

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

Doc Type: Working Group Document**Title: Revised proposal for encoding the Warang Citi script in the SMP of the UCS****Source: UC Berkeley Script Encoding Initiative (Universal Scripts Project)****Author: Michael Everson****Status: Liaison Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Replaces: N1958 (1999-01-29), N3411 (2008-04-08), N3668 (2009-08-05)****Date: 2011-12-31**

1. Introduction. The Warang Citi script is used to write the Ho language. Ho is a North Munda language, which family, together with the Mon-Khmer languages, makes up Austro-Asiatic. Warang Citi was devised by charismatic community leader Lako Bodra as part of a comprehensive cultural program, and was offered as an improvement over scripts used by Christian missionary linguists. Ho people live in the Indian states of Orissa and Jharkand. In Jharkand, they are found in Ranchi, Chaibasa and Jamshedpur, and in villages like Pardsa, Jaldhar, Tekasi, Tilupada, Baduri, Purtydigua, Roladih, Tupu Dana, Jetia, Dumbisai, Harira, Gitilpi, Karlajuri, Sarliburu, Narsana, Gidibas Kokcho, Lupungutu, Pandaveer, Jhinkapani, Kondwa. According to the SIL *Ethnologue*, there are 1,026,000 speakers of Ho.

Today, the Ho community can be characterized as still a primarily oral community, with an emergent literary tradition. Many Ho do not write their language in any form. Among those who do, there are different camps. One camp that promotes the use of Devanagari for Ho; another camp knows how to use Warang Citi, and expresses a strong sentiment, even a kind of reverence towards it. And there are some people who seem to fall into both camps. There are also people who use Latin letters to write Ho on an ad-hoc basis and for things like e-mail or SMS. The use of Warang Citi seems to be stronger in the eastern part of the diaspora, where books that use it have been printed. (See samples below.)

2. Structure. Warang Citi is a casing script, written from left to right. Unlike many other Indic scripts, it has features of an alphabet: vowels are written as full letters, with no vowel-modifiers. However, consonants may have an inherent vowel; it typically is pronounced [a] or [ɔ], and less often [ɛ], and never occurs in final position. Because the inherent vowels do not appear explicitly, there can be ambiguity in the reading of certain words. The script has no regular system of conjuncts nor an explicit virama. In the presentation here, the inherent vowel is transcribed *Aa* and the explicit vowel $\Gamma\mathcal{r}$ is transcribed $\mathring{A}\mathring{a}$.

The letter \mathcal{U} ω *HIYO* is used after a vowel to indicate lengthening or aspiration: $\Gamma\mathcal{U}$ $\mathcal{r}\omega$ $\mathring{A}H$ $\mathring{a}h$ (\bar{A} \bar{a}). Long *i* and *u* are represented by the individual letters $\mathcal{E}\mathcal{r}$ \bar{I} \bar{i} and $\mathcal{U}\mathcal{r}$ \bar{U} \bar{u} ; these do not appear to be equivalent to $\mathcal{F}\mathcal{U}$ $\mathcal{r}\omega$ *IH ih* or $\mathcal{L}\mathcal{U}$ ω *UH uh*. The letters $\mathcal{E}\mathcal{r}$ \bar{I} \bar{i} and $\mathcal{U}\mathcal{r}$ \bar{U} \bar{u} may also, in certain cases, be pronounced as short vowels or diphthongs ([ai] for long *i* or [au] for long *u*). The details of how (or if) length is handled orthographically in Warang Citi is unclear at this point.

As in the Brahmi-derived scripts, Warang Citi includes the OM sign \mathcal{O} , which occasionally appears in words, such as $\mathcal{V}\mathcal{O}\mathcal{E}$ *šoms*, representing Hindi *śvāms* ‘breathing’. The script does not include a diacritic mark for anusvara as in Devanagari, but rather has a separate grapheme, \mathcal{V} *NGAA*.

The script has four so-called “ligatures” ʎ YA (V + V M + M), ʎ YO (ʎ + ʎ A + A), ʎ II (ʎ I + ʎ I), and ʎ UU (ʎ U + ʎ U). None of these “ligatures” are decomposable, and so are encoded separately. The digraph hb occurs occasionally and is used for [b] or [w].

3. Digits and numbers. Warang Citi now makes use of a zero, though this may not have been the case previously; no zero is given in Pinnow 1972 or Zide 1996, but more recent materials use zero, as for instance in the date ʎW-ʎ0-ʎ00M ‘15-10-2003’ (see Figure 6). Note that the higher digits may have fallen into disuse. In the date, for example, 118EA WARANG CITI NUMBER TEN is not used, but rather 118E1 WARANG CITI DIGIT ONE and 118E0 WARANG CITI DIGIT ZERO. No symbols for 100 or 1000 are known, though Pinnow postulated that they may exist.

4. Sorting. Collation order is as in the code chart, with casing pairs interfiling as in Latin. An exception here is U+118FF WARANG CITI OM, which does not file like ZV zV OM om, but rather comes at the beginning of the alphabet, although it is placed last in the code chart so that the alphabet itself can have a constant offset. The order is as given in Figure 5, which is considered authoritative. In the code chart all the capital letters are given, then all the small letters, as in ASCII, Armenian, Cyrillic, Deseret, Glagolitic, and Greek.

ϕ om < V M <<< v m < ʎ A <<< ʎ a < ʎ I <<< ʎ i < ʎ U <<< ʎ u < ʎ P <<< ʎ p < ʎ Y <<< ʎ y < ʎ Ī <<< ʎ ī < ʎ Ū <<< ʎ ū < ʎ E <<< ʎ e < ʎ O <<< ʎ o < ʎ Ñ <<< ʎ ñ < ʎ G <<< ʎ g < ʎ K <<< ʎ k < ʎ Ñ̃ <<< ʎ ñ̃ < ʎ J <<< ʎ j < ʎ C <<< ʎ c < ʎ N <<< ʎ n < ʎ D <<< ʎ d < ʎ T <<< ʎ t < ʎ Ñ <<< ʎ ñ < ʎ D̃ <<< ʎ d̃ < ʎ Õ <<< ʎ õ < ʎ M <<< ʎ m < ʎ B <<< ʎ b < ʎ P <<< ʎ p < ʎ H <<< ʎ h < ʎ L <<< ʎ l < ʎ R <<< ʎ r < ʎ R̃ <<< ʎ r̃ < ʎ S <<< ʎ s < ʎ V <<< ʎ v

5. Character names. The name of the script is ʎoʎʎ ʎʎʎ Hbârη citi ‘undying writing’, where ʎo hb is a digraph representing /w/. The name *Varang Kshiti* found in some literature is a sanskritized form. Letter names are traditional. Pinnow 1972 gives the following transcriptions. After those are given the recommended spellings for English and French. The name VIYO for the last letter ʎs [v] has been constructed; it is an original ligature of ʎs A and ʎo HIYO.

| | | | | | | | | | | | |
|------------------|------|------|------------------|-----|------------------|-----|------|------|------|------|------|
| ṅāʎ | NGAA | NGÁʎ | ʎεʎ | E | ʎEʎ | ʎεř | ENN | ʎENN | puʎ | PU | PUʎ |
| ʎa | A | ʎAʎ | ʎoʎ | O | ʎOʎ | ʎođ | ODD | ʎODD | hiyo | HIYO | HIYO |
| wüʎ ⁱ | WI | Ūʎ | ʎang | ANG | ʎANG | ʎεʎ | TTE | TTEʎ | hōlo | HOLO | HOLO |
| ʎuʎ ^u | YU | YUʎ | gaʎ | GA | GAʎ | nun | NUNG | NUNG | hōr | HORR | HORR |
| ʎyaʎ | YA | YAʎ | koʎ | KO | KOʎ | daʎ | DA | DAʎ | har | HAR | HAR |
| ʎyoʎ | YO | YOʎ | ʎeñ | ENY | ʎEÑ | ʎat | AT | ʎAT | śu: | SSUU | ĀŪ |
| ʎi: | II | ʎĪ | yüʎ ^j | YUJ | YŪʎ ^J | ʎam | AM | ʎAM | sī: | SII | SĪ |
| ʎu: | UU | ʎŪ | ʎüc | UC | ʎÜC | bu | BU | BU | viyo | VIYO | VIYO |

6. Punctuation. Pinnow states explicitly that European punctuation is used as in English. The following characters have been observed: . , : ... ‘ ’ “ ” - — ? and !. In some fonts the dots tend to be drawn rather high—but this is still FULL STOP, not MIDDLE DOT. The apostrophe ’ (RIGHT SINGLE QUOTATION MARK) is used very frequently in words ending in -ʎ -ʎ, as in ʎʎʎʎ ʎ rāʎāʎ.

