ISO/IEC JTC1/SC2/WG2 N4347
L2/12-xxxx
2012-10-08

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

Doc Type: Working Group Document
Title: Proposal to Encode Nepal Himalayish Scripts in ISO/IEC 10646
Author: Pat Hall, Language Technology Kendra and Kathmandu University, Nepal
Status: Individual Contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Date: 2012-10-08

1. Purpose

This proposal brings together a number of previous proposals for scripts used to write the Himalayish Tibeto-Burman languages of Nepal, to unify them within a single proposed code block. The proposals and other documents being unified are:

N3649 L2/09-192 2009-05-04 *Preliminary proposal for encoding the Rañjana script in the SMP of the UCS*
N3692 L2/09-325 2009-09-28 *Roadmapping the scripts of Nepal*
N3962 2011-01-25 *Preliminary Proposal to Encode the Jenticha Script in ISO/IEC 10646.*
N3963 2011-01-25 *Preliminary Proposal to Encode the Tikamuli Script in ISO/IEC 10646.*
N4018 2011-04-13 *Introducing the Khambu Rai Script.*
N4019 2011-04-13 *Introducing the Khema Script for Writing Gurung.*
N4028 2011-05-31 *Proposal to Encode the Jenticha Script in ISO/IEC 10646.*
N4036 2011-05-01 *Introducing the Magar Akkha Script.*
N4037 2011-05-01 *Introducing the Kirat Rai Script.*
N4038 2011-05-03 *Preliminary Proposal to Encode the Prachalit Nepal Script in ISO.IEC 10646.*
N4140 2011-09-29 *Introducing a Script for Writing Dhimal.*
N4184 2012-01-05 *Proposal to Encode the Newar Script in ISO.IEC 10646.*
N4322 2012-02-05 *Proposal for the Nepaalalipi script in the UCS*

This proposal also considers other languages currently written using extended forms of Devanagari; the writing of these languages has been analysed in detail by Michael Noonan in his 2003 paper *Recent Adaptations of the Devanagari Script for the Tibeto-Burman Languages of Nepal* . He covers the writing of Gurung as well as the writing of Chantyal, Tamang, Sherpa, Thangmi and Limbu. He has collected together the phonology of all these languages as well as showing how they can be written in Devanagari. The use of Devanagari must be considered a passing phase, just as it has been for Limbu written in Devanagari until the encoding of their Sirijanga script in Unicode and the production of appropriate fonts.

This proposal is intended to gain agreement that a unification of the writing of all these languages and other languages in the Himalayish group within Nepal would be

beneficial to the 4 million members of the linguistic communities concerned. It would also be able to embrace antiquarian writing of other languages, notably Sanskrit, with the addition of extra characters not required for the living languages, in much the same way as the Devanagari code block has been extended for this purpose.

Following agreement in principle a full proposal will be developed in concert with all the language communities in Nepal and other interested parties, as well as experts in Nepal currently involved in documenting the languages of the country and in that context devising methods of transcribing them.

2. Name for the code block

We propose that the code block is named Himalayish.

There has been considerable debate about the naming of a proposed code block for the Prachalit script used by the Newar community – the original proposal from Pandey (N4038) named the block “Prachalit”, but the revised proposal (N4184) renamed it “Newar” while the alternative proposal by Manandhar and colleagues (N4322) proposed the name “Nepaalalipi”. This led to considerable debate, with “Nepaalalipi” being deprecated because it contains the word “lipi” which means “script”. Much of the debate focused on the need for an English term appropriate for the scripts, languages and community being served by the code block.

In this case the scripts and languages belong to the Tibeto-Burman speaking communities of Nepal living predominantly in the hills and mountains in the north of country. In Nepal the communities are often divided into three sections:

- the Madeshi who live on the southern plains, the Terai, mostly speaking Indo-Aryan languages such as Maithili, Awadhi, Bhojpujri, and Tharu, with a few small Dravidian and Austroasiatic speaking communities.
- the Pahadi, also called Pahari, who live in the hills, the foothills of the Himalayas. “Pahad” or “Pahar” means mountain in Nepali, Hindi and other languages of the area. Ethnologue lists a Pahari subfamily of the Indo-Aryan family of languages, with many languages across the region from Kashmir to Assam. The Pahari group includes Nepali, claimed by around half the population of Nepal as their mother tongue.
- the Himali who live in the mountains, and speak Tibeto-Burman languages. Ethnologue shows a language subfamily Himalayish within the Tibeto-Burman family of languages, which covers all of the Tibeto-Burman languages of Nepal, but also languages in Tibet and a few languages in India

It thus does seem that “Himalayish”, or simply “Himal” or “Himali”, would be a suitable name to the code block. To avoid confusion with the current proprietary uses of “Himal” and “Himali”, I suggest “Himalayish”.

3. The Scripts and Languages

In order to unify the various proposals we need to align the proposed character sets, and will use the English names given to these for this purpose. The English names give simple phonetic transliterations of the characters, and we will also consider the Devanagari equivalents and the full phonology expressed in ipa where appropriate.

Appendix 1 gives a table of the proposed character sets for each of the script-language pairs for all the proposals listed above, plus the languages covered by Noonan indicating what needs to be represented in the writing by reference to the phonology given by him. The phonological requirements of the Newar language have been added by drawing upon Hale and Shrestha (2005).

What this gives us is a superset that covers all scripts and languages in the set being considered.

Of the Himalayish linguistic communities who would be supported with this proposal, only the Newars have a mature written tradition with a variety of styles of writing that have evolved so that they fit comfortably with the modern orthography of the language as it is spoken and written today. Limbu has a script dating back several hundred years is used in limited domains but is now encoded in Unicode while Lepcha is in a similar position but is predominantly an Indian language of Sikkim.

The rest of the proposals listed above have newly created writing, with some of them now being used for printing newspapers while others have yet to be used. For a language that is currently unwritten, the procedure recommended by UNESCO (Robinson and Gadelli 2003) is to base this on a thorough phonemic analysis of the spoken language and then extend a locally dominant writing system to fit the new language; this is precisely what the analysis by Noonan reports, leading to the adaptation of Devanagari for writing the languages Gurung, Tamang, Sherpa, Thangmi and Chantyal. By contrast the writing described in the ISO/SC2/WG2 proposals for Magar, Gurung, Bantawa, Sunuwar and Dhimal does not appear to have been based on phonological analysis though the authors will have been very familiar with the spoken language and thus influenced by this while also being influenced by writing systems used in communities around them. The status the writing of these languages is discussed in the sections below.

For more details of the scripts and languages we refer you to the many ISO/SC2/WG2 documents cited, with their scholarly discussion and many examples of the writing. Not all the scripts have been documented with signs, punctuation, and numerals, though clearly all scripts need these. Appendix 2 contains extracts from the existing proposals of the characters proposed. As will be seen, these are sometimes visually very distinct, even though the underlying functions are the same. This point was put very succinctly by Michael Everson in document N3692(page 2) in discussing the various scripts or styles of writing used for the Newar's language, Nepal Bhasa.

> **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 Gaelic and Fraktur variants, however, is probably applicable, which is why the recommendations here have been made.

3.1 The Prachalit and Nepaalalipi and Ranjana scripts for Nepal Bhasa

The 1991 census listed 825,458 Newars with mother tongue Nepal Bhasa (also known as Newari), making it the 6th largest linguistic community of the country. The writing of the Newars has been considered in five recent ISO WG2 documents:

Michael Everson 2009-05-04 *Preliminary proposal for encoding the Rañjana script in the SMP of the UCS* N3649 L2/09-192

Michael Everson 2009-09-28 *Roadmapping the scripts of Nepal* N3692 L2/09-325
Anshuman Pandey 2011-05-03 *Preliminary Proposal to Encode the Prachalit Nepal Script in ISO.IEC 10646*. N4038
Anshuman Pandey 2012-01-05 *Proposal to Encode the Newar Script in ISO.IEC 10646*. N4184 L2/12-003
Devdass Manandhar, 2012-02-05 *Proposal for the Nepaalalipi script in the UCS* N4322 L2/12-120

See Appendix A2.1 for character examples.

The significant point to note is the inclusion in N4322 of breathy consonants that need to be distinguished from the sequence of the consonant followed by a <ha> since there exist words with substantially different meanings which only differ in their written form in the use of the these breathy consonants. N4184 dismisses these breathy consonants as conjuncts that have written incorrectly, but this is clearly a misunderstanding. The independent analysis of the writing of Chantyal using Devanagari given by Noonan provides further evidence here.

N4322 proposed that where vowels could be written in two distinct forms the pre-composed forms should be excluded, but in encodings of Brahmi derived scripts the practice has been established that both representations are included, as is done in N4184, and this established practice will be followed with the ambiguity resolved in normalisation and collation algorithms.

3.2 The Akkha script for the Magar Language

The Magar are the 7th largest linguistic community in Nepal, with 770,116 mother tongue speakers recorded in the 2001 census. But unlike the Newars, the Magar have no written tradition, though today they do recognise the importance of writing and have looked back in history to claim Brahmi as their original writing, to replace the use of Devanagari which has been be used by default. This is described in

Anshuman Pandey 2011-05-01 *Introducing the Magar Akkha Script*. N4036

which contains extracts from a larger document setting out the Magar case for the script they call “Akkha” (which means “script”). See Appendix A2.2 for an extract from this document.

