

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 Latin characters for Osage in the UCS
Source: Michael Everson, Herman Mongrain Lookout, Cameron Pratt
Status: Individual Contribution
Date: 2014-07-30
Replaces: N4548 (2014-02-20)

1. Introduction. This document requests the addition of Latin characters for the Osage language to the UCS.

1.1. Background. A variety of ad-hoc Latin orthographies and transcriptions have been used for the Osage language over the past 210 years, having been devised by people for reasons such as bible verse translation, language documentation, and study. Most of the writing systems were prepared by second-language Osage speakers. One of the earliest sources is dated circa 1851, attributed to Father Ponziglione, a Jesuit priest who ministered to the Osage.

2. The development of the contemporary Osage orthography. In 2004, the 31st Council of the Osage Nation passed a resolution initiating the Osage Language Program. Soon after, Herman Mongrain Lookout—known as “Mogri”—was hired as the director, and was afforded office space in downtown Pawhuska, Oklahoma. At that time no standard orthography existed for Osage, and students were exhorted to “spell it like you hear it”, which, naturally, led to conflicting conventions differing from student to student and lacking linguistic robustness. Mogri developed a Latin-based orthography which used all capitals: SHOⁿ-KE ‘dog’, TSI ‘house’, TA ‘deer’, KE ‘turtle’. It soon became felt that students’ knowledge of the English orthographic values of these letters interfered with good pronunciation of Osage, and by 2006, working with advanced learners, a practical orthography had been designed whose letters were modifications or fusions of Latin letters. The words above were written ḤO[^].ḶḶ ṣḡke ‘dog’, ṖḶ tsi ‘house’, ḶḶ ta ‘deer’, ḶḶ ke ‘turtle’ with these new letters.

The new Osage orthography was taken up with alacrity by students and teachers, and has been consistently and regularly used throughout the Osage Nation since 2006. In 2012 a number of linguists and language-speakers raised some issues about the repertoire of characters employed, on the basis of their experience using it. Cameron Pratt and Dylan Herrick produced some documents analysing the use of the orthography and outlining questions which they felt should be examined and addressed.

At the same time graphic artist Ryan RedCorn and more recently Jessica Harjo had explored Osage typography and noticed a number of features which they felt to be somewhat problematic. Issues of character encoding and of orthographic reform began to be discussed with Michael Everson late in 2012 and early in 2013, and finally a three-day working seminar was hosted by the Osage Language Department on 12–14 February 2014, at which questions of linguistic issues, graphic design and

fonts, and character encoding were discussed at length. Participants in the seminars included Mogri Lookout (who devised the orthography), linguist Cameron Pratt, teachers and curriculum developers Janice Carpenter and Stephanie Rapp, graphic designers Jessica Harjo and Ryan RedCorn, governmental representative Geoffrey StandingBear, and other members of the Osage community.

3. Osage consonants. The table below shows the relation between Osage consonantal phonology. The first column for each consonant is the transcription used in Quintero 2009, the second IPA transcription, and the third Osage orthography.

	Labial		Dental		Palatal-alveolar	Velar	Glottal
Glottalized stops	pʰ pʰ ɸʰ		cʰ tsʰ ʈʰ			kʰ kʰ ɰʰ	(ʔ)(ʔ)–
Lax voiceless stops	p p ɸ		t t ɖ		č tʃ ɕ	k k ɣ	
Affricates	c ts ʈ						
Preaspirated/tense voiceless stops	hp ʰp ɸʰ		ht ʰt ɖʰ		hč ʰtʃ ɕʰ	hk ʰk ɰʰ	
Voiced stops	br bɹ ɹ		hc ʰts ʈ				
(Post)aspirated stops	pš pʃ ɸʃ		ch tsʰ ʈ			kš kʃ ɰʃ	
	px px ɸʰ		tx tx ɖʰ			kx kx ɰʰ	
Voiceless/tense fric.			s s ɕ		š ʃ ɕ	x x ɰ	h h ɰ
Voiced/lax fricatives			z z ʒ		ž ʒ ʒ	ɣ ɣ ɰ	
Palatalized						kʲ kʲ ɰʲ	hʲ hʲ ɰʲ
Nasals	m m ɸ		n n ɲ				
Approximants	w w ɹ		l l ɕ				
			ð ð ɸ				

4. Osage vowels. Oral and nasal vowels are distinguished by means of a diacritic dot:

Λ ʌ A a, ʌ ʌ Ə ə, ʌ ʌ E e, ʌ ʌ I i, O o O o, U u U u; ʌ ʌ A ą, ʌ ʌ Ə ę, ʌ ʌ I ı, ʌ ʌ O ɔ

This diacritic dot is an existing a combining mark, U+0358 COMBINING DOT ABOVE RIGHT. Three diphthongs use a similar dot inside the character (this is not a combining character and these); there is also one oral diphthong:

ʌ ʌ aɪ aɪ, ʌ ʌ Eɪ eɪ, ʌ ʌ Oɪ oɪ; ʌ ʌ Ai ai

Long vowels are indicated by means of U+0304 COMBINING MACRON ABOVE:

ʌ ʌ Ā ā, ʌ ʌ Ē ē, ʌ ʌ Ī ī, ʌ ʌ Ō ō, ʌ ʌ Ū ū; ʌ ʌ Ā ą, ʌ ʌ Ī ı, ʌ ʌ Ō ɔ

Where pitch accent needs to be indicated U+0301 COMBINING ACUTE ACCENT is used:

ʌ ʌ Á á, ʌ ʌ É é, ʌ ʌ Í í, ʌ ʌ Ó ó, ʌ ʌ Ú ú; ʌ ʌ Á ą, ʌ ʌ Í ı, ʌ ʌ Ó ɔ
 ʌ ʌ Áı áı, ʌ ʌ Éı éı, ʌ ʌ Óı óı; ʌ ʌ Áı áı

Where pitch accent needs to be indicated in conjunction with vowel length, U+030B COMBINING DOUBLE ACUTE ACCENT is used:

ʌ ʌ Ā́ ā́, ʌ ʌ Ḗ ḗ, ʌ ʌ Ī́ ī́, ʌ ʌ Ṓ ṓ, ʌ ʌ Ū́ ū́; ʌ ʌ Ā́ ą́, ʌ ʌ Ī́ ı́, ʌ ʌ Ṓ ɔ́

5. Reforms of previous orthography. Five reforms were agreed at the February 2014 meeting, which have been implemented in the present proposal. The first of these was the introduction of case pairs. The examples of the introduction of case into Old Hungarian by young learners, and of the ramifications of the re-introduction of case into the existing Cherokee encoding were discussed at considerable length, and during the February meetings everyone including Mogri Lookout and the two graphic designers was quite certain that the orthography was “mature” enough to acquire and

implement case, in order to offer a robust new typographic tradition. (In particular the experience of the Cherokees changing the encoding model was taken quite seriously; the Osage accepted that any changes now needed to be permanent and stable.)

Another reform was in the special representation of the “pre-aspirate” consonant class. These consonants represent original geminates in proto-Dhegiha, and are either pronounced as geminates, or as pre-aspirated consonants. The original representation was either $\mathcal{S}\mathcal{P}$ *hp*, $\mathcal{S}\mathcal{D}$ *ht*, $\mathcal{S}\mathcal{P}$ *hc*, $\mathcal{S}\mathcal{H}$ *hč*, $\mathcal{S}\mathcal{K}$ *hk* or $\mathcal{S}\mathcal{P}$ *hp*, $\mathcal{S}\mathcal{D}$ *ht*, $\mathcal{S}\mathcal{P}$ *hc*, $\mathcal{S}\mathcal{H}$ *hč*, $\mathcal{S}\mathcal{K}$ *hk*; the reform replaced these with unitary characters \mathfrak{P} *hp*, \mathfrak{D} *ht*, \mathfrak{C} *hc*, \mathfrak{C} *hč*, \mathfrak{K} *hk*.

The third reform involved the abolition of two ligatures (\mathfrak{S} *sts* for $\mathcal{S}\mathcal{P}$ *s-ts* and \mathfrak{K} *sk* for $\mathcal{S}\mathcal{K}$ *s-k*). Two other characters which represent palatalized (but evidently not phonemic) consonants \mathcal{S}^{\prime} [hʲ] and \mathfrak{K} [kʲ] have been retained as *s hy* and *k ky*.

The fourth reform was in the representation of nasality. Instead of representing nasality in the nasal vowels Λ^{\wedge} *a*, \mathcal{O}^{\wedge} *i*, \mathcal{O}^{\wedge} *o* differently from its representation in nasal diphthongs Δ *aj*, \mathcal{O} *ej*, \mathcal{O} *oj*, an intrinsic dot, to the top left or internally, is now used for all of them: \mathcal{A} *a*, \mathcal{I} *i*, \mathcal{O} *o*, \mathcal{A} *aj*, \mathcal{O} *ej*, \mathcal{O} *oj*.

The final reform was modification of the glyph for \mathfrak{X} *x/y* and splitting it into two characters \mathfrak{X} *x* and a new \mathfrak{Y} *y*. Compare $\mathfrak{X}\mathcal{O}\mathcal{P}\mathcal{A}$ *xp̄ce* ‘cedar’ with $\mathfrak{Y}\mathcal{O}\mathcal{P}\mathcal{A}$ *yp̄ce* ‘sloppy’.

While the community was confident in the reforms agreed, in February they understood the permanency of the UCS encoding process and have taken the intervening months to test out the reforms following this preliminary proposal for encoding. It has been recognized that the reforms respond appropriately to criticisms and queries about the orthography.