7. Conjunct forms. A small number of conjunct forms are known, most of these represent doubled consonants. Pinnow calls them “ligatures” and says:

Sonstige Ligaturen kommen in den älteren Büchern gar nicht vor; in den späteren warden gelegentlich Ligaturen gebraucht, besonders zur Umschrift von Hindi-Wörtern. Die Ligaturen bestehen einfach darin, daß des jeweils zweite der ohne durch Vokal getrennten Konsonantenzeichen über das erste gesetzt wird, z.B. b(a)rⁿ, eigentlich brⁿ... Da die Ligaturen nur selten vorkommen, sind Druck und Schrift des Alphabets relativ einfach.

‘Other ligatures do not appear at all in the older books; in later ones ligatures were occasionally used, especially for the transliteration of Hindi words. The ligatures consist of the simple positioning of the second of two consonants which are not separated by a vowel on top of the first, e.g. *b(ar)*’, actually *br*... Since the ligatures occur only rarely, printing and writing the alphabet are relatively simple.’

Often the upper letter in the conjunct form is drawn upside down. It appears that this is the more common method of drawing a conjunct; where $\text{𑄚} rra$ contrasts with $\text{𑄛} rra$ it may be a result of an error in typography, or it may be an instance of free variation. When capitalized, only the lower letter takes a capital shape, and is reduced in size. In the table of contents of the book *Bha butu woṅga buru* (1984), the poem “*Bha Āṭeḍā kana Diṣuma Bano*” is written $\text{𑄚} \text{𑄛} \text{𑄜} \text{𑄝} \text{𑄞} \text{𑄟} \text{𑄠} \text{𑄡}$, but on p. 49 it is written with a conjunct in the first word, as $\text{𑄚} \text{𑄛} \text{𑄜} \text{𑄝} \text{𑄞} \text{𑄟} \text{𑄠} \text{𑄡}$ “*BHA ĀṬEḌĀ KANA DIṢUMA BANO*”. *Bha butu woṅga buru* shows the following conjunct forms:

| | | | |
|--------------------------------|-------------------------------------------------------------------------------------------------------|--------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $\text{𑄚} \text{𑄛}$ <i>Ām</i> | $\text{𑄚} \text{𑄛} \text{𑄜}$ <i>Āmde</i> | $\text{𑄛} \text{𑄜}$ <i>Tsa</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>Tsiṅa</i> |
| $\text{𑄛} \text{𑄜}$ <i>cca</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>Ūccuḍa</i> alongside $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>Ūccuḍa</i> | $\text{𑄛} \text{𑄜}$ <i>Bha</i> | $\text{𑄛} \text{𑄜}$ <i>Bha</i> and $\text{𑄛} \text{𑄜}$ <i>bha</i> alongside $\text{𑄚} \text{𑄛}$ <i>Bha~Baha</i> |
| $\text{𑄛} \text{𑄜}$ <i>ṭṭa</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>Ūṭṭuṭ</i> ; $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>āṭṭala</i> | $\text{𑄛} \text{𑄜}$ <i>ppa</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>juppuāḥa</i> (<i>juppuāḥa</i>) |
| $\text{𑄛} \text{𑄜}$ <i>ṭra</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>ūḍarāṣatrā</i> | $\text{𑄛} \text{𑄜}$ <i>rra</i> | $\text{𑄛} \text{𑄜}$ <i>purra</i> ; $\text{𑄛} \text{𑄜}$ <i>porra</i> ; $\text{𑄛} \text{𑄜}$ <i>Torra</i> , $\text{𑄛} \text{𑄜}$ <i>torra</i> ; $\text{𑄛} \text{𑄜}$ <i>haserra</i> |
| $\text{𑄛} \text{𑄜}$ <i>nda</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>Sippinduri</i> alongside $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>Sindurii</i> | $\text{𑄛} \text{𑄜}$ <i>rra</i> | $\text{𑄛} \text{𑄜}$ <i>parra</i> ; $\text{𑄛} \text{𑄜}$ <i>hirra</i> |
| $\text{𑄛} \text{𑄜}$ <i>nta</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>Ānanta</i> | $\text{𑄛} \text{𑄜}$ <i>lla</i> | $\text{𑄛} \text{𑄜}$ <i>sulla</i> ; |
| $\text{𑄛} \text{𑄜}$ <i>tta</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>Paruṣottama</i> | $\text{𑄛} \text{𑄜}$ <i>lla</i> | $\text{𑄛} \text{𑄜}$ <i>olla</i> alongside $\text{𑄛} \text{𑄜}$ <i>olla</i> |
| $\text{𑄛} \text{𑄜}$ <i>tra</i> | $\text{𑄛} \text{𑄜} \text{𑄝}$ <i>triyu</i> | | |

The choice of a conjunct form does not appear to be predictable. There are two possible encoding approaches.

One possibility is to encode a character (for instance at U+118FE) 𑄛 WARANG CITI SIGN VIRAMA, so that users can specify the conjunct form they prefer to have. Scripts like Chakma and Meetei Mayek have similar characters. This character would no visible appearance; where a font does not have the correct glyph, the dotted box 𑄛 or some other representation could be given to the user to indicate this.

Another option is the use of U+200D ZERO-WIDTH JOINER, so that users can specify the discretionary ligatures they prefer to have.

Because the forms are not “typographic” ligatures, but structural conjunct forms, the proposal author prefers the VIRAMA character approach over use of ZWJ. Since users of Warang Citi are likely to be familiar with Devanagari and Oriya scripts, it may be more prudent sense to use the familiar VIRAMA-model, since the “discretionary ligatures” are in fact conjuncts complete with a suppressed inherent vowel.

8. Linebreaking. Letters and digits behave as in Latin and other alphabetic scripts. U+118FF 𑄛 WARANG CITI OM behaves like a letter.

9. Unicode Character Properties.

- 118A0;WARANG CITI CAPITAL LETTER NGAA;Lu;0;L;;;;;N;;;118C0;
- 118A1;WARANG CITI CAPITAL LETTER A;Lu;0;L;;;;;N;;;118C1;
- 118A2;WARANG CITI CAPITAL LETTER WI;Lu;0;L;;;;;N;;;118C2;
- 118A3;WARANG CITI CAPITAL LETTER YU;Lu;0;L;;;;;N;;;118C3;
- 118A4;WARANG CITI CAPITAL LETTER YA;Lu;0;L;;;;;N;;;118C4;
- 118A5;WARANG CITI CAPITAL LETTER YO;Lu;0;L;;;;;N;;;118C5;
- 118A6;WARANG CITI CAPITAL LETTER II;Lu;0;L;;;;;N;;;118C6;
- 118A7;WARANG CITI CAPITAL LETTER UU;Lu;0;L;;;;;N;;;118C7;
- 118A8;WARANG CITI CAPITAL LETTER E;Lu;0;L;;;;;N;;;118C8;
- 118A9;WARANG CITI CAPITAL LETTER O;Lu;0;L;;;;;N;;;118C9;
- 118AA;WARANG CITI CAPITAL LETTER ANG;Lu;0;L;;;;;N;;;118CA;
- 118AB;WARANG CITI CAPITAL LETTER GA;Lu;0;L;;;;;N;;;118CB;
- 118AC;WARANG CITI CAPITAL LETTER KO;Lu;0;L;;;;;N;;;118CC;

