Script-hybrid and abbreviated CJK letters

Gen Kojitani(genkjtn@gmail.com)   2023 June 23

Overview

This document is a list of Unicode unregistered script-hybrid and abbreviated CJK letters, and was created to supplement discussions that arose in connection with my UTC **L2/23-139** document and to list some additional characters beyond those proposed in L2/23-139 The following is a list of some additional characters other than those proposed in **L2/23-139**.

For each character, we will list its shape, which language it is used in, and how it is used, followed by images of actual examples of its use. Note that since CJK characters of this type appear mainly in handwritten documents, many of the examples are taken from handwritten documents like comic, signage of shops or the geba characters(ゲバ文字) that appear on university billboards set up by students engaged in leftist political activities in Japan. The list also contain Sawndip letters for Bouyei(布依) language spoken in China.

Some of these script-hybrid or abbreviated CJK characters are commonly used in everyday life and are used in the same way as normal CJK ideographs, but each character has a character set other than CJK characters (katakana, hiragana, Latin, Hangul, etc.) as components or use brush strokes of shapes not used in normal CJK characters. It is therefore debatable whether these should be treated in the same way as normal CJK ideographic characters.

1

List of characters

This list includes the letter form and information of each character. The font used in the letter form field was created by Gen Kojitani(the author of this document) and is licensed under the IPA Font License.

Table 1: list of characters

No.	letter form	characters information
1	仸	• □亻A • used in Bouyei(布依) language spoken in Guizhou(貴州省), China. • from 《布依方塊古文字》第 64 頁第二欄第 1 條 • pronounced as "eis" [ʔɯi⁵]
2	呐	• □口A • used in Bouyei(布依) language spoken in Guizhou(貴州省), China. • from 《布依方塊古文字》第 64 頁第二欄第 1 條 • pronounced as "nduiz" [ʔdui²]
3	齒A	• □齒A • used in Bouyei(布依) language spoken in Guizhou(貴州省), China. • from 《布依方塊古文字》第 212 頁第二欄第 6 條 • pronounced as "nyaih" [ŋai⁶]
4	呢	• □口e • used in Bouyei(布依) language spoken in Guizhou(貴州省), China. • from 《布依方塊古文字》第 103 頁第一欄第 2 條 • pronounced as "heal" [xɯə¹]
5	呑	• □天e • used in Bouyei(布依) language spoken in Guizhou(貴州省), China. • from 《布依方塊古文字》第 324 頁第一欄第 2 條 • pronounced as "ih" [ʔi⁶]
6	呑e	• □□天门e • used in Bouyei(布依) language spoken in Guizhou(貴州省), China. • from 《布依方塊古文字》第 270 頁第一欄第 2 條 • pronounced as "roomc" [zo:m³]
7	灯	• □火o • used in Bouyei(布依) language spoken in Guizhou(貴州省), China. • from 《布依方塊古文字》第 229 頁第一欄第 2 條 and 第 6 條 • pronounced as "os" [ʔo⁵] or "ous" [ʔəu⁵]

8	庆	• ▢广K • Japanese abbreviation for U+6176 慶 • mainly used to describe Keio University(慶應大学), a Japanese university
9	庅	• ▢广O • U+61C9 應 • mainly used to describe Keio University(慶應大学), a Japanese university
10	柸	• ▢木キ • Japanese abbreviation for U+6A5F 機 • A dot may be written in the top right-hand corner like Fig.12. • the usage is very common in Japan
11	枋	• ▢木き • Japanese abbreviation for U+6A5F 機
12	訧	• ▢言ギ • Japanese abbreviation for U+8B70 議 • the usage is very common in Japan
13	訳	• ▢言R • Japanese abbreviation for U+8AD6 論 • derived from Geba letters (leftist usage)
14	宨	• ▢宀R • Japanese abbreviation for U+5BEE 寮 • mainly used for university signage • derived from Geba letters (leftist usage)
15	団	• ▢口▢エン • Japanese abbreviation for U+5712 園
16	訮	• ▢言コ • Japanese abbreviation for U+8B1B 講
17	凃	• ▢艹ド • Japanese abbreviation for U+5802 堂

18	塈	• ▢▢▢▢コン▢クリ土 • Japanese abbreviation for word "コンクリート(concrete)"
19	塎	• ▢▢▢ンメセ土 • Japanese abbreviation for word "セメント(cement)"
20	毨	• ▢▢▢モタル • Japanese abbreviation for word "モルタル(mortar)"
21	袜	• ▢ヤ▢キハ • Japanese ligature of "ヤキハ(yakiba)" • often found in sword books of the Edo era
22	冮	• ▢冫ヱ • Japanese ligature of "ニエ(nie)" • often found in sword books of the Edo era • not to be confused with U+51AE 冮
23	囗	• ▢口도 • Korean abbreviation for U+5716 圖
24	お	• Japanese abbreviation for U+524D 前 • derived from cursive form • the usage is very common in Japan
25	砥	• ▢石K • Japanese abbreviation for U+78BA 確 • derived from Geba letters (leftist usage)
26	訨	• ▢言N • Japanese abbreviation for U+8A8D 認 • derived from Geba letters (leftist usage)
27	杊	• ▢木コ • Japanese abbreviation for U+69CB 構 • derived from Geba letters (leftist usage)

仸eis(依壮侗語漢字)
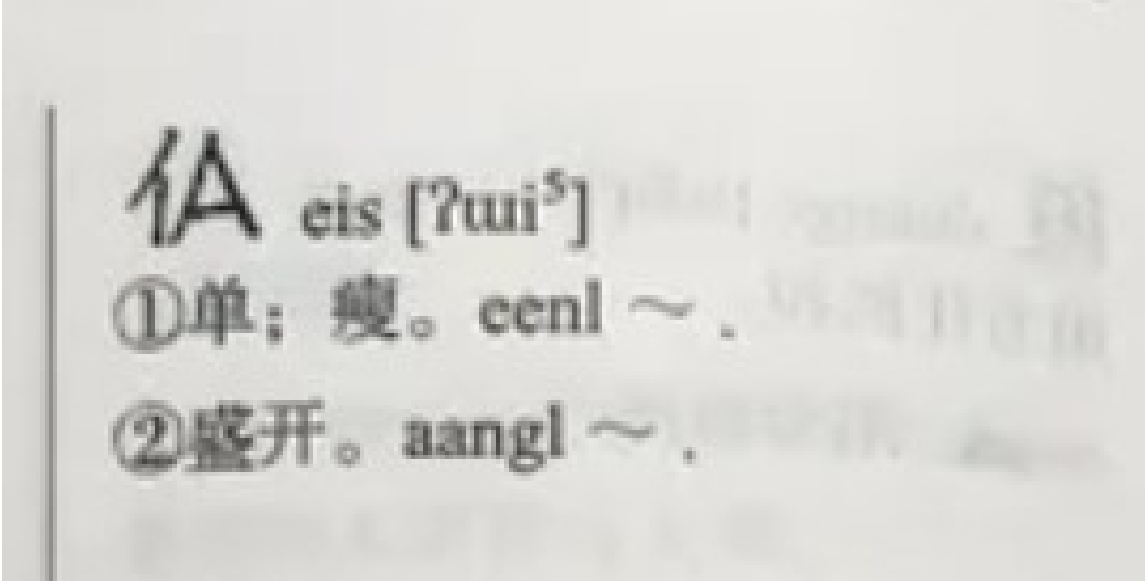
Fig. 1: "□仸 A eis"
from "布依方塊古文字", 覃家道 and 王定才編著, 中国新発現語言研究叢書, 民族出版社, August 2020

Samples

叺nduiz (依壮侗語漢字)

Fig. 2: "􀀀口A nduiz"
from "布依方块古文字", 覃家道 and 王定才編著, 中国新発現語言研究叢書, 民族出版社, August 2020

齘nyaih (依壯侗語漢字)

齘 nyaih [ȵai⁶]
细嚼。

～ leg. 嚼喂孩子。

Fig. 3: "齒A nyaih"
from "布依方塊古文字". 覃家道 and 王定才編著. 中国新発現語言研究叢書. 民族出版社. August 2020

呹 heal (依壮侗語漢字)

呹 heal [xɯə¹]
重复别人的话语。
gaais ～ haausgul. 别学我说话。

Fig. 4: "▨口 e heal"
from "布依方塊古文字", 覃家道 and 王定才編著, 中国新発現語言研究叢書, 民族出版社, August 2020

卷ih (依壯侗語漢字)

卷 ih [ʔi⁶]
天剛蒙蒙亮。

mbenl dams ～ . 天剛蒙蒙亮。

Fig. 5: "☐天 e ih"
from "布依方塊古文字", 覃家道 and 王定才編著, 中国新発現語言研究叢書, 民族出版社, August 2020

秔roomc (依壯侗語漢字)

秔 roomc [zoːm³]
早。

hadt ngaaiz 〜 . 大清早。

Fig. 6: "□□□天门e"
from "布依方塊古文字", 覃家道 and 王定才編著, 中国新発現語言研究叢書, 民族出版社, August 2020

灳os/ous (依壯侗語漢字)

火O os [ʔo⁵]
糊燋。
haul 〜 . 烧燋味。

火O ous [ʔəu⁵]
煳燋。
roongl 〜 . 煮煳。

Fig. 7: "☐火O os/ous"
from "布依方塊古文字", 覃家道 and 王定才編著, 中国新発現語言研究叢書, 民族出版社, August 2020

