

Title: Discussion on IRGN2414 among Woo Chi Chung, William Nelson and John Jenkins on Slack

Source: Eiso Chan

Time: 2019.10.23.

Oct 16th

SyaoranHinata (aka Woo Chi Chung) 3:21 PM

About IRGN2414 (2019.10.02) by Alexander Zapryagaev:

In fact, the shape “𠄎” pronounces “jì”, it means pig snout; the shape “𠄎” pronounces “xuě” in the Second stage simplifications. Their shapes, their pronunciations and their meanings are all different.

SyaoranHinata 6:37 PM

So I think “jì” and “xuě” should have different code point.

Oct 17th

John Jenkins 2:58 AM

This document is all about UAX #45, so it should have been submitted to the UTC, not the IRG. There's nothing in it for the IRG to do. It's definitely too late to add UTC-01005 to Extension G, and IVSs are not handled by the UTC directly. The process for registering them is described in UTS #37. Do we have an email address for Alexander Zapryagaev?

William Nelson 4:19 AM

I think IRG was the correct venue for IRGN2414 because the author wants IRG to reconsider its previous decision on this issue.

However, the argument in IRGN2414 is basically invalid. In IRGN2345, the IRG Principles and Procedures (PnP) document explains that: “The non-cognate rule does not apply to characters that have identical glyphs even if the characters are historically unrelated.”

In this case, U+5F50 𠄎 cannot be disunified because the representative glyph for China (G0-6566) and UTC-01005 are identical.

William Nelson 4:30 AM

In my opinion, there may be a valid argument to encode U+2E95 𠄎 as a new unified ideograph (not U+5F50 𠄎) because its history (alternate transcription of 𠄎, separate character from 𠄎) and its shape are distinct. This would be similar to TCA’s recent action to disunify U+5C6E 𠄎 and U+2F878 𠄎 to distinguish “plant sprout” and “left hand” uses.

Japan’s IVD collection treats all these shapes as unified with U+5F50 𠄎

5F50

𠄎

𠄏

𠄐

E0102
Moji_Joho
MJ011225

E0103
Moji_Joho
MJ011226

E0104
Moji_Joho
MJ011227

Oct 21st

SyaoranHinata 12:56 PM

In fact, the shape “𠄏/𠄐” comes from “又” (yòu), it means a hand, a right hand.

The shape “𠄎” comes from “彘” (jì), it means pig snout. The unification between “𠄏/𠄐” and “𠄎” violates the Noncognate Rule.

彘 = 𠄎 ≠ 𠄏 = 𠄐

William Nelson 4:04 PM

The actual shapes of G0-6566 and UTC-01005 are identical, so the noncognate rule does not apply for U+5F50 𠄎. It's the same situation as 芸 (ウン) and 芸 (ゲイ). IRG clarified this rule in January 2019. Please see IRG PnP version 12 draft 3.

Oct 22nd

SyaoranHinata 1:15 AM

The situation is different. The shape of "芸(ウン)" and "芸"(ゲイ) are totally the

same, but "𠄎" and "𠄎/𠄎" are NOT. If you wanna express pig snout, but you wrote the shape "𠄎/𠄎", that is 錯別字 -- just like you wanna express "soup" but you spelled "soap".

William Nelson 1:23 AM

I agree with you about U+2E95 𠄎

Shall we propose a new U-source ideograph for this shape to UTC?

John Jenkins 2:40 AM

You mean for 𠄎/𠄎? The shape of U+2E95 is already covered in the URO.

William Nelson 3:42 AM

U+2E95 𠄎 (with the long stem) and U+5F50 𠄎 are unifiable components (UCV) but they can be encoded separately due to the non-cognate rule and their non-identical actual shapes.

We could ask IRG directly if they would encode these characters separately, referencing them by their JMJ source identifiers, although I doubt that Japan would want to separate them.

William Nelson 6:39 AM

If IRG discusses IRGN2414, maybe Eiso or the UTC representative can ask about this issue.

Eiso Chan 7:22 AM

@William Nelson I can do that.

William Nelson 11:08 AM

漢字源: 「𠃉」は、「又(=て)」の変形。

3 《𠃉(𠃊・𠃋)部》けい・けい

がしら・いのこがしら

「𠃉」は、動物の頭の形。これを部首として、「𠃉」を含む漢字が集められる。「𠃉」は、「又(=て)」の変形。冠になるときは、特に「けいがしら」「いのこがしら」と呼ぶ。「けいがしら」の名称は、「𠃉」の音から。「いのこがしら」の名称は、「𠃊」が、豚やイノシシの頭を表したことから。

(End of Document)