

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

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG
Title: UK Response to Consolidated Comments on IRGN2155 IRG Working Set 2015v2
Source: Andrew West
Status: Member Body Contribution
Action: For consideration by IRG
Date: 2016-10-01

This document provides the UK responses to the consolidated comments on IRG Working Set 2015 ([IRGN2155ConsolidatedCommentsPage01-28.pdf](#) and [IRGN2155ConsolidatedCommentsPage29-67.pdf](#)). The responses refer only to those characters submitted by the UK (see IRGN2107). The UK responses to other comments on UK-submitted characters are in [IRGN2155 UK ReviewP1.pdf](#) and [IRGN2155 UK ReviewP2.pdf](#). (2016-07-13).

1. Unifications

S/N	Ref.	Glyph	Comment	UK Response
00719	UTC-02953	垸	H: Unifiable with U+580D 垸	<p>Do not unify as it is not cognate with U+580D 垸, and the presence or absence of the dot is significant in distinguishing the two characters. U+580D is read tù (with 兔 phonetic), but UTC-02953 is read wǎn (with 免 phonetic), and thus not cognate with U+580D. See 国家基础地理信息系统 1:5 万地名数据库 (1:50000 Placename database of national basic geographical information system) for Liwan Village 李垸村 in Hubei.</p> <p>See also article from <i>Hubei Daily</i> 湖北日报 2015-03-04 page 14 which discusses the correct glyph and reading for the character used in the name of Yuwan Village (see image below).</p>
01416	UTC-02632	揠	H: Duplicate of U+2BF4A.	<p>As discussed in IRGN2108 Andrew WG2N4682.pdf the glyph for U+2BF4A is a mistake. U+2BF4A is the simplified form of U+3A6D 揠, and its correct glyph should be 揠. 01416 (UTC-02632) is a different, non-cognate character (variant of U+632C 揠), and is not unifiable with U+2BF4A. Therefore keep in IRG WS2015.</p>
02995	UTC-01508	稷	J: Unifiable with U+7A37 稷.	<p>This is not a "minor stroke difference", but a completely different component on the bottom right. Unification is not appropriate in this case.</p>

S/N	Ref.	Glyph	Comment	UK Response
03332	UTC-01401	𦉰	SAT: Unifiable with U+2627C 哭.	This character has been withdrawn. See IRGN2155 UK ReviewP1.pdf p. 10.
03590	UTC-01962	崇	J: Unifiable with U+8417 崇.	03590 (UTC-01962) is the simplified form of U+799C 崇 (yíng 'a kind of ancient sacrifice'), and therefore not cognate with U+8417 (corrupt form of U+7B56 策 cè). The presence or absence of the dot is significant in this case. Do not unify, per PnP draft 2.1.3: "Ideographs with different glyph shapes that are unrelated in historical derivation (non-cognate characters) are not unified no matter how similar their glyph shapes may be."
03611	UTC-01967	蓼	H: Unifiable with U+26C9E 蓼.	Agree. U+26C9E only has T-source and U-source references. It would be useful if U+26C9E could also be given a HDZ G-source reference with the glyph shape of UTC-01967.
03707	UTC-01911	蛭	J: Unifiable with U+86CF 蛭.	03707 (UTC-01911) is the simplified form of U+86F5 蛭 (xīng, in Ancient Chinese 丁蛭 'dragonfly'), and therefore not cognate with U+86CF (chēng 'razor clam'). Do not unify, per PnP draft 2.1.3: "Ideographs with different glyph shapes that are unrelated in historical derivation (non-cognate characters) are not unified no matter how similar their glyph shapes may be."
03774	UTC-01412	衡	J: Unifiable with U+22545 衡.	This character has been withdrawn. See IRGN2155 UK ReviewP1.pdf p. 10.