<u>広for 慶, 広for 應</u>

This abbreviation is mainly used for Japanese university name "慶應大学."
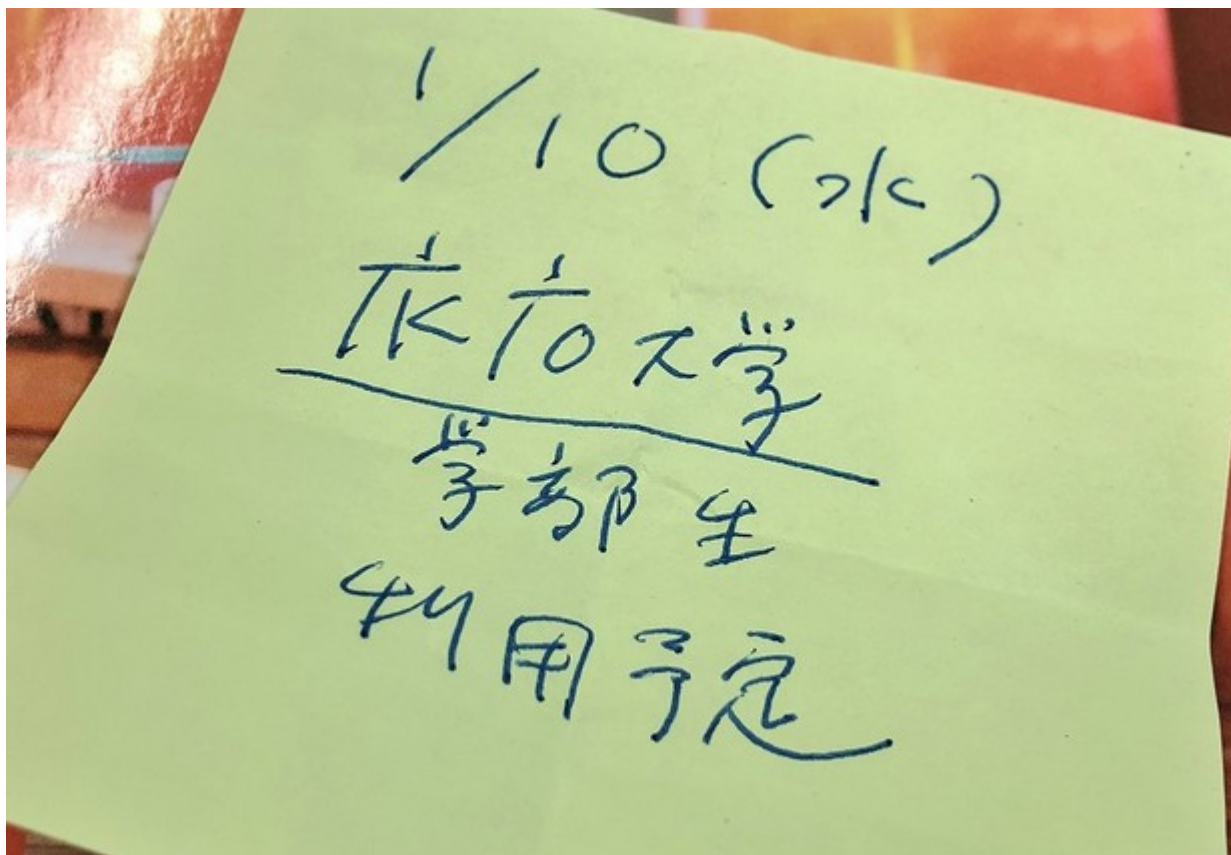
Fig. 8: "慶應大学"

This picture is posted on Twitter by @ketokate on 10 January, 2018 URL :
https://twitter.com/ketokate/status/951049551393177600

杵 for 機

Fig. 9: "自動販売機コーナー"
this picture is posted on Instagram by leon036 at Nakano Broadway, Tokyo on 17 September 2021 URL: https://www.instagram.com/p/CT6BYQwlfhD/
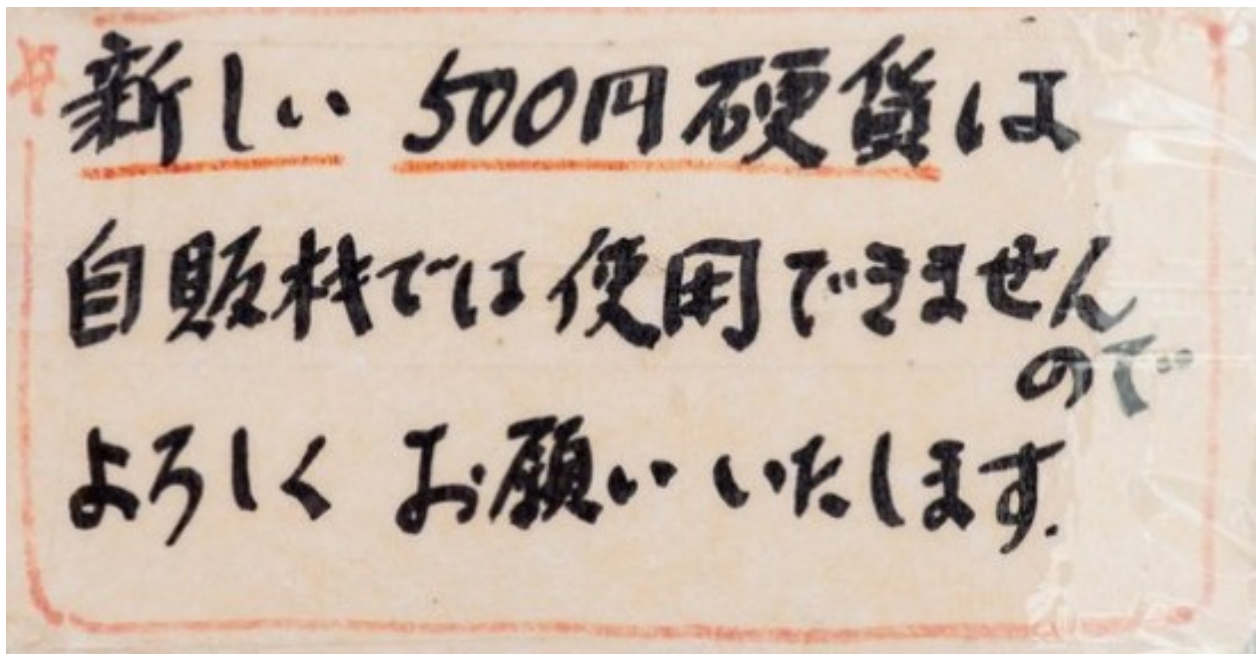
Fig. 10: "自販機では使用できませんので"
this picture is posted on Twitter by @kirnura on 4 October 2022 URL: https://twitter.com/kirnura/status/1577237089841659905

Fig. 11: "自動販売機"
this picture is posted on Twitter by @inchorin on 26 July 2017 URL: https://twitter.com/inchorin/status/890157700125933568

13

Fig. 12: "販売機"

枡 for 機

Fig. 13: "耕耘機　脱穀機　籾摺機　防除機"
This picture is posted on Twitter by @hakkaku_culture on 13 May, 2018 URL :
https://twitter.com/hakkaku_culture/status/995334947077435393

