

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 for encoding the Warang Citi script in the BMP 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)****Date: 2009-08-05**

**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.

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 Chiti seems to be stronger in the eastern part of the diaspora, where books that use it have been printed. There is no quantitative data on who uses Warang Citi, or how often, or where. Nevertheless, the community of Ho speakers numbers more than a million, widely dispersed. Even if only 10% were users of Warang Citi, that is still many times more people than most languages of the world have as their total number of speakers. The absence of Warang Chiti in the UCS presents a true barrier for Ho to enter the digital age with its own unique writing system intact.

**2. Structure.** Warang Citi is written from left to right, with the vowel characters in the order they are spoken; complex vowel placement and variant forms found in Brahmi-derived scripts does not obtain in Warang Citi. Letters have an inherent vowel, typically pronounced [a], but sometimes [o] or [e]. The letter  $\text{ᱠ}$   $\text{ᱡ}$  HA is used after a vowel to indicate lengthening:  $\text{ᱠᱠ}$   $\text{ᱡᱡ}$  AH ah ( $\bar{A}$   $\bar{a}$ ). Long *i* and *u* are described as “ligatures” and evidently  $\text{ᱠᱡ}$   $\text{ᱡᱠ}$   $\text{ᱠᱢ}$   $\text{ᱢᱠ}$   $\text{ᱠᱣ}$   $\text{ᱣᱠ}$  are written, not  $\text{ᱠᱡ}$   $\text{ᱡᱠ}$  and  $\text{ᱠᱢ}$   $\text{ᱢᱠ}$ . These are not, however, divisible typographical ligatures.

**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  $\text{ᱥᱟᱦᱮᱵᱽᱨᱮᱵᱽᱫᱽᱯᱩᱨ}$  ‘15-10-2003’ (see Figure X). No symbols for 100 or 1000 are known, though Pinnow postulated that they may exist.

**4. Collating order.** Collation order is as in the code chart, with casing pairs interfiling as in Latin. U+AB5F WARANG CITI OM files like  $\text{ᱠᱢᱣᱟ}$   $\text{ᱠᱢᱣᱟ}$   $\text{ᱠᱢᱣᱟ}$ .

**5. Character names.** The name of the script is  $\text{ᮊᮧᮒᮓ ᮘᮞᮞᮓ}$  *Hbarŋ citi* ‘undying writing’, where  $\text{ᮊᮧ}$  *hb* is a diagraph 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  $\text{ᮊᮓ}$  [v] has been constructed; it is an original ligature of  $\text{ᮊᮓ}$  A and  $\text{ᮊᮧ}$  HIYO. Other original ligatures are  $\text{ᮊᮓ}$  YA [ʔ] ( $\text{ᮊᮓ} + \text{ᮊᮓ}$  *m + m*);  $\text{ᮊᮓ}$  YO [y], [e], [i] ( $\text{ᮊᮓ} + \text{ᮊᮓ}$  *a + a*);  $\text{ᮊᮓ}$  II [i], [ai], [ii] ( $\text{ᮊᮓ} + \text{ᮊᮓ}$  *i + i*); and  $\text{ᮊᮓ}$  UU [u], [au], [uu] ( $\text{ᮊᮓ} + \text{ᮊᮓ}$  *u + u*). None of these ligatures are decomposable.

ᮊᮓʔ	NGAA	NGÁ’	ʔi:	II	’î	ʔɛ̃	ENN	’ENN	puʔ	PU	PU’
ʔa	A	’A	ʔu:	UU	’Û	ʔɔ̃	ODD	’ODD	hiyɔ̃	HIYO	HIYO
wüʔi	WI	Ū’I	ʔang	ANG	’ANG	ʔɛ̃ʔ	TTE	TTE’	hɔ̃ɔ̃	HOLO	HOLO
ʔuʔu	YU	YU’U	gaʔ	GA	GA’	nuŋ	NUNG	NUNG	hɔ̃ɔ̃	HORR	HORR
ʔɛ̃ʔ	E	’E’	kɔ̃ʔ	KO	KO’	daʔ	DA	DA’	har	HAR	HAR
ʔɔ̃ʔ	O	’O’	ʔɛ̃ñ	ENY	’EÑ	ʔat	AT	’AT	śu:	SSUU	ÇÛ
’yaʔ	YA	YA’	yü’j	YUJ	YÛJ	ʔam	AM	’AM	sī:	SII	SĪ
’yɔ̃ʔ	YO	YO’	ʔüc	UC	’ÛC	bu	BU	BU	viyɔ̃	VIYO	VIYO