6. Encoding model. Differences between this proposal and the previous proposal in N4548 include the decomposition of the nasal vowels to base letters + U+0358, the addition of \mathfrak{C} *c* distinct from \mathfrak{C} *cʰ*, and the following unifications:

\mathcal{O} = U+004F LATIN CAPITAL LETTER O	\mathcal{o} = U+006F LATIN SMALL LETTER O
\mathcal{U} = U+0055 LATIN CAPITAL LETTER U	\mathcal{u} = U+0075 LATIN SMALL LETTER U
$\mathcal{\Lambda}$ = U+0245 LATIN CAPITAL LETTER TURNED V	$\mathcal{\lambda}$ = U+02BC LATIN SMALL LETTER TURNED V

These unifications follow from the decision to encode Osage characters as extensions to the Latin script rather than to encode Osage as a separate script on its own. Both possibilities were discussed at meetings of WG2 and the UTC, where opinions were somewhat divided amongst experts as to which of the two options were more advantageous to the user community. At meetings held in Pawhuska with the stakeholders on 28–29 July 2014, it was determined that the Latin extensions option was the most advantageous. The user community explicitly wishes to have their orthography participate in the full range of Latin typography, and judges that industry implementation of both fonts and related aspects of encoding will be expedited if the characters are considered to be what they are in fact historically: letters derived directly from Latin bases.

For the purpose of identifying potential confusables, note should be taken of letters which are similar to other Latin letters but which have not been unified with them. In the code chart below on page 14 the names list contains cross-references to these characters.

follow ξ ; \mathfrak{a} should follow s ; \mathcal{D} and \mathfrak{D} and \mathcal{P} and \mathfrak{P} and \mathcal{P} and \mathfrak{P} should follow τ ; \mathfrak{z} should follow \mathfrak{m} ; \mathfrak{h} and Ψ should follow χ ; \mathfrak{z} should follow z ; \mathfrak{z} should follow z .