�詴 for 議, 訳 for 論
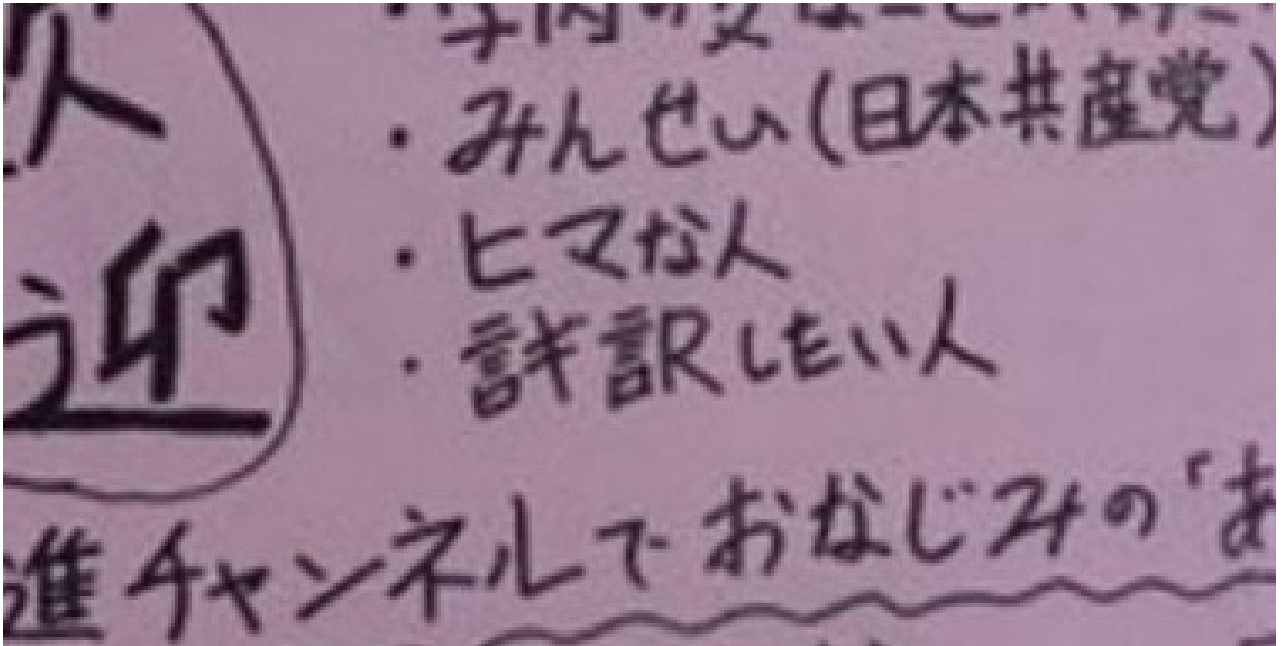
Fig. 14: "議論したい人"
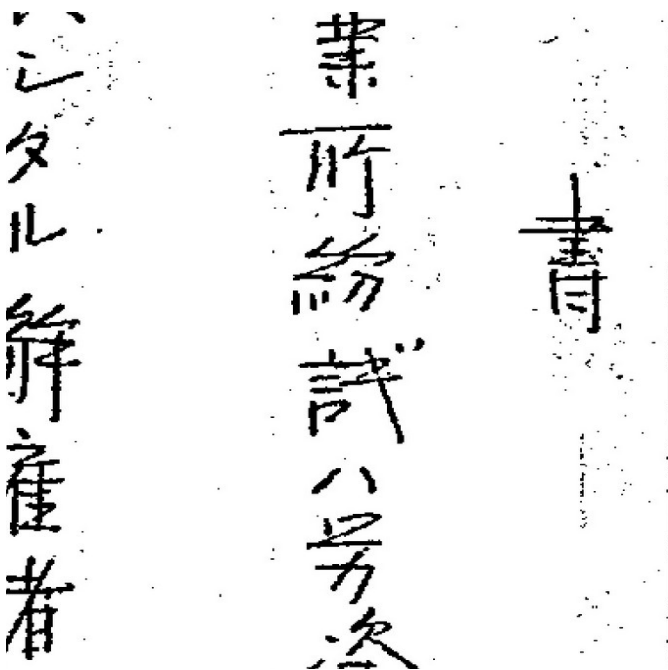*This picture is posted on Twitter by @fletsadslman on 15 August, 2018 URL :*
https://twitter.com/fletsadslman/status/1029627996741615616

Fig. 15: "紛議ハ"
This picture is posted on Twitter by @JUMANJIKYO on 4 October, 2021 URL :
https://twitter.com/JUMANJIKYO/status/1444813052537311234

宀 for 寮

Fig. 16: "吉田寮見学案内"and "吉田寮の"
this picture is posted on Twitter by @CNaka163 at Kyoto University on 30 November 2018 URL: https://twitter.com/CNaka163/status/1068327942449582081

Fig. 17: "駒場には寮を"
this picture is posted on Twitter by @sukalar223_ACK at Tokyo University on 11 August 2022 URL: https://twitter.com/sukalar223_ACK/status/1557577365282508800

Fig. 18: "熊野寮前"
this picture is posted on Twitter by @Baron_Fukushima at Kyoto University on 19 March 2023 URL: https://twitter.com/Baron_Fukushima/status/1637338390747303937

<u>圀 for 園</u>

Fig. 19: "幼稚園"

This picture is posted on Twitter by @rzeka_52 on 4 May, 2018 URL : https://twitter.com/rzeka_52/status/992219223324639233
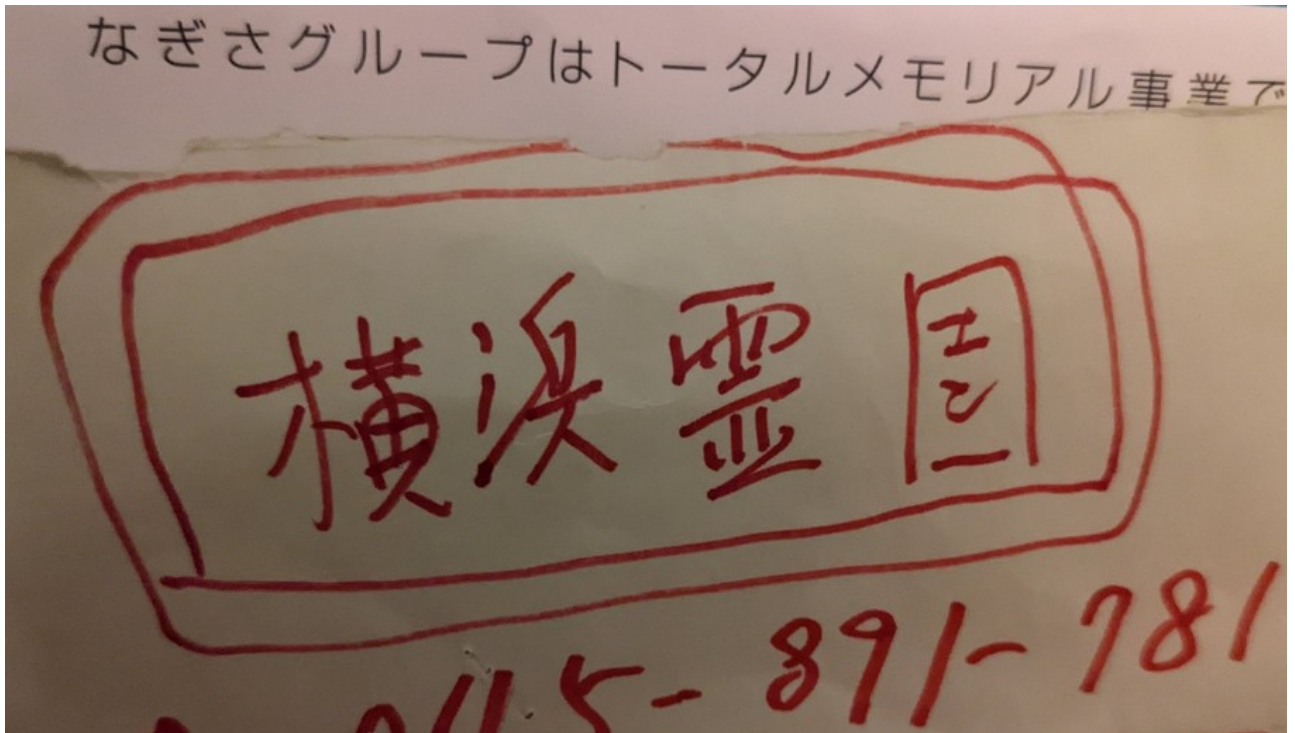
Fig. 20: "横浜霊園"

This picture is posted on Twitter by @yuki_523 on 14 October, 2022 URL :
https://twitter.com/yuki_523/status/1580711176048037888

Fig. 21: "...園"
This picture is posted on Twitter by @kabenomudagaki1 on 20 June, 2021 URL :
https://twitter.com/kabenomudagaki1/status/1406577565993631756

Fig. 22: "横山園"
This picture is posted on Twitter by skn0825 on 15 March, 2015 URL :
https://www.instagram.com/p/0PZa6vQYsn/
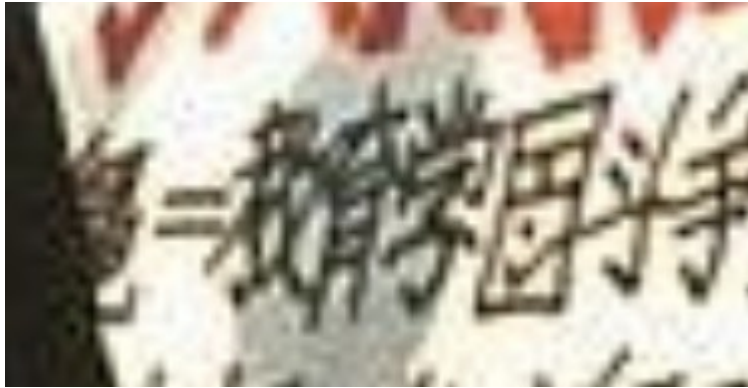
Fig. 23: "教育学園斗争"
This picture is posted on Twitter by @mosukoke on 16 July, 2015 URL :
https://twitter.com/mosukoke/status/621598335317639168

Fig. 24: "多摩川園子供プール"
This picture is posted on Twitter by @coeos_ on 13 July, 2018, from " 学童新聞(昭和25 年8 月7 日号)" URL :
https://twitter.com/coeos_/status/1017753059164807168

Fig. 25: "公園"
This picture is posted on Twitter by @prime371 on 18 August, 2018, URL :
https://twitter.com/prime371/status/1030781597245534208

