ISO/IEC JTC1/SC2/WG2 N3692

L2/09-325 2009-09-28

Universal Multiple-Octet Coded Character Set International Organization for Standardization Organisation internationale de normalisation Международная организация по стандартизации

Doc Type: Working Group DocumentTitle:Roadmapping the scripts of NepalSource:UC Berkeley Script Encoding Initiative (Universal Scripts Project)Author:Michael EversonStatus:Liaison ContributionDate:2009-09-28

Analysis of the Nepalese scripts is slightly complicated because different font styles have different names, and some writers equate name with script identity. Analysis of the chief features suggests that a number of these can be unified on structural grounds.

1 *RAÑJANA* encompasses: *Rañjana* (Figure 1, 2, 3) *Wartu* (Figure 4)

This is the most complex of the scripts, in particular because it is used for Tibetan, which has introduced some innovations from Tibetan script itself which do not occur in other Brahmic scripts. (An example is the use of an explicit "*head-ra*" alongside regular *ra*.) One unique feature is the long or short mark used for $-\bar{a}$; compare \mathbf{a} ka, \mathbf{a} kā and \mathbf{a} ga, \mathbf{a} gā. Named scripts which should be unified here are Rañjana (Tibetan " \mathbf{a} lañdza), and Wartu (Tibetan \mathbf{a} wartu). Stylistically, Wartu is much the same as Rañjana but with a wiggly headline on many of the letters. See Figures 3 and 4.

2 PRACALIT

There are many scripts in this group, mostly distinguished by their headlines. The major difference between the PRACALIT scripts and the RAÑJANA scripts is the way in which *-e* and *-ai* are made by changing the top bar (see Figure 20). Two major varieties are distinguished, and there is not yet enough evidence available to determine whether or not it is appropriate to encode them separately from one another. At this stage, at least we can distinguish:

2.1 PRACALIT flat-headed script, which encompasses:

Pracalit (Figures 5, 6, 7) *Pācūmol* (Figure 8) *Himmol* (Figure 9) *Kummol* (Figure 10)

2.2 BHUJIMMOL curve-headed script, which encompasses:

Bhujimmol (Figures 11, 12, 13) *Golmol* (Figure 14) *Kvemmol* (Figure 15) *Litumol* (Figure 16) Documentary evidence for the popularity or use of these so-called "nine scripts" (the eight PRACALIT scripts plus Rañjana) is rare except for Rañjana, Golmol and current Pracalit script in which the majority of Nepalese manuscripts and inscriptions were written from the 15th to 18th centuries CE. Golmol is said to have been introduced in Nepal by Bhaumagupta (ca. late 6th century CE.). The heyday of its use was, however, the early medieval period, i.e., the 13th-15th century CE. Though it was widely used in the 13th century, even then its use was not universal.

Traditionally, Rañjana was restricted to copying sacred texts—Hindu as well as Buddhist. It was rarely, if ever, used for secular purposes. More recently, however, Newars have used Ranjana as an ornamental script for invitations, very formal manuscripts and so on, and until about 1900 used it for their Sanskrit and Newari manuscripts. In the 1990s it was common to see posters and banners advertising Buddhist lectures, new year's parties and so forth with the heading in *Rañjana* and body in Newa Lipi (*Pracalit*) which only Newars could read.

Encoding considerations. It should first be said that some members of the user community have criticized the idea of unifying these "scripts". It may be that this is a misunderstanding of the UCS; the analogy of the Latin script with its **SAULC** and **Fraktur** variants, however, is probably applicable, which is why the recommendations here have been made.

A complete inventory of the shaping behaviour of any of the PRACALIT scripts (apart from *Pracalit* itself) has been unavailable, but it seems unlikely that any of them will exhibit behaviour that should warrant anything but the standard VIRAMA encoding model ($\overline{\mathbf{A}}$ KA + VIRAMA + $\overline{\mathbf{A}}$ SHA = $\overline{\mathbf{A}}$ KSHA) to be used for them. It has been suggested, however, that the Tibetan stacking model ($\overline{\mathbf{A}}$ KA + $\overline{\mathbf{A}}$ SUBJOINED-SHA = $\overline{\mathbf{A}}$ KSHA) could be used for the RAÑJANA scripts (*Rañjana* and *Wartu*). It does not seem that there is any benefit on either structural or practical grounds to doing so. Structurally, Rañjana's behaviour is classically Brahmic; $\overline{\mathbf{A}}$ is a stack: $\overline{\mathbf{A}}$ is a conjunct. Using the Tibetan model would seem to be artificial. Further discussion of the options and elaboration on this point is expected during the proposal process itself.

