

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

**Doc Type:** Working Group Document  
**Title:** Preliminary proposal to encode the Osage script in the UCS  
**Source:** Michael Everson, Herman Mongrain Lookout, Cameron Pratt  
**Status:** Individual Contribution  
**Date:** 2014-02-20

**1. Introduction.** This document requests the addition of the Osage script 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 Osage script.** 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 which differed from student to student and which lacked 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 ultimately based on Latin letters or fusions of Latin letters, but which were treated as a different script entirely. The words above were written ᏚᏊᏃᏉ ‘dog’, ᏚᏊᏃᏉ ‘house’, ᏚᏊᏃᏉ ‘deer’, ᏚᏊᏃᏉ ‘turtle’.

The Osage script was take 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 script 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 script), 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. Structure of the Osage script.** The Osage script is a casing alphabet which is written from left to right.

**3.1. 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	pʷ pʷ pʷ		cʷ tsʷ ɬʷ			kʷ kʷ kʷ	(ʔ)(ʔ)–
Lax voiceless stops	p p p		t t ɖ		č tʃ ɕ	k k k	
Affricates	c ts ɬ						
Preaspirated/tense voiceless stops	hp ʰp p̥		ht ʰt ɖ̥		hč ʰtʃ ɕ̥	hk ʰk k̥	
Voiced stops	br bɹ ɹ						
(Post)aspirated stops	pš pʃ pʃ̥		ch tsʰ ɬʰ			kš kʃ kʃ̥	
	px px p̥ʰ		tx tx ɖʰ			kx kx k̥ʰ	
Voiceless/tense fric.			s s ɕ		š ʃ ɕ	x x ɰ	h h ʂ
Voiced/lax fricatives			z z ʒ		ž ʒ ʒ	ɣ ɣ ʁ	
Palatalized						kʲ kʲ k̥	hʲ hʲ ʂ
Nasals	m m ɱ		n n ɳ				
Approximants	w w ɹ		l l ɭ				
			ð ð ɳ				

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

Λ Λ A a, Ω Ω E e, Ω Ω I i, O o O o, U u U u; Λ Λ A a, Ω Ω I i, O o O o

This diacritic is not a combining mark and cannot stand alone; three diphthongs use a similar dot inside the character (also not a combining character); there is also one oral diphthong:

Λ Λ aḡ aḡ, Ω Ω Eḡ eḡ, O o Oḡ oḡ; Λ Λ Ai ai

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

Λ̄ Λ̄ Ā ā, Ω̄ Ω̄ Ē ē, Ω̄ Ω̄ Ī ī, Ō Ō Ō Ō Ū Ū Ū Ū; Λ̄ Λ̄ Ā ā, Ω̄ Ω̄ Ī ī, Ō Ō Ō Ō

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

Λ̇ Λ̇ Á á, Ω̇ Ω̇ É é, Ω̇ Ω̇ Í í, Ó ó Ó ó, Ú ú Ú ú; Λ̇ Λ̇ Á á, Ω̇ Ω̇ Í í, Ó ó Ó ó  
 Á Á Áí áí, É É Éí éí, Ó Ó Óí óí; Λ̇ Λ̇ Áí áí

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

Λ̈ Λ̈ Á̈ á̈, Ω̈ Ω̈ É̈ é̈, Ω̈ Ω̈ Í̈ í̈, Ó̈ Ó̈ Ó̈ ó̈, Ú̈ Ú̈ Ú̈ ú̈; Λ̈ Λ̈ Á̈ á̈, Ω̈ Ω̈ Í̈ í̈, Ó̈ Ó̈ Ó̈