**6. Punctuation.** Pinnow states explicitly that European punctuation is used as in English.

**7. Linebreaking.** Letters and digits behave as in Latin and other alphabetic scripts. U+AB5F WARANG CITI OM behaves like a letter.

### 8. Unicode Character Properties.

```

AB00;WARANG CITI CAPITAL LETTER NGA;Lu;0;L;;;;N;;;AB20;
AB01;WARANG CITI CAPITAL LETTER A;Lu;0;L;;;;N;;;AB21;
AB02;WARANG CITI CAPITAL LETTER WI;Lu;0;L;;;;N;;;AB22;
AB03;WARANG CITI CAPITAL LETTER YU;Lu;0;L;;;;N;;;AB23;
AB04;WARANG CITI CAPITAL LETTER E;Lu;0;L;;;;N;;;AB24;
AB05;WARANG CITI CAPITAL LETTER O;Lu;0;L;;;;N;;;AB25;
AB06;WARANG CITI CAPITAL LETTER YA;Lu;0;L;;;;N;;;AB26;
AB07;WARANG CITI CAPITAL LETTER YO;Lu;0;L;;;;N;;;AB27;
AB08;WARANG CITI CAPITAL LETTER II;Lu;0;L;;;;N;;;AB28;
AB09;WARANG CITI CAPITAL LETTER UU;Lu;0;L;;;;N;;;AB29;
AB0A;WARANG CITI CAPITAL LETTER ANG;Lu;0;L;;;;N;;;AB2A;
AB0B;WARANG CITI CAPITAL LETTER GA;Lu;0;L;;;;N;;;AB2B;
AB0C;WARANG CITI CAPITAL LETTER KO;Lu;0;L;;;;N;;;AB2C;
AB0D;WARANG CITI CAPITAL LETTER ENY;Lu;0;L;;;;N;;;AB2D;
AB0E;WARANG CITI CAPITAL LETTER YUJ;Lu;0;L;;;;N;;;AB2E;
AB0F;WARANG CITI CAPITAL LETTER UC;Lu;0;L;;;;N;;;AB2F;
AB10;WARANG CITI CAPITAL LETTER ENN;Lu;0;L;;;;N;;;AB30;
AB11;WARANG CITI CAPITAL LETTER ODD;Lu;0;L;;;;N;;;AB31;
AB12;WARANG CITI CAPITAL LETTER TTE;Lu;0;L;;;;N;;;AB32;
AB13;WARANG CITI CAPITAL LETTER NUNG;Lu;0;L;;;;N;;;AB33;
AB14;WARANG CITI CAPITAL LETTER DA;Lu;0;L;;;;N;;;AB34;
AB15;WARANG CITI CAPITAL LETTER AT;Lu;0;L;;;;N;;;AB35;
AB16;WARANG CITI CAPITAL LETTER AM;Lu;0;L;;;;N;;;AB36;
AB17;WARANG CITI CAPITAL LETTER BU;Lu;0;L;;;;N;;;AB37;
AB18;WARANG CITI CAPITAL LETTER PU;Lu;0;L;;;;N;;;AB38;
AB19;WARANG CITI CAPITAL LETTER HIYO;Lu;0;L;;;;N;;;AB39;
AB1A;WARANG CITI CAPITAL LETTER HOLO;Lu;0;L;;;;N;;;AB3A;
AB1C;WARANG CITI CAPITAL LETTER HORR;Lu;0;L;;;;N;;;AB3B;
AB1C;WARANG CITI CAPITAL LETTER HAR;Lu;0;L;;;;N;;;AB3C;
AB1D;WARANG CITI CAPITAL LETTER SSUU;Lu;0;L;;;;N;;;AB3D;
AB1E;WARANG CITI CAPITAL LETTER SII;Lu;0;L;;;;N;;;AB3E;
AB1F;WARANG CITI CAPITAL LETTER VIYO;Lu;0;L;;;;N;;;AB3F;
AB20;WARANG CITI SMALL LETTER NGA;Ll;0;L;;;;N;;;AB00;AB00;
AB21;WARANG CITI SMALL LETTER A;Ll;0;L;;;;N;;;AB01;AB01;
AB22;WARANG CITI SMALL LETTER WI;Ll;0;L;;;;N;;;AB02;AB02;
AB23;WARANG CITI SMALL LETTER YU;Ll;0;L;;;;N;;;AB03;AB03;
AB24;WARANG CITI SMALL LETTER E;Ll;0;L;;;;N;;;AB04;AB04;
AB25;WARANG CITI SMALL LETTER O;Ll;0;L;;;;N;;;AB05;AB05;
AB26;WARANG CITI SMALL LETTER YA;Ll;0;L;;;;N;;;AB06;AB06;
AB27;WARANG CITI SMALL LETTER YO;Ll;0;L;;;;N;;;AB07;AB07;
AB28;WARANG CITI SMALL LETTER II;Ll;0;L;;;;N;;;AB08;AB08;
AB29;WARANG CITI SMALL LETTER UU;Ll;0;L;;;;N;;;AB09;AB09;
AB2A;WARANG CITI SMALL LETTER ANG;Ll;0;L;;;;N;;;AB0A;AB0A;
AB2B;WARANG CITI SMALL LETTER GA;Ll;0;L;;;;N;;;AB0B;AB0B;
AB2C;WARANG CITI SMALL LETTER KO;Ll;0;L;;;;N;;;AB0C;AB0C;
AB2D;WARANG CITI SMALL LETTER ENY;Ll;0;L;;;;N;;;AB0D;AB0D;
AB2E;WARANG CITI SMALL LETTER YUJ;Ll;0;L;;;;N;;;AB0E;AB0E;
AB2F;WARANG CITI SMALL LETTER UC;Ll;0;L;;;;N;;;AB0F;AB0F;





















































