Pandey correctly points out that Akkha is Brahmi, but I suggest that it would not be appropriate for the Magar community to use this script in its ancient form – over the past two thousand years both scripts and languages have evolved significantly and the Magar people deserve to benefit from this evolution. The Akkha scripts of the Magar people needs to be further developed as a font for the Himalayish code block, this code block will cover everything that they need.

3.3 The Khambu Rai and Kirat Rai scripts for the Bantawa language

This is a relatively large linguistic community of mother tongue speakers of the Bantawa language, 371,056 in the 2001 census. Two scripts are described in:

Anshuman Pandey 2011-04-13 *Introducing the Khambu Rai Script*. N4018
Anshuman Pandey 2011-05-01 *Introducing the Kirat Rai Script*. N4037

See Appendix A2.3 for examples of the scripts.

Pandey points out that Khambu Rai, the first of these, is in one-to-one correspondence with Devanagari, though it does also include three conjuncts <tra>, <ksha> and <gya> frequently treated as distinct letters of the Devanagari alphabet within the Nepalese educational system. Khambu Rai must be viewed as an alternative font for Devanagari, though it would be useful to know how long it has been in use and for what purposes.

Kirat Rai is quite different, and though Pandey says it probably evolved from Khambu Rai, it is hard to see this. As can be seen from the Appendix it is very close to the Tikamuli script for Sunuwar in what has been chosen to create characters for, the only significant differences being that Tikamuli includes a glottal stop and a trill.

3.4 The Jenticha and Tikamuli scripts for the Sunuwar language

The Sunuwar (Sunwar) is a relatively small linguistic community of 26,611 mother tongue speakers, definitely large enough to warrant a distinct style of writing. The writing of Sunuwar is reported in three documents covering two scripts.

Anshuman Pandey N3962 2011-01-25 *Preliminary Proposal to Encode the Jenticha Script in ISO/IEC 10646*.

Anshuman Pandey N3963 2011-01-25 *Preliminary Proposal to Encode the Tikamuli Script in ISO/IEC 10646*.

Anshuman Pandey N4028 2011-05-31 *Proposal to Encode the Jenticha Script in ISO/IEC 10646*.

See Appendix A2.4 for examples of the scripts.

Pandey reports Sunuwar as an official language of Sikkim in India, though Ethnologue records no Sunuwar community in India. Pandey gives two alternative names, Mukhiya and Koits, but neither of these names appear in Ethnologue, though Ethnologue does record a small language Koi spoken in eastern Nepal. This discrepancy needs further investigation, the community in Sikkim must use some other name or some other closely related language.

Jenticha is the older of the two scripts, though created relatively recently in 1942. It is reported as the official script in Sikkim – see note above for the open question here. Tikamuli has clearly been created recently, though no dates or sources are given.

The two scripts do have significant differences. Tikamuli is much closer to Devanagari in its choice of consonants, and we must wonder about the influence here. Both include a trill and glottal stop. They both only have a single set of vowels, not distinguishing between the independent and dependent forms, with Jenticha preferring the independent forms while Tikamula prefers the dependent forms.

This choice of only one form of the vowels makes us wonder whether the distinction is necessary in any Brahmi script – modern fonts could distinguish what needs to be rendered from context and select the independent or dependent form as appropriate.

3.5 The Khema script for the Gurung language

The Gurung is a large linguistic community with 338,925 mother tongue speakers recorded in the 2001 census. Ethnologue documents two distinct dialects not recognised in the 2001 census. A single document:

Anshuman Pandey N4019 2011-04-13 *Introducing the Khema Script for Writing Gurung*.

introduces the recently created Khema script, see Appendix A2.5 for examples of the Khema script. A second script is noted, but no details are given.

Here we have an independent phonology documented by Noonan and have compared the two within the table in the Appendix. The Khema script has the usual diphthongs not required by the phonology, and more consonants than required, indicating perhaps an undue influence of Devanagari in the Khema script.

Taking the phonology as setting the requirements for writing, we see that there is a need for four tone markers, but otherwise the needs of Gurung are well catered for by the requirement laid down by other scripts and languages.

3.6 The Dham script for Dhimal.

The Dhimal linguistic community is relatively small at 17,308 in the 2001 census. A recently created scripts is recorded in:

Anshuman Pandey 2011-09-29 *Introducing a Script for Writing Dhimal*. N4140

See Appendix A2.6 for examples of the script.

The total alignment with Devanagari makes this appear to be a proposal for an alternative font for Devanagari, though clearly if such a font existed it could equally well be mapped to the Himalayish code block.

4. Unwritten languages

4.1 Writing Tamang

At 1,179,145 mother tongue speakers, Tamang is the largest of the Himalayish language communities in Nepal. With their strong Buddhist traditions, there has been a movement to write Tamang in the Tibetan script, but equally well for administrative purposes writing in Devanagari has been acceptable. The writing of Tamang in Devanagari is described by Noonan, who notes the need for 4 tones.

4.2 Writing Thangmi (Thami)

The Thami or Thangmi linguistic community has 18,991 speakers recorded in the 2001 census, and has been written in an extension of Devanagari for the purposes of documenting the language. From Noonan's phonology we see that it needs fewer vowels, most of the standard consonants, and a glottal stop.

4.3 Writing Chantyal

The Chantyal language community is small at 5,912 speakers. Its writing needs are similar to those of Nepal Bhasa including the breathy nasals and other breathy consonants not present in Devanagari. As noted by Noonan, there are distinct words with the breathy consonants and with the consonant followed by the <ha>, making it not possible to represent the breathy consonant by a conjunct of the consonant and <ha>.

4.4 Writing Sherpa

Sherpa with 31,500 mother tongue speakers would fit well within the Himalayish code block, requiring 2 tone markers in addition to most of the standard characters. A scholar from East Asia has been developing a script for the Sherpa community, but we do not have access to details of this.

5. Other languages

5.1 Issues regarding the writing of Limbu

Limbu with 333,633 mother tongue speakers has a traditional writing system Sirijanga which has been coded in Unicode but there must be some questions about this encoding. It is Brahmi writing system, encoded using a limited form of the subjoined model of Tibetan, without a virama but with a similar character sa-i which could take on the full virama role (Boyd Michailovsky, private communication). The potential need for this has been raised by the recent proposal

Anshuman Pandey 2011-01-14 *Proposal to Encode the Letters GYAN and TRA for Limbu in the UCS* ISO/IEC JTC1/SC2/WG2 N3975

where the characters <gya> and <tra> would normally be rendered as a conjunct, not currently possible for Limbu as it currently stands.

5.2 Writing of non-Nepalese languages

As has been noted in the documents on Nepal Bhasa, the Nepal Lipi scripts have been used for Sanskrit, Maithili, and even Tibetan. A number of the characters shown in the Appendix for Nepal Lipi are clearly there only for this purpose, notably diphthongs and syllabic liquids (vocalics). A group of the consonants (TTA to NNA) are not used in Nepal Lipi, but are needed for other Himalayish languages and should be included in the main part of the code block.

6. The proposed Code Block

A code block of just under 100 characters is needed.

Independent Vowels: A, AA, I, II, U, UU, E, O. Dependent forms subject to review.

Diphthongs: AI, AU and dependent forms, subject to review.

Syllabic liquids: VOCALIC R, VOCALIC RR, VOCALIC L, VOCALIC LL and dependent forms included for old documents in Sanskrit etc.

Consonants: KA, KHA. GA, GHA, NGA, NGHA, CA, CHA, JA, JHA, NYA/NJA, NYHA/NJHA. TTA, TTHA, DDA, DDHA, NNA, NNHA, TA, THA, DA, DHA, NA, NHA, PA, PHA, BA, BHA, MA, MHA, YA, YHA, RA, RHA, RRA, LA, LHA, VA/WA, VHA, SHA, SSA, SA, HA, glottal stop

Various signs: CANDRABINDU, VISARGA, VIRAMA, TRILL, 4 TONES; and for Sanskrit etc: ANUSVARA, AVAGRAHA

Invocations: OM and other invocations and graphics as required.

Punctuation: DANDA, DOUBLE DANDA, WORD BREAK; and if really required: COMMA, FULL STOP.

Digits: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

7. References

Anderson, Deborah (2012) *Liaison Report, Script Encoding Initiative, UC Berkeley*. Document ISO/IEC JTC1/SC2/WG2 N4220 2012-02-12

BIS (1991) *Indian Standard Indian Script Code for Information Interchange – ISCII*. IS 12194: 1991. Bureau of Indian Standards, New Delhi.

Cahill, Michael and Elke Karan (2008) *Factors in Designing Effective Orthographies for Unwritten Languages,* SIL Electronic Working Papers 2008-001, February 2008

Everson, Michael ; Rick McGowan, Ken Whistler, and V.S. UMAmaheswaran (2012) *Snapshot of Pictorial view of Roadmaps to BMP, SMP, SIP, TIP and SSP* Document ISO/IEC JTC 1/SC 2/WG 2 N4186 2012-02-16 Downloadable from http://std.dkuug.dk/jtc1/sc2/wg2/docs/n4186.pdf

Everson, Michael (2000) *Newari code table and names list* http://www.evertype.com/standards/tai/newari.pdf (accessed 2012/02/15)

Everson, Michael (2009b) *Roadmapping the scripts of Nepal.* International Organisation for Standardisation document ISO/IEC JTC1/SC2/WG2 N3692 downloadable from http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3692.pdf (accessed 2012/02/15) also from http://www.evertype.com/formal.html

Hale, Austin and Kedar P. Shrestha (2005) *Newar*. Volume 256 in Languages of the World/Meterials. Munich and Newcastle: Lincom Europa

International Phonetic Association (1999) *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet.* Cambridge University Press

Kasah, Sharad Junior (2009) Paper distributed at meeting of the Nepal Lipi Guthi in Kathmandu, March 2010.

