

Unicode Technical Committee
Expert Contribution

L2/98-142
April 15, 1998

Title:	Comments on Text for Yi PDAM #14
Source:	Ken Whistler
Status:	Expert Contribution
Action:	For the consideration of UTC and NCITS/L2
References:	WG2 N1608 (= L2/97-153), WG2 N1415 (= X3L2/96-099), WG2 N1481 (= L2/97-020)
Distribution:	UTC and NCITS/L2 members

Analysis of Yi PDAM #14 (and WG2 N1608)

Document WG2 N1608 contains the high quality Yi text which was used as the basis for preparing PDAM #14 for the Yi script. The Yi names list in that document contains a number of significant technical errors. To assist UTC and L2 in developing a balloting position and comments on the Yi PDAM, I have analyzed the Yi proposal in detail. This contribution summarizes the technical errors that it contains.

Registration Error

The main error in PDAM 14 is a registration problem that apparently resulted from misplacing YI SYLLABLE HXE at U+A3D2 instead of at U+A17D. This resulted in shifting up all the names between U+A17D and U+A3D2 inclusive, so that every character name in that range is off-by-one with respect to the chart of character glyphs.

This must be fixed by moving YI SYLLABLE HXE back to U+A17D and moving down each character from U+A17D to U+A3D1 by one.

Typos in Character Names

WG2 N1608 had a number of typos in character names. Most of these were corrected in the preparation of the PDAM text, but one was not. (For brevity, here and subsequently, I will leave off the "YI SYLLABLE" in referring to Yi character names.)

N1608 values		Corrected values	
Code	Name	Code	Name
A147	HAP	A147	NAP (uncorrected in PDAM)
A191	UP	A192	JUP
A271	MGAT	A272	NGAT
A289	NAZP	A28A	NZAP
A2A3	NZEX	A2A4	NREX
A3D1	V0P	A3D2	VOP

Yi PDAM #14 Analysis

Ordering Problems

Analysis of the Yi data shows that the intended order of characters was by initial consonant, then vocalic nucleus, then tone, with the following suborders:

Initial consonants

b bb c ch d dd f g gg h hm hn hl hx j jj k l m mg n nb nd
ng nz nr nj ny p q r rr s ss sh t v w x y z zz zh (0)

Vocalic nuclei

i ie a uo o e u ur y yr

Tones

-t -x - -p

The problem in the PDAM text, derived from N1608 (and its predecessor documents), is that the intended order got stirred in various ways. Here is the list of ordering problems I have identified.

1. U+A147 NAP belongs immediately after U+A22B NA
2. U+A141 HNUT belongs immediately after U+A146 HNEP
3. Starting at the consonant "l", there is a recurrent pattern of placing the "i" vocalic nuclei at the end of a consonant range or even in the middle of some other consonant, instead of at the start of the relevant consonant. This should be corrected, as shown here. (For simplicity of reference, I refer to the code values in PDAM 14, preceding the fix required for the big off-by-one chunk in the middle of the repertoire.)

- U+A1EE LIT .. U+A1F1 LIP move to preceding U+A1CF LIET
- U+A223 MIT .. U+A226 MIP move to preceding U+A1F2 MIET
- U+A2DE NIT .. U+A2E1 NIP move to preceding U+A227 NIET
- U+A2FC PIT .. U+A2FF PIP move to preceding U+A2E2 PIEX
- U+A318 QIT .. U+A31B QIP move to preceding U+A300 QIET
- U+A348 RAT .. U+A34D RUO move to preceding U+A31C RUOP
- U+A39F SIT .. U+A3A5 SIEP move to preceding U+A34E SAT
- U+A3BB TIT .. U+A3BF TIEX move to preceding U+A3A6 TAT

These problems resulted from alphabetization issues derived from the bizarre Yi orthography of WG2 N1415, issues which were not corrected when the orthography itself was fixed in WG2 N1481.