8. Digits and punctuation. Osage makes use of standard European digits and punctuation.

9. Unicode Character Properties.

104B0;LATIN CAPITAL LETTER TURNED V WITH VERTICAL BAR;Lu;0;L;;;;N;;;104B1;
104B1;LATIN SMALL LETTER TURNED V WITH VERTICAL BAR;Ll;0;L;;;;N;;;104B0;;104B1
104B2;LATIN CAPITAL LETTER TURNED V WITH CENTRE DOT;Lu;0;L;;;;N;;;104B3;
104B3;LATIN SMALL LETTER TURNED V WITH CENTRE DOT;Ll;0;L;;;;N;;;104B2;;104B2
104B4;LATIN CAPITAL LETTER LAMBDA;Lu;0;L;;;;N;;;104B5;
104B5;LATIN SMALL LETTER LAMBDA;Ll;0;L;;;;N;;;104B4;;104B4
104B6;LATIN CAPITAL LETTER BRA;Lu;0;L;;;;N;;;104B7;
104B7;LATIN SMALL LETTER BRA;Ll;0;L;;;;N;;;104B6;;104B6
104B8;LATIN CAPITAL LETTER CHA;Lu;0;L;;;;N;;;104B9;
104B9;LATIN SMALL LETTER CHA;Ll;0;L;;;;N;;;104B8;;104B8
104BA;LATIN CAPITAL LETTER CHA WITH STROKE;Lu;0;L;;;;N;;;104BB;
104BB;LATIN SMALL LETTER CHA WITH STROKE;Ll;0;L;;;;N;;;104BA;;104BA
104BC;LATIN CAPITAL LETTER OSAGE E;Lu;0;L;;;;N;;;104BD;
104BD;LATIN SMALL LETTER OSAGE E;Ll;0;L;;;;N;;;104BC;;104BC
104BE;LATIN CAPITAL LETTER OSAGE E WITH CENTRE DOT;Lu;0;L;;;;N;;;104BF;
104BF;LATIN SMALL LETTER OSAGE E WITH CENTRE DOT;Ll;0;L;;;;N;;;104BE;;104BE
104C0;LATIN CAPITAL LETTER HA;Lu;0;L;;;;N;;;104C1;
104C1;LATIN SMALL LETTER HA;Ll;0;L;;;;N;;;104C0;;104C0
104C2;LATIN CAPITAL LETTER HA WITH VERTICAL STROKE;Lu;0;L;;;;N;;;104C3;
104C3;LATIN SMALL LETTER HA WITH VERTICAL STROKE;Ll;0;L;;;;N;;;104C2;;104C2
104C4;LATIN CAPITAL LETTER OSAGE I;Lu;0;L;;;;N;;;104C5;
104C5;LATIN SMALL LETTER OSAGE I;Ll;0;L;;;;N;;;104C4;;104C4
104C6;LATIN CAPITAL LETTER K WITH LOW HOOK;Lu;0;L;;;;N;;;104C7;
104C7;LATIN SMALL LETTER K WITH LOW HOOK;Ll;0;L;;;;N;;;104C6;;104C6
104C8;LATIN CAPITAL LETTER K WITH CURL;Lu;0;L;;;;N;;;104C9;
104C9;LATIN SMALL LETTER K WITH CURL;Ll;0;L;;;;N;;;104C8;;104C8
104CA;LATIN CAPITAL LETTER K WITH LOW HOOK AND STROKE;Lu;0;L;;;;N;;;104CB;
104CB;LATIN SMALL LETTER K WITH LOW HOOK AND STROKE;Ll;0;L;;;;N;;;104CA;;104CA
104CC;LATIN CAPITAL LETTER LA;Lu;0;L;;;;N;;;104CD;
104CD;LATIN SMALL LETTER LA;Ll;0;L;;;;N;;;104CC;;104CC
104CE;LATIN CAPITAL LETTER MA;Lu;0;L;;;;N;;;104CF;
104CF;LATIN SMALL LETTER MA;Ll;0;L;;;;N;;;104CE;;104CE
104D0;LATIN CAPITAL LETTER NA;Lu;0;L;;;;N;;;104D1;
104D1;LATIN SMALL LETTER NA;Ll;0;L;;;;N;;;104D0;;104D0
104D2;LATIN CAPITAL LETTER O WITH CENTRE DOT;Lu;0;L;;;;N;;;104D3;
104D3;LATIN SMALL LETTER O WITH CENTRE DOT;Ll;0;L;;;;N;;;104D2;;104D2
104D4;LATIN CAPITAL LETTER OSAGE PA;Lu;0;L;;;;N;;;104D5;
104D5;LATIN SMALL LETTER OSAGE PA;Ll;0;L;;;;N;;;104D4;;104D4
104D6;LATIN CAPITAL LETTER OSAGE PA WITH STROKE;Lu;0;L;;;;N;;;104D7;
104D7;LATIN SMALL LETTER OSAGE PA WITH STROKE;Ll;0;L;;;;N;;;104D6;;104D6
104D8;LATIN CAPITAL LETTER SA;Lu;0;L;;;;N;;;104D9;
104D9;LATIN SMALL LETTER SA;Ll;0;L;;;;N;;;104D8;;104D8
104DA;LATIN CAPITAL LETTER SHA;Lu;0;L;;;;N;;;104DB;
104DB;LATIN SMALL LETTER SHA;Ll;0;L;;;;N;;;104DA;;104DA
104DC;LATIN CAPITAL LETTER TA;Lu;0;L;;;;N;;;104DD;
104DD;LATIN SMALL LETTER TA;Ll;0;L;;;;N;;;104DC;;104DC
104DE;LATIN CAPITAL LETTER TA WITH STROKE;Lu;0;L;;;;N;;;104DF;
104DF;LATIN SMALL LETTER TA WITH STROKE;Ll;0;L;;;;N;;;104DE;;104DE
104E0;LATIN CAPITAL LETTER TSA;Lu;0;L;;;;N;;;104E1;
104E1;LATIN SMALL LETTER TSA;Ll;0;L;;;;N;;;104E0;;104E0
104E2;LATIN CAPITAL LETTER TSA WITH STROKE;Lu;0;L;;;;N;;;104E3;
104E3;LATIN SMALL LETTER TSA WITH STROKE;Ll;0;L;;;;N;;;104E2;;104E2
104E4;LATIN CAPITAL LETTER TSA WITH LOW STROKE;Lu;0;L;;;;N;;;104E5;
104E5;LATIN SMALL LETTER TSA WITH LOW STROKE;Ll;0;L;;;;N;;;104E4;;104E4
104E6;LATIN CAPITAL LETTER DHA;Lu;0;L;;;;N;;;104E7;
104E7;LATIN SMALL LETTER DHA;Ll;0;L;;;;N;;;104E6;;104E6
104E8;LATIN CAPITAL LETTER WA;Lu;0;L;;;;N;;;104E9;
104E9;LATIN SMALL LETTER WA;Ll;0;L;;;;N;;;104E8;;104E8
104EA;LATIN CAPITAL LETTER KHA;Lu;0;L;;;;N;;;104EB;
104EB;LATIN SMALL LETTER KHA;Ll;0;L;;;;N;;;104EA;;104EA
104EC;LATIN CAPITAL LETTER GHA;Lu;0;L;;;;N;;;104ED;
104ED;LATIN SMALL LETTER GHA;Ll;0;L;;;;N;;;104EC;;104EC
104EE;LATIN CAPITAL LETTER ZA;Lu;0;L;;;;N;;;104EF;
104EF;LATIN SMALL LETTER ZA;Ll;0;L;;;;N;;;104EE;;104EE
104F0;LATIN CAPITAL LETTER ZHA;Lu;0;L;;;;N;;;104F1;
104F1;LATIN SMALL LETTER ZHA;Ll;0;L;;;;N;;;104F0;;104F0

10. References.

Quintero, Carolyn. 2009. *Osage dictionary*. Norman: University of Oklahoma Press. ISBN 978-0-8061-3844-2

Osage Nation Language Department. ᎠᎩᎠᎩᎠ ᎠᎩ [wa.ža.že i.e]: *Osage Language Beginner Lessons*.

11. Acknowledgements. This project was made possible in part by support from the organization Friends of Osage Language, Inc., and in part by support from the Language Department of the Osage Nation.