Lewis, M. Paul (ed.), 2009. *Ethnologue: Languages of the World, Sixteenth edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com/

Michailovsky, Boyd and Michael Everson (2002) *Revised proposal to encode the Limbu script in the UCS.* Document ISO/IEC JTC1/SC2/WG2 N2410, International Organization for Standardization. Downloadable from http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2410.pdf

Noonan, Michael (2003) *Recent Adaptations of the Devanagari Script for the Tibeto-Burman Languages of Nepal*. https://pantherfile.uwm.edu/noonan/www/Papers.html (last accessed 29th May 2010)

Robinson, Clinton and Karl Gadelii, (2003) *Writing Unwritten Languages*. UNESCO downloadable from http://portal.unesco.org/education/en/ev.php-URL_ID=28300&URL_DO=DO_TOPIC&URL_SECTION=201.html

Shakya, Rabison (2002) *Alphabet of the Nepalese Script*. Patan, Nepal: Motiraj Shakya and Sanunani Shakya

UNESCO (2003) *Language Vitality and Endangerment*, UNESCO, Paris

Unicode Consortium (2001) *Roadmaps* ISO/IEC JTC1/SC2/WG2 N2383 downloadable from http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2383.pdf accessed 2012/02/14

Unicode Consortium (2012) *Roadmap to the SMP* http://www.unicode.org/roadmaps/smp/

Vajrachaarya, Suwarn (2009) *Towards Encoding Nepal Scripts: A Report and Preliminary Proposal*. Distributed at meeting of the Nepal Lipi Guthi in Kathmandu, March 2010.

Vajracharya, Suwarn (2012) *Feedback on the Proposal by Anshuman Pandey to Encode the Newar Script in ISO/IEC 10646 dated January 6 2012*. Communication on email discussion group.. February 8th, 2012.

Wikipedia (20123) articles on Pahari languages, http://en.wikipedia.org/wiki/Pahari_languages

Yadava, Yogendra P. (2003) Language Chapter 4. *Population Monograph, volume 1,* Kathmandu: Central Bureau of Statistics.2003 and UNFPA

Yadava, Yogendra P. (2007) Linguistic Diversity in Nepal Perspectives on Language Policy International Seminar on *Constitutionalism and Diversity in Nepal*, CNAS, 22-24 August,

Appendix 1.

Aligning and comparing the proposed encodings for the writing of languages in the Himalayish group in Nepal.

			Li	NB				Ma	Ba	Ba	Su	Su	Dh	Gu	Gu	Sh	Ta	Th	Ch
Language			Limbu	Nepal Bhasa				Magar	Bantawa		Sunuwar		Dhimal	Gurung		Sherpa	Tamang	Thangmi	Chantyal
Script: name or phonemic inventory			phonemic	Prachalit	Prachalit	Ranjana	phonemic	Akkha	KhambuRai	KiratRai	Jentichа	Tikamuli	Dham	phonemic	Khema	phonemic	phonemic	phonemic	phonemic
Source: ISO/SC2/WG2 Nxxxx document or Noonan (2003) or Hale and Shrestha (2005)			Noonan	4184	4322	3649	Hale et al	4036	4018	4037	4028	3963	4140	Noonan	4019	Noonan	Noonan	Noonan	Noonan
Semantics																			
English	**Naga**	**ipa**																	
Independent Vowels																			
A	अ	ʌ		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓
AA	आ	a	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓
I	इ	i	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓
II	ई	iː	✓	✓		✓	✓	✓	✓				✓				✓		✓
U	उ	u	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓
UU	ऊ	uː	✓	✓		✓	✓	✓	✓				✓				✓		
E	ए	e	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓
O	ओ	o	✓	✓	✓	✓			✓		✓		✓	✓	✓	✓	✓	✓	✓
Dependent forms																			
AA	ा			✓	✓	✓			✓	✓		✓	✓						
I	ि			✓	✓	✓			✓	✓		✓	✓						
II	ी			✓		✓			✓				✓						
U	ु			✓	✓	✓			✓	✓		✓	✓						
UU	ू			✓		✓			✓				✓						
E	े			✓	✓	✓			✓	✓		✓	✓						
O	ो			✓	✓	✓			✓	✓		✓	✓						
Diphthongs																			
AI	ऐ	ʌy		✓		✓	✓	✓	✓				✓		✓				
AU	औ	ʌw		✓		✓	✓		✓				✓		✓				
Dependent forms																			
AI	ै			✓	✓	✓	✓		✓	✓			✓						

			Li	NB				Ma	Ba	Ba	Su	Su	Dh	Gu	Gu	Sh	Ta	Th	Ch
AU	◌ौ			✓	✓	✓	✓		✓	✓			✓						
Syllabic liquids																			
VOCALIC R	ऋ	r̥		✓	✓	✓			✓										
VOCALIC RR	ॠ	r̥ː		✓	✓	✓			✓										
VOCALIC L	ऌ	l̥		✓	✓	✓													
VOCALIC LL	ॡ	l̥ː		✓	✓	✓													
Dependent forms																			
VOCALIC R	◌ृ	r̥		✓	✓	✓			✓										
VOCALIC RR	◌ॄ	r̥ː		✓		✓			✓										
VOCALIC L	◌ॢ	l̥		✓		✓													
VOCALIC LL	◌ॣ	l̥ː		✓		✓													
Consonants																			
KA	क	k	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
KHA	ख	kʰ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
GA	ग	g		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
GHA	घ	gh		✓	✓	✓	✓	✓	✓	✓		✓	✓		✓			✓	✓
NGA	ङ	ŋ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NGHA					✓														
CA	च	c	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CHA	छ	cʰ	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
JA	ज	j		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
JHA	झ	jh		✓	✓	✓	✓	✓	✓	✓		✓	✓		✓			✓	✓
NYA/NJA	ञ	ɲ	✓	✓	✓	✓		✓	✓	✓			✓			✓			✓
NYHA/NJHA		ɲh			✓														✓
TTA	ट	ʈ		✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
TTHA	ठ	ʈʰ		✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
DDA	ड	ɖ		✓	✓	✓			✓	✓		✓	✓	✓	✓	✓		✓	
DDHA	ढ	ɖh		✓	✓	✓			✓	✓		✓	✓		✓			✓	
NNA	ण	ɳ		✓	✓	✓			✓			✓	✓				✓		
NNHA																			
TA	त	t	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
THA	थ	tʰ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DA	द	d		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
DHA	ध	dh		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓
NA	न	n	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
NHA		nh			✓		✓												✓
PA	प	p	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PHA	फ	pʰ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
BA	ब	b		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
BHA	भ	bh		✓	✓	✓	✓	✓	✓	✓		✓	✓		✓			✓	✓
MA	म	m	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
MHA		mh			✓		✓												✓
YA	य	y	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
YHA																			
RA	र	r		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
RHA		rh			✓		✓		✓										✓
RRA									✓										
LA	ल	l	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
LHA		lh																	✓
VA/WA	व	w/v	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
VHA																			
SHA	श	ʃ		✓	✓	✓			✓	✓	✓		✓			✓			

			Li	NB				Ma	Ba	Ba	Su	Su	Dh	Gu	Gu	Sh	Ta	Th	Ch
SSA	ष	ʂ		✓	✓	✓			✓			✓	✓						
SA	स	s	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
HA	ह	ɦ	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
HHA											✓								
glottal stop	ʔ	ʔ	✓								✓	✓						✓	
KSSA									✓										
JNYA									✓										
TRA									✓										
Various signs																			
CANDRABINDU	◌ँ			✓	✓	✓				✓	✓								
ANUSVARA	◌ं			✓		✓			✓				✓						
VISARGA	◌ः			✓	✓	✓			✓		✓	✓	✓						
- extra long												✓							
VIRAMA	◌्			✓	✓	✓				✓	✓	✓	✓						
AVAGRAHA				✓															
TRILL											✓	✓							
Invocations																			
OM	ॐ			✓	✓	✓													
ANJI/SIDDHI				✓	✓														
SVASTI				✓															
FLOWER					॥※॥														
YIG MGO						✓													
Punctuation																			
DANDA				✓	✓	✓													
DOUBLE DANDA				✓	✓	✓													
COMMA				✓															
FULL STOP				✓															
WORD BREAK				✓	✓														
Digits																			
0	०			✓	✓	✓				✓	✓	✓							
1	१			✓	✓	✓				✓	✓	✓							
2	२			✓	✓	✓				✓	✓	✓							
3	३			✓	✓	✓				✓	✓	✓							
4	४			✓	✓	✓				✓	✓	✓							
5	५			✓	✓	✓				✓	✓	✓							
6	६			✓	✓	✓				✓	✓	✓							
7	७			✓	✓	✓				✓	✓	✓							
8	८			✓	✓	✓				✓	✓	✓							
9	९			✓	✓	✓				✓	✓	✓							

