

|         |  |
|---------|--|
| Source: | CheonHyeong Sim (沈天珩, aka TianHeng Shen), Ziheng Gao (高子恒) |
| Title:  | Proposal to Disunify U+5CC0 (𪗇 & 𪗈)                        |
| Status: | Individual Contribution on IRG #62                         |
| Action: | To be considered by IRG and multiple sources               |

## Background

U+5CC0 looks like the following picture in the CodeCharts of Unicode 15.1:



Fig.1 U+5CC0 in Unicode 15.1

And the following picture shows the CodeCharts of Unicode 16.0 Alpha Review, after the horizontal extension by J-source:



Fig.2 U+5CC0 in Unicode 16.0 (Alpha Review)

Additionally, from Unicode 3.1 (March, 2001), U+5CC0 has a compatibility character U+2F879 with its source TF-2662:

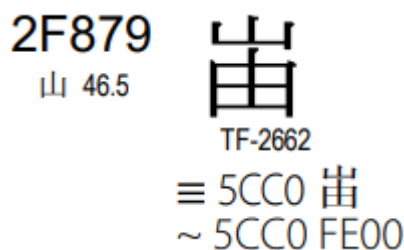


Fig.3 U+2F879 in Unicode 15.1

Obviously, U+5CC0 and U+2F879 mixed the two shapes 𪗇 and 𪗈. These two shapes are now considered as variants, but actually they are phonetically, graphically and semantically different, and need to be disunified. In [IRGN2210](#), Eiso Chan once raised this

issue to TCA but TCA disagreed to disunify this character. However, this is not an issue solely related to TCA. Please see the following description.

## Properties of 岫 and 岫

For the former one, in IRGN2210, Eiso Chan gave out some evidences related to G-source and K-source to show that 岫 is a variant of 岫. Due to space constraints, the evidences in IRGN2210 will not be repeated here. In this proposal, I would like to add evidences related to T-source and KP-source.

Although 教育部異體字字典 does not include 岫 as a character entry, in 形體資料表 column of 岫, 岫 appears as its variant:

國家教育研究院 維護

部首查詢 | 筆畫查詢 | 單字查詢 | 複合查詢 | 注音查詢 | 漢語拼音查詢 | 倉頡碼查詢 | 四角號碼查詢 | 附收字查詢

A01110 **岫** 解說 異體字 研訂瀏覽 回查詢結果

異體字

岫 岫

解說

|      |   |    |                 |
|------|---|----|-----------------|
| 字號   | A01110  | 正字 | 【岫】 山 - 05 - 08 |
| 說文釋形 | 大徐本：岫，山穴也。从山，由聲。𡵓，籀文，从宀。<br>(似又切)<br>段注本：岫，山有穴也。从山，由聲。𡵓，籀文，从宀。<br>(似又切)                 |    |                 |
| 字樣說明 | 右半作「由」：筆順如「田」，中豎筆上須出頭，寫法參「由」字。  |    |                 |
| 注音   | ㄒㄩˋ   |    |                 |
| 漢語拼音 | xiù   |    |                 |
| 釋義   | 1. 山洞、巖洞。段注本《說文解字·山部》：「岫，山有穴也。」晉·張協〈雜〉詩一〇首之九：「荒庭寂以閉，幽岫峭且深。」宋·李師中〈詠松〉詩：「半依岫倚雲端，獨上亭亭耐歲寒。」 |    |                 |

形體資料表  
重訂直音篇

岫  
音  
襍  
山  
有  
岫  
上  
洞

卷六·山部·頁474·右

Fig.4 岫 in 教育部異體字字典

It matches the statement in IRGN2210.

In KP0, the characters are sorted by the Korean pronunciation in the DPRK order, and we could see that 岫 is located between 岫 (pronounced “su”) and 帥 (pronounced “su”), so its pronunciation should undoubtedly be “su”.

zi 組字 聽 搜字 字音 字碼 字源 譯名 隨緣  
Compose Search Phonology Encoding Origin Transl. Random

U G H M T J K P V B S 篆

KP0 KP1

KPS 9566 (DPRK Standard Korean Graphic Character Set for Information Interchange) 국규 9566 國規 9566

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 |
| 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 |
| 89 | 90 | 91 | 92 | 93 | 94 |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |

少 巢 所 掃 搔 昭 梳 沼 消 溯 瀟 炤 燒 甦 疏 疎 瘡 笑 篠  
6701 6702 6703 6704 6705 6706 6707 6708 6709 6710 6711 6712 6713 6714 6715 6716 6717 6718 6719

簫 素 紹 蔬 蕭 蘇 訴 迨 遡 邵 銷 韶 騷 俗 屬 束 涑 粟 續 謾  
6720 6721 6722 6723 6724 6725 6726 6727 6728 6729 6730 6731 6732 6733 6734 6735 6736 6737 6738 6739