The Rañjana scripts differ in many ways from Tibetan. Both of them (and all the PRACALIT scripts) have unique Brahmic voiced aspirates \mathfrak{A} *gha*, \mathfrak{A} *jha*, \mathfrak{A} *dha*, \mathfrak{A} *dha*, rather than the Tibetan stacked \mathfrak{F} , \mathfrak{F}

Figures 17, 18, and 19 show examples of Rañjana consonant clusters. These are not simple stacks or even complex consonant conjuncts: they are monograms representing seed syllables, and should probably be encoded as such. Some of these may also be "example" clusters, such as *kkhgghňa* (*ka* + *kha* + *ga* + *gha* + *ňa*), and three different examples of *cchjjhña* (*ca* + *cha* + *ja* + *jha* + *ña*). Although it is in principle possible to predict what a string of a Tibetan base + *n* subjoins will look like, it wouldn't be possible to do the same for Rañjana; in Figure 18, for instance, several of the conjunct monograms have more than one rendering. It is unlikely that any rendering engine should be expected to render these dynamically; rather, they ought to be font-specific special ligatures, or, as mentioned above, encoded as monograms. (It should be pointed out that these are typically not set in type, but hand-drawn.)

Bibliography

Gu Wenyi & Shi Xueli. 1995. *Gańs-can mkhas pa'i phyag bris sna tshogs phyogs bsdus rin chen phreń ba* (= Zang wen shu fa hui cui). Lanzhou : Kan-su'u mi rigs dpe skrun khaň : Kan-su'u Źiň-chen Źin-hwa dpe khaň gis bkram pa. ISBN 7-5421-0048-3

Hemraj Shakyavansha. 1985. Nepalese Alphabets = Nepāl lipi samgraha. Seventh Edition.

- Shakya, Rabison. 2002. Nepāl lipi Varņamālā = Alphabet of the Nepalese script. Patan: Motirāj Shakya. ISBN 99933-34-36-7
- Tshe-tan Zabs-drun. 1992. Landzai thig rtsai dka gnad gsal byed dba brgya ldan pai glog sgron zes bya ba dpe ris dan bcas pa. ISBN 7-5420-0248-1

Acknowledgements.

This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley) in respect of Rañjana and Pracalit encoding. Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

Figures

| ৰ | व | বি | वी | ৰ | ৰু | ৰু | ৰু | व् | र्वे | ব | वि | đ | ৰা |
|--------|---------|---------|---------|--------|-----------|----------|---------|------------------|----------|-----------|----------|---------|----------|
| क-ka | का-kā | कि-ki | की-ki | कु-ku | कू-kū | क्-kı | कू-kī | के-ke | कै-kai | को-ko | कौ-kau | क-kam | कः-kaḥ |
| द्ध | द्ध | बि | द्वी | द्ध | द्ध् | द्धू | द्यू | द्ध | द्व | द्ध | द्ध | į | द्वर |
| रव-kha | रवा-khā | रिव-khi | रवी-khi | खु-khu | स्वू-khū | स्वृ-khr | रवॄ-khŗ | स्वे-khe | रवै-khai | स्वो-kho | रवौ-khau | खं-kham | स्वः-kaḥ |
| শ্ | ગ્ | P | ગી | শ্ | শু | J | J | ব্ | ē | (ग | म | تْا | শৃ |
| শ–ga | गा-gā | गि–gi | गी-gi | गु-gu | गू-gū | ıı́−ði | ııॄ−gī | गे-ge | गै-gai | गो-ko | गौ-kau | गं-kam | गः-kaḥ |
| দ্ব্য | স্থ্র | স্থি | স্থ্র | ব্ব | স্কু | র্ম্ব | रूगू | দ্ব | ্যন্ত | স্থ | স্থ্র | ক্ষ্ | স্কৃ |
| घ-gha | धा-ghā | घि-ghi | घी-ghī | घु-ghu | ਬ੍ਰ-ghū | ਬ੍ਰੋ-ghr | घॄ-ghŗ | घे-ghe | घै-ghai | घो-gho | घौ-ghau | घं-gham | घ:-ghaḥ |
| द्ध | द्रग् | ব্ধি | ङ्ग् | ক্র | রু | রু | ୢ୷ | ં જુથ | ক্ট | दुग् | ক্ত্ৰ্য | हुंग | ব্ৰু |
| ন্ট-na | জ্ঞা-nā | ङি-৸ | র্জা-ni | ङु-nu | ङू-nû | रङ्-nr | ş-hr | ङे-ne | ਤੈ-hai | ङो-no | ङो-nau | ङं-nam | ङ:-nah |

Figure 1. Consonants with vowel matras in Rañjana script from Shakya 2002. Note the particular shapes of $-\bar{a}$, $-\bar{c}$, -e, -ai, -o, and -au.