訁 for 講, 吢 for 堂
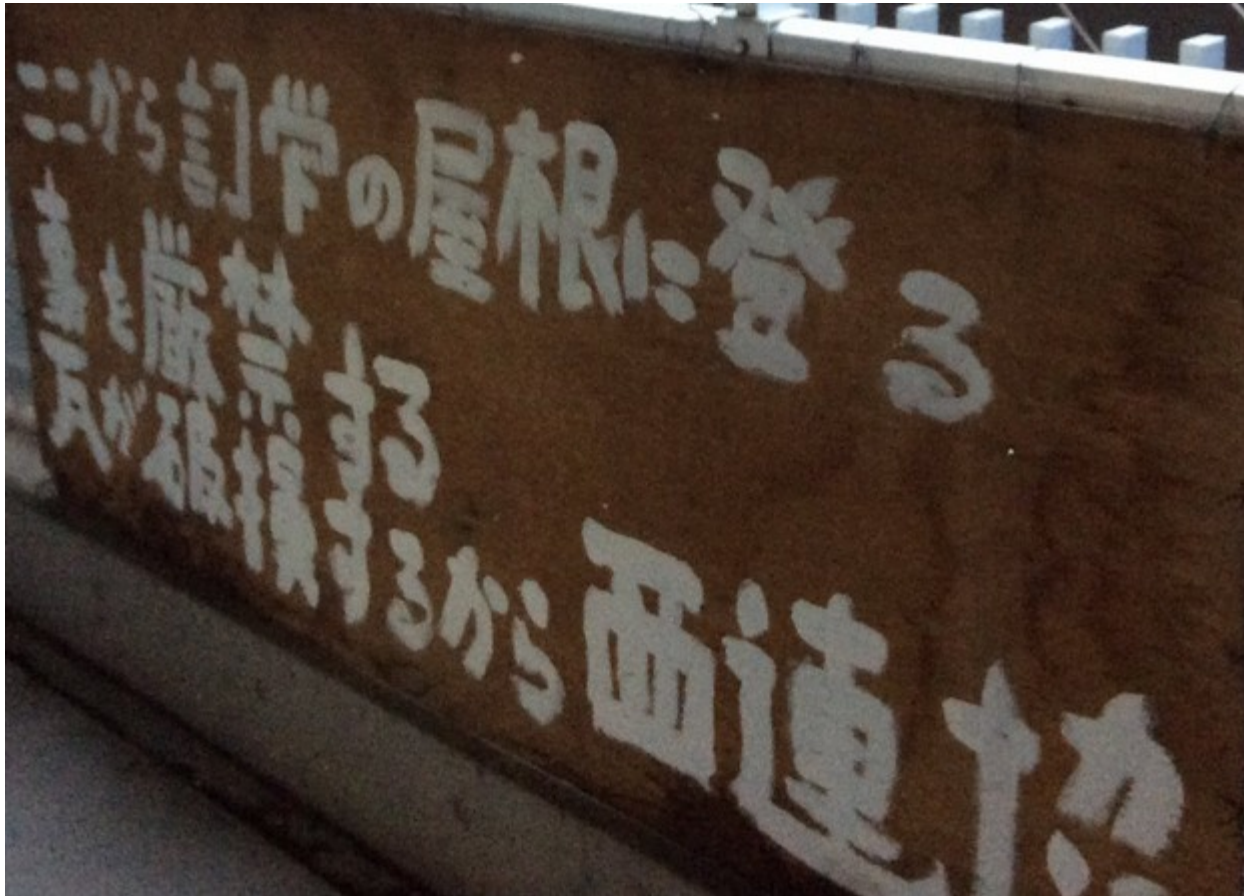
Fig. 26: "ここから講堂の屋根に登る事を厳禁する"

This picture is posted on Twitter by @atsatsatsa on 29 April, 2016, at Kyoto University URL :
https://twitter.com/atsatsatsa/status/725864175033966592
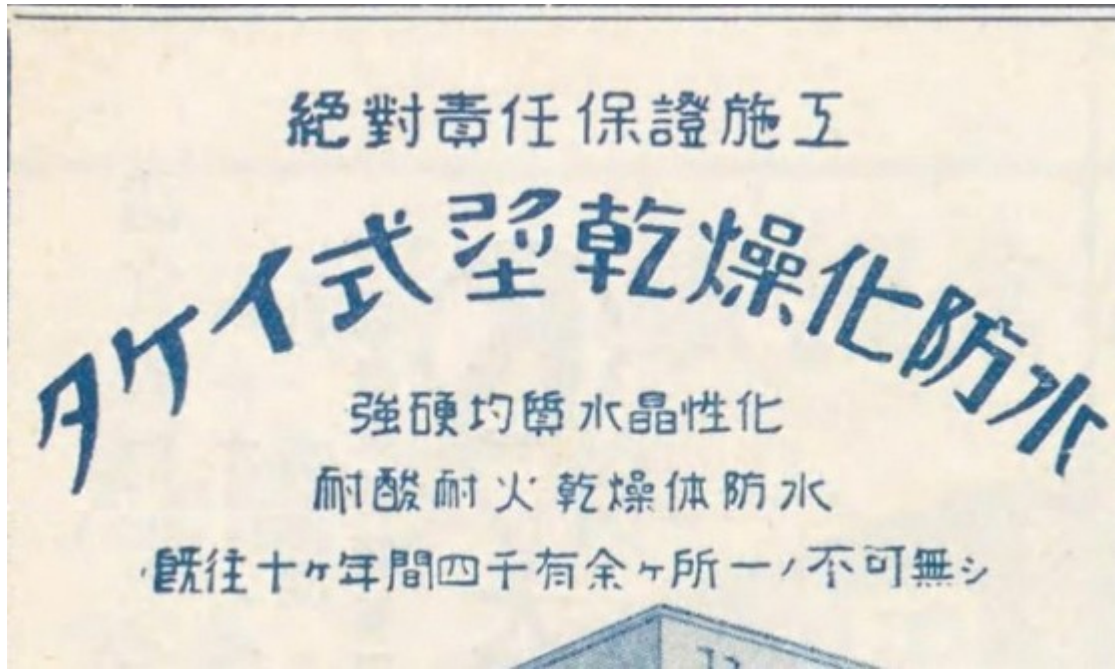
翌、坐、甩

Fig. 27: "タケイ式コンクリ土乾燥化防水"

のであつて決して発倒な技術的の頭悩らぬのである。

前號の「第一圖」の成績によると混凝た試験に在ては坴量の少ない一•五配其混凝材の粒度係數を高く四•二五のも凝と夫よりも坴量の多い一•三配合で其度係數の低い二•八〇を用ひた混凝とは其んぢ同じであることを示してゐる。
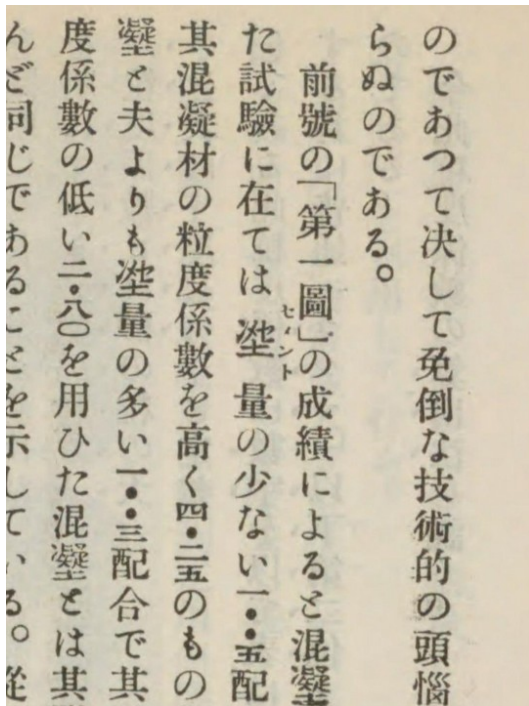
Fig. 28: "セメン土量の少ない" and *"夫よりもセメン土量の多い"*
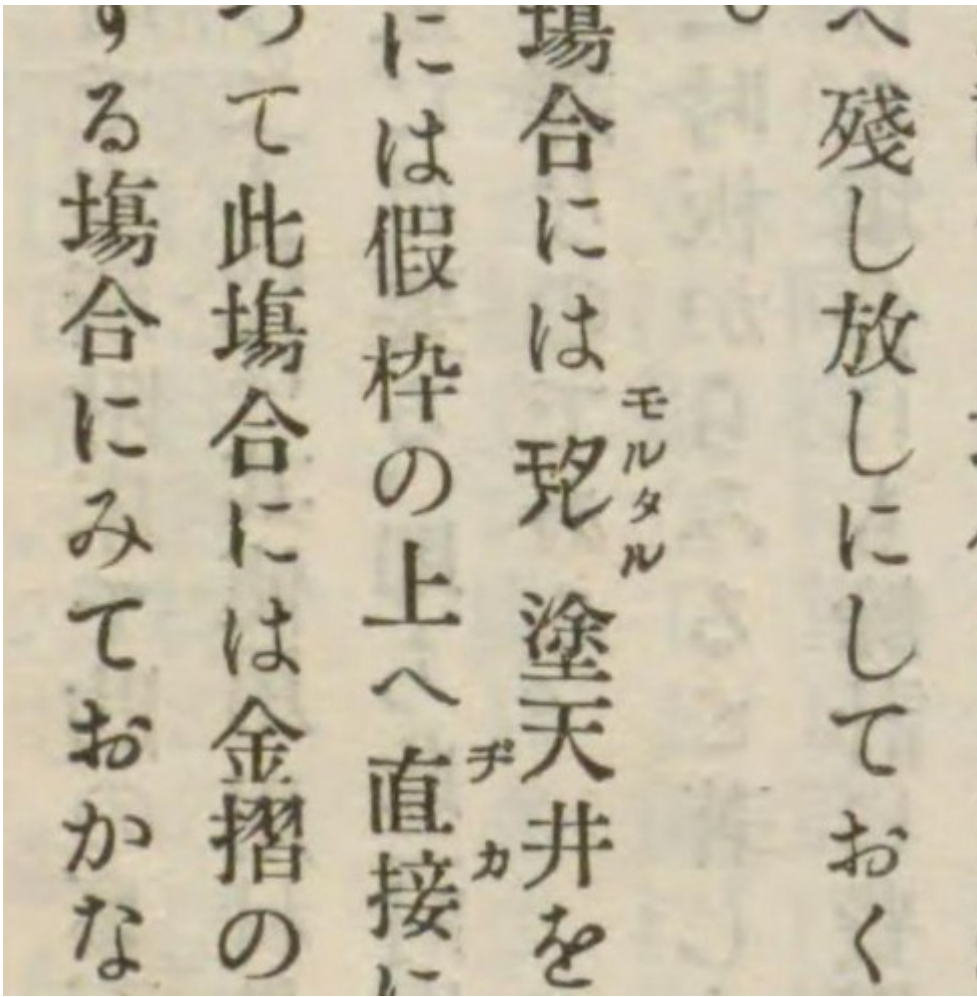
Fig. 29: "モルタル塗天井を"