118AD;WARANG CITI CAPITAL LETTER ENY;Lu;0;L; ; ; ; ;N; ; ; ;118CD;
118AE;WARANG CITI CAPITAL LETTER YUJ;Lu;0;L; ; ; ; ;N; ; ; ;118CE;
118AF;WARANG CITI CAPITAL LETTER UC;Lu;0;L; ; ; ; ;N; ; ; ;118CF;
118B0;WARANG CITI CAPITAL LETTER ENN;Lu;0;L; ; ; ; ;N; ; ; ;118D0;
118B1;WARANG CITI CAPITAL LETTER ODD;Lu;0;L; ; ; ; ;N; ; ; ;118D1;
118B2;WARANG CITI CAPITAL LETTER TTE;Lu;0;L; ; ; ; ;N; ; ; ;118D2;
118B3;WARANG CITI CAPITAL LETTER NUNG;Lu;0;L; ; ; ; ;N; ; ; ;118D3;
118B4;WARANG CITI CAPITAL LETTER DA;Lu;0;L; ; ; ; ;N; ; ; ;118D4;
118B5;WARANG CITI CAPITAL LETTER AT;Lu;0;L; ; ; ; ;N; ; ; ;118D5;
118B6;WARANG CITI CAPITAL LETTER AM;Lu;0;L; ; ; ; ;N; ; ; ;118D6;
118B7;WARANG CITI CAPITAL LETTER BU;Lu;0;L; ; ; ; ;N; ; ; ;118D7;
118B8;WARANG CITI CAPITAL LETTER PU;Lu;0;L; ; ; ; ;N; ; ; ;118D8;
118B9;WARANG CITI CAPITAL LETTER HIYO;Lu;0;L; ; ; ; ;N; ; ; ;118D9;
118BA;WARANG CITI CAPITAL LETTER HOLO;Lu;0;L; ; ; ; ;N; ; ; ;118DA;
118BB;WARANG CITI CAPITAL LETTER HERR;Lu;0;L; ; ; ; ;N; ; ; ;118DB;
118BC;WARANG CITI CAPITAL LETTER HAR;Lu;0;L; ; ; ; ;N; ; ; ;118DC;
118BD;WARANG CITI CAPITAL LETTER SSUU;Lu;0;L; ; ; ; ;N; ; ; ;118DD;
118BE;WARANG CITI CAPITAL LETTER SII;Lu;0;L; ; ; ; ;N; ; ; ;118DE;
118BF;WARANG CITI CAPITAL LETTER VIYO;Lu;0;L; ; ; ; ;N; ; ; ;118DF;
118C0;WARANG CITI SMALL LETTER NGAA;Ll;0;L; ; ; ; ;N; ; ; ;118A0; ;118A0
118C1;WARANG CITI SMALL LETTER A;Ll;0;L; ; ; ; ;N; ; ; ;118A1; ;118A1
118C2;WARANG CITI SMALL LETTER WI;Ll;0;L; ; ; ; ;N; ; ; ;118A2; ;118A2
118C3;WARANG CITI SMALL LETTER YU;Ll;0;L; ; ; ; ;N; ; ; ;118A3; ;118A3
118C4;WARANG CITI SMALL LETTER YA;Ll;0;L; ; ; ; ;N; ; ; ;118A4; ;118A4
118C5;WARANG CITI SMALL LETTER YO;Ll;0;L; ; ; ; ;N; ; ; ;118A5; ;118A5
118C6;WARANG CITI SMALL LETTER II;Ll;0;L; ; ; ; ;N; ; ; ;118A6; ;118A6
118C7;WARANG CITI SMALL LETTER UU;Ll;0;L; ; ; ; ;N; ; ; ;118A7; ;118A7
118C8;WARANG CITI SMALL LETTER E;Ll;0;L; ; ; ; ;N; ; ; ;118A8; ;118A8
118C9;WARANG CITI SMALL LETTER O;Ll;0;L; ; ; ; ;N; ; ; ;118A9; ;118A9
118CA;WARANG CITI SMALL LETTER ANG;Ll;0;L; ; ; ; ;N; ; ; ;118AA; ;118AA
118CB;WARANG CITI SMALL LETTER GA;Ll;0;L; ; ; ; ;N; ; ; ;118AB; ;118AB
118CC;WARANG CITI SMALL LETTER KO;Ll;0;L; ; ; ; ;N; ; ; ;118AC; ;118AC
118CD;WARANG CITI SMALL LETTER ENY;Ll;0;L; ; ; ; ;N; ; ; ;118AD; ;118AD
118CE;WARANG CITI SMALL LETTER YUJ;Ll;0;L; ; ; ; ;N; ; ; ;118AE; ;118AE
118CF;WARANG CITI SMALL LETTER UC;Ll;0;L; ; ; ; ;N; ; ; ;118AF; ;118AF
118D0;WARANG CITI SMALL LETTER ENN;Ll;0;L; ; ; ; ;N; ; ; ;118B0; ;118B0
118D1;WARANG CITI SMALL LETTER ODD;Ll;0;L; ; ; ; ;N; ; ; ;118B1; ;118B1
118D2;WARANG CITI SMALL LETTER TTE;Ll;0;L; ; ; ; ;N; ; ; ;118B2; ;118B2
118D3;WARANG CITI SMALL LETTER NUNG;Ll;0;L; ; ; ; ;N; ; ; ;118B3; ;118B3
118D4;WARANG CITI SMALL LETTER DA;Ll;0;L; ; ; ; ;N; ; ; ;118B4; ;118B4
118D5;WARANG CITI SMALL LETTER AT;Ll;0;L; ; ; ; ;N; ; ; ;118B5; ;118B5
118D6;WARANG CITI SMALL LETTER AM;Ll;0;L; ; ; ; ;N; ; ; ;118B6; ;118B6
118D7;WARANG CITI SMALL LETTER BU;Ll;0;L; ; ; ; ;N; ; ; ;118B7; ;118B7
118D8;WARANG CITI SMALL LETTER PU;Ll;0;L; ; ; ; ;N; ; ; ;118B8; ;118B8
118D9;WARANG CITI SMALL LETTER HIYO;Ll;0;L; ; ; ; ;N; ; ; ;118B9; ;118B9
118DA;WARANG CITI SMALL LETTER HOLO;Ll;0;L; ; ; ; ;N; ; ; ;118BA; ;118BA
118DB;WARANG CITI SMALL LETTER HERR;Ll;0;L; ; ; ; ;N; ; ; ;118BB; ;118BB
118DC;WARANG CITI SMALL LETTER HAR;Ll;0;L; ; ; ; ;N; ; ; ;118BC; ;118BC
118DD;WARANG CITI SMALL LETTER SSUU;Ll;0;L; ; ; ; ;N; ; ; ;118BD; ;118BD
118DE;WARANG CITI SMALL LETTER SII;Ll;0;L; ; ; ; ;N; ; ; ;118BE; ;118BE
118DF;WARANG CITI SMALL LETTER VIYO;Ll;0;L; ; ; ; ;N; ; ; ;118BF; ;118BF
118E0;WARANG CITI DIGIT ZERO;Nd;0;L; ; ; ; ;0;0;0;N; ; ; ; ;
118E1;WARANG CITI DIGIT ONE;Nd;0;L; ; ; ; ;1;1;1;N; ; ; ; ;
118E2;WARANG CITI DIGIT TWO;Nd;0;L; ; ; ; ;2;2;2;N; ; ; ; ;
118E3;WARANG CITI DIGIT THREE;Nd;0;L; ; ; ; ;3;3;3;N; ; ; ; ;
118E4;WARANG CITI DIGIT FOUR;Nd;0;L; ; ; ; ;4;4;4;N; ; ; ; ;
118E5;WARANG CITI DIGIT FIVE;Nd;0;L; ; ; ; ;5;5;5;N; ; ; ; ;
118E6;WARANG CITI DIGIT SIX;Nd;0;L; ; ; ; ;6;6;6;N; ; ; ; ;
118E7;WARANG CITI DIGIT SEVEN;Nd;0;L; ; ; ; ;7;7;7;N; ; ; ; ;
118E8;WARANG CITI DIGIT EIGHT;Nd;0;L; ; ; ; ;8;8;8;N; ; ; ; ;
118E9;WARANG CITI DIGIT NINE;Nd;0;L; ; ; ; ;9;9;9;N; ; ; ; ;
118EA;WARANG CITI NUMBER TEN;No;0;L; ; ; ; ;10;N; ; ; ; ;
118EB;WARANG CITI NUMBER TWENTY;No;0;L; ; ; ; ;20;N; ; ; ; ;
118EC;WARANG CITI NUMBER THIRTY;No;0;L; ; ; ; ;30;N; ; ; ; ;
118ED;WARANG CITI NUMBER FORTY;No;0;L; ; ; ; ;40;N; ; ; ; ;
118EE;WARANG CITI NUMBER FIFTY;No;0;L; ; ; ; ;50;N; ; ; ; ;
118EF;WARANG CITI NUMBER SIXTY;No;0;L; ; ; ; ;60;N; ; ; ; ;
118F0;WARANG CITI NUMBER SEVENTY;No;0;L; ; ; ; ;70;N; ; ; ; ;
118F1;WARANG CITI NUMBER EIGHTY;No;0;L; ; ; ; ;80;N; ; ; ; ;
118F2;WARANG CITI NUMBER NINETY;No;0;L; ; ; ; ;90;N; ; ; ; ;
118FF;WARANG CITI OM;Lo;0;L; ; ; ; ;N; ; ; ; ;