12. Figures.

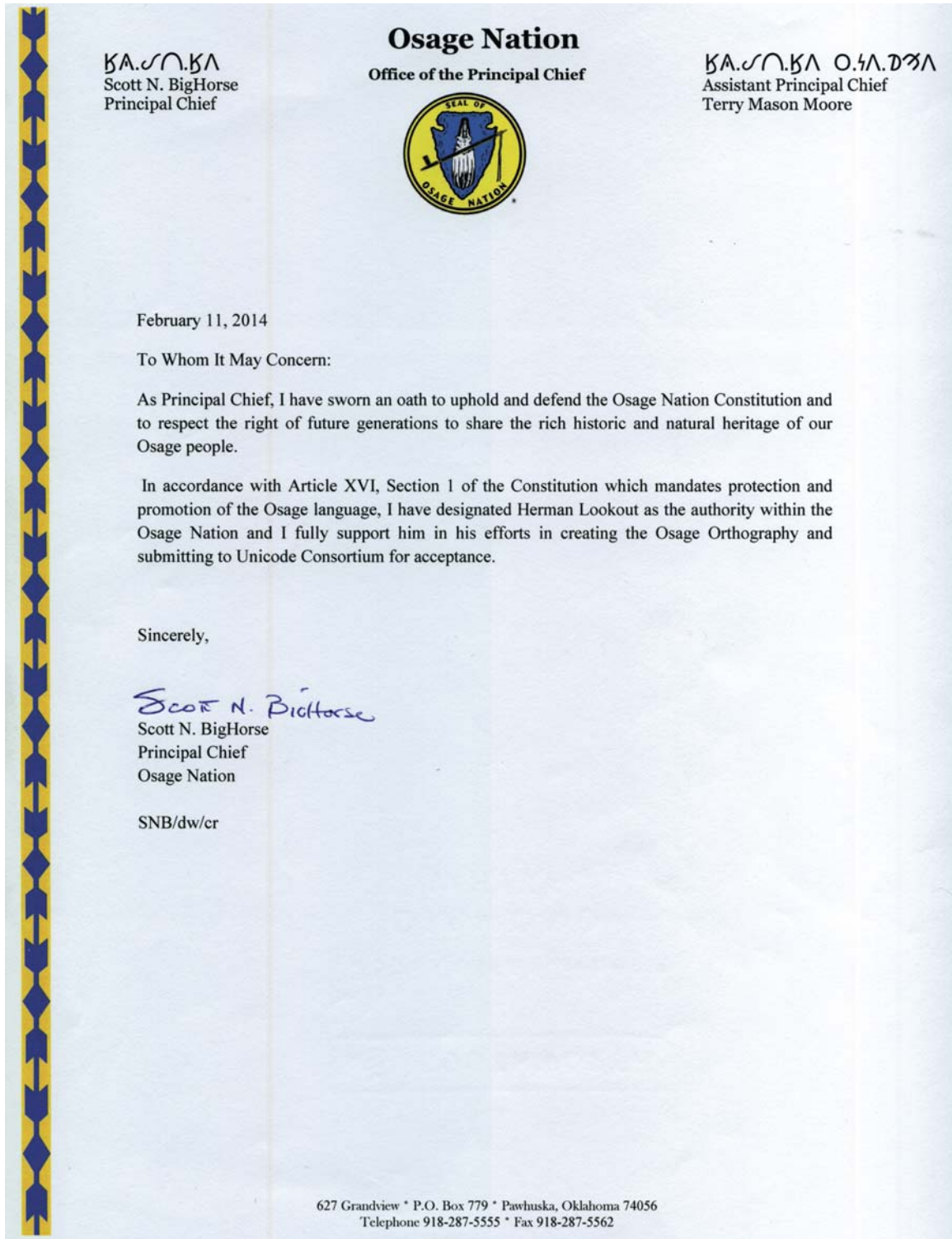


Figure 1. Letter from the Principal Chief of the Osage Nation detailing Herman Mongrain Lookout’s authority in the matter of Osage orthography and the UCS encoding.

Osage Orthography

Primaries

- Λ 'awe' sound as in the word, **aw**esome.
- ∩ long e sound as in the word, sl**ee**p.
- α short e sound as in the word, **e**gg.
- λ short u sound as in word, b**uh**!
- long o sound as in the word, **oh**.
- U 'oo' sound as in the word, **boot**.
- ^ nasal sound
- Δ blend of Λ and ∩^
- ⊙ blend of ○ and ∩^
- ⊘ blend of α and ∩^
- Λ̂ blend of Λ and ∩

Modifiers

- Ɔ similar to the 'X' sound, but guttural.
- ⊂ english S
- ⊄ english H
- ⊆ english W
- ⊇ english L
- ⊈ english M
- ⊉ english N
- ⊊ english Z
- ⊋ between the 'P' and 'B'
- ⊌ between the 'K' and 'G'
- ⊍ between the 'T' and 'D'
- ⊎ blend of T and H as in the word, **the**se.
- ⊏ blend of S and H as in the word, **sh**ip.
- ⊐ blend of C and H as in the word, **ch**ip.
- ⊑ blend of Z and H as in the word, **treas**ure.
- ⊒ blend of T and S together.
- ⊓ blend of P medial and R.
- ⊔ blend of S and K.
- ⊕ blend of K and Y.
- ⊖ blend of H and Y.
- ⊗ blend of S and D



Figure 2. Description of the earlier Osage orthography.



Figure 5. Exterior of the Osage Language Department showing the script in use. The text in red says *Ważaże ie* ‘Osage language’.



Figure 6. The Osage Language Department in Pawhuska, Oklahoma.