**4. 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 script was “mature” enough to acquire and



**7. Digits and punctuation.** Osage makes use of standard European digits and punctuation.

**8. Unicode Character Properties.**

TO BE SUPPLIED

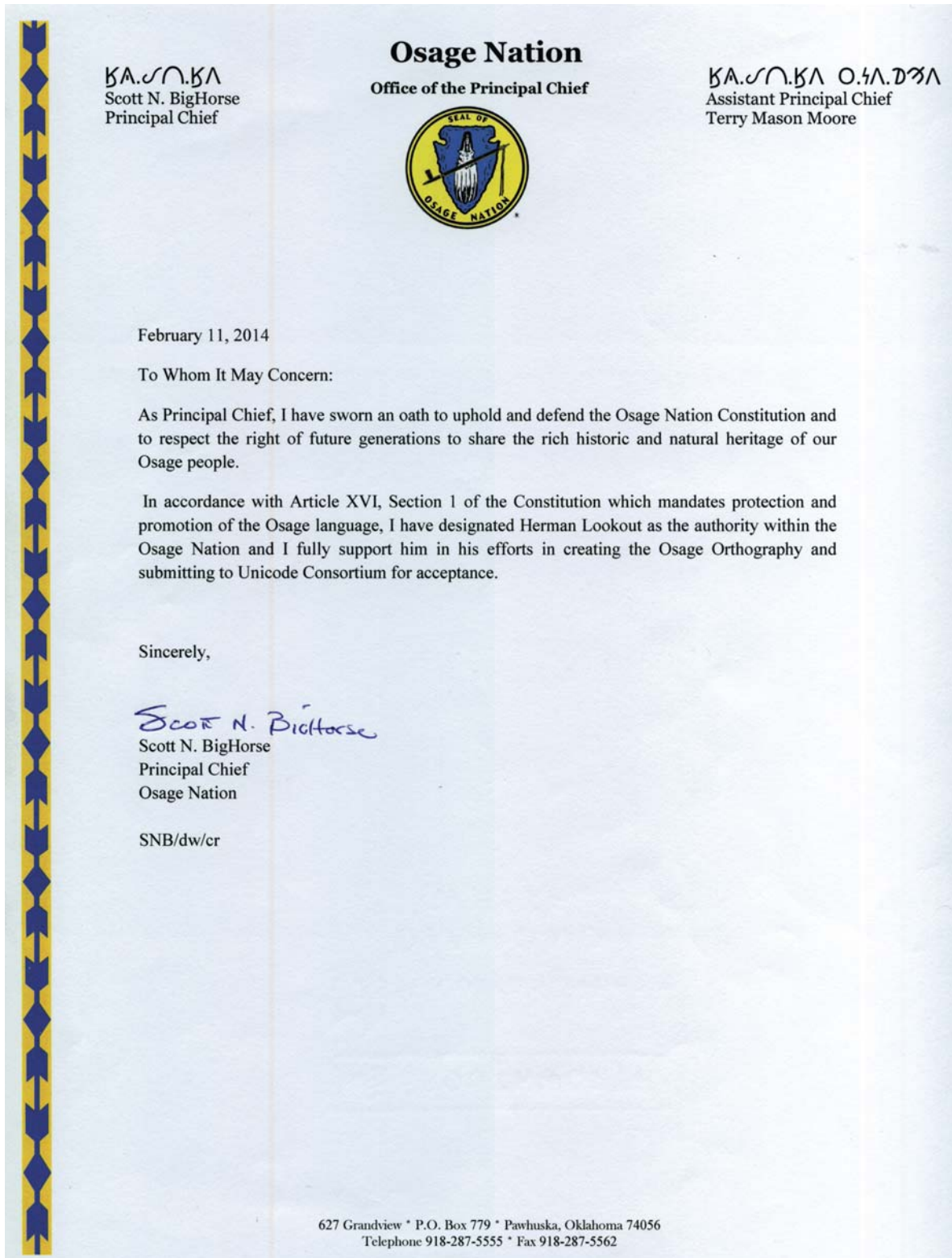
**9. References.**

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

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

**10. 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.

11. 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, **egg**.
- λ short u sound as in word, **buh!**
- 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 ∩^
- Λ∩ blend of Λ and ∩

