

Title: Feedback to IRG N2717, IRG N2731 and IRG N2738
Source: Wang Xieyang(王谢杨)
Status: Individual Contribution
Action: For consideration by IRG

We cannot unreasonably reject the encoding request of script-hybrid ideographs in historical materials, nor can we directly reject the encoding request of script-hybrid ideographs created in the future. Therefore, for a possible new script, we should not only consider what we see in front of us. Due to the reform of the writing system in various countries and the spread of Western culture in the East, the number of script-hybrid ideographs in historical documents are unlikely to be as less as those presented in IRG N2637 and IRG N2717. We have found four script-hybrid ideographs composed by Bopomofo and ideograph in a famous book written by a renowned paleontologist Mr. Tang Lan(唐兰). The four script-hybrid ideographs were not included in IRG N2637 or IRG N2717. Two of them are cited by a Singaporean expert and one of them have different glyphs between the two books.

因为我们不能无理地拒绝编码历史文献中的汉字-表音文字结合体文字，也不可能直接拒绝编码未来创造出来的汉字-表音文字结合体文字，所以对于一个可能的文种，我们不应该只考虑眼前之所见。受历史上各国文字改革及西方文化在东方传播的影响，历史文献中汉字-表音文字结合体文字的数量、种类不大可能只有 IRG N2637 和 IRG N2717 两个提案展示的这么多。我们在著名古文字学家唐兰先生的著作中发现了四个并未被 IRG N2637 和 IRG N2717 收录的汉字-注音符号混合体文字。其中的两个被一位新加坡学者引用了，且被引用的两字中有一个在两本书中的字形有差异。

第三，這種文字，對於固有文化，完全保存。

1. 把舊形聲字，改成新形聲字，有益而無害。因形聲字的聲母，本只注音，所以鴻洪同有洪水的意義，聲母作夆、塢或共，是不拘的，那末，改為注音新法，當然無害，只要拚出字音，看見形母，當然可以認識的。

例如：

松、柏、杞、梓寫做

控, 控, 杞, 杞 暫依現行注音符號

只要懂得拚音，決不會感覺不便，尤其是梓字，原來的辛聲，現在已不能做聲母，這樣一改，就便利多了。

2. 古文化不致因文字改革而湮滅，因為除了形聲字外，

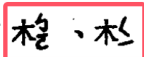
④ 唐兰曾提出新形声字草案，主张保留汉字的形，一般是左形右声，且只许两合。他把文字分成字母及形声字，字母以形母为主，即义母，至多约五百字左右；形声字的声母，用注音符号拼出，此注音符号约四十个，现行的注音符号须被包括在字母内；凡字的声调，在声母四侧，注以符号，例如柏、杞写作 ，一边是形符一边是注音，所以名为注音文字。详见《古文字学导论》下册第78页至第84页《研究古文字和创造新文字》，1965年香港太平书局重印本。

Fig.2 (新加坡)谢世涯：新中日简体字研究，北京：语文出版社，1992年6月第2版，page15

There has been no discussion on the following key issues in IRG yet:
IRG 还没有对以下关键问题进行过任何讨论：

1. The approximate number of script-hybrid ideographs in historical materials
1. 历史文献中汉字-表音文字结合体文字的大概数量

If the number of script-hybrid ideographs that can be encoded is indeed very small, is it worth changing IRG PnP for it? Will it be more suitable for them to be handled by WG2 or UTC?

If there are possibly many script-hybrid ideographs that can be encoded, it is obviously necessary to systematically study them before encoding.

如果可编码的汉字-表音文字结合体文字数量确实很少，那么为其修改 IRG PnP 是否值得？如果是这样，是不是更适合由 WG2 或者 UTC 来处理？

如果可编码的汉字-表音文字结合体文字数量可能很多，那么显然有必要在正式编码前先对其进行系统研究。

2. The unification rule between of script-hybrid ideographs and other Han ideographs
2. 汉字-表音文字结合体文字和其他汉字的统合问题

For example, can the following three sets of character shapes be unified to each other?

X 也 and ¹

T and 

主 and 

If they can be unified, then what is the standard of the unification? Is the standard similarity of glyphs? Even if we choose to unify script-hybrid ideographs and other Han ideographs according to the similarity between glyphs, it should be written down first, discussed, and then we can reach a consensus, do the encoding work in practice.

¹ It was created in the modern time.

比如，以下三组字形是否可以统合？

𠃉X也和𠃉乂也
 𠃉门T和𠃉门工
 𠃉木主 and 𠃉木キ

如果应该统合，那么统合的标准是什么？字形相近吗？即使以字形相近作为统合标准，我们也应该先形成文件，讨论一下，各方同意之后再推进实际的编码流程。

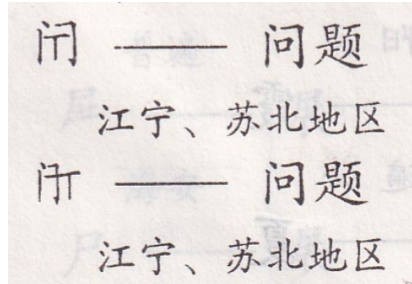


Fig.3 𠃉门 T found in 江苏方言总汇

3. How to determine the glyphs of script-hybrid ideographs in regular script as well as the representative glyphs of script-hybrid ideographs in code charts?

For example, can I normalize the glyph of 𠃉广 K to 𠃉广 卜?

3.汉字-表音文字结合体文字的楷体字形、码表中中汉字-表音文字结合体文字的代表字形应该如何确定？

例如，我能把 𠃉广 K 的字形正形为 𠃉广 卜吗？

4. The conditions for determining whether a character is a script-hybrid ideograph

For example, is it necessary to have Han ideograph components? If it is, it should be written down in a document, too. I noticed that there are some characters composed purely by kana in IRG N2637. If it is not necessary to have Han ideograph components, then can we encode the E (variant of 医) included in IRG N2731 as Han ideograph? How about the characters composed purely by Bopomofo, which are listed in Fig.5?

4.认定一个字是汉字-表音文字结合体文字的条件

比如，拥有汉字部件是必须的吗？如果是，相关的条款也应该先形成文件以供讨论。我注意到 IRG N2637 中列出了一些纯假名组成的字符，如果不必须拥有汉字部件，那么 IRG N2731 文件中作为“医”字异体字使用的 E 能否作为汉字编码？图 5 (Fig.5) 中列举的仅由注音符号组成的字符呢？

20	𠃉	<ul style="list-style-type: none"> • 𠃉モタル • Japanese abbreviation for word “モルタル(mortar)”
21	𠃉	<ul style="list-style-type: none"> • 𠃉ヤキハ • Japanese ligature of “ヤキハ(yakiba)” • often found in sword books of the Edo era

Fig.4 Characters combined by kana only found in IRG N2637