Appendix 2. Extracts from the WG2 proposals

A2.1 Ranjana N3649, Newar N4184, and Nepaalalipi N4322

A2.1.1 Ranjana

Independent vowels

11800 𑐀 RANJANA LETTER A
11801 𑐁 RANJANA LETTER AA
11802 𑐂 RANJANA LETTER I
11803 𑐃 RANJANA LETTER II
11804 𑐄 RANJANA LETTER U
11805 𑐅 RANJANA LETTER UU
11806 𑐆 RANJANA LETTER VOCALIC R
11807 𑐇 RANJANA LETTER VOCALIC RR
11808 𑐈 RANJANA LETTER VOCALIC L
11809 𑐉 RANJANA LETTER VOCALIC LL
1180A 𑐊 RANJANA LETTER E
1180B 𑐋 RANJANA LETTER AI
1180C 𑐌 RANJANA LETTER O
1180D 𑐍 RANJANA LETTER AU

Consonants

1180E 𑐎 RANJANA LETTER KA
1180F 𑐏 RANJANA LETTER KHA
11810 𑐐 RANJANA LETTER GA
11811 𑐑 RANJANA LETTER GHA
11812 𑐒 RANJANA LETTER NGA
11813 𑐓 RANJANA LETTER CA
11814 𑐔 RANJANA LETTER CHA
11815 𑐕 RANJANA LETTER JA
11816 𑐖 RANJANA LETTER JHA
11817 𑐗 RANJANA LETTER NYA
11818 𑐘 RANJANA LETTER TTA
11819 𑐙 RANJANA LETTER TTHA
1181A 𑐚 RANJANA LETTER DDA
1181B 𑐛 RANJANA LETTER DDHA
1181C 𑐜 RANJANA LETTER NNA
1181D 𑐝 RANJANA LETTER TA
1181E 𑐞 RANJANA LETTER THA
1181F 𑐟 RANJANA LETTER DA
11820 𑐠 RANJANA LETTER DHA
11821 𑐡 RANJANA LETTER NA
11822 𑐢 RANJANA LETTER PA
11823 𑐣 RANJANA LETTER PHA
11824 𑐤 RANJANA LETTER BA
11825 𑐥 RANJANA LETTER BHA
11826 𑐦 RANJANA LETTER MA
11827 𑐧 RANJANA LETTER YA
11828 𑐨 RANJANA LETTER RA
11829 𑐩 RANJANA LETTER LA
1182A 𑐪 RANJANA LETTER WA
1182B 𑐫 RANJANA LETTER SHA
1182C 𑐬 RANJANA LETTER SSA
1182D 𑐭 RANJANA LETTER SA
1182E 𑐮 RANJANA LETTER HA

Various signs

1182F ◌𑐯 RANJANA SIGN CANDRABINDU
11830 ◌𑐰 RANJANA SIGN ANUSVARA
11831 ◌𑐱 RANJANA SIGN VISARGA

Dependent vowel signs

11832 ◌𑐲 RANJANA VOWEL SIGN AA
11833 ◌𑐳 RANJANA VOWEL SIGN I
11834 ◌𑐴 RANJANA VOWEL SIGN II
11835 ◌𑐵 RANJANA VOWEL SIGN U
11836 ◌𑐶 RANJANA VOWEL SIGN UU
11837 ◌𑐷 RANJANA VOWEL SIGN VOCALIC R
11838 ◌𑐸 RANJANA VOWEL SIGN VOCALIC RR
11839 ◌𑐹 RANJANA VOWEL SIGN VOCALIC L
1183A ◌𑐺 RANJANA VOWEL SIGN VOCALIC LL
1183B ◌𑐻 RANJANA VOWEL SIGN E
1183C ◌𑐼 RANJANA VOWEL SIGN AI
1183D ◌𑐽 RANJANA VOWEL SIGN O
1183E ◌𑐾 RANJANA VOWEL SIGN AU

Various signs

1183F ◌𑐿 RANJANA VIRAMA

Digits

11840 𑑀 RANJANA DIGIT ZERO
11841 𑑁 RANJANA DIGIT ONE
11842 𑑂 RANJANA DIGIT TWO
11843 𑑃 RANJANA DIGIT THREE
11844 𑑄 RANJANA DIGIT FOUR
11845 𑑅 RANJANA DIGIT FIVE
11846 𑑆 RANJANA DIGIT SIX
11847 𑑇 RANJANA DIGIT SEVEN
11848 𑑈 RANJANA DIGIT EIGHT
11849 𑑉 RANJANA DIGIT NINE

Punctuation

1184B 𑑋 RANJANA OM
1184C 𑑌 RANJANA YIG MGO
1184D 𑑍 RANJANA DANDA
1184E 𑑎 RANJANA DOUBLE DANDA
1184F 𑑏 RANJANA VIRAM

A2.1.2 Newar N4184

	1140	1141	1142	1143	1144	1145	1146	1147
0	11400	11410	11420	11430	11440	11450		
1	11401	11411	11421	11431	11441	11451		
2	11402	11412	11422	11432	11442	11452		
3	11403	11413	11423	11433	11443	11453		
4	11404	11414	11424	11434	11444	11454		
5	11405	11415	11425	11435	11445	11455		
6	11406	11416	11426	11436	11446	11456		
7	11407	11417	11427	11437	11447	11457		
8	11408	11418	11428	11438	11448	11458		
9	11409	11419	11429	11439		11459		
A	1140A	1141A	1142A	1143A				
B	1140B	1141B	1142B	1143B				
C	1140C	1141C	1142C	1143C				
D	1140D	1141D	1142D	1143D				
E	1140E	1141E	1142E	1143E				
F	1140F	1141F	1142F	1143F				

The script is also known as 'Nepal Lipi', 'Prachalit Lipi', 'Newah Akhah', and 'Newari'.

Independent vowels

11400 𑐀 NEWAR LETTER A
11401 𑐁 NEWAR LETTER AA
11402 𑐂 NEWAR LETTER I
11403 𑐃 NEWAR LETTER II
11404 𑐄 NEWAR LETTER U
11405 𑐅 NEWAR LETTER UU
11406 𑐆 NEWAR LETTER VOCALIC R
11407 𑐇 NEWAR LETTER VOCALIC RR
11408 𑐈 NEWAR LETTER VOCALIC L
11409 𑐉 NEWAR LETTER VOCALIC LL
1140A 𑐊 NEWAR LETTER E
1140B 𑐋 NEWAR LETTER AI
1140C 𑐌 NEWAR LETTER O
1140D 𑐍 NEWAR LETTER AU

Consonants

1140E 𑐎 NEWAR LETTER KA
1140F 𑐏 NEWAR LETTER KHA
11410 𑐐 NEWAR LETTER GA
11411 𑐑 NEWAR LETTER GHA
11412 𑐒 NEWAR LETTER NGA
11413 𑐓 NEWAR LETTER CA
11414 𑐔 NEWAR LETTER CHA
11415 𑐕 NEWAR LETTER JA
11416 𑐖 NEWAR LETTER JHA
11417 𑐗 NEWAR LETTER NYA
11418 𑐘 NEWAR LETTER TTA
11419 𑐙 NEWAR LETTER TTHA
1141A 𑐚 NEWAR LETTER DDA
1141B 𑐛 NEWAR LETTER DDHA
1141C 𑐜 NEWAR LETTER NNA
1141D 𑐝 NEWAR LETTER TA
1141E 𑐞 NEWAR LETTER THA
1141F 𑐟 NEWAR LETTER DA
11420 𑐠 NEWAR LETTER DHA
11421 𑐡 NEWAR LETTER NA
11422 𑐢 NEWAR LETTER PA
11423 𑐣 NEWAR LETTER PHA
11424 𑐤 NEWAR LETTER BA
11425 𑐥 NEWAR LETTER BHA
11426 𑐦 NEWAR LETTER MA
11427 𑐧 NEWAR LETTER YA
11428 𑐨 NEWAR LETTER RA
11429 𑐩 NEWAR LETTER LA
1142A 𑐪 NEWAR LETTER VA
1142B 𑐫 NEWAR LETTER SHA
1142C 𑐬 NEWAR LETTER SSA
1142D 𑐭 NEWAR LETTER SA
1142E 𑐮 NEWAR LETTER HA

Dependent vowel signs

1142F 𑐯 NEWAR VOWEL SIGN AA
11430 𑐰 NEWAR VOWEL SIGN I
11431 𑐱 NEWAR VOWEL SIGN II
11432 𑐲 NEWAR VOWEL SIGN U
11433 𑐳 NEWAR VOWEL SIGN UU
11434 𑐴 NEWAR VOWEL SIGN VOCALIC R
11435 𑐵 NEWAR VOWEL SIGN VOCALIC RR
11436 𑐶 NEWAR VOWEL SIGN VOCALIC L
11437 𑐷 NEWAR VOWEL SIGN VOCALIC LL
11438 𑐸 NEWAR VOWEL SIGN E
11439 𑐹 NEWAR VOWEL SIGN AI
1143A 𑐺 NEWAR VOWEL SIGN O
1143B 𑐻 NEWAR VOWEL SIGN AU