```

AB30;WARANG CITI SMALL LETTER ENN;Ll;0;L; ; ; ; ;N; ; ;AB10;AB10  
 AB31;WARANG CITI SMALL LETTER ODD;Ll;0;L; ; ; ; ;N; ; ;AB11;AB11  
 AB32;WARANG CITI SMALL LETTER TTE;Ll;0;L; ; ; ; ;N; ; ;AB12;AB12  
 AB33;WARANG CITI SMALL LETTER NUNG;Ll;0;L; ; ; ; ;N; ; ;AB13;AB13  
 AB34;WARANG CITI SMALL LETTER DA;Ll;0;L; ; ; ; ;N; ; ;AB14;AB14  
 AB35;WARANG CITI SMALL LETTER AT;Ll;0;L; ; ; ; ;N; ; ;AB15;AB15  
 AB36;WARANG CITI SMALL LETTER AM;Ll;0;L; ; ; ; ;N; ; ;AB16;AB16  
 AB37;WARANG CITI SMALL LETTER BU;Ll;0;L; ; ; ; ;N; ; ;AB17;AB17  
 AB38;WARANG CITI SMALL LETTER PU;Ll;0;L; ; ; ; ;N; ; ;AB18;AB18  
 AB39;WARANG CITI SMALL LETTER HIYO;Ll;0;L; ; ; ; ;N; ; ;AB19;AB19  
 AB3A;WARANG CITI SMALL LETTER HOLO;Ll;0;L; ; ; ; ;N; ; ;AB1A;AB1A  
 AB3B;WARANG CITI SMALL LETTER HARR;Ll;0;L; ; ; ; ;N; ; ;AB1B;AB1B  
 AB3C;WARANG CITI SMALL LETTER HAR;Ll;0;L; ; ; ; ;N; ; ;AB1C;AB1C  
 AB3D;WARANG CITI SMALL LETTER SSUU;Ll;0;L; ; ; ; ;N; ; ;AB1D;AB1D  
 AB3E;WARANG CITI SMALL LETTER SII;Ll;0;L; ; ; ; ;N; ; ;AB1E;AB1E  
 AB3F;WARANG CITI SMALL LETTER VIYO;Ll;0;L; ; ; ; ;N; ; ;AB1F;AB1F  
 AB40;WARANG CITI DIGIT ZERO;Nd;0;L; ;0;0;0;N; ; ; ; ;  
 AB41;WARANG CITI DIGIT ONE;Nd;0;L; ;1;1;1;N; ; ; ; ;  
 AB42;WARANG CITI DIGIT TWO;Nd;0;L; ;2;2;2;N; ; ; ; ;  
 AB43;WARANG CITI DIGIT THREE;Nd;0;L; ;3;3;3;N; ; ; ; ;  
 AB44;WARANG CITI DIGIT FOUR;Nd;0;L; ;4;4;4;N; ; ; ; ;  
 AB45;WARANG CITI DIGIT FIVE;Nd;0;L; ;5;5;5;N; ; ; ; ;  
 AB46;WARANG CITI DIGIT SIX;Nd;0;L; ;6;6;6;N; ; ; ; ;  