妹
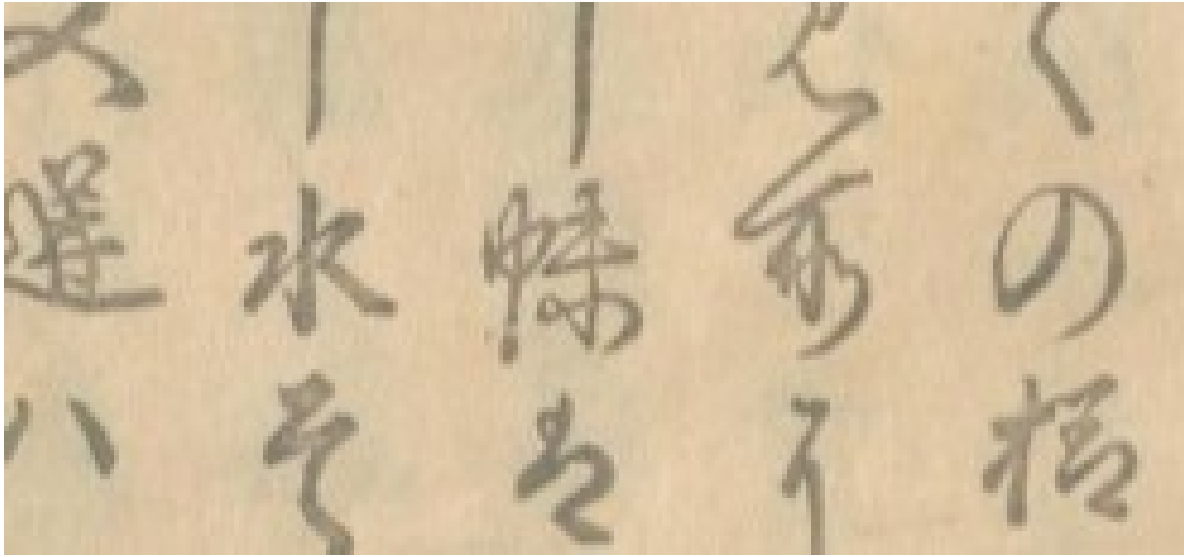
Fig. 30: "ヤキハ"

from 黒庵 [著]『解紛記 3 巻』[1],慶長 12 [1607]. p.38, 国立国会図書館デジタルコレクション
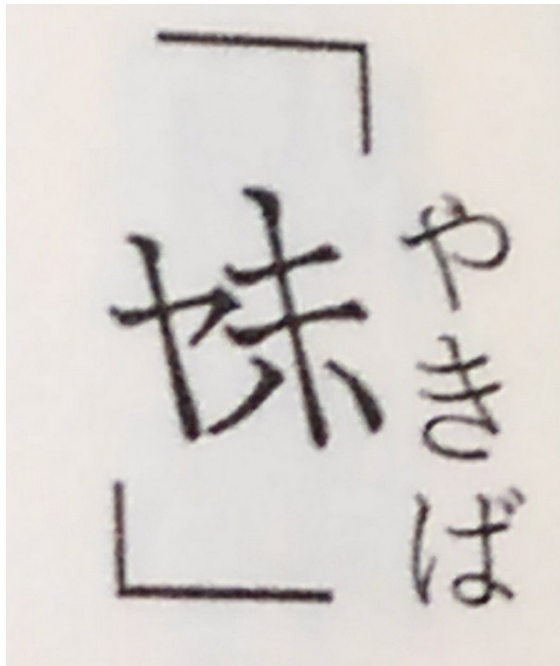*https://dl.ndl.go.jp/pid/2544318 (参照 2023-06-19)*

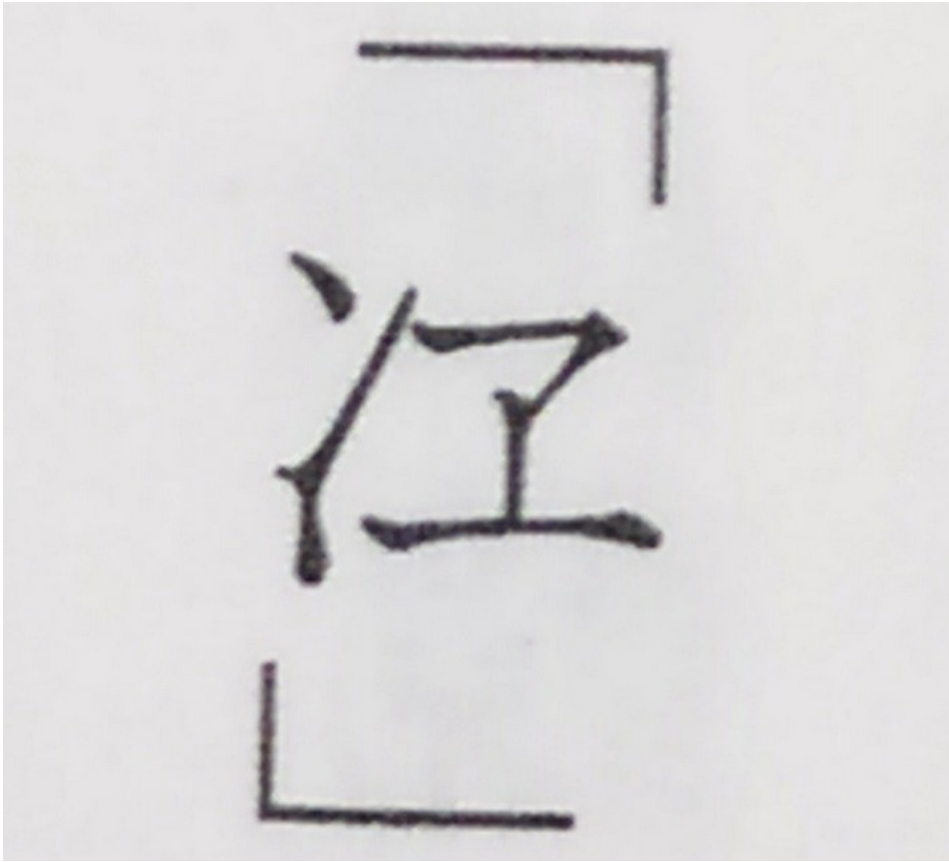*Fig. 31:* "ヤキハ"

from 笹原宏之『日本の漢字』（岩波新書、2006）

汙

Fig. 32: "ニヱ"
from 笹原宏之『日本の漢字』（岩波新書、2006）

<u>囻 for 圖</u>

Fig. 33: "韓国道路地圖"
This picture is posted on Twitter by @XUAUOaeBookxpCU on 13 July, 2020 URL :
https://twitter.com/XUAUOaeBookxpCU/status/1282674520004816896

Fig. 34: "仁川案内圖"

This picture is posted on Twitter by @mutsuosa on 2 October, 2019, from "仁川市統計年報", 韓国国会図書館蔵 1964 URL :
https://twitter.com/mutsuosa/status/1179085111502884864
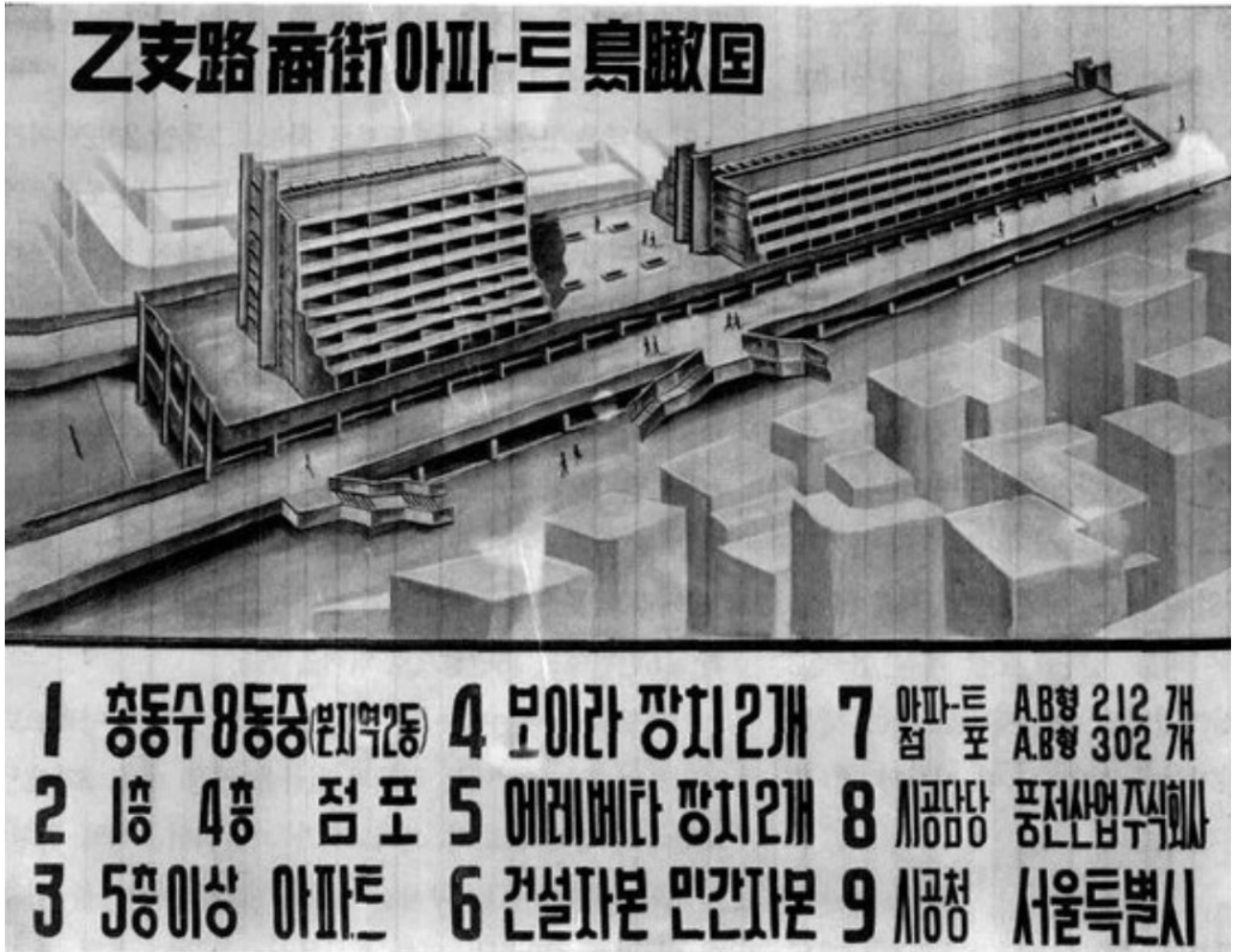
Fig. 35: "乙支路 商街 아파-트 鳥瞰圖"
This picture is posted on Twitter by @JUMANJIKYO on 4 October, 2021 URL :
https://twitter.com/JUMANJIKYO/status/1444813052537311234

Fig. 36: "案内圖"
This picture is posted on Twitter by @RCAVictorCo on 3 October, 2021 URL :
https://twitter.com/RCAVictorCo/status/1444563142814547970
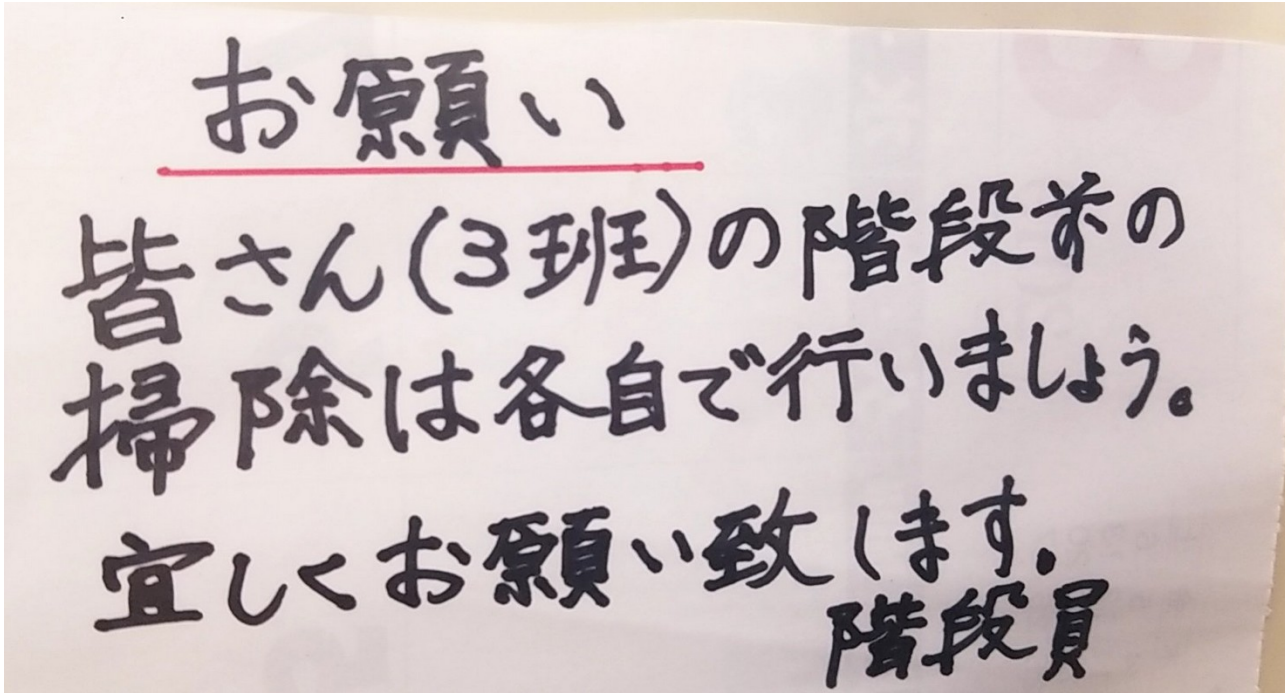
<u>お for 前</u>

Fig. 37: "階段前の"
This picture is posted on Twitter by @Palf_s on 19 March, 2020 URL : https://twitter.com/Palf_s/status/1240352325891284992
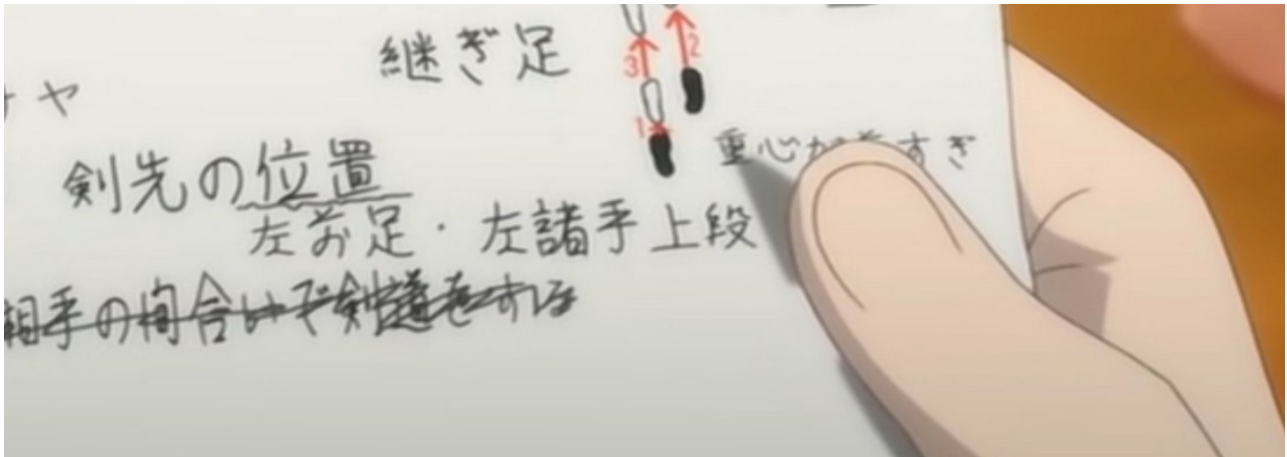

Fig. 38: "左前足・左諸手上段"
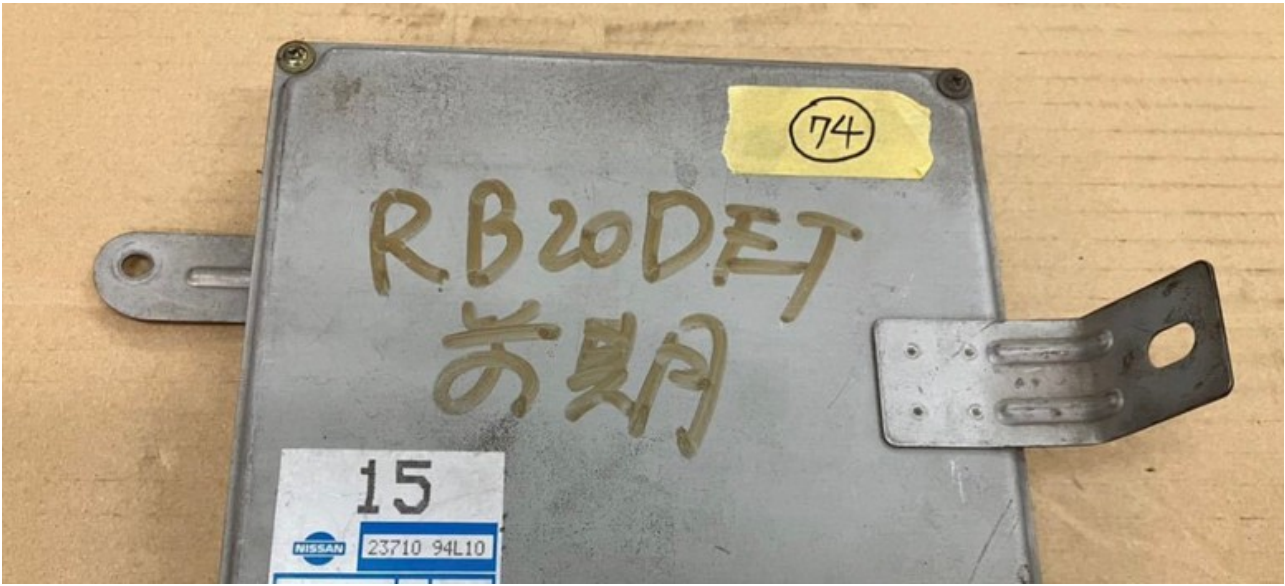from "Bamboo Blade" Episode 22, AIC ASTA, broadcast on 3 March, 2008

Fig. 39: "RB20DET 前期"
This picture is posted on Twitter by @taku_laurel33 on 17 October, 2022 URL :
https://twitter.com/taku_laurel33/status/1581973893056966657

Fig. 40: "駅前"
This picture is posted on Instagram by nukoaz on 7 December, 2019 URL :
https://www.instagram.com/p/B5xPxRoB_8g/
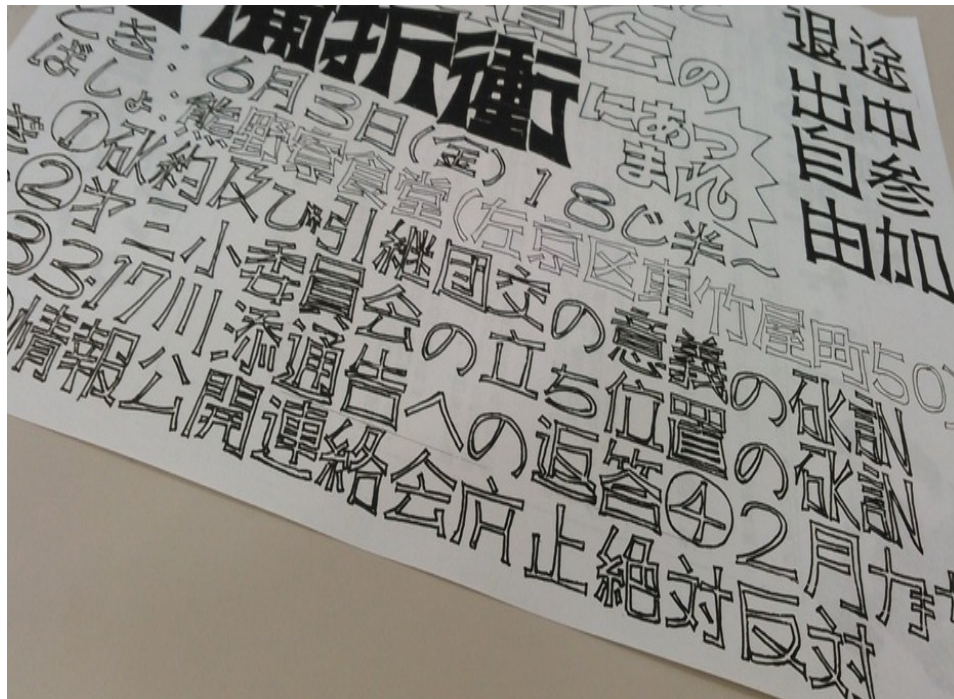
�<u>砅 for 確, 訊 for 認</u>

Fig. 41: "確約及び" and "意義の確認"

This picture is posted on tumblr on 17 December, 2016, from "熊野寮自治会ビラ", URL :
https://moji.tumblr.com/post/154570024094/%E7%86%8A%E9%87%8E%E5%AF%AE
%E8%87%AA%E6%B2%BB%E4%BC%9A%E3%83%93%E3%83%A9-via

Fig. 42: "入稿確認!! メール受領"

This image is posted on Twitter by @10boi__ on 20 August, 2020 URL :
https://twitter.com/10boi__/status/1296296804082831360

Fig. 43: "確認"
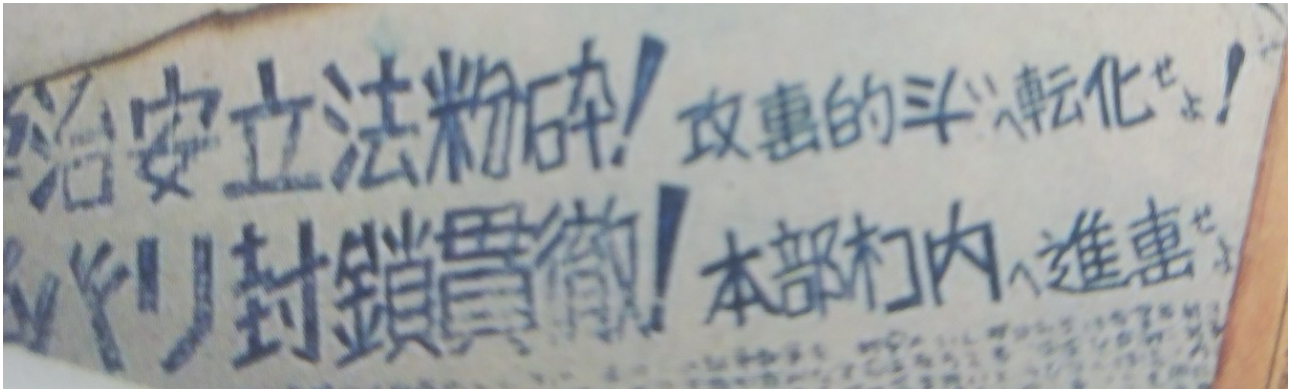*from 蜂矢真郷　著『国語文字史の研究』8（2005, 泉書院）*

杣 for 構

Fig. 44: "本部構内へ進垂せよ"

This picture is posted on Twitter by @jyunku on 3 November, 2020, from "京都大学 大学文書館だより Kyoto University Archives Newsletter #28, 京大闘争関係資料" URL : https://twitter.com/jyunku/status/1323377721183928320
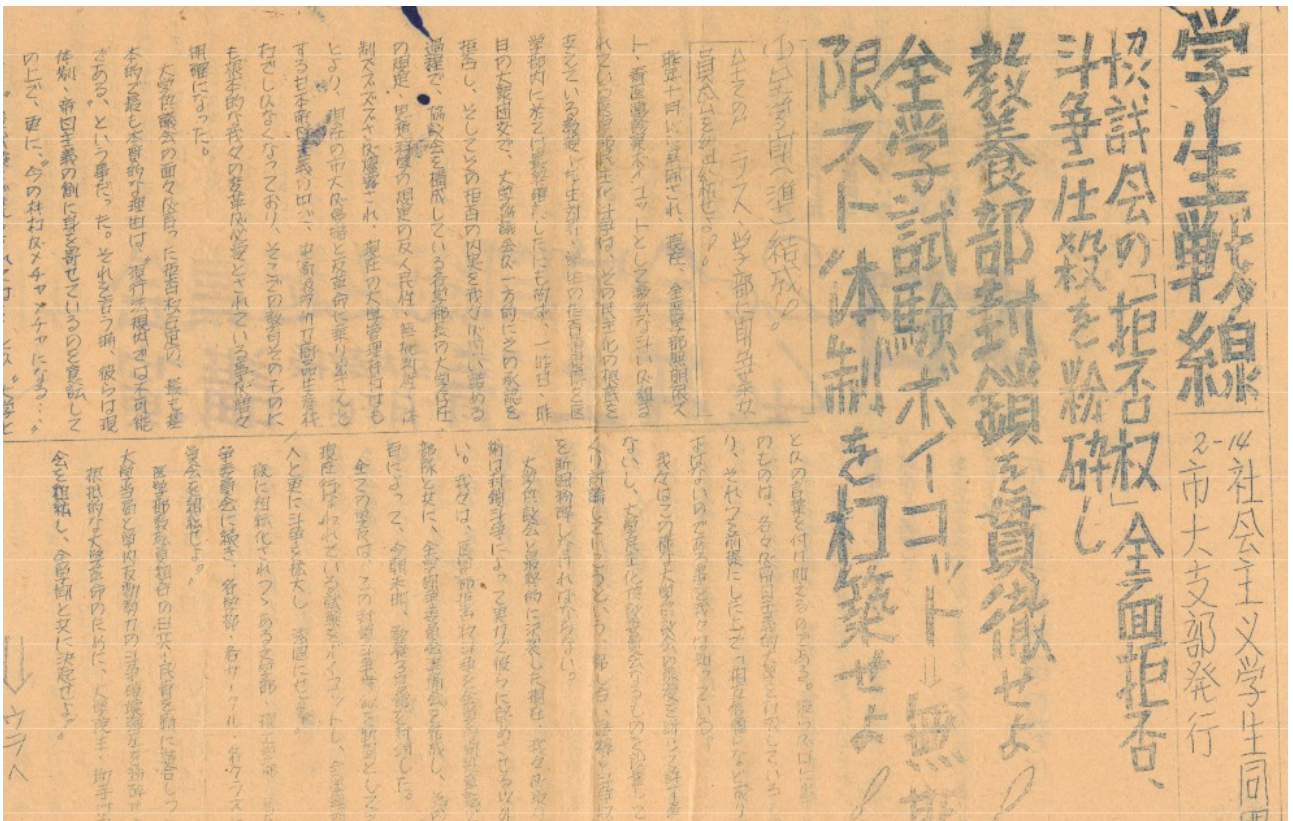
Fig. 45: "限スト体制を構築せよ！"

from '69 社学同市大支部, "学生戦線2.14", 大阪市立大学 , URL : https://assert.jp/wordpress/data/bira/196901/sgs19690214.pdf

Properties

All characters proposed follow the CJK Ideographs' properties.

U+XXXXX; <CJK Ideograph-XXXXX>;Lo;0;L;;;;;N;;;;;

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по Стандартизации

Doc Type:	Working Group Document
Title:	Feedback on hybrid characters
Source:	Kushim JIANG (姜兆勤)
Status:	Individual Contribution
Action:	For consideration by Gen KOJITANI and IRG
Date:	2023–10–05

The issues related to the hybrid characters contains following materials:
和 hybrid character 有关的议题包含如下材料：

- Online Review Tool, WS2021–00021 匞 <hc.jsecs.org/irg/ws2021/app/?id=00021>.
- Online Review Tool, WS2021–00029 㐤 <hc.jsecs.org/irg/ws2021/app/?id=00029>.
  - Based on the evidence, the left component should be analyzed as a Latin capital letter X and is therefore included in the list.
    依据字证，左侧构件应分析为大写拉丁字母 X，因此纳入本列表。
- Online Review Tool, WS2021–00718 鑅 <hc.jsecs.org/irg/ws2021/app/?id=00718>.
  - Sergeevich characterized the form as a combination of multiple Siddham grapheme clusters, and is therefore included in the list.
    Sergeevich 在审查意见中将该形刻画为多个悉昙形位簇，因此纳入本列表。
- Online Review Tool, WS2021–00760 乙 <hc.jsecs.org/irg/ws2021/app/?id=00760>.
  - Note that Eiso has overridden the relevant comments in IRG N2413 R2 by using contrary conclusion in its review comment (#11329) through a more careful analysis.
    请注意 Eiso 在审查意见中通过更细致的证据分析，使用相反的结论覆盖了 IRG N2413 R2 中与该字有关的意见。
- Online Review Tool, WS2021–00765 㗊 <hc.jsecs.org/irg/ws2021/app/?id=00765>.
- Online Review Tool, WS2021–00770 承 <hc.jsecs.org/irg/ws2021/app/?id=00770>.
- WANG Yifan (2019). *On Encoding Policy of Gongche Notations and Upcoming Para–ideographs (tentative)*. L2/19–346.
- Eiso CHAN (2019). *Request to clarify some IDS issues*. IRG N2413 R2.
- Gen KOJITANI (2023). *Script–hybrid and abbreviated CJK letters*. IRG N2637.
- Gen KOJITANI (2023). *Proposal to Encode Abbreviated Ideographs (Kanjis) in Japanese*. L2/23–139.
- Ken LUNDE. *Unihan Database Property "kStrange"*. UTN #43.
  - Informative. Please focus on Category H, Category K and Category O.
    资料性。请着重浏览 Category H、Category K 与 Category O。
- Andrew WEST. *BabelStone Fonts: BabelStone Han PUA*. <babelstone.co.uk/Fonts/PUA.html>
  - Informative. Please focus on the end of the repertoire.
    资料性。请着重浏览字汇结尾部分。

There are four levels of issues here.
这里有四个层面的问题。

The first level is whether the evidence meets the *IRG PnP* requirements. That is, whether the character is in Han script, whether the script is in Kaishu, and whether it has textuality.
第一层为字符是否符合 *IRG PnP* 要求。即文种是否为广义汉字、书体是否为楷书、是否具有文本性。

Textuality is easily recognized from the evidence, but the first two require additional consideration.
文本性从字证中容易获证，但前两者需要另外考察。

Han script does not require the language recorded, but requires the stroke system, the component system, and the cognition system be broadly consistent with Han ideographs. Therefore, if these hybrid characters belong to a writing system that uses Han ideographs, they can be generalized as belonging to Han script. So, the more important question is to determine whether these hybrid characters are or must be part of these writing systems.
广义现代汉文字对图符系统所记录的语言不作要求，但要求笔画系统、构件系统和理据系统与汉字大体一致。因此如果这些 hybrid character 处于使用广义汉字的书写系统中（方块壮文、方块布依文、方块苗文……），可以笼统地将其视作属于广义汉字。因此更重要的问题在于判断这些 hybrid character 是否属于或必须属于这些书写系统的一部分。

For example, should the pictographs in the text of Zhuang ideographs (mention in Liu, 2022; see **Figure 1**) be regarded as part of Han script? If not, the pictograph ℛ mentioned in L2/19–346 (page 9) should not be regarded as part of Han script either.
例如，刘建琼《云南方块壮字的构造》中提及的方块壮文中的"象形字"（参见 **Figure 1**）是否应当视作广义现代汉文字的一部分？若否，则 L2/19–346 所提象形字"ℛ"等也不应视作广义现代汉文字。

文字的遗存。但在云南宁蒗县壮族《麽经》文本中却看到了不少的象形字，如："▨"读作fu⁴²，意为"火"；"▨"读作huŋ⁴⁵，意为"猪肚子"；"▨"读作ȵɯŋ²⁴、"▨"读作ja:ŋ²⁴，都意为"粮食"；"▨"读作ma⁵⁵，意为"马"；"▨"，读作tha:u³¹，意为"桃子"；"▨"，读作ŋu:i²⁴，意为"山"。从这些实例可以看出，宁蒗壮族方块壮

Figure 1   Pictographs in Zhuang text
Liu J. (2022) "The structure of Zhuang characters of Yunnan", in *Journal of Honghe University*, 20(1).

Kaishu is one of the scripts that describes (narrow) Han ideographs. Thus, for one, non–Kaishu components will not be allowed. For example, 苏 in IRG N2637 (page 4, no. 24) directly uses non–Kaishu component and may not be allowed (Kaishu context cannot decide that the character belongs to Kaishu). For two, non–Han components will not be able to define Kaishu.
楷书是描述狭义现代汉文字的书体之一。从而，其一，非楷书的构件将不被允许。如 IRG N2637 中"苏"直接使用非楷书构件，可能不被允许（楷书上下文无法决定该字同属于楷书）。其二，非汉字的构件将无法定义楷书。

The second level is whether the evidence meets the *IRG PnP* requirements. That is, whether the evidence is reliable (confirmed by IRG), whether there are multiple pieces of evidence, whether function and cognition are clarified, whether context is included, and whether it is for justifiable public interest. These requirements are very clear, so there is only discussion in the resolutions and no additional clarification is needed.

第二层为字证是否符合 *IRG PnP* 要求。即字证是否可靠（由会议确认）、是否为多字证、是否包含功能与字理、是否包含上下文、是否出于正当公共利益。这些要求均非常明晰，因此只有决议上的讨论，没有需要额外说明的地方。

The third level is whether the abstract shape can be analyzed. If it can be determined that the character does use the grapheme cluster as a component , then the abstract shape can be analyzed, otherwise the abstract shape is unknown.

第三层为抽象构形是否可被分析。若能确定该字确实使用该形位簇作为构件，则可分析出抽象构形，否则抽象构形不明。

For example, L2/19–346 points out that 3 (now encoded as U+2CF36 丁) uses a component like digital three (or 丁 shape) to represent the image of a person, and more examples are given in Hu (2006). The abstract shape of the actual shape "3" or "丁" in these characters must therefore not be analyzed as the digital three or [丁] class, but rather as a hieroglyphic component.

如 L2/19–346 指出"3"（已编码为"丁"）使用的阿拉伯数字 3 的构件（或"了"形构件）实际为人象形，胡惠《方块壮字字体类型研究》中给出更多例字。这些字中的 3 形（或"了"形）的抽象构形因此不得分析为阿拉伯数字 3（或"了"类），而是某象形构件。

3	aemq	背	象人形背后背有东西，表示背的动作。	指示象事字
3	naengh	坐	"3" 表人形，"一" 表示人坐于其上	指示象事字
3	ndwn	站立	"、" 于上，表人直立	指示象事字
3	ngaem	低（頭）；頫（首）	"、" 于人形下表低头状	指示象事字

Figure 2　Pictographs in Zhuang text
Hu H. (2006) *A study on the type of Zhuang ideographs*, Master thesis, Guangxi University.

The last layer is whether the value for attributes can be defined, mainly in the RS and IDS information. Non–Kaishu components and non–Han components are unable to define radical and stroke information, while IDS may have to be extended, where each cell is a legal grapheme cluster.

最后一层为属性值是否可被定义，主要在于 RS 信息与 IDS 信息。非楷书构件与非广义汉字构件无法定义部首信息与笔顺信息；同时 IDS 可能须扩展，其中每个单元是一合法形位簇。

(End of Document)