S/N	Ref.	Glyph	Comment	UK Response
03877	UTC-02857	𪛗	SAT: Unifiable with U+27B0C 𪛗.	Agree.
04310	UTC-02926	𪛘	Unifiable with U+8598 𪛘.	<p>Do not unify.</p> <p>According to S.1.4.2 “Different relative position of components”, UTC-002926 is not unifiable with U+8598. Related examples of non-unification include:</p> <p>U+26F5F 𪛗 and U+2C7A4 𪛗</p> <p>U+8596 𪛗 and U+285F2 𪛗</p> <p>U+3C08 𪛗 and U+2E7DD 𪛗</p>

「垸」字引出的文字公案

玉垸村是明朝由江西迁居阳新的单一血缘村落。村里居民全为李氏宗亲，由赣入鄂已有400多年的历史。

阳新地处吴头楚尾，方言众多，缘于此，阳新语源集赣、湘方言与楚语于一体，形成了独特的“吴楚”方言。其“九板十三腔”被语言学家王力教授称之为研究古方言的“活化石”。

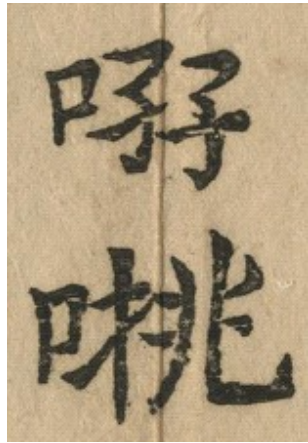
就说这玉垸的“垸”字吧，词典释文是“桥两旁靠近平地的地方”，从土从兔，音读 tù 形声。玉垸人说城里人把咱们的村名写错了，也读错了，应该是“土”旁加一个“兔”字，读“碗”音。但这个“土兔”字在字典怎么也查不出来，玉垸人就说，这个字，在《康熙字典》里才查得出，是古代的人声字。

<http://hbrb.cnhubei.com/html/hbrb/20150304/hbrb2573592.html>

2. Attributes

We have no objections to any proposed changes in attributes.

3. Font Design

S/N	Ref.	Glyph	Comment	UK Response
00045	UTC-02964	収	T: There's too much space below the glyph.	Agree. We will provide the editor with a new glyph image: 収
00046	UTC-02965	𠂇	T: There's too much space below the glyph.	Agree. We will provide the editor with a new glyph image: 𠂇
00122	UTC-02967	𠂇	T: The glyph should be put in the middle.	Agree. We will provide the editor with a new glyph image: 𠂇
00617	UTC-02917	𠂇	H: The middle component looks like (Hand, R64) rather than (Tree, R75), and the SC should be 9.	<p>The high resolution image shown below confirms that the glyph is 𠂇 𠂇, not 𠂇 𠂇. 00617 (UTC-02917) should therefore be unified with 00599 (G_Z1212406).</p>  <p>We will provide the editor with a new glyph image:</p>

S/N	Ref.	Glyph	Comment	UK Response
				挑
02119	UTC-01802	潇	H: The evidence shows that the upper right component should be (Bamboo, R118) rather than (Grass, R140)	Agree. We will provide the editor with a new glyph image: 潇
03489	UTC-02968	臨	J: the glyph in the evidence is designed to be full height.	Agree. We will provide the editor with a new glyph image: 臨
04290	UTC-01478	迨	J: The enclosed component in the Old Hanzi glyph ... could be determined as "言".	The glyph follows the modernized form given in the printed source, and should not be changed.

4. Evidence

S/N	Ref.	Glyph	Comment	UK Response
00815	UTC-01347	𠂇	H: Evidence unclear.	This character has been withdrawn. See IRGN2155 UK ReviewP1.pdf p. 9.
01511	UTC-01432	𠂇	H: Evidence unclear.	See IRGN2155 UK ReviewP2.pdf p. 24 for additional evidence.
01514	UTC-01429	𠂇	H: Evidence unclear.	See IRGN2155 UK ReviewP2.pdf pp. 22-23 for additional evidence.
03617	UTC-01969	𠂇	H: Invalid evidence.	As discussed in the original submission (IRGN2107), there is a printing error for this character in <i>Hanyu Da Zidian</i> , but it is clear from the context that the

S/N	Ref.	Glyph	Comment	UK Response
				submitted glyph is correct.
03860	UTC-02918	誣	SAT: The shape of this evidence is a little ambiguous. It might be unified with 誣 U+46E9.	Unfortunately, the quality of the image is not very good. However, it is clear to us that the right hand side is 誣 not 亞, and these are not unifiable components. Therefore keep this character in IRG WS2015.
05414	UTC-02819	𪛗	H: The glyph is different from the evidence ... the appropriate phonetic component on the right should be 𪛗 as shown in the glyph submitted, not 𪛗 as shown in the evidence.	We concur that the evidence glyph is a mistake, and the submitted glyph is correct.