## Modifiers

- Ƴ similar to the 'X' sound, but guttural.
- Ꞛ english S
- ꞛ english H
- Ꞝ english W
- ꞝ english L
- Ꞟ english M
- ꞟ english N
- Ꞡ english Z
- ꞡ between the 'P' and 'B'
- Ꞣ between the 'K' and 'G'
- ꞣ between the 'T' and 'D'
- Ꞥ blend of T and H as in the word, **the**se.
- ꞥ blend of S and H as in the word, **sh**ip.
- Ꞧ blend of C and H as in the word, **ch**ip.
- ꞧ blend of Z and H as in the word, **treas**ure.
- Ꞩ blend of T and S together.
- ꞩ blend of ꞡ medial and R.
- Ɦ blend of S and K.
- Ɜ blend of K and Y.
- Ɡ blend of H and Y.
- Ɬ blend of S and D



Figure 2. Description of the earlier Osage orthography.

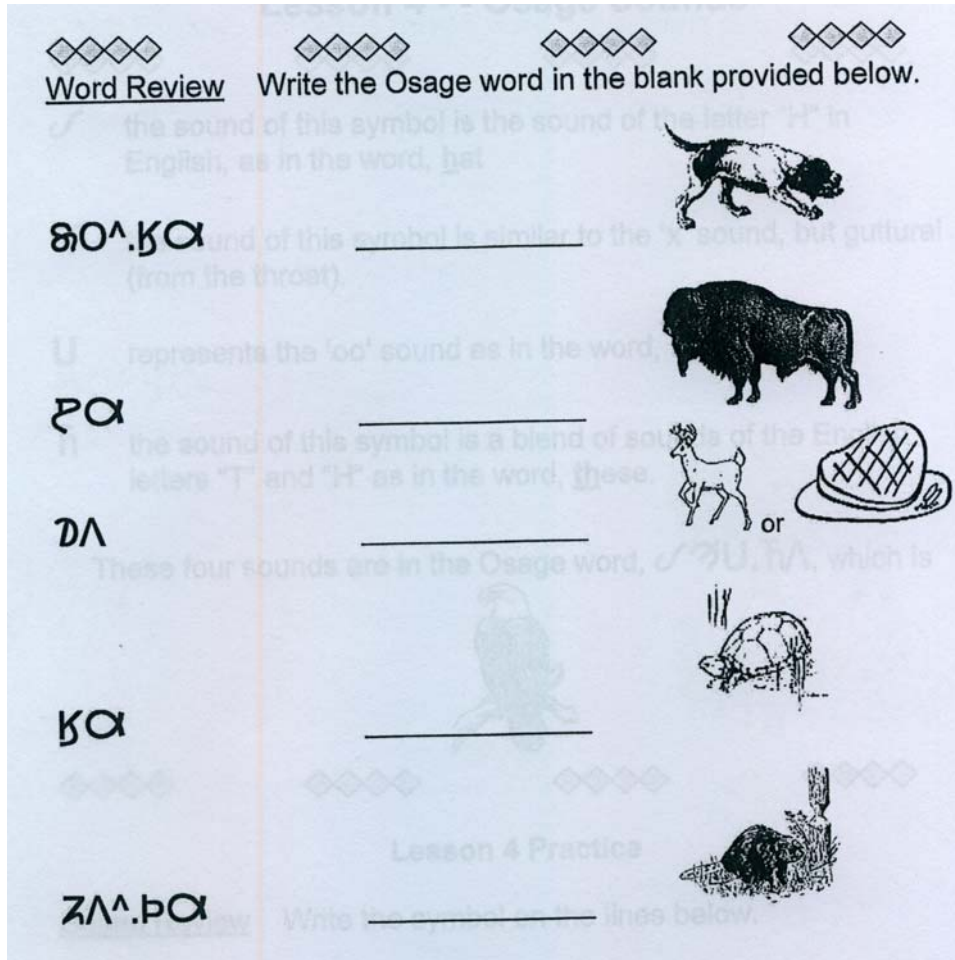


Figure 3. Example text from the beginning workbook for Osage language.