Figure 7. One of many drawings by children displayed in the Osage Language Department. Text reads: ᑕᐱ ᓂᓕᓂ ᑭᓕᓕᓕᓕ ᓕᓕᓕᓕ ᑭᓕᓕᓕᓕ ᑭᓕᓕᓕᓕ; in another font style ᑕᐱ ᓂᓕᓂ ᑭᓕᓕᓕᓕ ᓕᓕᓕᓕ ᑭᓕᓕᓕᓕ.



Figure 8. More drawings by children displayed in the Osage Language Department.

ᑭᑭᑭ ᑭᑭᑭᑭᑭ ᑭᑭ ᑭᑭᑭᑭᑭ
I want to address everyone as friends and relatives

ᑭᑭ ᑭᑭᑭ ᑭᑭᑭᑭᑭᑭ ᑭᑭ ᑭᑭᑭᑭᑭᑭ
I want to say a few words

ᑭᑭ ᑭᑭᑭᑭ ᑭᑭᑭᑭ ᑭᑭᑭᑭᑭᑭ
I am addressing you all.

ᑭᑭᑭ ᑭᑭᑭᑭᑭᑭ ᑭᑭ ᑭᑭᑭᑭᑭᑭᑭ
I want to address everyone as friends and relatives

ᑭᑭ ᑭᑭᑭ ᑭᑭᑭᑭᑭᑭᑭ ᑭᑭ ᑭᑭᑭᑭᑭᑭᑭᑭ
I want to say a few words

ᑭᑭ ᑭᑭᑭᑭ ᑭᑭᑭᑭ ᑭᑭᑭᑭᑭᑭᑭᑭ
I am addressing you all.

Figure 9. Above, text in the original standard “Official Osage font”; below, examples of Jessica Harjo’s experiments in quite graceful Osage typography, using the original orthography. The text reads: ᑭᑭᑭ ᑭᑭᑭᑭᑭ ᑭᑭ ᑭᑭᑭᑭᑭᑭᑭ. ᑭᑭ ᑭᑭᑭ ᑭᑭᑭᑭᑭᑭᑭ ᑭᑭ ᑭᑭᑭᑭᑭᑭᑭᑭ. ᑭᑭ ᑭᑭᑭᑭ ᑭᑭᑭᑭᑭᑭᑭ; in another font style: ᑭᑭᑭ ᑭᑭᑭᑭᑭᑭ ᑭᑭ ᑭᑭᑭᑭᑭᑭᑭᑭ. ᑭᑭ ᑭᑭᑭ ᑭᑭᑭᑭᑭᑭᑭ ᑭᑭ ᑭᑭᑭᑭᑭᑭᑭᑭᑭ. ᑭᑭ ᑭᑭᑭᑭ ᑭᑭᑭᑭᑭᑭᑭᑭᑭᑭ.



Figure 10. Typographic explorations in a heavy display face by Ryan RedCorn.

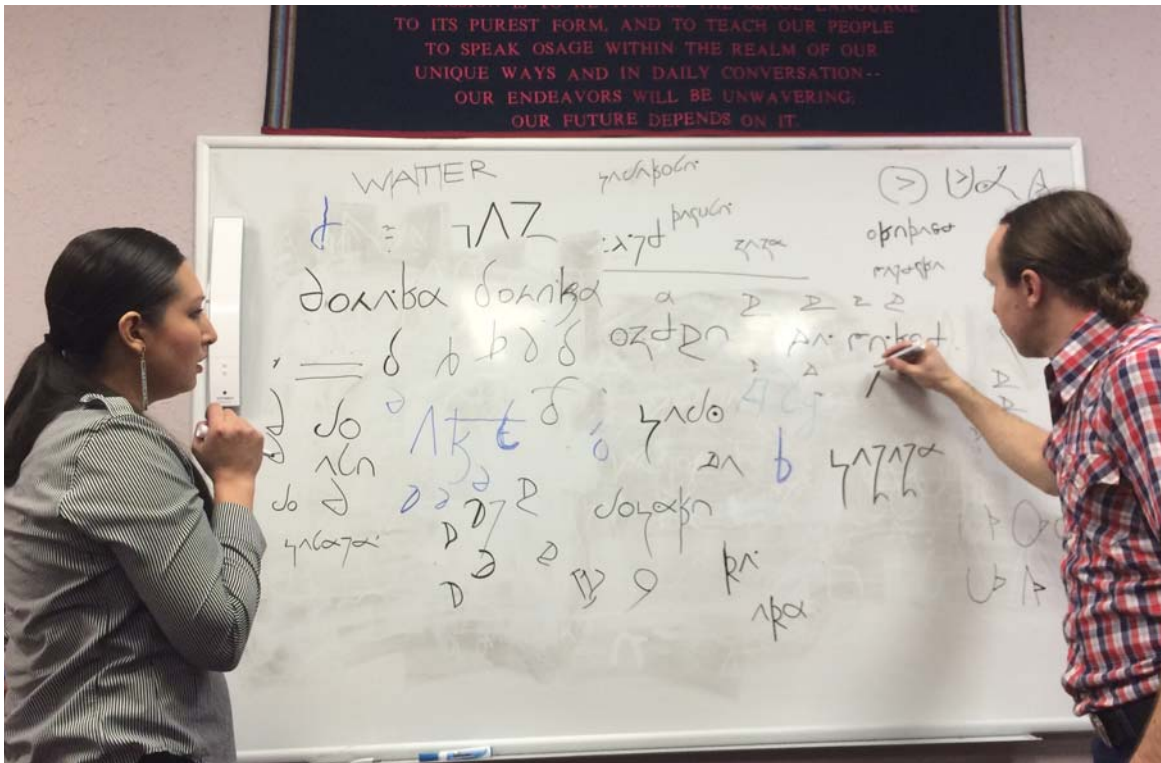


Figure 11. Jessica Harjo and Ryan RedCorn on the third day of the seminars in February 2014. Text here is casing, and the lower-case letters have ascenders and descenders.