 AB47;WARANG CITI DIGIT SEVEN;Nd;0;L; ;7;7;7;N; ; ; ; ;  
 AB48;WARANG CITI DIGIT EIGHT;Nd;0;L; ;8;8;8;N; ; ; ; ;  
 AB49;WARANG CITI DIGIT NINE;Nd;0;L; ;9;9;9;N; ; ; ; ;  
 AB4A;WARANG CITI NUMBER TEN;No;0;L; ; ; ; ;10;N; ; ; ; ;  
 AB4B;WARANG CITI NUMBER TWENTY;No;0;L; ; ; ; ;20;N; ; ; ; ;  
 AB4C;WARANG CITI NUMBER THIRTY;No;0;L; ; ; ; ;30;N; ; ; ; ;  
 AB4D;WARANG CITI NUMBER FORTY;No;0;L; ; ; ; ;40;N; ; ; ; ;  
 AB4E;WARANG CITI NUMBER FIFTY;No;0;L; ; ; ; ;50;N; ; ; ; ;  
 AB4F;WARANG CITI NUMBER SIXTY;No;0;L; ; ; ; ;60;N; ; ; ; ;  
 AB50;WARANG CITI NUMBER SEVENTY;No;0;L; ; ; ; ;70;N; ; ; ; ;  
 AB51;WARANG CITI NUMBER EIGHTY;No;0;L; ; ; ; ;80;N; ; ; ; ;  
 AB52;WARANG CITI NUMBER NINETY;No;0;L; ; ; ; ;90;N; ; ; ; ;  
 AB5F;WARANG CITI OM;Lo;0;L; ; ; ; ;N; ; ; ; ;

## 9. 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

**10. 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.

	AB0	AB1	AB2	AB3	AB4	AB5
0	 AB00	 AB10	 AB20	 AB30	 AB40	 AB50
1	 AB01	 AB11	 AB21	 AB31	 AB41	 AB51
2	 AB02	 AB12	 AB22	 AB32	 AB42	 AB52
3	 AB03	 AB13	 AB23	 AB33	 AB43	
4	 AB04	 AB14	 AB24	 AB34	 AB44	
5	 AB05	 AB15	 AB25	 AB35	 AB45	
6	 AB06	 AB16	 AB26	 AB36	 AB46	
7	 AB07	 AB17	 AB27	 AB37	 AB47	
8	 AB08	 AB18	 AB28	 AB38	 AB48	
9	 AB09	 AB19	 AB29	 AB39	 AB49	
A	 AB0A	 AB1A	 AB2A	 AB3A	 AB4A	
B	 AB0B	 AB1B	 AB2B	 AB3B	 AB4B	
C	 AB0C	 AB1C	 AB2C	 AB3C	 AB4C	
D	 AB0D	 AB1D	 AB2D	 AB3D	 AB4D	
E	 AB0E	 AB1E	 AB2E	 AB3E	 AB4E	
F	 AB0F	 AB1F	 AB2F	 AB3F	 AB4F	 AB5F