10. Bibliography.

- Pinnow, Heinz-Jürgen. 1972. "Schrift und Sprache in den Werken Lako Bodras im Gebiet der Ho von Singbhum (Bihar)", in *Anthropos* 67:822–857.
Zide, Norman. 1996. "Scripts for Munda languages", in Peter T. Daniels and William Bright, eds. *The world's writing systems*. New York; Oxford: Oxford University Press. ISBN 0-19-507993-0
[s.n.]. 2003. 𑒧𑒻𑒟𑒱 (𑒧𑒻𑒟𑒱). [s.l.]: [s.n.]. (See Figure 25.)

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

| | 118A | 118B | 118C | 118D | 118E | 118F |
|---|------------|------------|------------|------------|------------|------------|
| 0 | V 118A0 | ∩ 118B0 | ∪ 118C0 | ∞ 118D0 | 0 118E0 | ∪ 118F0 |
| 1 | ∩ 118A1 | ∪ 118B1 | ∞ 118C1 | ∪ 118D1 | ∞ 118E1 | ∩ 118F1 |
| 2 | F 118A2 | L 118B2 | f 118C2 | ∪ 118D2 | ∞ 118E2 | ∩ 118F2 |
| 3 | L 118A3 | ∩ 118B3 | ∪ 118C3 | ∞ 118D3 | ∞ 118E3 | |
| 4 | Y 118A4 | U 118B4 | y 118C4 | ∞ 118D4 | ∞ 118E4 | |
| 5 | H 118A5 | O 118B5 | H 118C5 | 6 118D5 | Z 118E5 | |
| 6 | E 118A6 | ∩ 118B6 | 7 118C6 | 9 118D6 | W 118E6 | |
| 7 | ∩ 118A7 | ◇ 118B7 | ∪ 118C7 | 0 118D7 | ∪ 118E7 | |
| 8 | ∇ 118A8 | U 118B8 | ∩ 118C8 | ∪ 118D8 | ∩ 118E8 | |
| 9 | Z 118A9 | ∩ 118B9 | 2 118C9 | ∞ 118D9 | ∩ 118E9 | |
| A | ∪ 118AA | ∩ 118BA | ∩ 118CA | ∩ 118DA | ∩ 118EA | |
| B | ∩ 118AB | 5 118BB | ∩ 118CB | 5 118DB | ∩ 118EB | |
| C | ∩ 118AC | T 118BC | 9 118CC | ∩ 118DC | X 118EC | |
| D | ∩ 118AD | U 118BD | ∩ 118CD | ∩ 118DD | ∩ 118ED | |
| E | E 118AE | ∩ 118BE | ∩ 118CE | ∩ 118DE | ∩ 118EE | |
| F | ∩ 118AF | ∩ 118BF | ∩ 118CF | ∩ 118DF | ∩ 118EF | ∩ 118FF |

Uppercase vowels

| | | |
|-------|---|---------------------------------|
| 118A0 | ∨ | WARANG CITI CAPITAL LETTER NGAA |
| 118A1 | ┘ | WARANG CITI CAPITAL LETTER A |
| 118A2 | ƒ | WARANG CITI CAPITAL LETTER WI |
| 118A3 | ┘ | WARANG CITI CAPITAL LETTER YU |
| 118A4 | Y | WARANG CITI CAPITAL LETTER YA |
| 118A5 | ƒ | WARANG CITI CAPITAL LETTER YO |
| 118A6 | E | WARANG CITI CAPITAL LETTER II |
| 118A7 | ┘ | WARANG CITI CAPITAL LETTER UU |
| 118A8 | ∇ | WARANG CITI CAPITAL LETTER E |
| 118A9 | Z | WARANG CITI CAPITAL LETTER O |

Uppercase consonants

| | | |
|-------|---|---------------------------------|
| 118AA | ∂ | WARANG CITI CAPITAL LETTER ANG |
| 118AB | ┘ | WARANG CITI CAPITAL LETTER GA |
| 118AC | 𐄀 | WARANG CITI CAPITAL LETTER KO |
| 118AD | 𐄁 | WARANG CITI CAPITAL LETTER ENY |
| 118AE | E | WARANG CITI CAPITAL LETTER YUJ |
| 118AF | 𐄂 | WARANG CITI CAPITAL LETTER UC |
| 118B0 | 𐄃 | WARANG CITI CAPITAL LETTER ENN |
| 118B1 | ┘ | WARANG CITI CAPITAL LETTER ODD |
| 118B2 | ┘ | WARANG CITI CAPITAL LETTER TTE |
| 118B3 | ┘ | WARANG CITI CAPITAL LETTER NUNG |
| 118B4 | U | WARANG CITI CAPITAL LETTER DA |
| 118B5 | O | WARANG CITI CAPITAL LETTER AT |
| 118B6 | 𐄄 | WARANG CITI CAPITAL LETTER AM |
| 118B7 | ◊ | WARANG CITI CAPITAL LETTER BU |
| 118B8 | U | WARANG CITI CAPITAL LETTER PU |
| 118B9 | U | WARANG CITI CAPITAL LETTER HIYO |
| 118BA | ∂ | WARANG CITI CAPITAL LETTER HOLO |
| 118BB | 5 | WARANG CITI CAPITAL LETTER HORR |
| 118BC | T | WARANG CITI CAPITAL LETTER HAR |
| 118BD | U | WARANG CITI CAPITAL LETTER SSUU |
| 118BE | ┘ | WARANG CITI CAPITAL LETTER SII |
| 118BF | 5 | WARANG CITI CAPITAL LETTER VIYO |

Lowercase vowels

| | | |
|-------|---|-------------------------------|
| 118C0 | ∨ | WARANG CITI SMALL LETTER NGAA |
| 118C1 | ┘ | WARANG CITI SMALL LETTER A |
| 118C2 | ƒ | WARANG CITI SMALL LETTER WI |
| 118C3 | ┘ | WARANG CITI SMALL LETTER YU |
| 118C4 | Y | WARANG CITI SMALL LETTER YA |
| 118C5 | ƒ | WARANG CITI SMALL LETTER YO |
| 118C6 | 7 | WARANG CITI SMALL LETTER II |
| 118C7 | 7 | WARANG CITI SMALL LETTER UU |
| 118C8 | ∇ | WARANG CITI SMALL LETTER E |
| 118C9 | z | WARANG CITI SMALL LETTER O |