贖 速 孫 巽 損 蓀 遜 滄 率 宋 棟 松 淞 訟 誦 送 頌 修 受 售  
6740 6741 6742 6743 6744 6745 6746 6747 6748 6749 6750 6751 6752 6753 6754 6755 6756 6757 6758 6759

嗽 囚 垂 壽 嫂 守 岫 岫 帥 愁 戍 手 授 搜 收 數 樹 殊 水 洙  
6760 6761 6762 6763 6764 6765 6766 6767 6768 6769 6770 6771 6772 6773 6774 6775 6776 6777 6778 6779

漱 燧 狩 獸 琇 璫 瘦 睡 秀 穗 豎 粹 綏 綏 繡  
6780 6781 6782 6783 6784 6785 6786 6787 6788 6789 6790 6791 6792 6793 6794

Fig.5 KPS 9566 on [zi.tools](http://zi.tools)

In KP1, the characters are sorted by the radical first; for the same radical, the characters are sorted by the residual strokes; and when the residual strokes are also the same, they are still sorted by pronunciation (Sim, 2022). We could see that 岫 is located between 岫 (pronounced “su”) and 岫 (pronounced “jeo”), so its pronunciation should very likely be “su” (the variant of 岫), and impossible to be “bang” (岫, the variant of 邦), since b comes before s.

3F0\_ 屨  
 3F1\_ 屨  
 3F2\_ 屨  
 3F3\_ 屨  
 3F4\_ 屨  
 3F5\_ 屨  
 3F6\_ 屨

Fig.6 [KPS\\_10721\\_reconstructed\\_by\\_CheonHyeong\\_Sim](#)

It also matches the statement in IRGN2210.

For the latter one, in IRGN2210, Eiso Chan also gave out some evidences related to G-source and T-source to show that 屨 is a variant of 邦; and for G-source, it is also an independent character used for a place name different from 邦. Due to space constraints, the evidences in IRGN2210 will not be repeated here. In this proposal, I would like to add evidences related to J-source.

In 《大漢和辭典》 (*Daikanwa* Dictionary), 屨 is interpreted as the ancient form of 邦, thus it is a variant of 邦.