**Uppercase vowels**

AB00	V	WARANG CITI CAPITAL LETTER NGA
AB01	𑄑	WARANG CITI CAPITAL LETTER A
AB02	𑄒	WARANG CITI CAPITAL LETTER WI
AB03	L	WARANG CITI CAPITAL LETTER YU
AB04	𑄔	WARANG CITI CAPITAL LETTER E
AB05	Z	WARANG CITI CAPITAL LETTER O
AB06	Y	WARANG CITI CAPITAL LETTER YA
AB07	𑄗	WARANG CITI CAPITAL LETTER YO
AB08	E	WARANG CITI CAPITAL LETTER II
AB09	𑄙	WARANG CITI CAPITAL LETTER UU

**Uppercase consonants**

AB0A	𑄚	WARANG CITI CAPITAL LETTER ANG
AB0B	𑄛	WARANG CITI CAPITAL LETTER GA
AB0C	𑄜	WARANG CITI CAPITAL LETTER KO
AB0D	𑄝	WARANG CITI CAPITAL LETTER ENY
AB0E	E	WARANG CITI CAPITAL LETTER YUJ
AB0F	𑄟	WARANG CITI CAPITAL LETTER UC
AB10	𑄠	WARANG CITI CAPITAL LETTER ENN
AB11	𑄡	WARANG CITI CAPITAL LETTER ODD
AB12	L	WARANG CITI CAPITAL LETTER TTE
AB13	𑄣	WARANG CITI CAPITAL LETTER NUNG
AB14	𑄤	WARANG CITI CAPITAL LETTER DA
AB15	O	WARANG CITI CAPITAL LETTER AT
AB16	𑄧	WARANG CITI CAPITAL LETTER AM
AB17	𑄨	WARANG CITI CAPITAL LETTER BU
AB18	𑄩	WARANG CITI CAPITAL LETTER PU
AB19	𑄪	WARANG CITI CAPITAL LETTER HIYO
AB1A	𑄫	WARANG CITI CAPITAL LETTER HOLO
AB1B	𑄬	WARANG CITI CAPITAL LETTER HERR
AB1C	T	WARANG CITI CAPITAL LETTER HAR
AB1D	𑄮	WARANG CITI CAPITAL LETTER SSUU
AB1E	𑄯	WARANG CITI CAPITAL LETTER SHI
AB1F	𑄰	WARANG CITI CAPITAL LETTER VIYO

**Lowercase vowels**

AB20	v	WARANG CITI SMALL LETTER NGA
AB21	𑄱	WARANG CITI SMALL LETTER A
AB22	𑄲	WARANG CITI SMALL LETTER WI
AB23	l	WARANG CITI SMALL LETTER YU
AB24	𑄴	WARANG CITI SMALL LETTER E
AB25	z	WARANG CITI SMALL LETTER O
AB26	y	WARANG CITI SMALL LETTER YA
AB27	𑄷	WARANG CITI SMALL LETTER YO
AB28	e	WARANG CITI SMALL LETTER II
AB29	𑄹	WARANG CITI SMALL LETTER UU

**Lowercase consonants**

AB2A	𑄚	WARANG CITI SMALL LETTER ANG
AB2B	𑄛	WARANG CITI SMALL LETTER GA
AB2C	𑄜	WARANG CITI SMALL LETTER KO
AB2D	𑄝	WARANG CITI SMALL LETTER ENY
AB2E	𑄞	WARANG CITI SMALL LETTER YUJ
AB2F	𑄟	WARANG CITI SMALL LETTER UC
AB30	𑄠	WARANG CITI SMALL LETTER ENN
AB31	𑄡	WARANG CITI SMALL LETTER ODD
AB32	l	WARANG CITI SMALL LETTER TTE
AB33	𑄣	WARANG CITI SMALL LETTER NUNG
AB34	𑄤	WARANG CITI SMALL LETTER DA