Lowercase consonants

| | | |
|-------|---|-------------------------------|
| 118CA | ∂ | WARANG CITI SMALL LETTER ANG |
| 118CB | ┘ | WARANG CITI SMALL LETTER GA |
| 118CC | 𐄀 | WARANG CITI SMALL LETTER KO |
| 118CD | 𐄁 | WARANG CITI SMALL LETTER ENY |
| 118CE | ε | WARANG CITI SMALL LETTER YUJ |
| 118CF | 𐄂 | WARANG CITI SMALL LETTER UC |
| 118D0 | 𐄃 | WARANG CITI SMALL LETTER ENN |
| 118D1 | ┘ | WARANG CITI SMALL LETTER ODD |
| 118D2 | ┘ | WARANG CITI SMALL LETTER TTE |
| 118D3 | ┘ | WARANG CITI SMALL LETTER NUNG |
| 118D4 | U | WARANG CITI SMALL LETTER DA |
| 118D5 | O | WARANG CITI SMALL LETTER AT |
| 118D6 | 𐄄 | WARANG CITI SMALL LETTER AM |
| 118D7 | ◊ | WARANG CITI SMALL LETTER BU |
| 118D8 | U | WARANG CITI SMALL LETTER PU |
| 118D9 | U | WARANG CITI SMALL LETTER HIYO |
| 118DA | ∂ | WARANG CITI SMALL LETTER HOLO |
| 118DB | 5 | WARANG CITI SMALL LETTER HORR |

| | | |
|-------|---|-------------------------------|
| 118DC | 8 | WARANG CITI SMALL LETTER HAR |
| 118DD | 8 | WARANG CITI SMALL LETTER SSUU |
| 118DE | 8 | WARANG CITI SMALL LETTER SII |
| 118DF | 8 | WARANG CITI SMALL LETTER VIYO |

Digits

| | | |
|-------|---|-------------------------|
| 118E0 | 0 | WARANG CITI DIGIT ZERO |
| 118E1 | 1 | WARANG CITI DIGIT ONE |
| 118E2 | 2 | WARANG CITI DIGIT TWO |
| 118E3 | 3 | WARANG CITI DIGIT THREE |
| 118E4 | 4 | WARANG CITI DIGIT FOUR |
| 118E5 | 5 | WARANG CITI DIGIT FIVE |
| 118E6 | 6 | WARANG CITI DIGIT SIX |
| 118E7 | 7 | WARANG CITI DIGIT SEVEN |
| 118E8 | 8 | WARANG CITI DIGIT EIGHT |
| 118E9 | 9 | WARANG CITI DIGIT NINE |

Numbers

| | | |
|-------|----|----------------------------|
| 118EA | 10 | WARANG CITI NUMBER TEN |
| 118EB | 20 | WARANG CITI NUMBER TWENTY |
| 118EC | 30 | WARANG CITI NUMBER THIRTY |
| 118ED | 40 | WARANG CITI NUMBER FORTY |
| 118EE | 50 | WARANG CITI NUMBER FIFTY |
| 118EF | 60 | WARANG CITI NUMBER SIXTY |
| 118F0 | 70 | WARANG CITI NUMBER SEVENTY |
| 118F1 | 80 | WARANG CITI NUMBER EIGHTY |
| 118F2 | 90 | WARANG CITI NUMBER NINETY |

Signs

| | | |
|-------|---|----------------|
| 118FF | 0 | WARANG CITI OM |
|-------|---|----------------|

Figures.

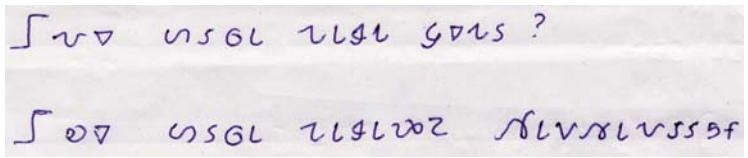



Figure 1. Sample of Warang Citi handwriting.

𑌒𑌓𑌔 𑌕𑌖𑌗 Hbar(a)η citi „Alphabet“

| Originalschr. Groß- Klein- Buchstabe | | Transliteration Devanāgarī Latein- Schrift | Bezeichnung der Grapheme Devanāgarī Phonetisch | | Phonement- sprechung |
|------------------------------------------------------------|---|--------------------------------------------------|---------------------------------------------------|-------|------------------------------|
| 1. 𑌒𑌓𑌔 𑌕𑌖𑌗 ip(a)n bor(o)η „(einfache) Vokale“ | | | | | |
| — | — | — | (a), (e), (o) | — | („inhärierend“) a , e , o |
| V | v | . | M m | इँ | ηāʔ m , ŋ |
| ┌ | ┐ | अ, आ | A a | अँ | ʔa a |
| f | f | इ | I i | विःइँ | wiʔi, wiʔi i |
| L | l | उ | U u | युःउँ | yuʔu u |
| 2. 𑌒𑌓𑌔 𑌕𑌖𑌗 sell ip(a)n bor(o)η „(gemischte) Vokale“ | | | | | |
| ∇ | ▽ | ए | E e | एँ | ʔεʔ e |
| Z | z | औ | O o | औःअँ | ʔɔʔ o |
| 3. 𑌒𑌓𑌔 𑌕𑌖𑌗 doʔi bor(o)η „Ligaturen“ | | | | | |
| Y | y | : | ʔ ʔ | यः | 'yaʔ ʔ |
| H | h | य | Y y | योँ | 'yoʔ y , e , i |
| E | e | इ, ई, ऐ | I ī | ईँ | ʔi: i , ai , ii |
| h | h | उ, ऊ, औ | ū | ऊँ | ʔu: u , au , uu |
| 4. 𑌒𑌓𑌔 𑌕𑌖𑌗 beʔn bor(o)η „Konsonanten“ | | | | | |
| ʒ | ʒ | ड. | Ń ɳ | ओइँ | ʔaŋ ŋ |
| 𑌒 | 𑌓 | ग | G g | गँ | gaʔ g |
| 𑌔 | 𑌕 | क | K k | कः | kaʔ k |
| 𑌖 | 𑌗 | ज | Ń ñ | विः | ʔεñ ñ |

Figure 2a. Description of the alphabet from Pinnow 1972.

| Originalschr. | | Transliteration | | Bezeichnung der Grapheme | | Phonement- sprechung |
|--------------------|---------------------|-----------------|--------------------|--------------------------|------------|-------------------------|
| Groß- Buchstabe | Klein- Buchstabe | Devanāgarī | Latein- Schrift | Devanāgarī | Phonetisch | |
| ए | ε | ज | J j | बिजू | yü'j | j |
| प | ϕ | ब | C c | बिबू | ʔüc | c |
| न | ~ | न | N n | बैरन | ʔεɳ̄ | n |
| ड | ɽ | ड | D d | ओडू | ʔɔ'd | d |
| ट | ɽ | ट | T t | टैःए | tεʔ | t |
| न | ~ | न | N n | नुडू | nun | n |
| ड | ~ | द | D d | दःऽ | daʔ | d |
| ओ | 6 | त | T t | औत | ʔat | t |
| म | १ | म | M m | ऽम | ʔam | m |
| ब | o | ब | B b | बू | bu | b |
| प | ~ | प | P p | पुःडू | puʔ | p |
| ह | 5 | ह | H h | ह्यी | hiyo | h , Vokaldehnung |
| ल | o | ल | L l | हली | holo | l |
| र | 5 | ड़ | R r | हड़ | hor | r |
| र | ४ | र | R r | हर | har | r |
| ष | ४ | ष, श | S s | षू | śu: | ś , s |
| स | ~ | स | S s | सी | sī: | s |
| हब | हब | हब | HB hb | — | — | b , w |

5. Zeichen für die heilige Silbe  OM

6. Zahlzeichen

| | | | | | | | | |
|------|----|----|------|----|----|----|----|----|
| १ | २ | ३ | ४ | ५ | ६ | ७ | ८ | ९ |
| १० | २० | ३० | ४० | ५० | ६० | ७० | ८० | ९० |
| aber | ११ | २३ | usw. | ५७ | | | | |

Figure 2b. Description of the alphabet from Pinnow 1972.