Figure 2. Consonants in Rañjana script from Shakya 2002.



Figure 3. Example text in Rañjana script from Gu & Shi 1995. Note that the shapes of *a*, *e*, and *o* are unrelated to one another, that the voiced aspirates are Brahmic, and that *l* and *l* are distinct.

99 G เง 5 3 W GN 51 2 h ς 21 22 5

Figure 4. Example text in Wartu script from Gu & Shi 1995. Note that the shapes of *a*, *e*, and *o* are unrelated to one another, that the voiced aspirates are Brahmic, and that *l* and *l* are distinct.

| ৰ্ণ | না | नि | नी | ৰ্ | ৰূ | কৃ | कृ | র্ণা | न | <mark>์ ค</mark> ิโ | ना | नं | ৰাঃ |
|--------|----------|---------|----------|--------|----------|--------|---------|------------|---------|---------------------|-----------|---------|--------------|
| क-ka | का-kā | कि-ki | की-ki | कु-ku | कू-kū | कृ-kr | कू-ki | के-ke | कै-kai | को-ko | कौ-kau | कं-kam़ | कः-kaḥ |
| ख | खा | रिव | र्स्वी | ধ্ব | ধ্বু | স্বৃ | ख़ॄ | र्व | र्ये | श्व | खा | ক্ষ | হ্ব ঃ |
| ख-kha | स्वा-khā | रिव-khi | स्वी-khi | खु-khu | स्वू-khū | खृ-khr | खॄ-khŗ | स्वे-khe | खै-khai | स्वो-kho | स्वौ-khau | खं-kham | खः-kaḥ |
| ا | শ | গি | হী | ว | الله ال | ə | ع | () | (স | <u>(১</u> | শৌ | হা | ৶ |
| ग-ga | गा-gā | गि-gi | गी-gi | गु-gu | गू-gū | ਸੁ-gr | ਸ਼ੵ−gi | गे-ge | गै-gai | गो-ko | गौ-kau | र्ग-kam | गः-kaḥ |
| ସ | ঘা | ঘি | ची | ସ୍ | ସୂ | ସୂ | ចុ | র্ঘ | ন্থ | ସା | चा | र्घं | ସଃ |
| ম-gha | घा-ghā | घि-ghi | घी-ghi | घु-ghu | घू-ghū | घृ-ghr | घॄ-ghr | घे-ghe | घै-ghai | घो-gho | घौ-ghau | घं-gham | घ:-ghaḥ |
| 2 | ଅ | S | දි | 2 | ୢୡ | Ş | Z | ຽ | S | 2 | 2 | Ž | 28 |
| ন্ড-na | ङा-nā | ন্ডি-ni | জী-াা | ङ्-nu | ङू-hū | ई-µt | ន្ទ្-nř | ङे-ne | ङै-hai | ङो-no | ङौ-nau | ङं-nam | ङ:-nah |

Figure 5. Consonants with vowel matras in Pracalit script from Shakya 2002. Note the horizontal head-line, and the particular shapes of -u, -e, -ai, -o, and -au.



Figure 6. Consonants in Pracalit script from Shakya 2002.



Pa Pha Ba Bha Ma Ya Ra La Wa प প্চ ब ਜ म य ₹ ल व Sha Sha Sa Ha Ksha Tra Jna হা G. स Ę T द्र

Figure 7. Alphabet in Pracalit script from Shakya 2002. Note that the shapes of *a*, *e*, and *o* are unrelated to one another, that the voiced aspirates are Brahmic, and that *l* and *l* are distinct.



Figure 8. Alphabet in Pācūmol script from Shakya 2002. Note that the shapes of *a*, *e*, and *o* are unrelated to one another, that the voiced aspirates are Brahmic, and that *l* and *l* are distinct.



Figure 9. Alphabet in Himmol script from Shakya 2002. Note that the shapes of a and e are unrelated to one another, though o is based on a, and that the voiced aspirates are Brahmic.

Pha

Ŧ

Sha

যা

Ba Bha

व

Sha

ष

भ

Sa

स

Ma

म

Ha

ē,

Ra La

र ल

Tra Jna

7 5

Wa

۹

Ya

य

Ksha

দ্ব



Figure 10. Alphabet in Kummol script from Shakya 2002. Note that the shapes of *a*, *e*, and *o* are unrelated to one another, and that the voiced aspirates are Brahmic.