Various signs

1143C 𑐼 NEWAR SIGN CANDRABINDU
= milaphuti
1143D 𑐽 NEWAR SIGN ANUSVARA
= sinhaphuti
1143E 𑐾 NEWAR SIGN VISARGA
• Also used for marking vowel length
= lyaphuti
1143F 𑐿 NEWAR SIGN VIRAMA
= tutipala, halant

Additions for Sanskrit

11440 𑑀 NEWAR SIGN AVAGRAHA

Invocation signs

11441 𑑁 NEWAR OM
11442 𑑂 NEWAR ANJI
• Written at the beginning of a text
• Represents Sanskrit “siddhirastu”
= bhin arambha, siddhi
11443 𑑃 NEWAR SVASTI

Punctuation

11444 𑑄 NEWAR DANDA
= dipu cinha
11445 𑑅 NEWAR DOUBLE DANDA
= nidho dipu cinha
11446 𑑆 NEWAR COMMA
• Phrase separator
11447 𑑇 NEWAR FULL STOP
• Indicates end of text section
11448 𑑈 NEWAR PADA SANDHI MARK
• Indicates word break at end of line

Digits

11450 𑑐 NEWAR DIGIT ZERO
11451 𑑑 NEWAR DIGIT ONE
11452 𑑒 NEWAR DIGIT TWO
11453 𑑓 NEWAR DIGIT THREE
11454 𑑔 NEWAR DIGIT FOUR
11455 𑑕 NEWAR DIGIT FIVE
11456 𑑖 NEWAR DIGIT SIX
11457 𑑗 NEWAR DIGIT SEVEN
11458 𑑘 NEWAR DIGIT EIGHT
11459 𑑙 NEWAR DIGIT NINE

A2.1.3 Nepaalalipi N4322

Code Point	Character	Character Name
		Vowel Letters
XX00	𑐀	NEPAALALIPI LETTER A
XX01	𑐁	NEPAALALIPI LETTER AA
XX02	𑐂	NEPAALALIPI LETTER I
		Vowel Diacritic
XX03	𑐃	NEPAALALIPI MARK AI
		Vowel Letter
XX04	𑐄	NEPAALALIPI LETTER U
		Vowel Diacritic
XX05	𑐅	NEPAALALIPI VOWEL MARK AU
		Vowel Letters
XX06	𑐆	NEPAALALIPI LETTER E
XX07	𑐇	NEPAALALIPI LETTER O
		Bā-ākha Letters
XX08	𑐈	NEPAALALIPI LETTER KA
XX09	𑐉	NEPAALALIPI LETTER KHA
XX0A	𑐊	NEPAALALIPI LETTER GA
XX0B	𑐋	NEPAALALIPI LETTER GHA
XX0C	𑐌	NEPAALALIPI LETTER NGA
XX0D	𑐍	NEPAALALIPI LETTER NGHA
XX0E	𑐎	NEPAALALIPI LETTER CA
XX0F	𑐏	NEPAALALIPI LETTER CHA
XX10	𑐐	NEPAALALIPI LETTER JA
XX11	𑐑	NEPAALALIPI LETTER JHA
XX12	𑐒	NEPAALALIPI LETTER NJA
XX13	𑐓	NEPAALALIPI LETTER NJHA
XX14	𑐔	NEPAALALIPI LETTER TA
XX15	𑐕	NEPAALALIPI LETTER TTA
XX16	𑐖	NEPAALALIPI LETTER THA
XX17	𑐗	NEPAALALIPI LETTER TTHA
XX18	𑐘	NEPAALALIPI LETTER DA
XX19	𑐙	NEPAALALIPI LETTER DDA
XX1A	𑐚	NEPAALALIPI LETTER DHA
XX1B	𑐛	NEPAALALIPI LETTER DDHA
XX1C	𑐜	NEPAALALIPI LETTER NA
XX1D	𑐝	NEPAALALIPI LETTER NNA
XX1E	𑐞	NEPAALALIPI LETTER NHA
XX1F	𑐟	NEPAALALIPI LETTER PA
XX20	𑐠	NEPAALALIPI LETTER PHA
XX21	𑐡	NEPAALALIPI LETTER BA
XX22	𑐢	NEPAALALIPI LETTER BHA
XX23	𑐣	NEPAALALIPI LETTER MA
XX24	𑐤	NEPAALALIPI LETTER MHA
XX25	𑐥	NEPAALALIPI LETTER YA
XX26	𑐦	NEPAALALIPI LETTER RA
XX27	𑐧	NEPAALALIPI LETTER RHA
XX28	𑐨	NEPAALALIPI LETTER LA
XX29	𑐩	NEPAALALIPI LETTER LHA
XX2A	𑐪	NEPAALALIPI LETTER WA
XX2B	𑐫	NEPAALALIPI LETTER SA
XX2C	𑐬	NEPAALALIPI LETTER SHA
XX2D	𑐭	NEPAALALIPI LETTER SSA
XX2E	𑐮	NEPAALALIPI LETTER HA
		Functional Vowel Modifier
XX2F	𑐯	NEPAALALIPI LONG SOUND MARK
		Vowel Modifier
XX30	𑐰	NEPAALALIPI NASAL SOUND MARK
		Vowel Diacritical Marks
XX31	𑐱	NEPAALALIPI VOWEL MARK AA
XX32	𑐲	NEPAALALIPI VOWEL MARK I
XX33	𑐳	NEPAALALIPI VOWEL MARK U
XX34	𑐴	NEPAALALIPI VOWEL MARK E
XX35	𑐵	NEPAALALIPI VOWEL MARK O
		Consonantal Mark
XX36	𑐶	NEPAALALIPI CONSONANTAL MARK
		Syllabic Mark
XX37	𑐷	NEPAALALIPI SYLLABIC MARK RRI
		Syllable Letters
XX38	𑐸	NEPAALALIPI SYLLABIC LETTER RRI
XX39	𑐹	NEPAALALIPI SYLLABIC LETTER RRII
XX3A	𑐺	NEPAALALIPI SYLLABIC LETTER LRRI
XX3B	𑐻	NEPAALALIPI SYLLABIC LETTER LRRII
		Starting Symbol
XX3C	𑐼	NEPAALALIPI SIGN OM
XX3D	𑐽	NEPAALALIPI SIGN SIDDHA
		Word Break Symbol
XX3E	𑐾	NEPAALALIPI SIGN KHAGWA SWAPU
		Different Signs
XX3F	𑐿	NEPAALALIPI SIGN
		Full Stop Marks
XX40	𑑀	NEPAALALIPI SIGN DIPU
XX41	𑑁	NEPAALALIPI SIGN DOUBLE DI
XX42	𑑂	NEPAALALIPI SIGN FLOWER

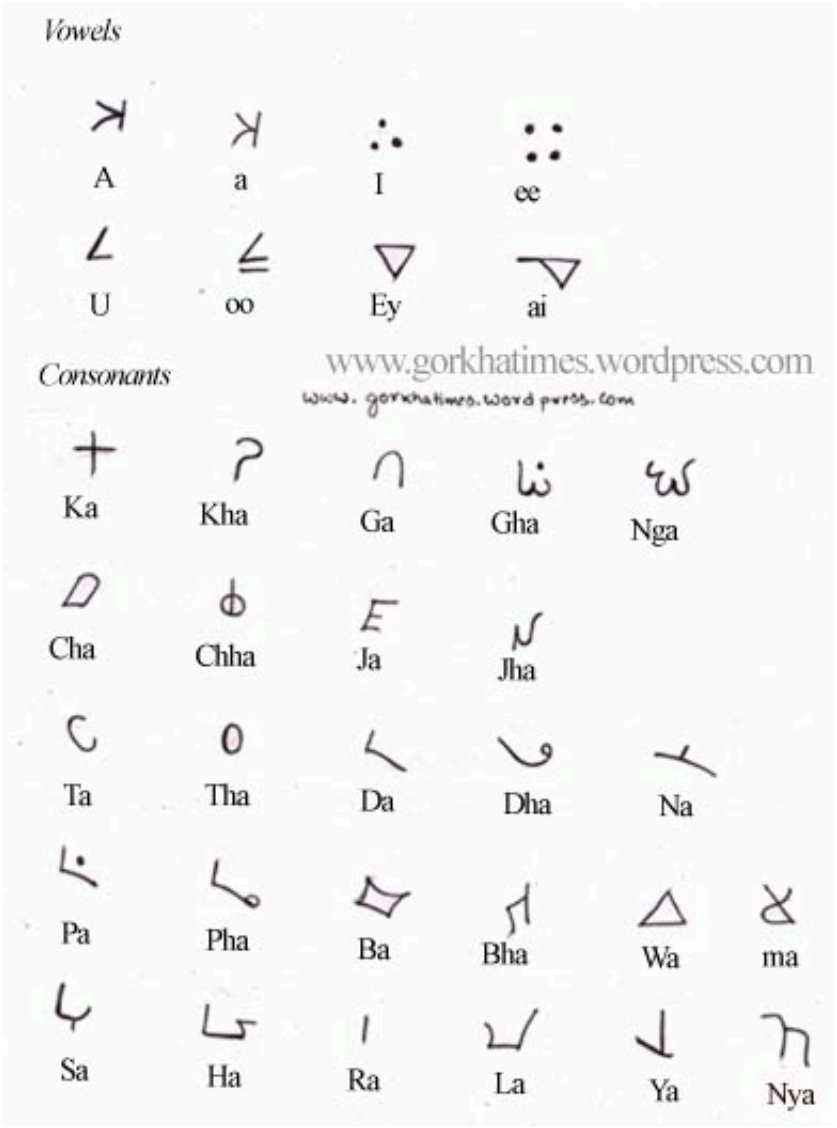
NOTE: XX42 NEPAALALIPI SI

Numerals

Code Point	Character	Character Name
XX50	𑑐	NEPAALALIPI DIGIT 0
XX51	𑑑	NEPAALALIPI DIGIT 1
XX52	𑑒	NEPAALALIPI DIGIT 2
XX53	𑑓	NEPAALALIPI DIGIT 3
XX54	𑑔	NEPAALALIPI DIGIT 4
XX55	𑑕	NEPAALALIPI DIGIT 5
XX56	𑑖	NEPAALALIPI DIGIT 6
XX57	𑑗	NEPAALALIPI DIGIT 7
XX58	𑑘	NEPAALALIPI DIGIT 8
XX59	𑑙	NEPAALALIPI DIGIT 9