Osage Orthography Keyboard  
Reference Chart

Keyboard	6
Osage Orthography	^
Keyboard	W E T Y U I O P [ ] \
Osage Orthography	ʎ α Ⴑ Ⴑ U A O Q P Z Ⴑ Ⴑ
Keyboard	A a S D H J K L ; ' ,
Osage Orthography	^ Δ Ⴑ Ⴑ Ⴑ λ Ⴑ Ⴑ Ⴑ ' ,
Keyboard	Z X C V B N M , . /
Osage Orthography	7 Ⴑ Ⴑ Ⴑ R Ⴑ Ⴑ Ⴑ . Ⴑ

Figure 4. Scheme for the first Osage-script keyboard layout.



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

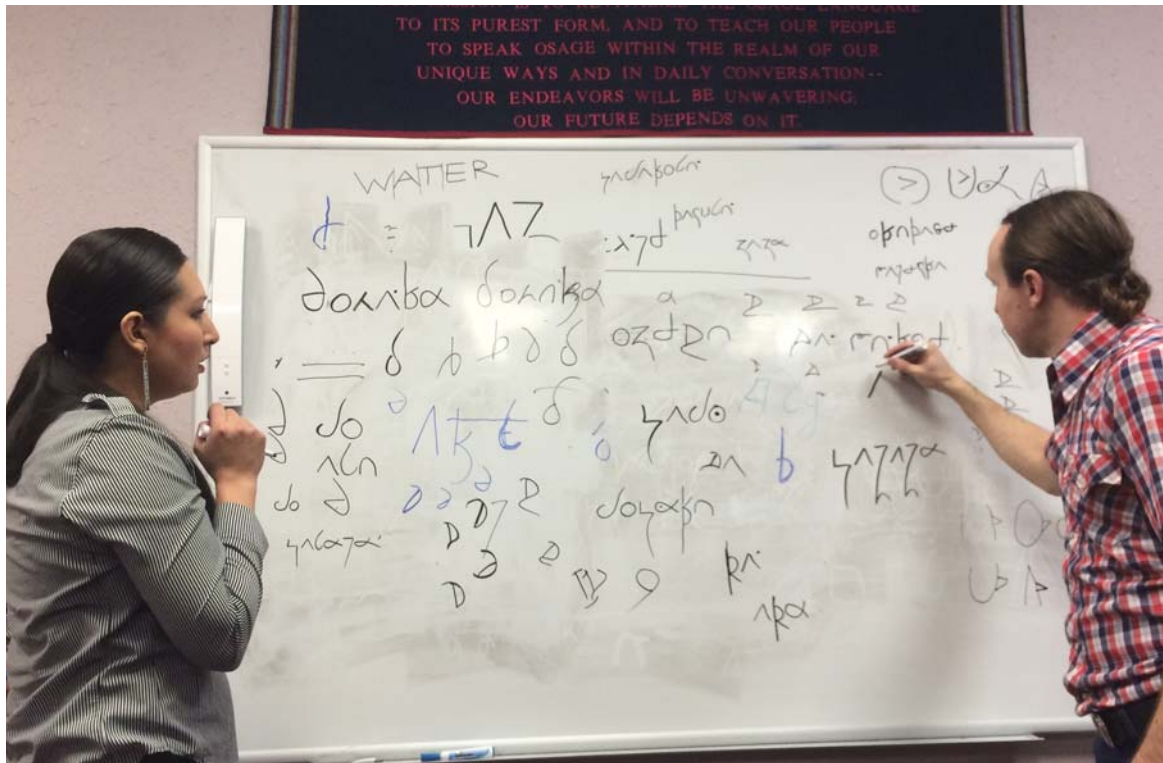


**Figure 6.** The Osage Language Department in Pawhuska, Oklahoma.









**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	Λ 104B0	Ɔ 104C0	Ʒ 104D0	α 104E0	Ɔ 104F0
1	Λ̇ 104B1	Ɔ̇ 104C1	Ʒ̇ 104D1	α̇ 104E1	Ɔ̇ 104F1
2	Λ̂ 104B2	Ɔ̂ 104C2	Ʒ̂ 104D2	α̂ 104E2	Ɔ̂ 104F2
3	Λ̃ 104B3	Ɔ̃ 104C3	Ʒ̃ 104D3	α̃ 104E3	Ɔ̃ 104F3
4	ƚ 104B4	Ɔ̇̂ 104C4	Ʒ̇̂ 104D4	α̇̂ 104E4	Ɔ̇̂ 104F4
5	ƚ̇ 104B5	Ɔ̇̂̇ 104C5		ƚ̇̂ 104E5	Ɔ̇̂̇ 104F5
6	ƚ̂ 104B6	Ɔ̂̇ 104C6		ƚ̂̇ 104E6	Ɔ̂̇ 104F6
7	ƚ̃ 104B7	Ɔ̃̇ 104C7		ƚ̃̇ 104E7	Ɔ̃̇ 104F7
8	α̇̂ 104B8	Ɔ̇̂ 104C8	Λ̇̂ 104D8	Ɔ̇̂ 104E8	Ʒ̇̂ 104F8
9	α̂̇ 104B9	Ɔ̂̇̇ 104C9	Λ̂̇̇ 104D9	Ɔ̂̇̇ 104E9	Ʒ̂̇̇ 104F9
A	α̂̇̂ 104BA	Ɔ̂̇̂ 104CA	Λ̂̇̂ 104DA	Ɔ̂̇̂ 104EA	Ʒ̂̇̂ 104FA
B	α̃̇̂ 104BB	Ɔ̃̇̂ 104CB	Λ̃̇̂ 104DB	Ɔ̃̇̂ 104EB	Ʒ̃̇̂ 104FB
C	α̃̇̂̇ 104BC	Ɔ̃̇̂̇ 104CC	ƚ̃̇̂̇ 104DC	Ɔ̃̇̂̇ 104EC	Ʒ̃̇̂̇ 104FC
D	ƚ̃̇̂̇̂ 104BD	Ɔ̃̇̂̇̂ 104CD	ƚ̃̇̂̇̂ 104DD	Ɔ̃̇̂̇̂ 104ED	
E	ƚ̃̇̂̇̂̇ 104BE	Ɔ̃̇̂̇̂̇ 104CE	ƚ̃̇̂̇̂̇ 104DE	Ɔ̃̇̂̇̂̇ 104EE	