5. Mistaken Unifications

The following five UK-submitted characters are listed as unified or withdrawn in IRGN2155CJKWSet2015_UnifiedWithdrawn.pdf. We believe that the unification of these characters is mistaken, and we request that these five characters are added back to IRG Working Set 2015.

S/N	Ref.	Glyph	UK Rationale for Keeping in WS2015
00553	UTC-02957	𡵓	Not cognate with U+20C3C 𡵓, and different actual glyph shapes (市 vs. 市), which is significant in this case as they are different phonetic elements.
02204	UTC-01462	𡵓	The difference in glyph shape between 02204 and U+24261 𡵓 is not minor, and the two bottom components should not be considered to be unifiable minor variants. We can find not precedent for unifying characters with these two components, and the occurrence of both characters as separate entries in <i>Zhōnghuá Zìhǎi</i> 中华字海 p.974 argues against unification.
02343	UTC-02856	𡵓	02343 is the traditional character corresponding to U+2C312 𡵓, and traditional and simplified characters should not be unified. The components 將 and 將 are not unifiable, as evidenced by these pairs: <ul style="list-style-type: none"> • U+5968 獎 and U+596C 獎 • U+848B 蔣 and U+8523 蔣 • U+8780 蜚 and U+87BF 蜚 • U+91A4 醬 and U+91AC 醬 • U+9535 鏘 and U+93D8 鏘 • U+9CC9 鰐 and U+9C42 鰐 • U+21E83 蔣 and U+5D88 蔣 • U+2B98B 𡵓 and U+2B994 𡵓 • U+2BF63 擗 and U+646A 擗 • U+2C9C0 蹕 and U+8E61 蹕

S/N	Ref.	Glyph	UK Rationale for Keeping in WS2015
03555	UTC-01950	苍	<p>03555 is not cognate with U+82B2 苍, and the characters have different actual glyph shapes. 03555 is the simplified form of U+83D5 薈 lún "a type of plant", whereas U+82B2 苍 is a variant form of U+82B1 花 huā "flower". Unifying 苍 and 苍 is a very bad idea as it will mean that U+82B2 will be treated as the simplified form of U+83D5, and so applying a <i>ToTraditional</i> operation to U+82B2 will result in it changing to an unrelated non-cognate character.</p> <p>Moreover, 苍 and 苍 do not have the same glyph form in China, and the glyph forms are not interchangeable. In China the simplified component 仑 lún is written with the slanted stroke not going through the bent stroke (乚); whereas the bottom component of U+82B2 苍 is written with the slanted stroke going through the bent stroke (乚). Therefore 03555 and U+82B2 have distinctly different glyph shapes in China.</p>
05051	UTC-02343	鬚	<p>05051 is the simplified character corresponding to U+9B17 鬚, and simplified characters cannot be unified with the corresponding traditional form. The component 𦍋 in 05051 is a simplified component, and not unifiable with the non-simplified component 𦍋 in U+9B17, as evidenced by these pairs:</p> <ul style="list-style-type: none"> • U+748A 璚 and U+2B7A9 璚 (Ext. D) • U+4761 𧆒 and U+2C94A 𧆒 (Ext. E) • U+93CB 𧆒 and U+2CB6E 𧆒 (Ext. E) • U+29798 𧆒 and U+2CCCCF 𧆒 (Ext. E). <p>Note that 04374 (UTC-02093) 𧆒, which has the same component 𦍋, was explicitly not unified with U+409D 𧆒 at IRG46.</p>