Code from XX43 to XX4Fand XX5A to XX5F are vacant space

A2.2 Magar N4036

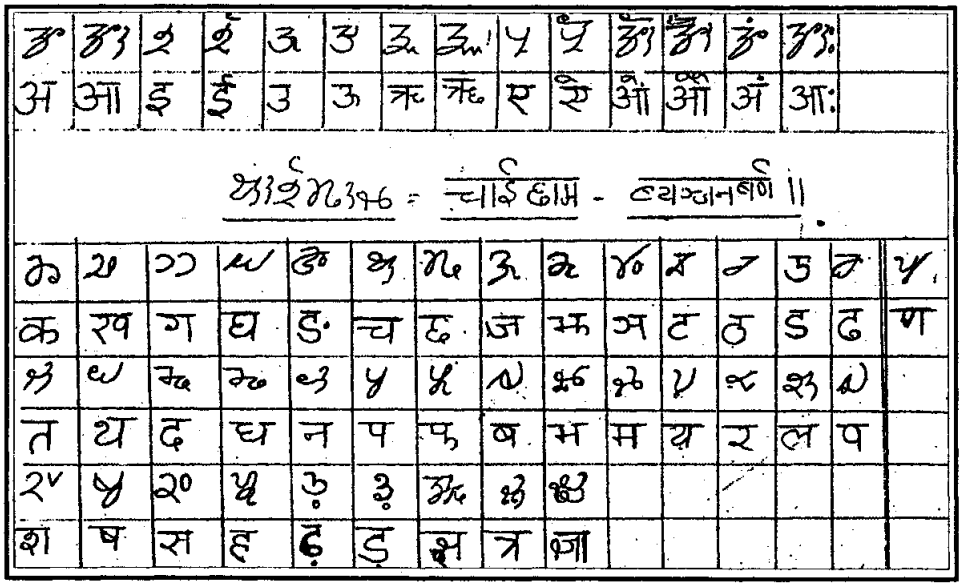


A2.3 Khambu Rai N4018 and Kirat Rai N4037

A2.3.1 Khambu Rai N4018

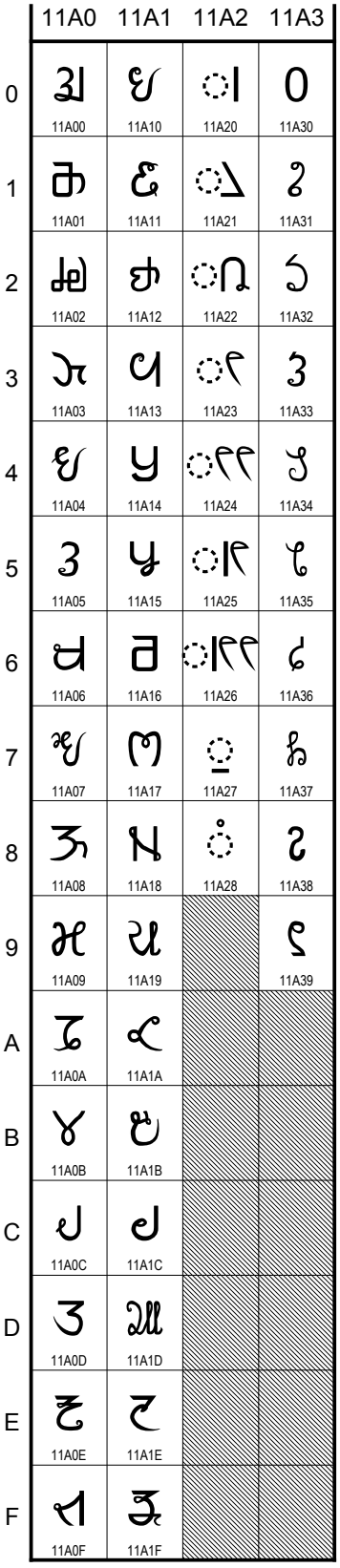
7.09 Rai Script of Khambu-Rai:

The Khambu-Rai languages of 38 sub-tribes is written in Khambu-Rai script developed by Late Kripasalyan Rai in 1981-82 from the Devnagari script. The Khambu-Rai language (Bantawa language) is taught in schools up to the primary level ever since the Khambu-Rai language was recognized as one of the official languages of Sikkim in 1997.



The Khambu-Rai script has the same letters as that of Devnagari with 14 vowels, 13 vowel signs, and 38 consonants (Rai, Krishnadas in Lipi Sangalo (ed) by Limboo, R.B. 1998:54-59).

A2.3.2 Kirat Rai N4037



Vowel

11A00 KIRAT RAI LETTER A

Consonants

11A01 KIRAT RAI LETTER KA
11A02 KIRAT RAI LETTER KHA
11A03 KIRAT RAI LETTER GA
11A04 KIRAT RAI LETTER GHA
11A05 KIRAT RAI LETTER NGA
11A06 KIRAT RAI LETTER CA
11A07 KIRAT RAI LETTER CHA
11A08 KIRAT RAI LETTER JA
11A09 KIRAT RAI LETTER JHA
11A0A KIRAT RAI LETTER NYA
11A0B KIRAT RAI LETTER TTA
11A0C KIRAT RAI LETTER TTHA
11A0D KIRAT RAI LETTER DDA
11A0E KIRAT RAI LETTER DDHA
11A0F KIRAT RAI LETTER TA
11A10 KIRAT RAI LETTER THA
11A11 KIRAT RAI LETTER DA
11A12 KIRAT RAI LETTER DHA
11A13 KIRAT RAI LETTER NA
11A14 KIRAT RAI LETTER PA
11A15 KIRAT RAI LETTER PHA
11A16 KIRAT RAI LETTER BA
11A17 KIRAT RAI LETTER BHA
11A18 KIRAT RAI LETTER MA
11A19 KIRAT RAI LETTER YA
11A1A KIRAT RAI LETTER RA
11A1B KIRAT RAI LETTER LA
11A1C KIRAT RAI LETTER VA
11A1D KIRAT RAI LETTER SHA
11A1E KIRAT RAI LETTER SA
11A1F KIRAT RAI LETTER HA

Vowel signs

11A20 KIRAT RAI VOWEL SIGN AA
11A21 KIRAT RAI VOWEL SIGN I
11A22 KIRAT RAI VOWEL SIGN U
11A23 KIRAT RAI VOWEL SIGN E
11A24 KIRAT RAI VOWEL SIGN AI
11A25 KIRAT RAI VOWEL SIGN O
11A26 KIRAT RAI VOWEL SIGN AU

Various signs

11A27 KIRAT RAI SIGN VIRAMA
11A28 KIRAT RAI SIGN CANDRABINDU

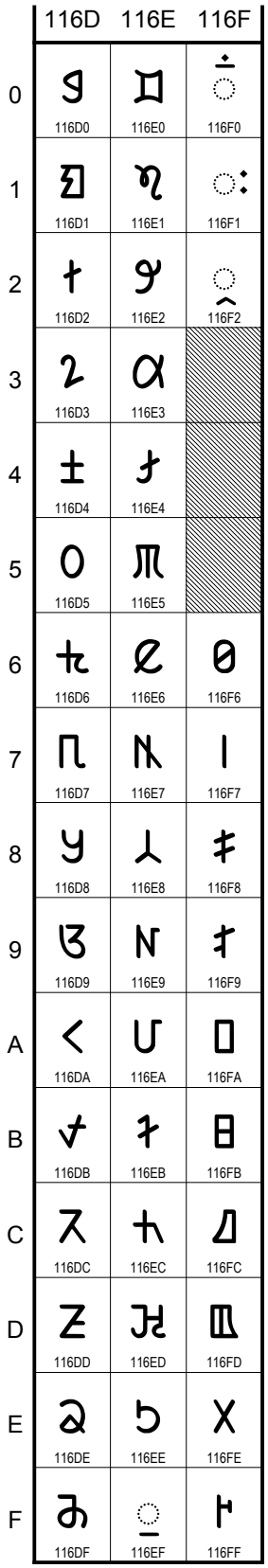
Digits

11A30 KIRAT RAI DIGIT ZERO
11A31 KIRAT RAI DIGIT ONE
11A32 KIRAT RAI DIGIT TWO
11A33 KIRAT RAI DIGIT THREE
11A34 KIRAT RAI DIGIT FOUR
11A35 KIRAT RAI DIGIT FIVE
11A36 KIRAT RAI DIGIT SIX
11A37 KIRAT RAI DIGIT SEVEN
11A38 KIRAT RAI DIGIT EIGHT
11A39 KIRAT RAI DIGIT NINE

Figure 1: Proposed code chart and names list for the Kirat Rai script.

