Subject: Comments regarding Section 3 in IRG N2482: Proposal to improve IRG process
(3. Concretize the process of defining non-cognate characters)
Date: 2021.09.15.; 2021.09.16. (b)
Authors: KIM Kyongsok (KR)
Status: Individual contribution
Relevant docs: IRG N2482

1. Background

- Section 3 of IRG N2482 is shown below:

ISO/IEC JTC1/SC2/WG2/**IRGN2482**
2020-09-02

Doc Type:	Ideographic Rapporteur Group Document
Title:	Proposal to improve IRG process
Source:	Wang Xieyang (王谢杨)
Status:	Individual Contribution
Action:	For consideration by IRG
Date:	2020-09-02

3. Concretize the process of defining non-congnate characters.

Here is a part of a chart in the book named 《韩国汉文古文献异形字研究》（吕浩 著, 上海: 上海人民出版社, 2013 年 12 月）

主形字	韓國漢文古文獻異形字	敦煌文獻異形字
備	偹、俻、僃、俻、脩、俻、偹、偹、偹	俻、俻、俻
邊	邉、邊、邊、邊、邊、邊	邊、邊、邉、邊
稱	稱、稱、稱、稱、稱、稱、稱、稱、稱、稱、稱、稱	稱、稱、稱、稱
處	處、處、處、處、處、處、處、処、竅	處、處、胤、處
辭	辤、辭、辭、辭、舜、辭、辭、辭	辭、辤、辤、舜
帶	帶、帶、帶、帶、帶、帶	帶
殿	殿、殿、殿、殿、殿、殿、殿、殿、殿、殿	壂、殿、殿、壂
發	發、發、發、發、發、發、發、發、発、發、發、發、發、發、發、發	發、發、發、發、發、發、發、茇、發
歸	埽、歸、歸、歸、歸、歸、歸、歸、歸、帰、歸、歸、歸、歸、歸、埽、歸、歸、歸	埽、歸、歸、歸、歸、歸、隬、隬、歸、歸、歸、歸、埽、皈
龜	龜、龜、龜、龜、龜、龜、龜、龜、電、龜、龜、龜	龜、龜、龜、龜
號	號、號、號、號、號、號、號、号、號、號、號	號、號、號、號、號、號、號、号、号、号、号

Fig.3 吕浩: 韩国汉文古文献异形字研究, 上海: 上海人民出版社, 2013 年 12 月, P27

This chart is just a tip of the iceberg. Many Han characters have a wide range of glyphs. It is very likely that one of the glyphs inherit one rare meaning of the character while other glyphs don't. Then even if the glyph has little difference from others, it is of great possibility that we disunify them. Actually, however, the users even don't care. Some characters are usually written as very similar glyphs in the practice and most people even don't think there is difference between them. In this case, the disunification is not very necessary.

In IRG PnP 2.1.3, it says

> 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.
> ......

> Because shape analysis alone may not tell non-cognateness or semantic differences, it is the submitter's responsibility to provide information and supporting evidence in order to invoke the non-cognate rule.

It doesn't say, however, which kind of information and evidence is needed in spercific. What's more, IRG PnP doesn't clarify how to deal with the mixing used simlar glyphs. Consideing the large quantity of Han characters' glyphs and the conclusion of our discussion of GDM-00241(囗木 旬) in IRG WS2021 and UK-10757(囗門門) in IRG WS2017, I suggest

> 1) Add one sentence to the quoted paragraph 1 that "Ideographs with unifiable glyph shapes should be considered congnate if they can be used without distinction in fact".

> 2) Add one sentence to the quoted paragraph 2 that "For unifiable ideographs, the information and supporting evidence provied by submitters should be able to clearly explain a) the pronunciation of the two ideographs have no historical derivation; b) the meaning of the two ideographs have no relationship.

2. The author of IRG N2482 suggests as follows:

2) Add one sentence to the quoted paragraph 2 that "For unifiable ideographs, the information and supporting evidence provided by submitters should be able to clearly explain
   a) the pronunciation of the two ideographs have no historical derivation;
   b) the meaning of the two ideographs have no relationship.

3. Comments regarding the above suggestions:

3.1 pronunciations of characters

information and supporting evidence provided by submitters should be able to clearly explain
   a) the pronunciation of the two ideographs have no historical derivation;
==>

   1) In some cases, the pronunciation of characters are helpful to decide whether or not to unify, but not always.

   2) Sometimes one may not know the pronunciation of characters (for example, characters in old documents).

   3) Some characters have several pronunciations. Japanese Kanji characters usually have several (sometimes more than 10) pronunciations. Even when the pronunciations of two slightly different glyphs are different, actually the two can be one and the same character.

   4) When two glyphs are used in different countries/regions, it is likely that pronunciations are different.
   - Furthermore, a submitter may have to contact an expert in other countries to know the pronunciations there.

   5) I understand what the author of N2482 is trying to incorporate in the IRG PnP. However, I am not sure if this is a resonable approach.

3.2 meanings of characters

information and supporting evidence provided by submitters should be able to clearly explain
   b) the meaning of the two ideographs have no relationship.
==>

   1) In some cases, the meaning of characters are helpful to decide whether or not to unify, but not always.

   2) Sometimes, one may not know the meaning of characters (for example, proper names such as personal names or place names).

   3) Some characters have several meanings. Even when the meanings of two slightly different glyphs are different, actually the two can be one and the same character.

   4) When two glyphs are used in different countries/regions, sometimes meanings may be different.
   - Furthermore, a submitter may have to contact an expert in other countries to know the meanings there.

   5) I understand what the author of N2482 is trying to incorporate in the IRG PnP. However, I am not sure if this is a resonable approach.

3.3 A conclusion

- Even when the pronunciations and meanings of two similar glyphs are different, the two glyphs may or may not be unifiable.

- Based on the discussions in 3.1 and 3.2 above, it may be desirable not to add the following sentence which seems too restrictive.

For unifiable ideographs, the information and supporting evidence provided by submitters should be able to clearly explain
   a) the pronunciation of the two ideographs have no historical derivation;
   b) the meaning of the two ideographs have no relationship.

* * *