 字
    - 人名・地名・専門用語・義務教育の漢文（文言文）

一 级 字 表

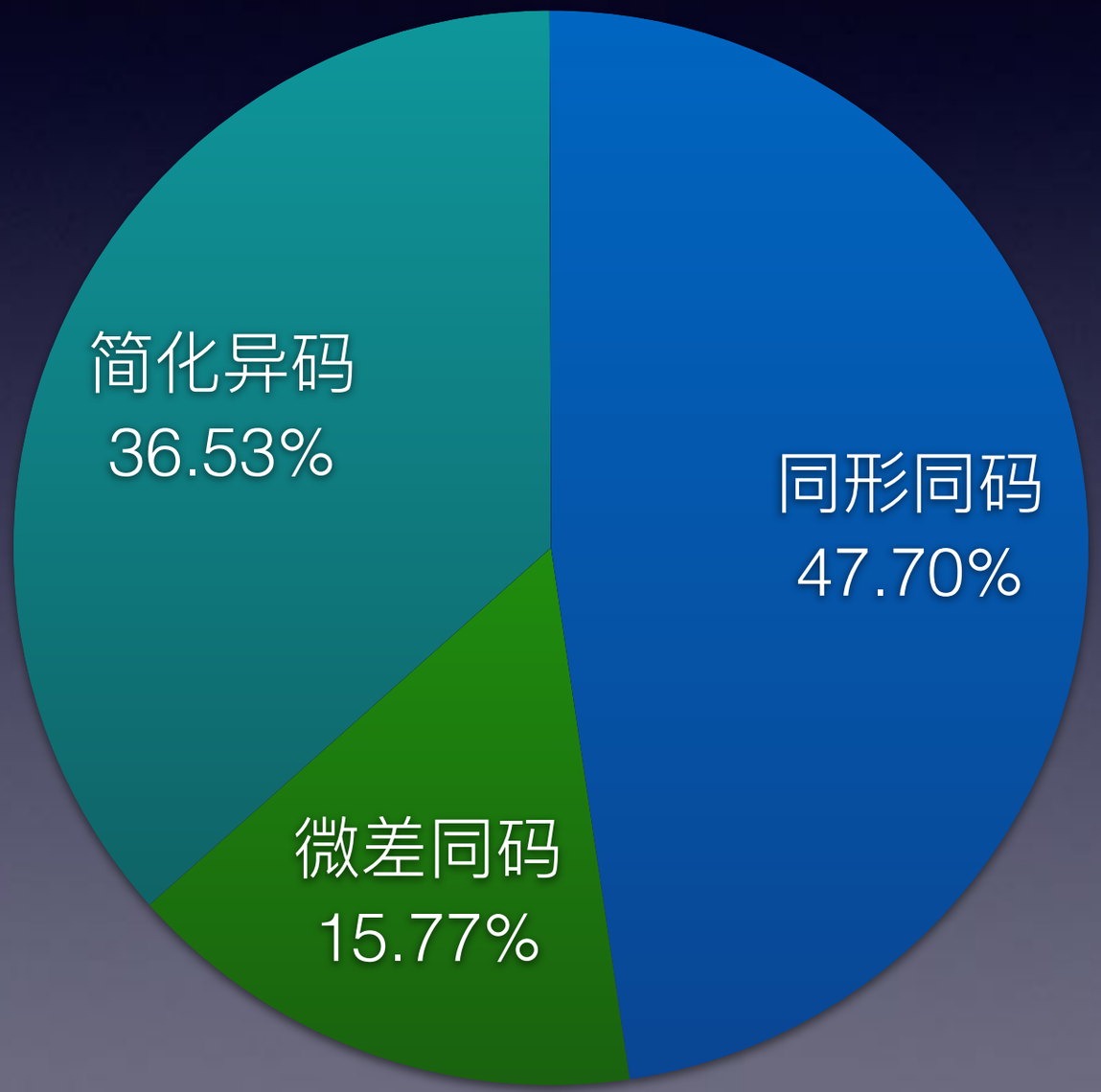
0001 一
0002 乙
0003 二
0004 十
0005 丁
0006 厂
0007 七
0008 卜
0009 八
0010 人
0011 入
0012 儿
0013 匕
0014 几
0015 九
0016 刁
0035 万
0036 上
0037 小
0038 口
0039 山
0040 巾
0041 千
0042 乞
0043 川
0044 亿
0045 个
0046 夕
0047 久
0048 么
0049 勺
0050 凡
0069 飞
0070 习
0071 叉
0072 马
0073 乡
0074 丰
0075 王
0076 开
0077 井
0078 天
0079 夫
0080 元
0081 无
0082 云
0083 专
0084 丐
0103 屯
0104 戈
0105 比
0106 互
0107 切
0108 瓦
0109 止
0110 少
0111 曰
0112 日
0113 中
0114 贝
0115 冈
0116 内
0117 水
0118 见
0137 斤
0138 爪
0139 反
0140 介
0141 父
0142 从
0143 仑
0144 今
0145 凶
0146 分
0147 乏
0148 公
0149 仓
0150 月
0151 氏
0152 勿
0171 认
0172 冗
0173 讥
0174 心
0175 尺
0176 引
0177 丑
0178 巴
0179 孔
0180 队
0181 办
0182 以
0183 允
0184 予
0185 邓
0186 劝