A2.4 Jenticha N4028 and Tikamuli N3963

A2.4.1 Jenticha N4028



Vowels

116D0 JENTICHA LETTER A
116D1 JENTICHA LETTER AA
116D2 JENTICHA LETTER I
116D3 JENTICHA LETTER U
116D4 JENTICHA LETTER E
116D5 JENTICHA LETTER O

Consonants

116D6 JENTICHA LETTER KA
116D7 JENTICHA LETTER KHA
116D8 JENTICHA LETTER GA
116D9 JENTICHA LETTER NGA
116DA JENTICHA LETTER CA
116DB JENTICHA LETTER JA
116DC JENTICHA LETTER TTA
116DD JENTICHA LETTER TTHA
116DE JENTICHA LETTER TA
116DF JENTICHA LETTER THA
116E0 JENTICHA LETTER DA
116E1 JENTICHA LETTER NA
116E2 JENTICHA LETTER PA
116E3 JENTICHA LETTER PHA
116E4 JENTICHA LETTER BA
116E5 JENTICHA LETTER MA
116E6 JENTICHA LETTER YA
116E7 JENTICHA LETTER RA
116E8 JENTICHA LETTER LA
116E9 JENTICHA LETTER VA
116EA JENTICHA LETTER SHA
116EB JENTICHA LETTER SA
116EC JENTICHA LETTER HA
116ED JENTICHA LETTER HHA
116EE JENTICHA LETTER GLOTTAL STOP

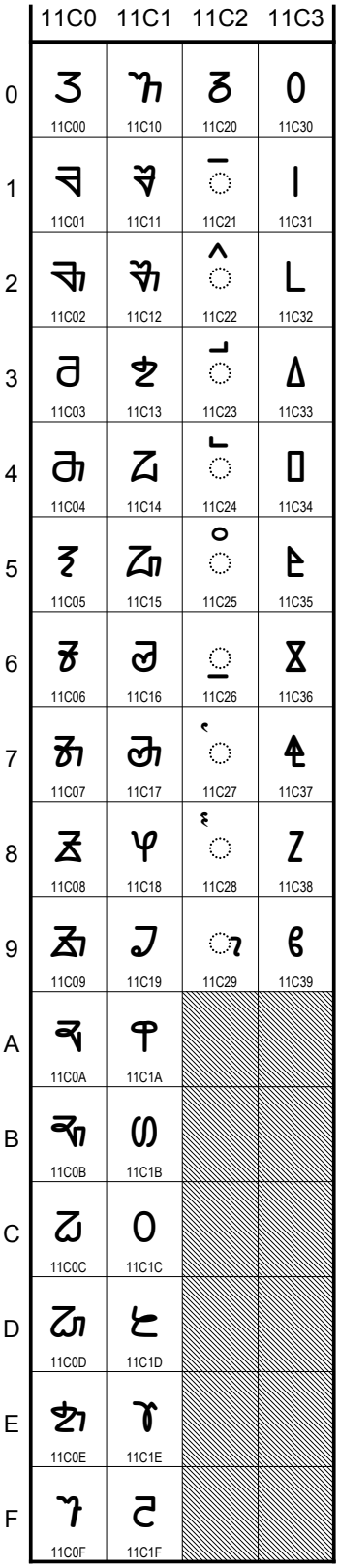
Various signs

116EF JENTICHA SIGN VIRAMA
= sangmilu, halant
116F0 JENTICHA SIGN CANDRABINDU
= taslathenk
116F1 JENTICHA VOWEL LENGTH MARK
= gyosh, dirghata
116F2 JENTICHA TRILL MARK
= sangrums

Digits

116F6 JENTICHA DIGIT ZERO
116F7 JENTICHA DIGIT ONE
116F8 JENTICHA DIGIT TWO
116F9 JENTICHA DIGIT THREE
116FA JENTICHA DIGIT FOUR
116FB JENTICHA DIGIT FIVE
116FC JENTICHA DIGIT SIX
116FD JENTICHA DIGIT SEVEN
116FE JENTICHA DIGIT EIGHT
116FF JENTICHA DIGIT NINE

A2.4.2 Tikamuli N3963



Vowel

11C00 𑐀 TIKAMULI LETTER A

Consonants

11C01 𑐁 TIKAMULI LETTER KA
11C02 𑐂 TIKAMULI LETTER KHA
11C03 𑐃 TIKAMULI LETTER GA
11C04 𑐄 TIKAMULI LETTER GHA
11C05 𑐅 TIKAMULI LETTER NGA
11C06 𑐆 TIKAMULI LETTER CA
11C07 𑐇 TIKAMULI LETTER CHA
11C08 𑐈 TIKAMULI LETTER JA
11C09 𑐉 TIKAMULI LETTER JHA
11C0A 𑐊 TIKAMULI LETTER TTA
11C0B 𑐋 TIKAMULI LETTER TTHA
11C0C 𑐌 TIKAMULI LETTER DDA
11C0D 𑐍 TIKAMULI LETTER DDHA
11C0E 𑐎 TIKAMULI LETTER NNA
11C0F 𑐏 TIKAMULI LETTER TA
11C10 𑐐 TIKAMULI LETTER THA
11C11 𑐑 TIKAMULI LETTER DA
11C12 𑐒 TIKAMULI LETTER DHA
11C13 𑐓 TIKAMULI LETTER NA
11C14 𑐔 TIKAMULI LETTER PA
11C15 𑐕 TIKAMULI LETTER PHA
11C16 𑐖 TIKAMULI LETTER BA
11C17 𑐗 TIKAMULI LETTER BHA
11C18 𑐘 TIKAMULI LETTER MA
11C19 𑐙 TIKAMULI LETTER YA
11C1A 𑐚 TIKAMULI LETTER RA
11C1B 𑐛 TIKAMULI LETTER LA
11C1C 𑐜 TIKAMULI LETTER VA
11C1D 𑐝 TIKAMULI LETTER SSA
11C1E 𑐞 TIKAMULI LETTER SA
11C1F 𑐟 TIKAMULI LETTER HA
11C20 𑐠 TIKAMULI LETTER GLOTTAL STOP

Dependent vowel signs

11C21 ◌𑐵 TIKAMULI VOWEL SIGN AA
11C22 ◌𑐶 TIKAMULI VOWEL SIGN I
11C23 ◌𑐷 TIKAMULI VOWEL SIGN U
11C24 ◌𑐸 TIKAMULI VOWEL SIGN E
11C25 ◌𑐹 TIKAMULI VOWEL SIGN O

