

**TITLE:** Comments on N4409R Proposal to Encode Variants for Siddham Script  
**SOURCE:** Andrew Glass (Microsoft) via Deborah Anderson (UC Berkeley)  
**STATUS:** Expert Contribution  
**DATE:** 10 October 2013  
**ACTION:** For review by ISO/IEC JTC1/SC2/WG2

The proposed full vowel forms,

U+1115E0 SIDDHAM LETTER VARIANT FORM A

U+1115E1 SIDDHAM LETTER VARIANT FORM B

U+1115E2 SIDDHAM LETTER II VARIANT FORM A and

U+1115E3 SIDDHAM LETTER U VARIANT FORM A,

should only be employed when the symbols are used logographically to denote particular Bodhisattvas or to reflect a supra-phonetic value. Outside of these contexts, the proposed characters should not be used. Instead, users should rely on fonts to reflect the glyphs of the text.

The rationale is that paleographic distinctions are best handled at the font level. If users were to use logograms derived from paleographic variants (i.e., U+1115E0-U+1115E3) for paleographic variants that were not used logographically, it would likely lead to expectations that other or all paleographic distinctions should be made at the encoding level. Such an expectation would quickly become untenable.

The combining vowels

U+1115E4 SIDDHAM VOWEL SIGN U VARIANT FORM A

U+1115E5 SIDDHAM VOWEL SIGN UU VARIANT FORM A

appear only in the seed syllable huṃ/huṃ̣ in the documentation of N4407R. Is it true that the vowel signs U and UU only appear with the base HA?

If so, I would be open to encoding symbols consisting of the full syllable for huṃ and huṃ̣, with the same proviso as above, namely that it would be best not to encode paleographic variation.

However, if the distinction between the variants of long and short U dependent vowels is used symbolically with bases other than HA, then this information would be good to know, so a case for a combining diacritical mark can be properly assessed.