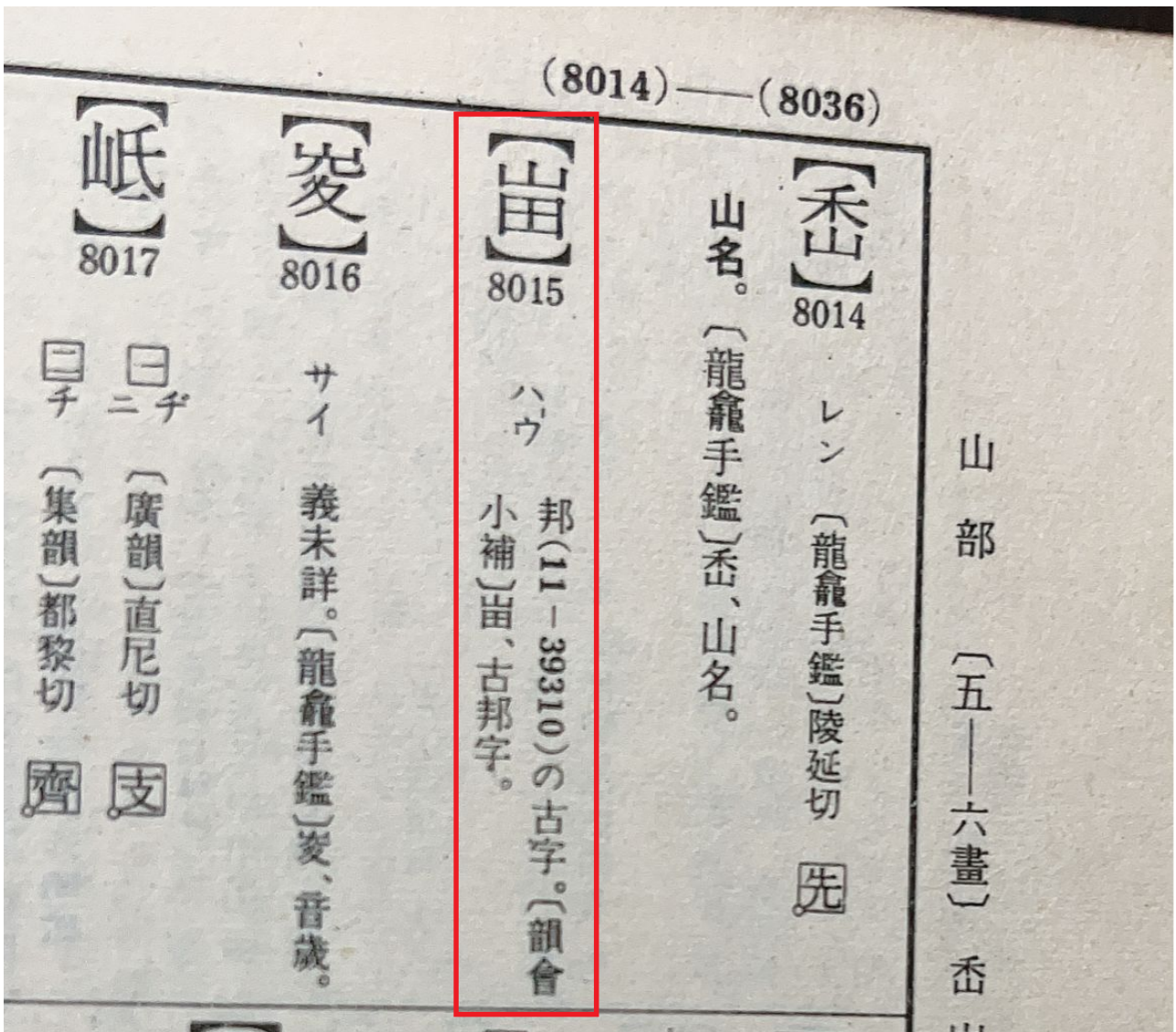


Fig.7 《大漢和辭典》 P3646 (i.e. 第4冊 P244)

It also matches the statement in IRGN2210.

By the way, K-source also includes 岷 as KC-07223; but in their [CJK Ideographs Search System](#), the given provenance and context in [the page of KC-07223](#) are not for 岷 but for 岷 (U+31844, KC-07242). Maybe the database has some malpositions, and I hope that the Korean experts could give out the right provenance and context for 岷, to show us whether it is also the variant of 邦 or not.

## Summary

To briefly sum up, we can list 岷 and 岷 in multiple sources as the following table: (Note that, red indicates the current sources for U+5CC0, including the newly horizontally-extended J-source in Unicode 16.0 Alpha Review)

|   | G-source                                    | T-source | J-source   | K-source | KP-source              |
|---|---|----------|------------|----------|------------------------|
| 畠 | GE-3735                                     | TF-2662  | JMJ-068030 | K0-617C  | KP0-E3E3<br>(KP1-3F3E) |
| 畠 | GKX-1594.81<br>GHZR-10791.16<br>(GDM-?????) | T4-2634  | JMJ-010468 | KC-07223 | /                      |

## Suggestions on Disunifying

As mentioned above, 畠 and 畠 are phonetically, graphically and semantically different in multiple sources. Thus it is unreasonable to unify them.

Since K0 and KP0 are the basic character sets in South Korea and North Korea, the correspondence between them and Unicode are not able to be changed. We suggest to keep 畠 on U+5CC0 – just as in the initial Unicode version; and move 畠 to a new codepoint (maybe as a UNC, or maybe as a new G-source character to WS2024 as the place name).

In other words, TCA is suggested to move T4-2634 to a new codepoint, TF-2662 to U+5CC0, and change the source reference on U+2F879 to TU-2F879; Japan is suggested to move MJM-010468 to a new codepoint, and add MJM-068030 for U+5CC0; China and (South) Korea are suggested to do a horizontal extension on the new codepoint, or to submit 畠 as a new character.

(End of document)

# IRGN2676\_KIM\_resp1

(= k2680\_811\_IRGN2676\_U+5CC0\_KIM\_K)

Subject: KIM's response to IRGN2676Disunify5CC0.pdf (2024-03-09)

Date: 2024.03.13.

Status: Individual opinion

Authors: KIM Kyongsok

## 1. Background

- This is a response to IRGN2676Disunify5CC0.pdf (2024-03-09) by CheonHyeong Sim (沈天珩, aka TianHeng Shen) and Ziheng Gao (高子恒).

- IRGN2676 is much related with the issue #1 in IRGN2210R\_U+5CC0\_U+2F936\_\_Eiso\_CHAN dated on 2017-06-08 at IRG #48.

- At IRG #48, I could find the following comments in the editorial report.

*o IRGN2210 Dis-unification of 2 Characters*

*The editors reviewed all raised issues and will respond before next IRG meeting. For Issue 2, .... Need consideration*

- TCA seems to have said that TCA needed more time to investigate this issue. However, the author could not find a response from TCA after IRG #48 (please correct me if I am wrong).

- It seems that, after IRG #48, the issue of disunifying U+5CC0 mentioned in IRGN2210R has not been discussed.

- The current situation of two characters, 𪗇 U+5CC0 and 𪗈 U+2F879 (Compatibility char.) is well summarized in IRGN2676 and repeated below:

|   | G-source                                    | T-source | J-source   | K-source | KP-source              |
|---|---|----------|------------|----------|------------------------|
| 𪗇 | GE-3735                                     | TF-2662  | JMJ-068030 | K0-617C  | KP0-E3E3<br>(KP1-3F3E) |
| 𪗈 | GKX-1594.81<br>GHZR-10791.16<br>(GDM-?????) | T4-2634  | JMJ-010468 | KC-07223 | /                      |

5CC0  
山 46.5

𪗇  
GE-3735

𪗈  
T4-2634

𪗈  
JMJ-  
010468

𪗇  
K0-617C

𪗇  
KP0-E3E3



## 2. Review results

- The author reviewed the evidences and opinions in the two documents.
- It seems that, based on the evidences in the two documents, the suggestions in the two documents are quite logical and reasonable.

## 3. Suggestions

- Basically, the suggestions below are based on the suggestions in the two documents.

1) TCA could consider moving T4-2634 glyph and source reference in U+5CC0 to a new character (codepoint) 𪗇 (NOT to U+2F879, a compatibility char).

2) TCA could consider moving TF-2662 glyph and source reference in U+2F879 to U+5CC0.

3) TCA could consider changing the T source reference of U+2F879 to TU-2F879.

4) Japan could consider moving JMJ-068030 glyph and source reference to a new character (codepoint) mentioned in 1) above.

5) Japan could consider adding JMJ-010468 glyph and source reference to U+5CC0.

6) Rep. of Korea (src. ref. KC-07223) and China (src. ref. GKX-1594.81) could consider

6a) doing a horizontal extension on the new character (codepoint) mentioned in 1) above; or

6b) submitting 𪗇 as a new character.

7) IRG could consider removing UCV #94.

|    |        |  |  |   |  |  |                      |
|----|--------|--|--|---|--|--|----------------------|
| 94 | 𪗇<br>𪗇 |  |  | <small>092/192</small><br><b>5CC0</b><br><small>E-2635</small><br><small>E-0621</small> | <small>4-2634</small><br><small>4-0620</small> | <small>0-0170</small><br><small>0-0332</small> | <small>2F879</small> |
|----|--------|--|--|---|--|--|----------------------|



4. RE: the wrong information about KC-07223 in <https://www.koreanhistory.or.kr/newchar/>

- The evidence information about KC-07223 in koreanhistory.or.kr is NOT about KC-07223, but about a different character 擷 (KC-07242, U+31844).

- I will try to contact the manager of this web site. However, it may not be easy to get the correct evidence information since the evidence information is not updated these days.

|        |  |        |         |
|--------|--|--------|---------|
| 자형     |  <span style="float: right; border: 1px solid black; padding: 2px;">이체자 보기</span> |        |         |
| 부수     | 山  | 통합코드   | KC07223 |
| 획수     | 5  | 총획수    | 8       |
| KS코드   |  | JIS코드  |         |
| GB코드   |  | BIG5코드 |         |
| 한국뜻    |  |        |         |
| 한글음가   |  |        |         |
| 영어뜻    |  |        |         |
| 한어병음   |  |        |         |
| 일본어 훈독 |  |        |         |
| 사각호마   | 파자 (山/田)   |        |         |

# The following information is about KC-07242 (=U+31844), NOT about KC-07223.

|         |  |       |                           |
|---------|--|-------|---------------------------|
| 문자이미지파일 |  | 기관명   | 한국고전번역원                   |
| 사업차수    | 2008   | 관련DB  | 고전번역총서                    |
| 출전서명    | 용재집 (容齋集)  | 출전목록  | 3책                        |
| 출전위치    | [부]김석근(金@根)  | 출전파일명 | exm_UM_6JOc_JS_001318.xml |
| 전후문맥    | <pre> &lt;전력 type="전력" 식별자="" id=""&gt;무오 혼충(戊午魂忠)&lt;/전력&gt; &lt;관직 type="관직" 식별자="" id=""&gt;&lt;/관직&gt; &lt;봉작 id="" 식별자=""&gt;&lt;/봉작&gt;&lt;/관력&gt; &lt;타과 type="타과"&gt;&lt;/타과&gt; &lt;본관 id="296" 식별자=""&gt;안동(安東)&lt;/본관&gt; &lt;기타&gt;&lt;/기타&gt; &lt;/본인&gt; &lt;친족 가계="친족" 대수="1" 유형="부"&gt; &lt;인명 id="" type="성명" 식별자=""&gt;김석근(金{才+稟}根)&lt;/인명&gt; &lt;인명 type="자"&gt;&lt;/인명&gt; &lt;인명 type="호"&gt;&lt;/인명&gt; &lt;관력&gt; &lt;관직 type="관직" 식별자=""                     </pre> |       | 원전이미지파일명                  |

\* \* \*