Figure 12. Participants in the third day of seminars in February 2014. From left to right are Stephanie Rapp, Janis Carpenter, Michael Everson, Mogri Lookout, Jessica Harjo, Cameron Pratt, and Ryan RedCorn.

	104B	104C	104D	104E	104F
0	À 104B0	Á 104C0	Â 104D0	Ã 104E0	Ä 104F0
1	Å 104B1	Ā 104C1	Ă 104D1	Ȧ 104E1	Ⱥ 104F1
2	Ą 104B2	Ȧ 104C2	Ⱥ 104D2	Ȧ 104E2	
3	Ȧ 104B3	Ȧ 104C3	Ȧ 104D3	Ȧ 104E3	
4	Ȧ 104B4	Ȧ 104C4	Ȧ 104D4	Ȧ 104E4	
5	Ȧ 104B5	Ȧ 104C5	Ȧ 104D5	Ȧ 104E5	
6	Ȧ 104B6	Ȧ 104C6	Ȧ 104D6	Ȧ 104E6	
7	Ȧ 104B7	Ȧ 104C7	Ȧ 104D7	Ȧ 104E7	
8	Ȧ 104B8	Ȧ 104C8	Ȧ 104D8	Ȧ 104E8	
9	Ȧ 104B9	Ȧ 104C9	Ȧ 104D9	Ȧ 104E9	
A	Ȧ 104BA	Ȧ 104CA	Ȧ 104DA	Ȧ 104EA	
B	Ȧ 104BB	Ȧ 104CB	Ȧ 104DB	Ȧ 104EB	
C	Ȧ 104BC	Ȧ 104CC	Ȧ 104DC	Ȧ 104EC	
D	Ȧ 104BD	Ȧ 104CD	Ȧ 104DD	Ȧ 104ED	
E	Ȧ 104BE	Ȧ 104CE	Ȧ 104DE	Ȧ 104EE	
F	Ȧ 104BF	Ȧ 104CF	Ȧ 104DF	Ȧ 104EF	