AB35	𑄶	WARANG CITI SMALL LETTER AT
AB36	𑄷	WARANG CITI SMALL LETTER AM
AB37	𑄸	WARANG CITI SMALL LETTER BU
AB38	𑄹	WARANG CITI SMALL LETTER PU
AB39	𑄺	WARANG CITI SMALL LETTER HIYO
AB3A	𑄻	WARANG CITI SMALL LETTER HOLO
AB3B	𑄼	WARANG CITI SMALL LETTER HERR
AB3C	𑄽	WARANG CITI SMALL LETTER HAR
AB3D	𑄾	WARANG CITI SMALL LETTER SSUU
AB3E	𑄿	WARANG CITI SMALL LETTER SHI
AB3F	𑅀	WARANG CITI SMALL LETTER VIYO

**Digits**

AB40	0	WARANG CITI DIGIT ZERO
AB41	1	WARANG CITI DIGIT ONE
AB42	2	WARANG CITI DIGIT TWO
AB43	3	WARANG CITI DIGIT THREE
AB44	4	WARANG CITI DIGIT FOUR
AB45	5	WARANG CITI DIGIT FIVE
AB46	6	WARANG CITI DIGIT SIX
AB47	7	WARANG CITI DIGIT SEVEN
AB48	8	WARANG CITI DIGIT EIGHT
AB49	9	WARANG CITI DIGIT NINE

**Numbers**

AB4A	10	WARANG CITI NUMBER TEN
AB4B	20	WARANG CITI NUMBER TWENTY
AB4C	30	WARANG CITI NUMBER THIRTY
AB4D	40	WARANG CITI NUMBER FORTY
AB4E	50	WARANG CITI NUMBER FIFTY
AB4F	60	WARANG CITI NUMBER SIXTY
AB50	70	WARANG CITI NUMBER SEVENTY
AB51	80	WARANG CITI NUMBER EIGHTY
AB52	90	WARANG CITI NUMBER NINETY


**Sign**

AB5F	𑄿	WARANG CITI OM
------	---	----------------





Originalschr.		Transliteration		Bezeichnung der Grapheme		Phonement- sprechung
Groß- Buchstabe	Klein- Buchstabe	Devanāgarī	Latein- Schrift	Devanāgarī	Phonetisch	
ए	ε	ज	J j	बिज्	yü'j	j
प	ϕ	ब	C c	बिच्	ʔüc	c
न	~	रा	N n	त्रैरा	ʔεʔ	n
द	ɾ	ड	D d	ओड्	ʔɔ'd	d
ट	ɽ	ट	T t	टेःए	ʔεʔ	t
न	~	न	N n	नुड्	nun	n
द	~	द	D d	दःऽ	daʔ	d
त	6	त	T t	औत	ʔat	t
म	१	म	M m	ऽम	ʔam	m
ब	0	ब	B b	बु	bu	b
प	~	प	P p	पुःऽ	puʔ	p
ह	5	ह	H h	ह्यी	hiyo	h , Vokaldehnung
ल	७	ल	L l	ह्ली	holo	l
र	5	ड़	R r	हड़	hor	r
र	४	र	R r	हर	har	r
ष	४	ष,श	S s	षु	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 εφβ ϖζης!  
 Τφχθς Λιςιηχ θζςς! ιφχθς θεχς θζςς!!  
 ιφχθς ιοιςχβ ϖςςυ!!! ιφχθς ιηζεφβ ϖζης!!!!  
 θυθυ ον οια 1 θ Δ

⊕ ⊕ ϕ Ν ςν ον.