F	ƚ̃̇̂̇̂̇̂ 104BF	Ɔ̃̇̂̇̂̇̂ 104CF	α̃̇̂̇̂̇̂ 104DF	Ɔ̃̇̂̇̂̇̂ 104EF	

**Capital letters**

104B0	Λ	OSAGE CAPITAL LETTER A
104B1	Λ̇	OSAGE CAPITAL LETTER AN
104B2	Λ̈	OSAGE CAPITAL LETTER AI
104B3	Λ̉	OSAGE CAPITAL LETTER AIN
104B4	ᚱ	OSAGE CAPITAL LETTER BRA
104B5	ᚦ	OSAGE CAPITAL LETTER CHA
104B6	ᚦ̇	OSAGE CAPITAL LETTER EHCHA
104B7	ᚺ	OSAGE CAPITAL LETTER E
104B8	ᚺ̇	OSAGE CAPITAL LETTER EIN
104B9	ᚾ	OSAGE CAPITAL LETTER HA
104BA	ᚾ̇	OSAGE CAPITAL LETTER HYA
104BB	ᚾ̈	OSAGE CAPITAL LETTER I
104BC	ᚾ̉	OSAGE CAPITAL LETTER IN
104BD	ᚿ	OSAGE CAPITAL LETTER KA
104BE	ᚿ̇	OSAGE CAPITAL LETTER KYA
104BF	ᚿ̈	OSAGE CAPITAL LETTER EHKA
104C0	ᚺ	OSAGE CAPITAL LETTER LA
104C1	ᚺ̇	OSAGE CAPITAL LETTER MA
104C2	ᚺ̈	OSAGE CAPITAL LETTER NA
104C3	ᚺ̉	OSAGE CAPITAL LETTER O
104C4	ᚺ̊	OSAGE CAPITAL LETTER ON
104C5	ᚺ̋	OSAGE CAPITAL LETTER OIN
104C6	ᚺ̌	OSAGE CAPITAL LETTER PA
104C7	ᚺ̍	OSAGE CAPITAL LETTER EHPA
104C8	ᚺ̎	OSAGE CAPITAL LETTER SA
104C9	ᚺ̏	OSAGE CAPITAL LETTER SHA
104CA	ᚺ̐	OSAGE CAPITAL LETTER TA
104CB	ᚺ̑	OSAGE CAPITAL LETTER EHTA
104CC	ᚺ̒	OSAGE CAPITAL LETTER TSA
104CD	ᚺ̓	OSAGE CAPITAL LETTER EHTSA
104CE	ᚺ̔	OSAGE CAPITAL LETTER DHA
104CF	ᚺ̕	OSAGE CAPITAL LETTER U
104D0	ᚺ̖	OSAGE CAPITAL LETTER WA
104D1	ᚺ̗	OSAGE CAPITAL LETTER KHA
104D2	ᚺ̘	OSAGE CAPITAL LETTER GHA
104D3	ᚺ̙	OSAGE CAPITAL LETTER ZA
104D4	ᚺ̚	OSAGE CAPITAL LETTER ZHA

**Small letters**

104D8	λ	OSAGE SMALL LETTER A
104D9	λ̇	OSAGE SMALL LETTER AN
104DA	λ̈	OSAGE SMALL LETTER AI
104DB	λ̉	OSAGE SMALL LETTER AIN
104DC	ᚱ	OSAGE SMALL LETTER BRA
104DD	ᚦ	OSAGE SMALL LETTER CHA
104DE	ᚦ̇	OSAGE SMALL LETTER EHCHA
104DF	ᚺ	OSAGE SMALL LETTER E
104E0	ᚺ̇	OSAGE SMALL LETTER EIN
104E1	ᚾ	OSAGE SMALL LETTER HA
104E2	ᚾ̇	OSAGE SMALL LETTER HYA
104E3	ᚾ̈	OSAGE SMALL LETTER I
104E4	ᚾ̉	OSAGE SMALL LETTER IN
104E5	ᚿ	OSAGE SMALL LETTER KA
104E6	ᚿ̇	OSAGE SMALL LETTER KYA
104E7	ᚿ̈	OSAGE SMALL LETTER EHKA
104E8	ᚺ	OSAGE SMALL LETTER LA
104E9	ᚺ̇	OSAGE SMALL LETTER MA
104EA	ᚺ̈	OSAGE SMALL LETTER NA
104EB	ᚺ̉	OSAGE SMALL LETTER O
104EC	ᚺ̊	OSAGE SMALL LETTER ON
104ED	ᚺ̋	OSAGE SMALL LETTER OIN
104EE	ᚺ̌	OSAGE SMALL LETTER PA
104EF	ᚺ̍	OSAGE SMALL LETTER EHPA
104F0	ᚺ̎	OSAGE SMALL LETTER SA
104F1	ᚺ̏	OSAGE SMALL LETTER SHA
104F2	ᚺ̐	OSAGE SMALL LETTER TA

104F3	ᚺ̑	OSAGE SMALL LETTER EHTA
104F4	ᚺ̒	OSAGE SMALL LETTER TSA
104F5	ᚺ̓	OSAGE SMALL LETTER EHTSA
104F6	ᚺ̔	OSAGE SMALL LETTER DHA
104F7	ᚺ̕	OSAGE SMALL LETTER U
104F8	ᚺ̖	OSAGE SMALL LETTER WA
104F9	ᚺ̗	OSAGE SMALL LETTER KHA
104FA	ᚺ̘	OSAGE SMALL LETTER GHA
104FB	ᚺ̙	OSAGE SMALL LETTER ZA
104FC	ᚺ̚	OSAGE SMALL LETTER ZHA

## A. Administrative

1. Title

**Preliminary 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

**2014-02-20**

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)

**Yes.**

1b. Proposed name of script

**Osage.**

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

**74.**

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.

**No.**

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.**

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?