规范字与繁体字、异体字对照表

规范字		繁体字	异体字
0006	厂	(廠)	
0008	卜	~	
		(蔔)	
0012	儿	(兒)	
0014	几	~	
		(幾)	
0017	了	~	
		(瞭[1])	
0020	乃		[廼迺[2]]
0023	干	~	
		(乾[3])	[乹乾]
		(幹)	[榦]
0025	亏	(虧)	
0029	才	~	
		(纔)	
0034	与	(與)	
0035	万	~	
		(萬)	
0041	千	~	
		(韆)	
0044	亿	(億)	
0045	个	(個)	[箇]
0048	么	(麼[4])	

规范字		繁体字	异体字
0059	尸		[屍]
0065	卫	(衛)	
0069	飞	(飛)	
0070	习	(習)	
0072	马	(馬)	
0073	乡	(鄉)	
0074	丰	~	
		(豐)	
0076	开	(開)	
0081	无	(無)	
0082	云	~	
		(雲)	
0083	专	(專)	[耑[6]]
0084	丐		[匄匂]
0085	扎		[紮紮]
0086	艺	(藝)	
0090	厅	(廳)	
0094	区	(區)	
0095	历	(歷)	[歷歷]
		(曆)	[厤]
0099	匹		[疋]
0100	车	(車)	
0101	巨		[鉅[7]]

台湾と中国大陸の字形比較



source: http://www.unihan.com.cn/

JP KR CN TW

+U9AD3

中国本土の文字セット

- GB18030：強制。
  最新は2022版（2023/8 実施）＋第一回改訂票
  （2024/10 実施）
- Unicode（≒ ISO/IEC 10646:2020/
  Amd.1:2023 の GB 規格化、
  GB/T 13000 最新版は 2025/8 実施予定）