Letters for Osage orthography

104B0	Λ	LATIN CAPITAL LETTER TURNED V WITH VERTICAL BAR	104D5	Ɔ	LATIN SMALL LETTER OSAGE PA
104B1	λ	LATIN SMALL LETTER TURNED V WITH VERTICAL BAR = Osage ai	104D6	Ɔ̄	LATIN CAPITAL LETTER OSAGE PA WITH STROKE → A764 Ɔ̄ latin capital letter thorn with stroke
104B2	Λ̣	LATIN CAPITAL LETTER TURNED V WITH CENTRE DOT	104D7	Ɔ̄	LATIN SMALL LETTER OSAGE PA WITH STROKE = Osage pre-aspirate or geminate pa
104B3	λ̣	LATIN SMALL LETTER TURNED V WITH CENTRE DOT = Osage nasal ai	104D8	C	LATIN CAPITAL LETTER SA → 0043 C latin capital letter c
104B4	Λ	LATIN CAPITAL LETTER LAMBDA	104D9	ç	LATIN SMALL LETTER SA
104B5	λ	LATIN SMALL LETTER LAMBDA = Osage schwa → 019B λ̇ latin small letter lambda with stroke	104DA	Š	LATIN CAPITAL LETTER SHA
104B6	Ř	LATIN CAPITAL LETTER BRA → 01A6 Ř latin capital letter yr	104DB	š	LATIN SMALL LETTER SHA
104B7	ř	LATIN SMALL LETTER BRA → 0280 ř latin letter small capital r	104DC	Đ	LATIN CAPITAL LETTER TA → 0044 Đ latin capital letter d
104B8	Ġ	LATIN CAPITAL LETTER CHA	104DD	đ	LATIN SMALL LETTER TA
104B9	ġ	LATIN SMALL LETTER CHA	104DE	Đ̄	LATIN CAPITAL LETTER TA WITH STROKE → 00D0 Đ̄ latin capital letter eth
104BA	Ġ̄	LATIN CAPITAL LETTER CHA WITH STROKE	104DF	đ̄	LATIN SMALL LETTER TA WITH STROKE = Osage pre-aspirate or geminate ta
104BB	ġ̄	LATIN SMALL LETTER CHA WITH STROKE = Osage pre-aspirate or geminate cha	104E0	Ț	LATIN CAPITAL LETTER TSA
104BC	Ɔ	LATIN CAPITAL LETTER OSAGE E → 2C6D Ɔ latin capital letter alpha	104E1	ț	LATIN SMALL LETTER TSA
104BD	α	LATIN SMALL LETTER OSAGE E → 0251 α latin small letter alpha	104E2	Ț̄	LATIN CAPITAL LETTER TSA WITH STROKE
104BE	Ɔ̣	LATIN CAPITAL LETTER OSAGE E WITH CENTRE DOT	104E3	ț̄	LATIN SMALL LETTER TSA WITH STROKE = Osage pre-aspirate or geminate tsa
104BF	α̣	LATIN SMALL LETTER OSAGE E WITH CENTRE DOT = Osage nasal ei	104E4	Ț̄̂	LATIN CAPITAL LETTER TSA WITH LOW STROKE
104C0	Ș	LATIN CAPITAL LETTER HA → 0053 Ș latin capital letter s	104E5	ț̄̂	LATIN SMALL LETTER TSA WITH LOW STROKE = Osage aspirate tsa
104C1	ș	LATIN SMALL LETTER HA	104E6	Ѡ	LATIN CAPITAL LETTER DHA → 040B Ѡ cyrillic capital letter tshe
104C2	Ș̄	LATIN CAPITAL LETTER HA WITH VERTICAL STROKE	104E7	ѡ	LATIN SMALL LETTER DHA • also used in Unifon orthography
104C3	ș̄	LATIN SMALL LETTER HA WITH VERTICAL STROKE = Osage hya	104E8	Ț̄	LATIN CAPITAL LETTER WA
104C4	Ω	LATIN CAPITAL LETTER OSAGE I	104E9	ț̄	LATIN SMALL LETTER WA
104C5	ω	LATIN SMALL LETTER OSAGE I	104EA	Ѡ̄	LATIN CAPITAL LETTER KHA
104C6	Ɔ̆	LATIN CAPITAL LETTER K WITH LOW HOOK	104EB	ѡ̄	LATIN SMALL LETTER KHA
104C7	Ɔ̆̄	LATIN SMALL LETTER K WITH LOW HOOK = Osage ka	104EC	Ψ	LATIN CAPITAL LETTER GHA → 03A8 Ψ greek capital letter psi
104C8	Ɔ̆̇	LATIN CAPITAL LETTER K WITH CURL	104ED	ψ	LATIN SMALL LETTER GHA
104C9	Ɔ̆̇̄	LATIN SMALL LETTER K WITH CURL = Osage kya	104EE	Ț̄̇	LATIN CAPITAL LETTER ZA
104CA	Ɔ̆̇̄̂	LATIN CAPITAL LETTER K WITH LOW HOOK AND STROKE	104EF	z̄	LATIN SMALL LETTER ZA
104CB	Ɔ̆̇̄̂̄	LATIN SMALL LETTER K WITH LOW HOOK AND STROKE = Osage pre-aspirate or geminate ka	104F0	Ț̄̇̂	LATIN CAPITAL LETTER ZHA → 0224 Ț̄̇̂ latin capital letter z with hook → 2C6B Ț̄̇̂ latin capital letter z with descender
104CC	C	LATIN CAPITAL LETTER LA → 0047 G latin capital letter g	104F1	z̄̂	LATIN SMALL LETTER ZHA
104CD	c	LATIN SMALL LETTER LA			
104CE	Ɔ̇	LATIN CAPITAL LETTER MA			
104CF	Ɔ̇̄	LATIN SMALL LETTER MA			
104D0	Λ̣̇	LATIN CAPITAL LETTER NA			
104D1	λ̣̇̄	LATIN SMALL LETTER NA			
104D2	Ɔ̣̇	LATIN CAPITAL LETTER O WITH CENTRE DOT			
104D3	ọ̇	LATIN SMALL LETTER O WITH CENTRE DOT = Osage nasal oi			
104D4	Ɔ̣̇̄	LATIN CAPITAL LETTER OSAGE PA → 00DE Ɔ̣̇̄ latin capital letter thorn			

A. Administrative

1. Title

Proposal to encode the Osage script in the UCS

2. Requester's name

Michael Everson, Herman Mongrain Lookout, Cameron Pratt

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

Individual contribution.

4. Submission date

2014-07-30

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

No.

6b. More information will be provided later

Yes.

B. Technical – General

1. Choose one of the following:

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

No.

1b. Proposed name of script

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

Yes.

1d. Name of the existing block

A new block called Latin Extended-F.

2. Number of characters in proposal

50.

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 N4548

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?

Members of Osage governmental and educational authorities, and others: Debra Atterberry, Janis Carpenter, Jessica Harjo, Mary Linn, Herman Lookout, Ted Moore, Terry Mason Moore, Veronica Pipestem, Cameron Pratt, Billy Proctor, Stephanie Rapp, Ryan RedCorn, Ed Shaw, Geoffrey StandingBear, Joe Tillman, David Webb.

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?

Osage speakers and writers.

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

In modern use.

4b. Reference

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

Yes.

5b. If YES, where?

Various publications, many educational.

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

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

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. The script derives from Latin, but is not unifiable with Latin.

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?

No.

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)?

Yes.

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

No.

11c. If YES, reference

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

Yes.

11e. If YES, reference

Existing combining characters are used with Osage vowels to indicate length and pitch accent.

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?