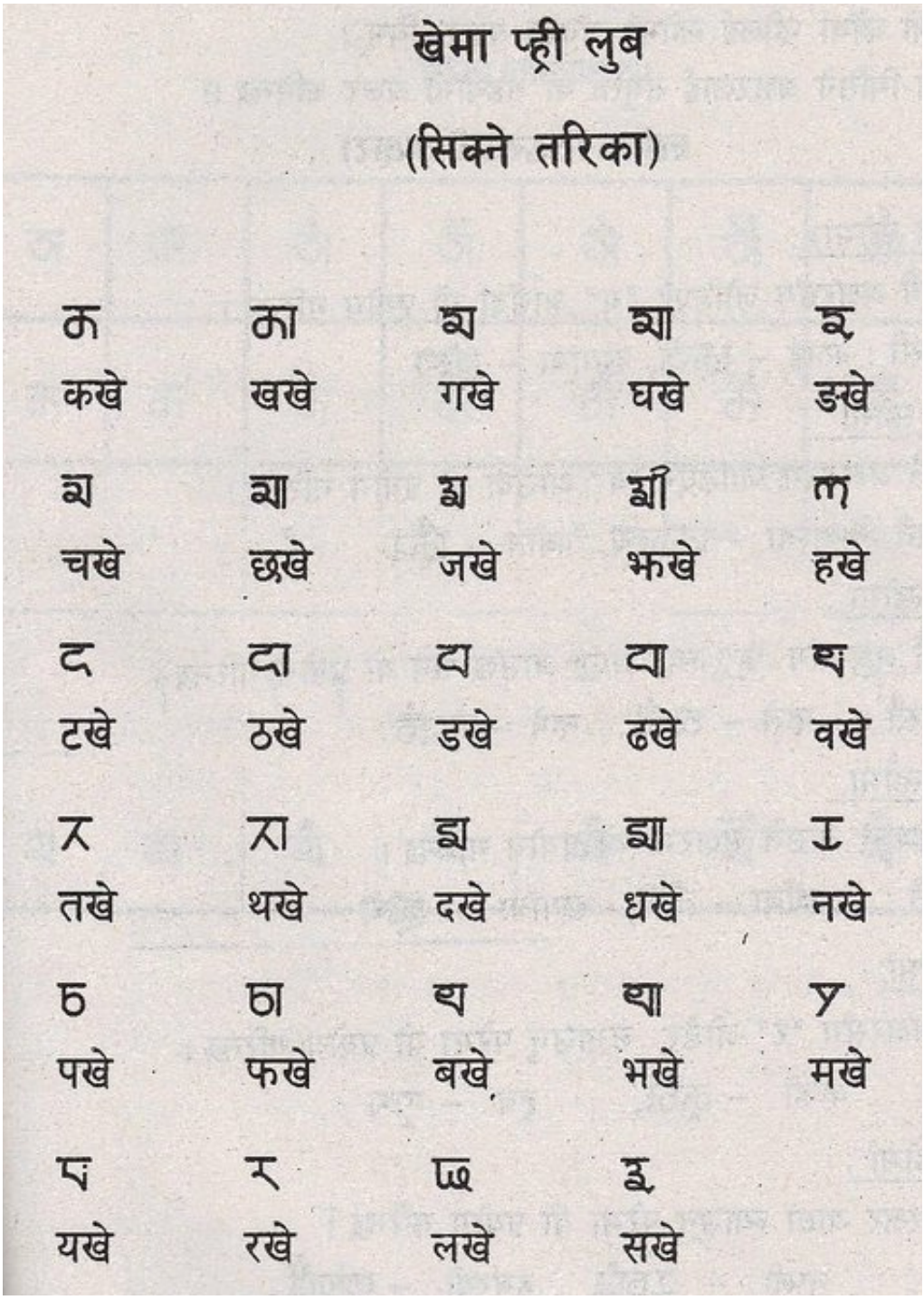
Various signs

11C26 ◌𑑂 TIKAMULI SIGN VIRAMA
11C27 ◌𑑃 TIKAMULI SIGN DIRGHA
11C28 ◌𑑄 TIKAMULI SIGN ATI DIRGHA
11C29 ◌𑑅 TIKAMULI SIGN TRILL

Digits

11C30 𑑐 TIKAMULI DIGIT ZERO
11C31 𑑑 TIKAMULI DIGIT ONE
11C32 𑑒 TIKAMULI DIGIT TWO
11C33 𑑓 TIKAMULI DIGIT THREE
11C34 𑑔 TIKAMULI DIGIT FOUR
11C35 𑑕 TIKAMULI DIGIT FIVE
11C36 𑑖 TIKAMULI DIGIT SIX
11C37 𑑗 TIKAMULI DIGIT SEVEN
11C38 𑑘 TIKAMULI DIGIT EIGHT
11C39 𑑙 TIKAMULI DIGIT NINE

A2.5 Khema N4019



खेमा प्ह्री लुब

(सिक्ने तरिका)

क	का	ग	गा	ङ
कखे	खखे	गखे	घखे	ङखे
च	चा	ज	झ	ञ
चखे	छखे	जखे	झखे	हखे
ट	टा	ड	ढा	ण
टखे	ठखे	डखे	ढखे	वखे
त	ता	द	धा	न
तखे	थखे	दखे	धखे	नखे
प	फा	ब	भा	म
पखे	फखे	बखे	भखे	मखे
य	र	ल	स	
यखे	रखे	लखे	सखे	

A2.6 Dhimal N4140

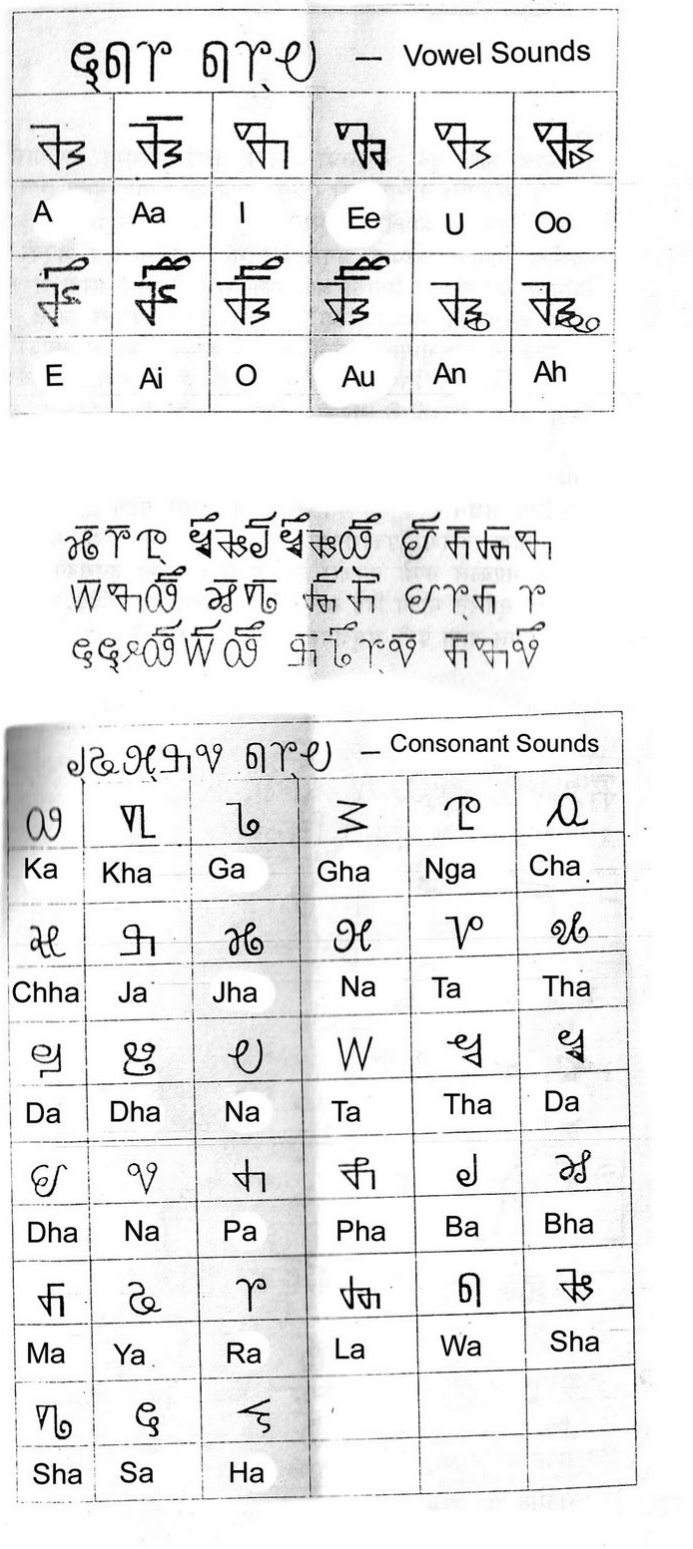


Figure 2: Vowel and consonant letters of Dhimal (from Dhimal 2008).

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html.
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Proposal to Encode Nepal Himalayish Scripts in ISO/IEC 10646*
2. Requester's name: *Patrick A V Hall*
3. Requester type (Member body/Liaison/Individual contribution): *individual contribution*
4. Submission date: *05 October 2012*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal:
   (or) More information will be provided later: ☐✔

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): ✔
      Proposed name of script: *Himalayish*
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block:
2. Number of characters in proposal: *98*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary B.1-Specialized (small collection) B.2-Specialized (large collection) *A*
   C-Major extinct D-Attested extinct E-Minor extinct
   F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols
4. Is a repertoire including character names provided? *NO*
   a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
   b. Are the character shapes attached in a legible form suitable for review? *YES*
5. Fonts related:
   a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
   either Sanir Karmacharya or Rabindra Pachhai both of whom have developed fonts
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
   to be provided in the full proposal in due course
6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *YES*
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *NO*
7. Special encoding issues:
   Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *NO*

8. Additional Information:
Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

[1] Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? ***YES***
   If YES explain ***Unifying proposals in N N3649, N3692, N3962, N3963, N4018, N4019, N4036, N4037, N4038, N4028, N4140, N4184, N4322***
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? ***YES***
   If YES, with whom? ***in person with: Nepal Lipi Guthi (centre for writing the language of the Newar) - Sharad Kasahchhen (Chair) Suwarn Vajracharya, Rabindra Pacchai***
   Authors of N4322: Devdass Manandhar, Bishnu Chitrakar, Sanir Karmacharya, Professor Tej Raj Kansakar
   By email with authority on Newar writing: Professor Kamal P Malla
   By email with encoding experts: Anshuman Pandey, Deborah Anderson, Michael Everson, Andrew West
   If YES, available relevant documents:
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? ***YES***
   Reference: ***Yadava, Yogendra P. (2003) Language Chapter 4. Population Monograph, volume 1, Kathmandu: Central Bureau of Statistics.2003 and UNFPA previous WG2 documents cited in C1 above***
4. The context of use for the proposed characters (type of use; common or rare) ***Common***
   Reference: ***previous WG2 documents cited in C1 above***
5. Are the proposed characters in current use by the user community? ***YES***
   If YES, where? Reference:
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? ***ideally***
   If YES, is a rationale provided?
   If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? ***YES***
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? ***NO***
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? ***some***
   If YES, is a rationale for its inclusion provided? ***indirectly***
   If YES, reference: ***N4038 and N4184***
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? ***NO***
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? ***YES***
    If YES, is a rationale for such use provided? ***indirectly***
    If YES, reference: ***all documents cited in C1 above***
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? ***YES***
    If YES, reference: ***all documents cited in C1 above***
12. Does the proposal contain characters with any special properties such as control function or similar semantics? ***NO***
    If YES, describe in detail (include attachment if necessary)
13. Does the proposal contain any Ideographic compatibility characters? ***NO***
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: