

Subject: KR Experts' comment RE: 𠄎 WS2021 SN 02280 (UTC-03193)
Date: 2022-03-11; 2022-03-16 (b)
Authors: SHIN Sanghyun, CHO Sungduk, CHOI Jinwook, KIM Kyongsok (KR)
Relevant documents: IRG N2516

Note: Paragraph 3.4) was added in a revised version (2022-03-16, KR_comment1b)

1. Background

1) UTC proposed the following char in WS2021:



WS2021 SN 02280 = UTC-03193

2) The relevant information in IRG ORT (Online Review Tool) is shown below:

<https://hc.jsecs.org/irg/ws2021/app/?id=02280>

The screenshot shows the IRG ORT interface for character 𠄎 (U+02280). The interface includes a header with navigation arrows and the character code point 02280. Below the header is a table with the character 𠄎 and its code point UTC-03193. The table also shows similar characters: 𠄎 (U+5655), 𠄎 (U+7233), 𠄎 (U+7232), and 𠄎 (U+3018A). The interface also displays attributes for the character, including SC, FS, and TC, and a table comparing expected and recorded counts for these attributes. A note at the bottom states: "Note: expected counts are determined algorithmically from the given IDS and may be inaccurate in edge cases."

Char	SC	FS	TC
𠄎 (U+7232)		?	?
𠄎 (U+3018A)		?	?

	Expected	Recorded
SC	/	10 Radical not found
FS	/	5 Radical not found
TC	0	14 Mismatch

3) The source "GK" for U+7233 is "GB12052-89, Korean Character Coded Character Set for Information Interchange".

- Row 72 of GB12052-89 is shown in Fig. 1 below and row 93 on page 24 of some unknown document is shown in Fig. 2 below.

- Row 93 in Fig. 2 seems a "virtual" row 93 of G1 (GB12345-90).

- GB12052-89 in Fig. 1 contains characters 𪛗 U+7233 and 𪛘 UTC-03193.

- Some unknown document in Fig. 2 contains 𪛗 U+7233 and 𪛙 U+5655.

中华人民共和国国家标准

GB 12052-89

信息交换用朝鲜文字编码字符集

Korean character coded character
set for information interchange

72区	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
01 19	𪛗	鹿	尔	𪛘	𪛙	𪛚	𪛛	𪛜	𪛝	𪛞	𪛟	𪛠	𪛡	𪛢	𪛣	𪛤	𪛥	𪛦	𪛧
20 39	𪛨	𪛩	𪛪	𪛫	𪛬	𪛭	𪛮	𪛯	𪛰	𪛱	𪛲	𪛳	𪛴	𪛵	𪛶	𪛷	𪛸	𪛹	𪛺
40 59	𪛻	𪛼	𪛽	𪛾	𪛿	𪜀	𪜁	𪜂	𪜃	𪜄	𪜅	𪜆	𪜇	𪜈	𪜉	𪜊	𪜋	𪜌	𪜍
60 79	𪜎	𪜏	𪜐	𪜑	𪜒	𪜓	𪜔	𪜕	𪜖	𪜗	𪜘	𪜙	𪜚	𪜛	𪜜	𪜝	𪜞	𪜟	𪜠
80 94	𪜡	𪜢	𪜣	𪜤	𪜥	𪜦	𪜧	𪜨	𪜩	𪜪	𪜫	𪜬	𪜭	𪜮	𪜯	𪜰	𪜱	𪜲	𪜳

Fig. 1. Row 72 of GB12052-89 (GK)

93区	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
01/19	𪛗	鹿	尔	𪛘	𪛙	𪛚	𪛛	𪛜	𪛝	𪛞	𪛟	𪛠	𪛡	𪛢	𪛣	𪛤	𪛥	𪛦	𪛧
9	𪛨	𪛩	𪛪	𪛫	𪛬	𪛭	𪛮	𪛯	𪛰	𪛱	𪛲	𪛳	𪛴	𪛵	𪛶	𪛷	𪛸	𪛹	𪛺
40/59	𪛻	𪛼	𪛽	𪛾	𪛿	𪜀	𪜁	𪜂	𪜃	𪜄	𪜅	𪜆	𪜇	𪜈	𪜉	𪜊	𪜋	𪜌	𪜍
60/79	𪜎	𪜏	𪜐	𪜑	𪜒	𪜓	𪜔	𪜕	𪜖	𪜗	𪜘	𪜙	𪜚	𪜛	𪜜	𪜝	𪜞	𪜟	𪜠
80/94	𪜡	𪜢	𪜣	𪜤	𪜥	𪜦	𪜧	𪜨	𪜩	𪜪	𪜫	𪜬	𪜭	𪜮	𪜯	𪜰	𪜱	𪜲	𪜳

Fig. 2. Row 93 on page 24 of some unknown document
Probably a "virtual" row 93 of G1 (GB12345-90)

2. Review of 𪛑 U+5655 and 𪛒 U+7233 similar to 𪛓 WS2021 SN 02280 (UTC-03193)

- Chars 𪛑 U+5655 and 𪛒 U+7233 are similar to 𪛓 WS2021 SN 02280 (UTC-03193) proposed by UTC in WS2021.

- Therefore, before reviewing 𪛓 WS2021 SN 02280 (UTC-03193), similar chars 𪛑 U+5655 and 𪛒 U+7233 are reviewed in 2.1 and 2.2, respectively.

2.1 Review of 𪛑 U+5655



1) 𪛑 U+5655 is a variant of 𪛒 U+5645

2) The Korean reading of 𪛒 U+5645 is "wi (위)" and its meaning is "lie (거짓말)".

3) The Korean reading of 𪛑 U+5655 as an ordinary CJK char is "wi" ("위") and its meaning is "lie" ("거짓말").

4) KR could not find a document showing that 𪛑 U+5655 is used as a Idu/Gugyeol char.

5) However, supposing that 𪛑 U+5655 is "really" used as an Idu/Gugyeol char., the reading of 𪛑 U+5655 might be "hago (하고)".

6) In Gugyeol, "hago (하고)" is represented as "ㄴ 口".

- Gugyeol char "ㄴ" is a short form of the original char 爲 U+7232 (reading is "ha (하)") as Gugyeol.

- Gugyeol char "口" is a short form of the original char 古 U+53E4 (reading id "go (고)").

- The meaning of "hago (하고)" is "do (and) ...".

7) Although "ㄴ 口" (hago (하고), 2 chars) is used in Gugyeol, it is not evident that 𪛑 U+5655 (𪛑爲口) is used as Gugyeol. Evidences must be provided to claim that 𪛑 U+5655 (𪛑爲口) is read as "hago (하고)" in Gugyeol.

8) It is not evident that 𪛑 U+5655 (𪛑爲口) can be read as "go (고)".

2.2 Review of 𪛗 U+7233

1) 隱 U+96B1 is used to represent a syllable-final consonant “ㄴ (n sound)” in Idu/Gugeol.

2) It "seems" that the short forms of the left component of char 隱 (hide), "𠃉" (mound), are derived in two different ways, N1 and N2, as shown in Fig. 3.

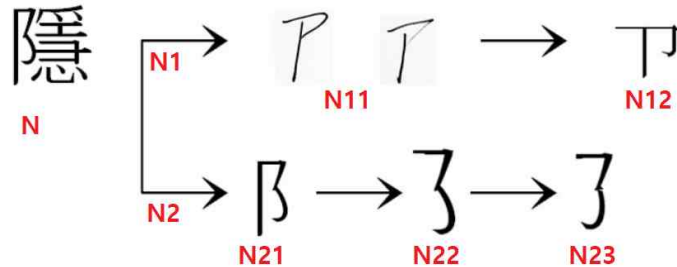


Fig. 3. Short forms of 隱 (hide, reading: Eun, 숨길 은)

3) In general, although there is some kind of consistency in Gugyeol and in Idu, they are **not totally systematic**.

Gugyeol and Idu are used and developed gradually for a long time by many people in different regions. In other words, Gugyeol and Idu were NOT developed by one person or by a group of people at some specific time.

Therefore, a person P1 at time T1 in region R1 may follow convention C1 which is **slightly different** from convention C2 that a person P2 at time T2 in region R2 follows.

4) The bottom component 𠃊 (will be referred to as "N12". see Fig. 3) in UTC-03193 is NOT a variant of the following SEAL radicals.

2F19 𠃊 KANGXI RADICAL SEAL, ≈ 5369 𠃊
2E8B 𠃊 CJK RADICAL SEAL • form used at bottom → 353E 𠃊

5) Short form N22 (and N23 too) in Fig. 3 seems derived from N21 in Fig. 3 by omitting a vertical stroke. N23 is very similar to 了 (U+4E86, reading: 료/요 (ryo/yo), meaning: complete).

6) Short form N23 should NOT be considered a wrong form written by a mistake.

7) According to Yuseo Pilji (儒胥必知)" (see Appendix 1), Idu 𪛗 U+7233 is read as "han" ("ᄒᆞᆫ" in Old Hangul, "한" in Modern Hangul).

8) Gugyeol 爲隱 (two chars) is read as "han (ᄒᆞᆫ)" (a single syllable). In Gugyeol, 爲 is read as "ha (ᄒᆞ)" and 隱 is read as syllable-final "n (ㄴ)".

9) KR could not find a document showing 𪛗 U+7233 other than Yuseo Pilji (儒胥必知)".

3. Review of 𠄎, WS2021 SN 02280 (UTC-03193)

1) Short forms N12 and N23 in Fig. 3 are very different and, therefore, in principle, 𠄎 UTC-03193 (whose bottom component is N12) should NOT be unifiable with 𠄎 U+7233 (whose bottom component is N23).

2) KR tried to find a document showing that 𠄎 UTC-03193 is really used, not in a standard or a dictionary. However, KR failed to find such a document.

3) Therefore, if an evidence (a document) showing that 𠄎 UTC-03193 is really used can be provided, KR suggests that 𠄎 UTC-03193 be encoded as a new char in UCS separately from 𠄎 U+7233.

4) When we talk about normalization or unification, there must be two distinct "real" characters found in real documents, not just in a standard or a dictionary.

In case of 𠄎 UTC-03193, this "whole" character is not a real character found in an actual document although the bottom component is found/used in other characters.

In other words, 𠄎 UTC-03193 as a whole is NOT a real character found in real documents, but a character "composed" of two components each of which is really used in different characters but not in one character.

4. A tentative conclusion

- IRG members need to try to find an evidence (a document) showing that 𠄎 UTC-03193 is really used to be able to encode 𠄎 UTC-03193 in UCS (i.e., a vertical extension).

* * *

Appendix 1. Two versions of book "Yuseo Pilji (儒胥必知)" showing 學 U+7233

- There are at least 11 versions of book "Yuseo Pilji (儒胥必知)".
- The pages of two versions containing 學 U+7233 are shown below.
- The first version is wood-cut and the second version is hand-written (calligraphy).

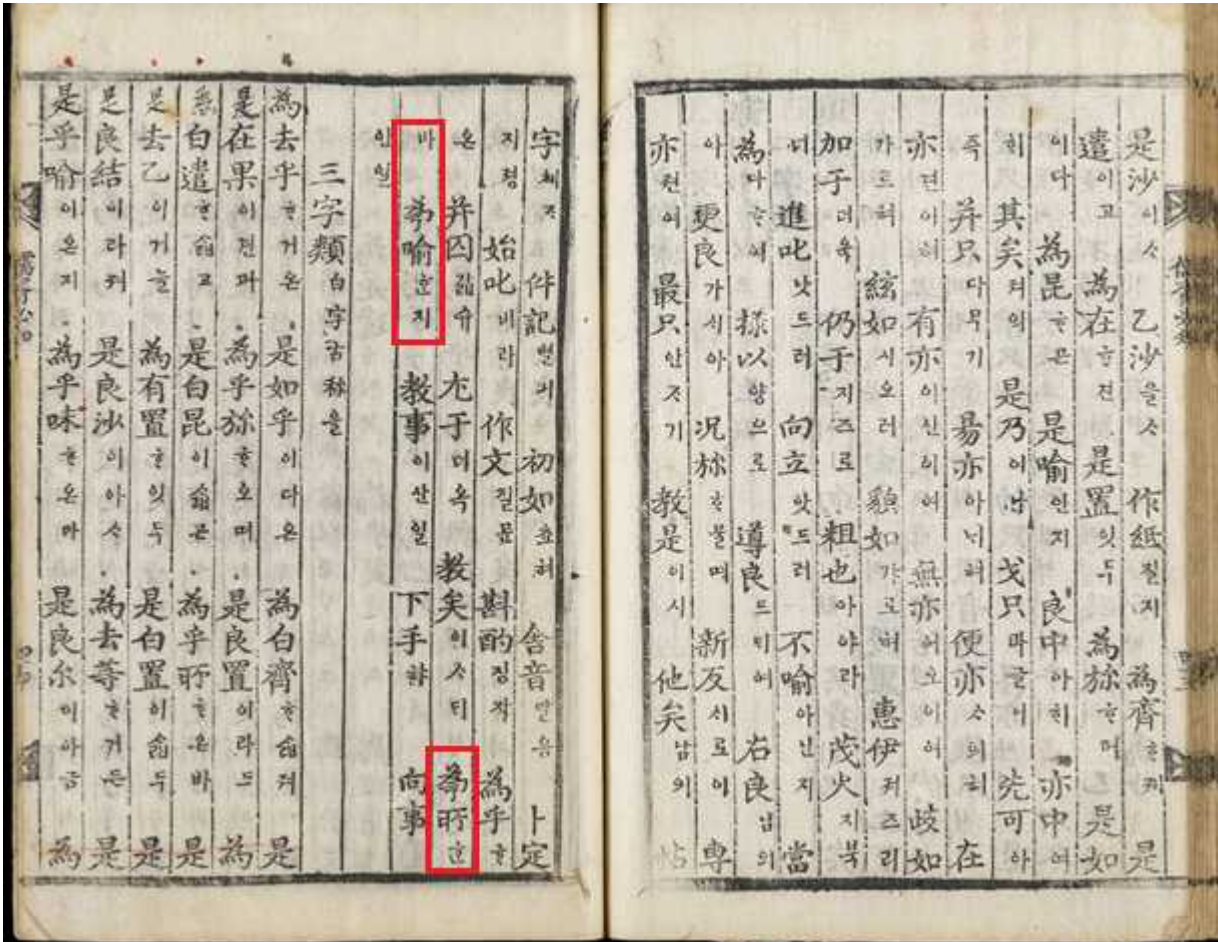


Fig. A1a. 儒胥必知_04_051.jpg

良結 <small>이온지</small> 乎喻 <small>이온지</small> 良結 <small>이온지</small>	去己 <small>이거늘</small> 是良泐 <small>이아스</small> 為乎味 <small>이온마</small> 是良尔 <small>이아근</small> 為 <small>이온</small>	白遣 <small>이순고</small> 是白昆 <small>이순고</small> 為有置 <small>이순드</small> 是白置 <small>이순드</small> 是 <small>이순드</small>	在果 <small>이견과</small> 為乎旅 <small>이오며</small> 是良置 <small>이라드</small> 為乎所 <small>이순바</small> 是 <small>이순바</small>	為去乎 <small>이거온</small> 是如乎 <small>이다온</small> 為白齊 <small>이순저</small> 是 <small>이순저</small>	三字類 <small>白字古絲을</small> 安일	始叱 <small>비타</small> 作文 <small>질문</small> 斟酌 <small>김작</small> 為乎 <small>이</small>	并囚 <small>이순</small> 尤于 <small>더우</small> 教矣 <small>이서티</small> 為乎所 <small>이</small>	為乎喻 <small>이순지</small> 教事 <small>이산일</small> 下手 <small>하</small> 向事
-------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------	--------------------------------	---------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------	------------------------------------------------------------------------------

Fig. A1b. 儒胥必知_07_051a