| ণী | নী | নি | ளி | ð | Ð | ð | ð | ମି | শ | ሻ | নী | ຄໍ | କୃଣ୍ଣ |
|-------|---------|---------|---------|--------|---------|----------|----------|----------|----------|-------------|---------|----------|---------|
| क-ka | का-kā | कि-ki | की-ki | कु-ku | कू-kū | क्-kr | कू-kr | के-ke | कै-kai | को-ko | कौ-kau | कं-kam़ | कः-kaḥ |
| ন্থ | ક્લી | হি | ક્ષી | ন্থ | ন্থ | ક્ર્યુ | ş | ঞ্জ | ন্থ | হৰ | জ্বী | ર્સ્વ | ন্থ |
| ख-kha | रवा-khā | रिव-khi | रवी-khi | खु-khu | रवू-khū | स्वृ-khr | स्वॄ-khŗ | स्वे-khe | रवै-khai | स्वो-kho | खौ-khau | रवं-kham | रवः-kaḥ |
| 51 | থা | গি | ઝી | হা | হা | ર્ગ | ปรี | (51 | ন্থ | <u>(</u> 2] | শ্র | สํ | ଧ୍ୟହ |
| ग-ga | गा-gā | गि-gi | गी-gi | गु-gu | गू-gū | ıı⊥-di | -ਸੂ-gī | गे-ge | गै-gai | गो-ko | गौ-kau | र्ग-kam | गः-kaḥ |
| ধ্ | ধ্য | ধি | ধ্বী | ध् | ଜ୍ୟୁ | ସ୍ | 2 | ধ্ব | ন্থ | থা | ধ্বী | ର୍ଷ | ଣ୍ବହ |
| घ-gha | घा-ghā | ঘি-ghi | घी-ghi | घु-ghu | ਬ੍ਰ-ghū | 된-ghr | 뜉-ghĩ | घे-ghe | घै-ghai | घो-gho | घौ-ghau | घं-gham | घ:-ghaḥ |
| 3 | 3 | B | ধ্ব | 5 | 5 | Le la | San | 3 | 3 | 31 | ব্র | z | 38 |
| ङ-na | a ङा-nā | ন্তি-ni | ङी-ni | ङ्-nu | डू-hū | रङ्-nr | ङ्ग-hř | ङे-ne | ङै-nai | ङो-no | ङौ-nau | ङं-nam | ङ:-nah |

Figure 11. Consonants with vowel matras in Bhujimmol script from Shakya 2002. Note the curved a head-line, and the particular shapes of $-\bar{a}$, -u, -e, -ai, -o, and -au.



Figure 12. Consonants in Bhujimmol script from Shakya 2002.





Sha Sha Sa Ha Ksha श ष स ह ज्ञ —

Figure 13. Alphabet in Bhujimmol script from Shakya 2002. Note that the shapes of a and e are unrelated to one another, though o is based on a, and that the voiced aspirates are Brahmic.



Figure 14. Alphabet in Golmol script from Shakya 2002. Note that the shapes of *a*, *e*, and *o* are unrelated to one another, and that the voiced aspirates are Brahmic.



Figure 15. Alphabet in Kvemmol script from Shakya 2002. Note that the shapes of *a*, *e*, and *o* are unrelated to one another, and that the voiced aspirates are Brahmic, and that *l* and *l* are distinct.





Figure 16. Alphabet in Litumol script from Shakya 2002. Note that the shapes of a and e are unrelated to one another, though o is based on a, and that the voiced aspirates are Brahmic, and that l and l are distinct.



Figure 17. Examples of monograms text in Rañjana script from Gu & Shi 1995. These read in transliteration from left to right as: *kkhgghna*, *cchjjhña*, *tthddhna*, *pphbbhma*, *yrlwa*, *ssha*.



Figure 18. Example text in Rañjana script from Gu & Shi 1995. These read in transliteration from left to right in rows as: *J*, *kkhgghna*, *cchjjhña*, *cchjjhña*, *cchjjhña ttddhna*, *ttddhna*, *tthddhna*, *pphbbhma*, *yrlva*, *śşsha hksmlwryain* (*ha kşa ma la va ra yain*), *madagalain* ~ *maindagala*



Figure 19a. Monogram in Rañjana script from Tshe-tan Zabs-drun 1992.



Figure 19b. Example text in Rañjana script from Tshe-tan Zabs-drun 1992. The text of the syllable is from the Kalacakra Vajra mantra. The green elements in this rendering should read *kşa ma* but it is difficult to parse as it is the lowest element that ought to be the *ma*. The *la wa ra ya* (yellow, white, red, black) are quite clear.



Figure 20. Summary chart comparison of Rañjana, Pracalit, and Bhujimmol from Gu & Shi 1995; These are taken from Figures 1, 5, and 11.



Figure 21. Summary chart comparison of Rañjana, Wartu, Pracalit, Pācūmol, Himmol, Kummol, Bhujimmol, Golmol, Kvemmol, and Litumol, from Gu & Shi 1995; These are taken from Figures 3, 4, 7, 8, 9, 10, 13, 14, 15, and 16.