

Feedback on Siddham proposal (WG2 N4294)

suzuki toshiya, Hiroshima University, Japan

This document is the feedback to the proposal to encode Siddham script, WG2 N4294, I propose to invite the comments from the experts working for the digitization of Siddham materials, about the appropriate handling of the variants.

The current proposal (WG2 N4294) handles the variant shapes as;

Unification The encoding for Siddham is to serve as a unifying block for all regional variants of the script, such as ‘Siddhamātṛkā’ and ‘Kuṭīla’. The representative glyphs are based upon Japanese forms of Siddham characters on account of active usage of the script by Japanese Buddhist communities.

(WG2 N4294 p.1)

It seems that the ‘regional’ means the shape difference between Chinese, Japanese and Korean materials

The ‘formal’ and ‘informal’ styles of letters are shown in figures 12–17. The forms of Siddham letters used in Korea (see figures 21–26) differ significantly from those of Japanese and Chinese Siddham. However, given the close relationships between the regional forms and their identities as ‘Siddham’, it is practical to unify these variants with the proposed script block.

(WG2 N4294 p.2)

The idea to unify the regional variants would be reasonable. Comparing the vowel characters in Japanese and Korean, the significant shape differences are observed in a, aa, u, uu. However, if these differences are considered as per writing material variants (e.g. plate- versus pole-brushing) or regional variants (China, Korea and Japan), encoding them as different characters may induce the inconvenience in the coded text interchange (e.g. when a Siddham text from Chinese material is digitized for Korean researchers community, Korean Siddham codepoint should be used?)

	a	aa	i	ii	u	uu	e	ai
J								
K								

Fig. 1: Comparison of referential glyphs in WG2 N4294 fig. 19 (J) and fig. 21 (K).

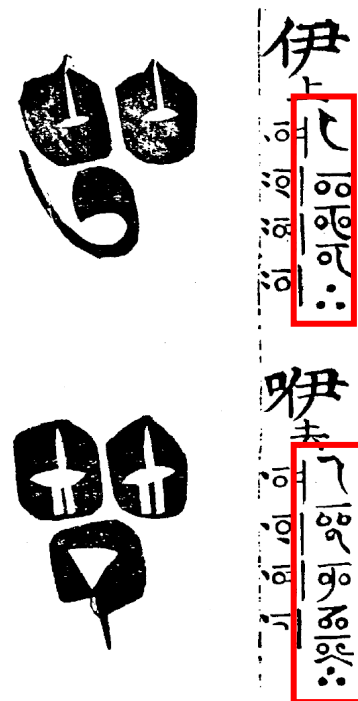
But, the materials referred by WG2 N4294, several character charts show the variants that are supposed to have different background. The evidences from Korean material are not so many, but considering that some similar variants are found in both of Korean and Japanese materials, they are supposed to be non-regional, and non-calligraphic variants. There is a possibility that the normalization to single referential glyph might induce the semantic impact.

	formal	brush	stroke order	variations
1				,
2				
3				
4				

WG2 N4294 fig. 12 (Stevens 1981:34)

6	5	4	3	2	1	番号
						文字 悉曇
						異体字 (別体)

WG2 N4294 fig. 19 (Kodama 1991:65)



WG2 N4294 fig. 21 (Lim 2005)

Fig. 2: Variants with shape differences that supposed to be non-regional

For example, Taisho-Sinshu-Daizokyo (大正新脩大藏經) was typesetted, but it used multiple glyphs for same phonetic value. There might be some distinction of the variants. As shown in Fig. 3, there are some documents using two different shapes for same phonetic value, and it is quite difficult to reproduce the shapes by the consideration of the surrounding context.

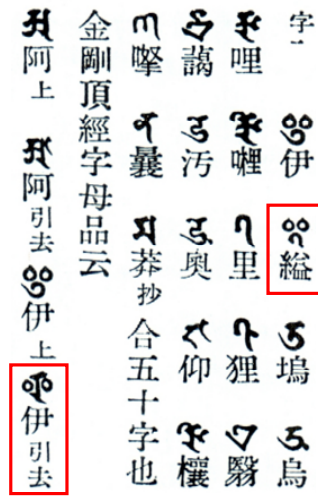


Fig. 3: Example of the document using two different shapes for same phonetic value, Sittanzou (悉曇藏) by Annen (安然), typesetted in Taisho Sinshu Daizokyo (大正新脩大藏經), No. 2702, volume 84.

In addition, although the origin of the Siddham script is not ideographic script, some Buddhist communities seem to relate some glyphs with semantic information. For example, 梵字必携 (Kodama 1991, referred in WG2 N4294) has a list between the Siddham glyph and the corresponding Ksitigarbha name. The stability of the list is commented to be unstable, but it is clear that the variant distinction is expected to digitize such materials.

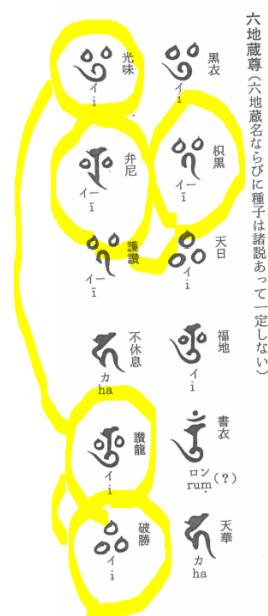


Fig. 4: Example of the distinguished usage of the variant shapes of the characters with same phonetic value, Kodama 1991, p. 233.

In summary, there might be different aspects about following issues:

- All variant shapes of the Siddham script could be recognized as the regional (or user community, or writing material) difference?
 - If not, how to distinguish the regional and non-regional?
- The normalization of the variant shapes of the Siddham script has no semantic impact in the digitization of existing Siddham materials?
 - The separated encoding of the variants can introduce some difficulty in the handling of the Siddham text?

I heard that Japanese national body had already sent the preliminary proposal (WG2 N4185) to SAT (the international expert group working for the digitization of Taisho Sinshu Daizokyo) and requested their comments, but only offered the above evidences (fig. 3 and 4) and the brief comment. It seems that SAT experts are too busy to summarize their proposal to CJK Unified Ideographs Extension F (to forthcoming IRG#39). For the smooth and wide utilization of the future standard of Siddham script encoding, I propose to wait their feedback until next WG2 meeting.

(end of document)