ηζθ ηλεθ ηλεθ—ελελε ελελε ϖλε χλε  
 εζ χυυζ ζιζθ εζς ου ηζιυ  
 υηχθθ οηχθθ εςχθθ ςιηθθ ςιθς χζςιι  
 οχης χυυζ ηθν ηςθν εςιυ.  
 Τζιζζ ϖςςχ ϖλεχ ςελεθ εςελε ιυ Δ'  
 ηυεζιυ θεεζιυ θυθυ ηζιυ εςελε  
 ηυ, θεχθθ ελεχθθ ελεχθθ ηςελεθ εςελεθ  
 υςυηη, ϖυυηηθθθθ εςεθ εεχς ηζ ⊕ ηοθης-  
 Κιηης χλελεθθθθ ϖφχλεθ εζεθ εζεθ  
 ⊕ ηοιςχβ εοηελευ Δ' εοηελε εοηελευ ελεςς

# ΗΥJΤ UZIJ

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

᳚— ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚, ᳚᳚,  
गा, का, आ, 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, २f.  
मि, वि, पि, हि, लि, डि, रि, षि, सि।

L— १L, १L, ६L, ६L, १L, १L, १L, १L, १L, १L,  
गु, कु, ञु, जु, चु, णु, डु, टु, नु, दु,  
६L, १L, ०L, १L, ७L, ७L, ७L, ४L, ४L, २L.  
तु, मु, वु, पु, हु, लु, डु, रु, षु, सु।

V— १V, १V, ६V, ६V, १V, १V, १V, १V, १V, १V,  
गो, को, जे, जे, चो, णो, डो, टो, नो, दो,  
६V, १V, ०V, १V, ७V, ७V, ५V, ४V, ४V, २V.  
ते, मो, वो, पो, हो, लो, डो, रो, षो, सो।


Z— १Z, १Z, ६Z, ६Z, १Z, १Z, १Z, १Z, १Z, १Z,  
गो, को, जो, जो, चो, णो, डो, टो, नो, दो,  
६Z, १Z, ०Z, १Z, ७Z, ७Z, ५Z, ४Z, ४Z, २Z.  
तो, मो, वो, पो, हो, लो, डो, रो, षो, सो।

Y— १Y, १Y, ६Y, ६Y, १Y, १Y, १Y, १Y, १Y, १Y,  
गः, कः, जः, जः, चः, णः, डः, टः, नः,  
१Y, ६Y, १Y, ०Y, १Y, ७Y, ७Y, ५Y, ४Y,  
दः, तः, मः, वः, पः, हः, लः, डः, रः,  
४Y, २Y.  
षः, सः।

H— १H, १H, ६H, ६H, १H, १H, १H, १H, १H, १H,  
गौ, कौ, जौ, जौ, चौ, णौ, डौ, टौ, नौ,  
१H, ६H, १H, ०H, १H, ७H, ७H, ५H, ४H,  
दौ, तौ, मौ, वौ, पौ, हौ, लौ, डौ, रौ,  
४H, २H.  
षौ, सौ।

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, अ्	।, अ	a
F	f	विऽ	ि	i
L	l	युऽ	ु	u
Y	y	यऽ	ः	?
H	h	यो	य	ea
E	e	इ	ई,ी	ii
T	t	उ	उ,ू	uu
R	r	ए	, ए	e
V	v	ो अ्	ो, ओ	o
<b>०३१ ०२४३ व्यन्जन वर्ण - Consonant</b>				
3	3	ोड्	ड	ng
1	1	ग्	ग्	g
9	9	क्	क्	k, ch
6	6	विञ्	ञ्	n
8	8	विज्	ज्	j
5	5	विच्	च्	ch, c
11	11	एण	ण	N
1	1	ोड्	ड	d
1	1	टेःए	ट	t
1	1	नुड	न	n
1	1	दऽ	द	d
6	6	ोत्	त्	t
4	4	ऽह्	म्	m
4	4	वू	व्	b
0	0	पुऽ	प्	p
U	U	ह्यो	ह	h
0	0	हलौ	ल्	l
5	5	हड	ड्	!
4	4	हऽ	र्	r
4	4	षू	ष्	sh
4	4	सी	स्	s

**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 sammiśraṇa yah varna banā hai” ‘from a ligature of ङ + उ this character is made’.





## A. Administrative

1. Title

**Proposal for encoding the Warang Citi script in the BMP 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

**2009-08-05**

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 and N3411.**

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 (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

**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?