エリックの多言語文字散歩

英語か米語か？簡体字か繁体字か？〈エリックの多言語文字散歩〉

みなさんが多言語対応を行う場合、対応すべき言語はいくつあるか？と言語数を把握したくなるかもしれません。そう…

ヒラギノフォント公式n…
1年前

53

GB18030-2022が来た！詳細編ー内容から読み取る本質〈エリックの多言…

前回の概要編では、じっくりとGB18030-2022の表紙を見ながら、中国におけるGBの定義、GB18030規格改定の…

ヒラギノフォント公式n…
1年前

14

GB 18030-2022が来た！概要編ー表紙から読み取る情報〈エリックの多言…

時は2023年、夏。中国の文字界隈でもっともホットな話題は、新しいGB 18030-2022の施行でしょう。みなさん…

ヒラギノフォント公式n…
1年前

35

「その多言語、必要ですか？」本当に必要な多言語対応を、街中の案内から考え…

いきなりですが、街角の商業施設でこんな看板を見かけら、皆さんはどう思いますか？ 上の図は、実在する看板の写真…

ヒラギノフォント公式n…
2年前

166

「脱・ツギハギ」の第一歩〈エリックの多言語文字散歩〉

前回の記事では街中で見かける案内板や注意書きなど、読み手を気遣うための中国語が、なぜツギハギのような乱れ…

ヒラギノフォント公式n…
2年前

29

その中国語の案内文、フォントがツギハギになっているかも？〈エリックの多言…

自己紹介こんにちは！文字散歩（文字を探しながら散歩すること！）が大好きなエリックです。 私は中国生まれです…

ヒラギノフォント公式n…
1年前

60

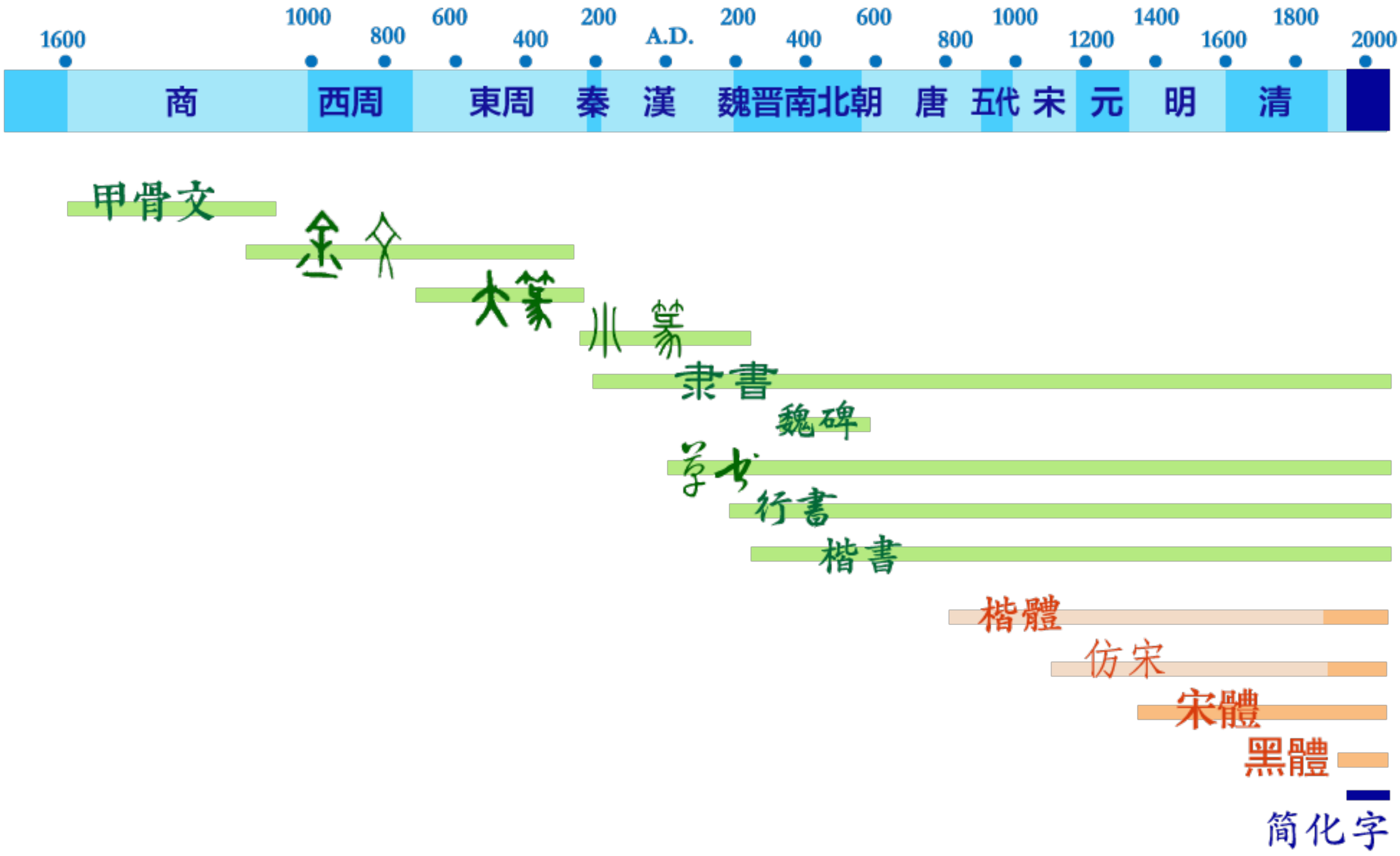
エリックの多言語文字散歩

GB 18030-2022が来た！概要編—表紙から読み取る情報〈エリックの多言語文字散歩〉

♡ 35

ヒラギノフォント公式note
2023年8月28日 10:31

時は2023年、夏。中国の文字界隈でもっともホットな話題は、新しいGB 18030-2022の施行でしょう。


1600
1000
800
600
400
200
A.D.
200
400
600
800
1000
1200
1400
1600
1800
2000
商
西周
東周
秦
漢
魏晋南北朝
唐
五代
宋
元
明
清
甲骨文
金文
大篆
小篆
隸書
魏碑
草书
行書
楷書
楷體
仿宋
宋體
黑體
简化字

宋体

宋体

宋体

宋体

黑体

黑体

黑体

黑体

黑体

楷体

楷体

楷体

光明日报
GUANGMING RIBAO
2023年5月15日 星期日 农历乙卯年五月二十 今日12版
光明网网址：http://www.gmw.cn 国内统一连续出版物号 CN 11-0026 代号1-16 总2729号

推动中国－中亚航船乘风破浪、勇毅前行

——中亚各界热切期待习近平主席出席第二届中国－中亚峰会

一块“老树根”，扎在秦岭一辈子

人民需要这样的科学家

研究真问题和真研究问题

他把一生献给山川大地

雪宝山的“神秘歌者”

我国新增10家国家考古遗址公园

仿宋

仿宋

仿宋

仿宋

北京市旅游局

京旅发〔2007〕106号

北京市旅游局关于第十届北京国际旅游文化节动漫形象征集活动的通知

各有关设计单位:

北京国际旅游文化节，是由北京市人民政府和中国国家旅游局主办，北京市旅游局、首旅集团共同承办的重要旅游文化节庆活动。北京国际旅游文化节影响日益加大，国际性、全民性、多样性特点越加显著，节庆本身已营造出大量商机被社会与企业所重视。国际旅游文化节已经成为北京一道靓丽的风景线，成为北京与世界各地沟通交流的桥梁，为千年古都增添了夺目的时代风采。

随着北京2008年奥运会的临近，北京展现在世界面前的将是勃勃的生机和前所未有的发展机遇。在“同一个世界，同一个梦想”的奥运主题口号下，北京国际旅游文化节作为北京旅游业的窗口，将更好地向全世界诠释奥林匹克精神，为促进北京旅游

中国本土における組版の話し

- 中国本土：ほとんど横組。古文書（古籍）の書籍だけ縦組。ポスター、タイトルなどは縦組多数。
- 横組のため、欧米化
- 禁則が必須→日本と同じだが、台湾と違う
- 本格な出版物では、いわゆる「アキ量調整」が必須。
- 傍点・傍線があるが、使用頻度が少ない

光明日報

第九十八號

中華人民共和國開國盛典

人民政協隆重揭幕

到會代表六百卅五人會場充盈莊嚴和諧的氣氛

毛主席致開幕詞劉少奇等發表講演

毛主席宣佈：人民政協意識中於是制定人民政協和中央人民政府的組織法、人民政協共同綱領，選舉人民政協的全國委員會和中央人民政府委員會，制定國旗、國歌，決定國都所在地以及年號。

毛主席致開幕詞

社論

慶賀人民政治協商會議開幕

中央中央政治局委員劉少奇講話

主席團及秘書長名單

特邀代表宋慶齡講話

國民黨革命委員會中央常務委員何香凝講話

公明日報・1949 年 9 月 22 日

光明日報

社論

滿懷信心迎接新任務

全國人民代表大會常委會舉行會議

一年來對蔣賊軍作戰的戰果

分別舉行交接儀式慶祝大會

北京市協商委員會決定成立公債推銷委員會

光明日報・1955年1月1日

関連規格

- JIS X 4051 日本語組版規則
- GB/T 15834—2011 标点符号用法
- GB/T 15835—2011 出版物上数字用法

GB/T 15835—2011 出版物上数字用法

General Rules for writing numerals in public texts

- 漢数字とアラビア数字

  七八人

  甲午年四月廿七日

  唐开元二年五月

- 千分空/千分撇

  0.35

  13,456.83698

  46 177.247 589

- 「〇」と「零」

  九十五点零六

  二〇〇一年十月

ICS 01. 40.10
A 19

GB

中华人民共和国国家标准

GB/T 15835—2011
代替 GB/T 15835—1995

出版物上数字用法

General rules for writing numerals in public texts

2011-07-29 发布　　2011-11-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会　发布

「縦中横」は要らない？

示例一

雪花牌BCD188型家用电冰箱容量是一百八十八升，功率为一百二十五瓦，市场售价两千零五十元，返修率仅为百分之零点一五。

示例二

海军J12号打捞救生船在太平洋上航行了十三天，于一九九〇年八月六日零时三十分返回基地。

GB/T 15834—2011 标点符号用法

General Rules for punctuation

点号

（句中）， 、 ； ：

（句末）。？！

标号

“” （） 《》 ——……—— / ·

禁则

ICS 01.140.10
A 19

GB

中华人民共和国国家标准

GB/T 15834—2011
代替 GB/T 15834—1995

标点符号用法

General rules for punctuation

2011-12-30 发布　　2012-06-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会　发布

引号

- いずれ " "　' ' は無し
- 横書き　“ ‘ ’ ”
- 縦書：大陸『「」』と台湾「『』」(⊢　¬)

GB/T 15834—2011 标点符号用法

General Rules for punctuation

- 禁則（避头尾）
- 行頭禁則
- 行末禁則
- ぶら下がり
- 分離禁止

アキ量調整

開明式

行末半角

全角式

半角式

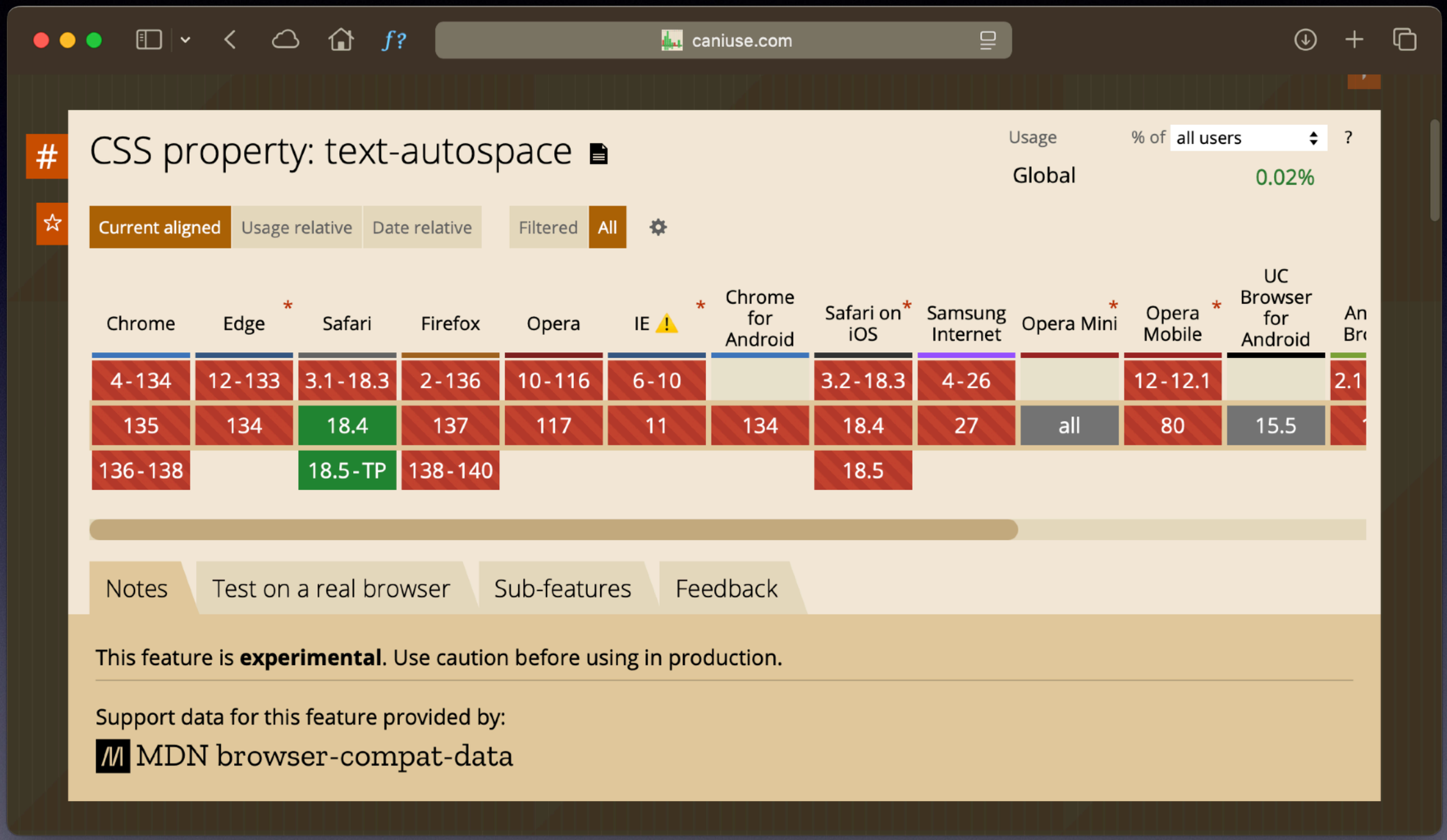
現代中国語の組版

近代に欧文に合わせる風潮の影響

- 古文書などの特例以外、ほとんど横書き
- 二字下げのインデントは大原則
- 約物は現代中国語には非常に重要
- GB には約物（及び禁則）と数字の規定のみ

最近の動向

text-autospace



codepen.io

`text-autospace`

Jen Simmons PRO + Follow

Settings Sign Up Log In

HTML

```
<main>
  <figure class="no-autospace">
    <figcaption>
      <code>text-autospace: no-autospace;</code>
    </figcaption>
    <div><p>Eric的播客TypeChat已经10年了? </p></div>
  </figure>

  <figure class="auto">
    <figcaption>
      <code>text-autospace: auto;</code>
    </figcaption>
    <div><p>Eric的播客TypeChat已经10年了? </p></div>
  </figure>
```

CSS

```
.no-autospace {
  text-autospace: no-autospace;
}
.auto {
  text-autospace: auto;
}
.normal {
  text-autospace: normal;
}
.ideograph-alpha {
```

JS

text-autospace: no-autospace;

Eric的播客TypeChat已经10年了?

text-autospace: auto;

Eric的播客TypeChat已经10年了?

text-autospace: normal; *(both ideograph-alpha & ideograph-numeric)*

Eric 的播客 TypeChat 已经 10 年了?

text-autospace: ideograph-alpha;

Eric 的播客 TypeChat 已经10年了?

text-autospace: ideograph-numeric;

Eric的播客TypeChat已经 10 年了?

Console Assets Comments ⌘ Keys ⌥ Fork Embed Export Share

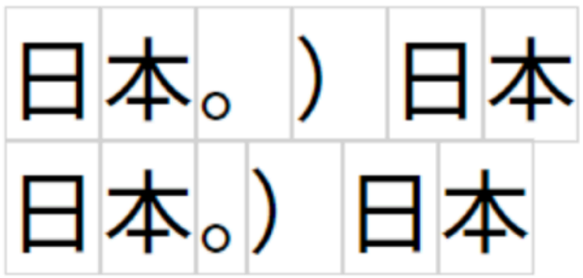
CJK punctuation kerning: `text-spacing-trim`

The following feature is under development, with the aim that it will be enabled by default in Chrome.

In Chinese, Japanese, and Korean, applying kerning between punctuation characters improves the readability and produces more visually pleasing typography. Most printed materials and word processors apply this kerning today.

For example, CJK periods and CJK close parentheses are usually designed to have glyph-internal spacings on the right half of the glyph spaces, so that each character has a constant advance.

But when these punctuation characters appear in a row, this glyph-internal spacings become excessive. In the following two examples, the second one is the correct typography; the right-half of the CJK period should be removed.



When punctuation characters appear in a row, the right-half of the CJK period should be removed.

The default behavior provides good results in general, but developers can use the text-spacing-trim property to choose different styles, or to disable it in some cases.

★ **Note:** The feature consists of two parts; one for line edges and another for adjacent characters. The latter is already available on Android and ChromeOS when the used font is Noto CJK, from Android 13 and ChromeOS 90. This feature complements it by supporting line edges, supporting more fonts and platforms, and supporting more edge cases such as when the fonts change.