1ZUJTF



ε...γ...δ εφθ ζητις !
 Τφχθς Λιλιτιχ θζςς ! ιφχθς θεχςς θζςς !!
 ιφχθς ιοιςχβ υςςυ!!! ιφχθς ιηζεφθ ζητις!!!!
 θυθυ ον οια Γ θ Δ
 ♡ ♡ ϑ Ν υν ον.

ηζθ ηιθυ ηιθυ—ειχθθ ειχθθ υιχ υθθ
 εζ υυυζ ζιζθ εζθ ϑυ ιηζιυ
 υιιθθ ουυθθ εςςθθ ριιθθ ριθθ χζςιι
 οχιις υυυζ ιθυ ηιθυ εςιιυ.
 Τζιζζ υςςυ υιιιχ ρειχ θςειχ ιυ Δ'
 ηυθζιυ θυθζιυ θυθυ ιηζιυ εςειυ
 ιθ, θςςυ θιιιχ θιιιχ ιηζιχ θςειχ
 υςυιιι, υυυιθθυ θςθ εθχς ηζ ♡ ιοθιις.
 Κιιις χιιιθθυ υφχις θζθθ θζθθ
 ♡ ιοιςχβ εουχιυ Δ' εουυς εουυθςυ ειιςς

υυιι υζιι

Seite 1 aus *Shar Hora* „Der Weg zum Himmel“ (Transliteration 6.1.4)

Figure 3. Sample text from Pinnow 1972.

᳚

| | | | |
|---|--------|-----|------------|
| ᳚ | ᳚ | बो | हं ᳚᳚ |
| ᳚ | ᳚ | हो | ल ᳚᳚ |
| ᳚ | ᳚ | हड़ | ह ᳚᳚ |
| ᳚ | ᳚ | हर | र ᳚᳚ |
| ᳚ | ᳚ | ष | ष ᳚᳚᳚᳚᳚ ᳚᳚ |
| ᳚ | ᳚ | सी | स ᳚᳚᳚ ᳚᳚ |
| | ᳚᳚ ᳚᳚᳚ | | ᳚᳚᳚᳚᳚ ᳚᳚᳚᳚ |
| | अप पाट | | तृतीय पाठः |

᳚᳚᳚᳚᳚ ᳚᳚᳚᳚᳚᳚
ककर ककहारा

V— ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚,
अं, गं, कं, जं, चं, णं, डं, टं,
᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚,
नं, दं, तं, मं, वं, पं, हं, लं, ङं,
᳚᳚, ᳚᳚, ᳚᳚.
रं, षं, सं।

᳚— ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚,
गा, का, जा, चा, णा, डा, टा, ना,
᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚,
दा, ता, मा, वा, पा, हा, ला, डा, रा,
᳚᳚, ᳚᳚,
षा, सा।

᳚— ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚,
गि, कि, जि, चि, णि, डि, टि, नि, दि, ति,

Figure 4a. Sample text with syllabary from Pinnow 1972.

4

8

५f, ०f, १f, ७f, ७f, ५f, ४f, ४f, ५f, ५f.
मि, वि, पि, हि, लि, डि, रि, षि, सि।

L— १L, १L, ५L, ६L, १L, १L, १L, १L, १L, १L,
ग, कु, ञ, जु, चु, ण, ड, ट, न, द,
६L, १L, ०L, १L, ७L, ७L, ७L, ४L, ४L, ५L.
त, म, व, प, ह, ल, ड, रु, ष, सु।

V— १V, १V, ५V, ६V, १V, १V, १V, १V, १V, १V,
ग, के, जे, जे, चे, णे, डे, टे, ने, दे,
६V, १V, ०V, १V, ७V, ७V, ७V, ४V, ४V, ४V, ५V.
ते, मे, वे, पे, हे, ले, डे, रे, षे, से।


Z— १Z, १Z, ५Z, ६Z, १Z, १Z, १Z, १Z, १Z, १Z,
गो, को, जो, जो, चो, णो, डो, टो, नो, दो,
६Z, १Z, ०Z, १Z, ७Z, ७Z, ७Z, ४Z, ४Z, ४Z, ५Z.
तो, मो, वो, पो, हो, लो, डो, रो, षो, सो।

Y— १Y, १Y, ५Y, ६Y, १Y, १Y, १Y, १Y, १Y, १Y,
गः, कः, ञः, जः, चः, णः, डः, टः, नः,
५Y, ६Y, १Y, ०Y, १Y, ७Y, ७Y, ७Y, ४Y, ४Y,
दः, तः, मः, वः, पः, हः, लः, डः, रः,
४Y, ५Y.
षः, सः।

H— १H, १H, ५H, ६H, १H, १H, १H, १H, १H, १H,
गै, कै, जै, जै, चै, णै, डै, टै, नै,
५H, ६H, १H, ०H, १H, ७H, ७H, ७H, ४H, ४H,
दै, तै, मै, वै, पै, है, लै, डै, रै,
४H, ५H.
षै, सै।

Seite 4 aus *Ho H(a)y(a)m P(a)ham Puti*

Figure 4b. Sample text with syllabary from Pinnow 1972.

अध्याय 5  **पाठ**
अध्ययनार्थ कुञ्जी Key to read the book
एव ०२४३ - स्वर वर्ण - Vowels
पाठ १२ - अक्षरें - Alphabets

| बड़ा Capitals | छोटा Small | उच्चारण Pronunciation | हिन्दी Hindi | अंग्रेजी English |
|--------------------------------|---------------|--------------------------|-----------------|---------------------|
| V | v | ड् | - | n |
| J | j | ऽ, अ | ।, अ | a |
| F | f | विःइ | ि | i |
| L | l | युःउ | ु | u |
| Y | y | यःऽ | ः | ? |
| H | h | यो | य | ea |
| E | e | इ | ई,ी | ii |
| T | t | उ | उ,ू | uu |
| A | a | ए | , ए | e |
| N | n | ओ | ो, ओ | o |
| व्यंजन वर्ण - Consonant | | | | |
| ठ | th | ोड् | ड | ng |
| ग | g | ए | ए | g |
| क | k, ch | क् | क् | k, ch |
| ज | j | ज् | ज् | j |
| च | ch, c | च् | च् | ch, c |
| ण | N | ण | ण | N |
| ड | d | ड | ड | d |
| त | t | ट | ट | t |
| न | n | नुड | न | n |
| द | d | दःऽ | द | d |
| त | t | त् | त | t |
| m | m | ऽह | म् | m |
| b | b | बू | ब् | b |
| p | p | पुःउ | प् | p |
| h | h | ह्यो | ह | h |
| r | r | हलौ | ल् | r |
| sh | sh | हड | ड् | sh |
| s | s | हऽ | र् | s |
| | | षू | ष् | |
| | | सी | स् | |

नोट :- ङ + उ के सम्मिश्रण से यह वर्ण बना है।

Figure 5. Alphabet chart from the book *शिशु हलाम*, 2003.

In the last line the letter ङ is given with the text “*ङ + उ ke sammishraṇa yah varna banā hai*” ‘from a ligature of ङ + उ this character is made’.



Warang Citi Alphabet and its History

-History of Warang Citi Alphabet-

Warang Citi is a script used in the Warang Citi region of Karnataka, India. It is a form of the Brahmic script family. The script is known for its unique characters and is used to write the Warang Citi language. The text in the image discusses the history and development of this script, mentioning its use in inscriptions and its relationship to other scripts like the Kannada script.

Warang Citi is a script used in the Warang Citi region of Karnataka, India. It is a form of the Brahmic script family. The script is known for its unique characters and is used to write the Warang Citi language. The text in the image discusses the history and development of this script, mentioning its use in inscriptions and its relationship to other scripts like the Kannada script.

Warang Citi is a script used in the Warang Citi region of Karnataka, India. It is a form of the Brahmic script family. The script is known for its unique characters and is used to write the Warang Citi language. The text in the image discusses the history and development of this script, mentioning its use in inscriptions and its relationship to other scripts like the Kannada script.

Warang Citi is a script used in the Warang Citi region of Karnataka, India. It is a form of the Brahmic script family. The script is known for its unique characters and is used to write the Warang Citi language. The text in the image discusses the history and development of this script, mentioning its use in inscriptions and its relationship to other scripts like the Kannada script.

Figure 7. Part of a newsletter in Warang Citi.

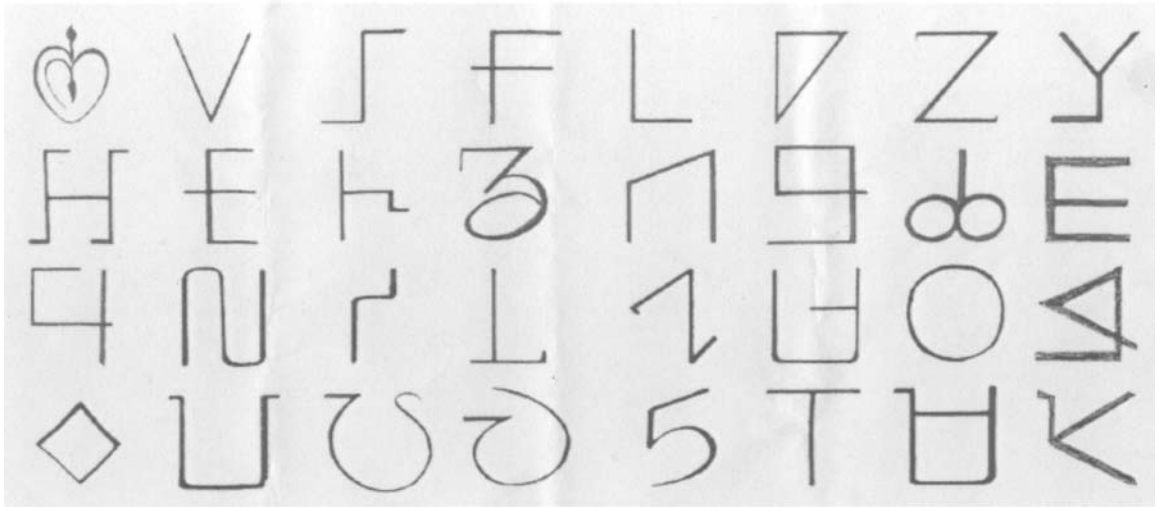


Figure 8. Warang Citi alphabet as written out by K. C. Naik.

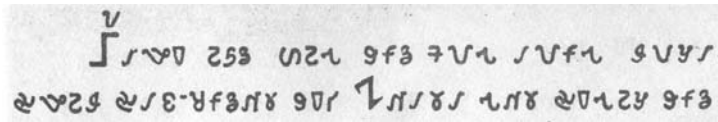


Figure 9. Text showing $\overset{\vee}{\Gamma}r\omega\theta$ *Āmāde*, with a $\Gamma\vee$ conjunct.

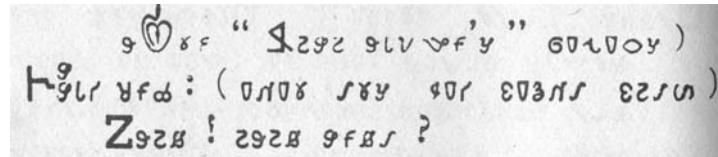


Figure 10. Text showing $\text{ጉ}^{\text{ጵ}}$ *Ūccuḍa*, with a $ጵጵ$ conjunct.

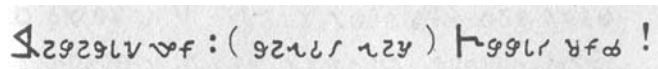


Figure 11. Text showing $\text{ጉ}^{\text{ጵ}}$ *Ūccuḍa*, without a $ጵ$ conjunct.

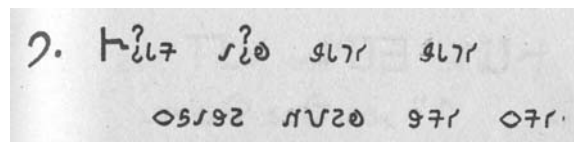


Figure 12. Text showing $\text{ጉ}^{\text{ጵ}}$ *Ūḥuṭ* and $r\text{ጵ}$ *āṭṭala*, with $\omega\omega$ conjuncts.

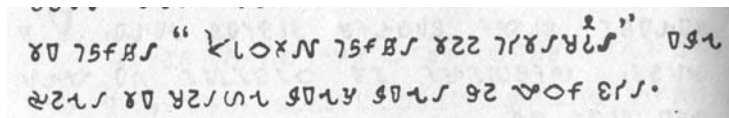


Figure 13. Text showing $r\text{ጵ}r\text{ጵ}$ *ūḍarāṣatrā*, with a $\omega\text{ጵ}$ conjunct.

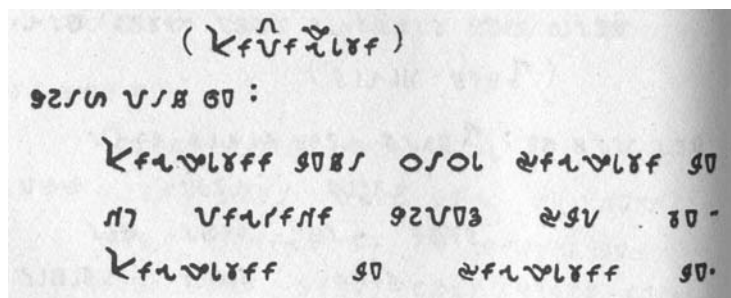


Figure 14. Text showing $\text{ጵ}^{\text{ጵ}}\text{ጵ}^{\text{ጵ}}\text{ጵ}^{\text{ጵ}}$ *Sippinduri*, with $\omega\omega$ and $\omega\text{ጵ}$ conjuncts.
Below is $\text{ጵ}^{\text{ጵ}}\text{ጵ}^{\text{ጵ}}$ *Sindurii*, which has no ጵ conjunct.

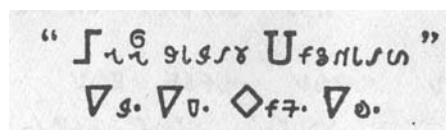


Figure 15. Text showing $\Gamma\text{ጵ}^{\text{ጵ}}$ *Ānanta*, with a $\omega\text{ጵ}$ conjunct.

ᄃᄃᄃ ᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃ ᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ
 (Prime Minister) ᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ
 ᄃᄃᄃ ᄃᄃᄃᄃᄃᄃ ᄃᄃᄃ ᄃᄃᄃᄃ ᄃᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ

Figure 16. Text showing ᄃᄃᄃᄃᄃ *Paruṣottama*, with a ᄃᄃ conjunct.

ᄃᄃᄃᄃ ᄃᄃᄃᄃ ᄃᄃᄃᄃ ᄃᄃ
 ᄃᄃᄃᄃ ᄃᄃᄃᄃ ᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ

Figure 17. Text showing ᄃᄃᄃᄃ *triyu*, with a ᄃᄃ conjunct.

ᄃᄃᄃ ᄃᄃᄃᄃ ᄃᄃ ᄃᄃ ᄃᄃ ᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ;

Figure 18. Text showing ᄃᄃᄃ *Tsiṅa*, with a ᄃᄃ conjunct.

ᄃᄃᄃ. ᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃ ᄃᄃᄃᄃ. ᄃᄃ

Figure 19. Text showing ᄃᄃᄃ *Bha~Baha*, without a ᄃᄃ conjunct.

ᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃᄃ
 “ᄃᄃᄃᄃᄃ ᄃᄃᄃ ᄃᄃᄃᄃᄃ” .

Figure 20. Text showing ᄃᄃᄃ *Bha*, with a ᄃᄃᄃ conjunct.

ᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃᄃ.

Figure 21. Text showing ᄃᄃᄃ *bha*, with a ᄃᄃᄃ conjunct.

ᄃᄃᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃᄃ. ᄃᄃᄃ ᄃᄃᄃ ᄃᄃᄃᄃᄃᄃ, ᄃᄃᄃᄃᄃᄃᄃᄃ

Figure 22. Text showing ᄃᄃᄃᄃᄃᄃᄃ *juppuāḥḥ*, (*juppuāḥ*), with a ᄃᄃᄃ conjunct.

ᄃᄃ ᄃᄃᄃᄃᄃᄃᄃ — ᄃᄃᄃᄃᄃᄃ ᄃᄃᄃᄃ ᄃᄃ ᄃᄃᄃᄃᄃᄃᄃ
 ᄃᄃᄃᄃᄃᄃ — ᄃᄃᄃᄃᄃ — ᄃᄃᄃᄃᄃ — ᄃᄃᄃᄃᄃ ᄃᄃᄃ ᄃᄃᄃᄃᄃ
 ᄃᄃᄃᄃᄃᄃᄃᄃᄃᄃ |

Figure 23. Text showing ᄃᄃᄃᄃ *parra* and ᄃᄃᄃᄃᄃ *porra*, each with a ᄃᄃ conjunct. It appears that the right-side-up conjunct form on ᄃᄃᄃᄃ *parra* is an error and should be ᄃᄃᄃᄃ *parra*.

ᄃᄃᄃᄃᄃ ᄃᄃ ᄃᄃᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃᄃ ᄃᄃᄃᄃ ᄃᄃ ᄃᄃᄃᄃᄃᄃᄃ ᄃᄃᄃᄃ.
 ᄃᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃ ᄃᄃᄃ ᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃᄃ. ᄃᄃᄃᄃᄃᄃᄃ
 ᄃᄃᄃᄃᄃᄃ ᄃᄃ ᄃᄃᄃᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃᄃ ᄃᄃᄃᄃᄃᄃᄃᄃ.

Figure 24. Text showing ᄃᄃᄃᄃ *olla* and ᄃᄃᄃᄃ *olla*, each with a ᄃᄃ conjunct. It appears that the right-side-up conjunct form on ᄃᄃᄃᄃ *olla* is an error and should be ᄃᄃᄃᄃ *olla*.

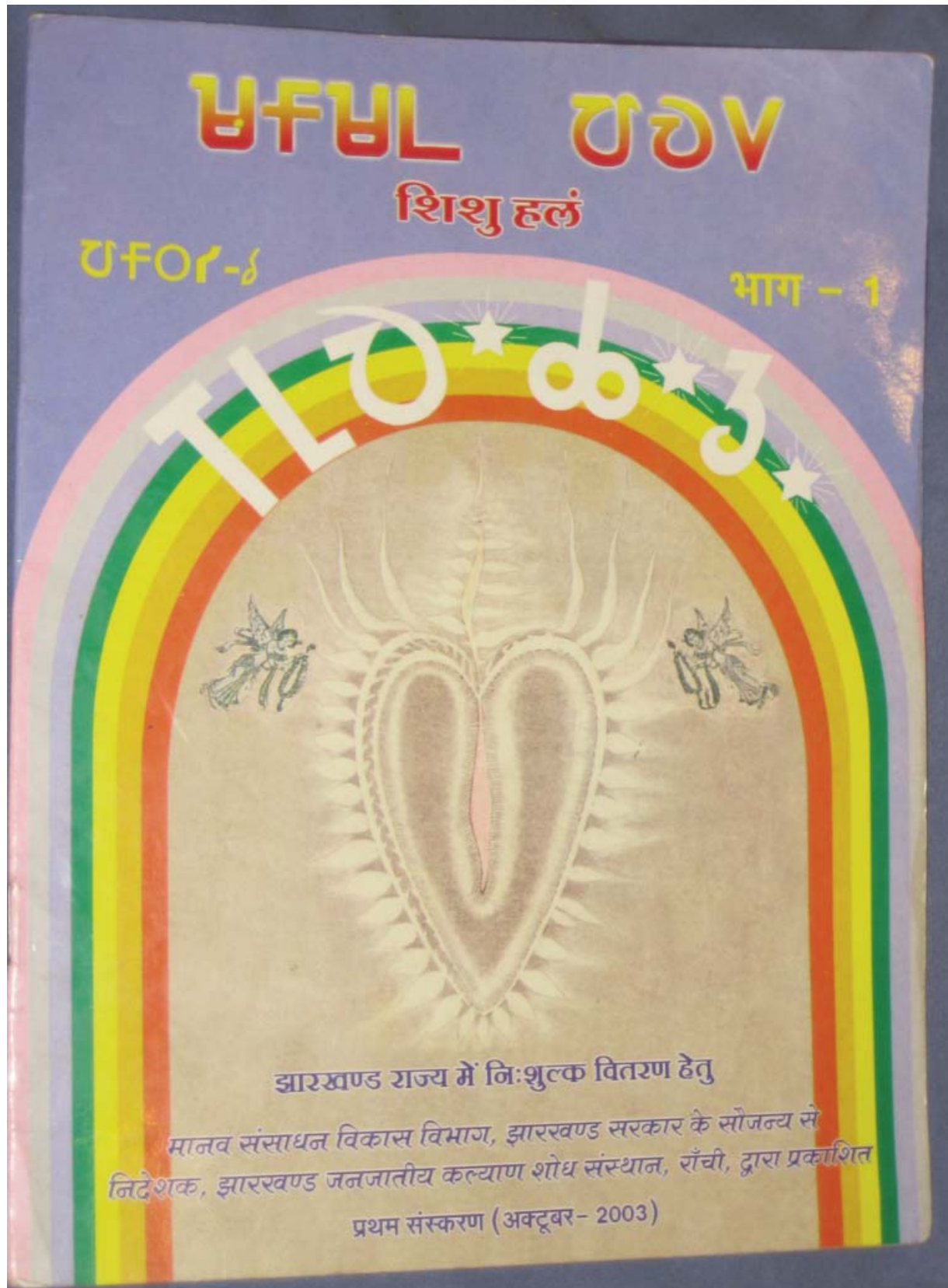


Figure 25. The cover of UFUL UDU (Şiřu halam).

A. Administrative

1. Title

Revised proposal for encoding the Warang Citi script in the SMP of the UCS

2. Requester's name

UC Berkeley Script Encoding Initiative (Universal Scripts Project)

3. Requester type (Member body/Liaison/Individual contribution)

Liaison contribution.

4. Submission date

2011-12-31

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

1b. Proposed name of script

Warang Citi.

1c. The proposal is for addition of character(s) to an existing block

No.

1d. Name of the existing block

2. Number of characters in proposal

84.

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category A.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. 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)?

Yes.

8. 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/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

Yes, in N1958, N3411, N3668.

2a. 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.

2b. If YES, with whom?

K. David Harrison, Gregory Anderson, Anna Pucilowski, Chandra Mohan Haibru, K.C. Naik Biruli.

2c. 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?

Speakers of the Ho language, whose population is 1,026,000, with 25%-50% literacy, according to the SIL *Ethnologue*. The *Ethnologue* mentions only the Oriya and Devanagari script areas.

4a. The context of use for the proposed characters (type of use; common or rare)

To write the Ho language.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

In India.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Contemporary use and accordance with the Roadmap.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?