

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 the addition of Cherokee characters to the UCS
Source: Michael Everson and Durbin Feeling
Status: Individual Contribution
Replaces: N4487
Date: 2014-02-25

1. Introduction. This document requests a change of the encoding model for Cherokee from unicameral to bicameral, and in addition requests the addition of an archaic syllable *mv* to the repertoire.

1.1. Background. In the early 1990s when Cherokee was being encoded in the UCS, communication was not what it is today. Indeed much of the correspondence done concerning the encoding was even done by *fax*—hard as it is to believe now. In any case, it was not known at the time, either by character encoders or by the Cherokee users themselves, that the casing pairs which were found in certain important documents were required for encoding. It took some years after encoding before UCS-encoded Cherokee text became prevalent. Now that practical use has been made of the encoded script, Cherokee users have recognized that it was a mistake not to encode Cherokee as a bicameral script. The change of the encoding model for Cherokee from unicameral to bicameral should be prioritized as highly important. *The user community has requested this change and has identified it as an urgent requirement.*

2. The “missing” letter *mv*. Even when Cherokee was being first encoded, it was recognized that Sequoyah’s document (illustrating his original character designs alongside handwritten versions of the typographic symbols being designed for them) contained an extra letter; back in 1994 this was mis-identified by Michael Everson as “archaic *hv*” (most likely due to the comparison given with *Q* *hv* in the *Cherokee Phoenix*, 1828-08-06—see Figure 2). The actual reading of the character is *mv*, a very rare syllable in Cherokee usually represented by *ma* in standard orthography. In consultation with the Cherokee Nation Language Program staff, the typographic form *G* has been chosen to answer to the way in which Sequoyah drew the character. See Figure 1.

3. Bicamerality in Cherokee. Both capital letters and small letters are found in the Cherokee New Testament (1860) as well as in other documents, printed with lead-type throughout the 19th century and into the 20th. It is likely that the practice of unicameral Cherokee is in part a result first of the limitations of manual typewriters, as well as of early 8-bit encodings which followed the typewriter conventions. For a period of time through the middle of the 20th century the script was used much less than it had been previously. At present it is enjoying a UCS-based renaissance. Currently the Cherokee Nation’s Office of Translation, Language Technology Department, and Curriculum Development staff are regularly forced to use fancy text in order to increase the font size of letters in sentence-initial position and representing personal and place-names. They find this disruptive of

their work-flow and impractical as fancy-text is less stable than plain text. Many database programs do not permit text in two point sizes in a database field, for instance.

3.1. Treatment of current Cherokee text. At present Cherokee text is encoded as caseless, with the property “Lo” (Letter Other) rather than “Lu” (Letter Upper) or “Ll” (Letter Lower) . There has been considerable discussion on this topic with the Cherokee user community and with representatives of both WG2 and the UTC. If the encoding model for Cherokee is to be changed from unicameral to bicameral, existing text has to be changed from “Lo” to either “Lu” or “Ll”, and the new characters assigned whatever case is left over. There are advantages and disadvantages to both changes.

3.1.1. Change of current Cherokee from “Lo” to “Ll”. Changing from “Lo” to “Ll” and adding a new “Lu” set of characters would seem to be the expected way of handling this problem, because in general it is upper-case characters which are added for Latin (casing pairs for African scripts using IPA, for example). If this were done, than Cherokee case-folding would be the same as Latin, Greek, Cyrillic, Armenian, Glagolitic, and Deseret: from upper-case to lower-case. In IDNA, current Cherokee characters (which are PVALID) would be made lower-case, and would fold to new upper-case characters (which would *not* be PVALID).

There are disadvantages to declaring existing Cherokee characters to be lower-case, however. Existing fonts draw Cherokee letters at the same size as Latin capitals. Despite what was suggested in N4487 (the preliminary proposal superseded by the present document) it does not make sense to retain Cherokee lower-case characters at Latin caps-height size and to make Cherokee capitals reach to the ascender and descender lines, because in the common scenario of mixed English and Cherokee text there would be essentially three sizes, Latin x-height, Latin/Cherokee caps-height, and super-large Cherokee caps. This would look pretty ridiculous.

Alternatively one could specify that fonts be changed so that existing Cherokee glyphs would be made, say, 68% smaller, and new caps added at Latin caps height. But this would mean two things. First, it would invalidate all current Cherokee fonts, and there is no real mechanism for comprehensively getting rid of them. And second, even if all of them could be got rid of and replaced with fonts that made the lower-case letters smaller, it would have the major implication of transforming the layout of every currently existing document.

3.1.2. Change of current Cherokee from “Lo” to “Lu”. Changing from “Lo” to “Lu” and adding a new “Ll” set of characters would have the implication that Cherokee case-folding would be *different* from that of Latin, Greek, Cyrillic, Armenian, Glagolitic, and Deseret: it would instead be from lower-case to upper-case. In IDNA, current Cherokee characters (which are PVALID) would be made upper-case, and would fold to new lower-case characters (which would *not* be PVALID). This would be exceptional for IDNA, but it does not break anything in current practice, since currently PVALID characters would remain PVALID. Browser developers and other developers having to deal with IDNA would have to take special note of this (in the event a user typed in a URL with lower-case Cherokee).

Implications for the user in a word-processing environment would be that selecting text and applying the “all caps” style to it would work oppositely from the way it does for Latin. Typed-in Latin small letters grow with “all caps” style, and typed-in Latin capital letters stay the same size. For Cherokee it would be the opposite: Typed-in Cherokee capital letters would *shrink* with “all caps” style

(conceptually “all smalls” though the keyboard command would remain the same), and typed-in Cherokee small letters would stay the same size.

Implications for Cherokee font display is made much easier with this model, however. To add lower-case letters to an existing font is (technically) easy, and designers would not have to change the metrics of existing upper-case letters. The formatting of old documents would not change: basically everything would just *work*,

The disadvantage of treating currently encoded text as capital would mean that all currently encoded documents will, one day look like SHOUTING to readers, as the use of lower-case letters becomes more widespread. The same could be said for legacy Latin text written in 6-bit encodings, and for legal text which often written in all caps (for whatever reason).

3.1.3. Conclusion. After weighing the advantages and disadvantages in 3.1.1 and 3.1.2, the user community has decided to request the solution found in 3.1.2. Existing content and fonts will not go away; the fonts will be around for a long time. Implementors have no control over that. But implementors *do* control the treatment of casing behaviour in a user agent and can ensure that it does the right thing. Accordingly, our request is that WG2 and the UTC make the Cherokee script bicameral by making the existing Cherokee letters uppercase, in order to provide better backward compatibility with existing implementations.

4. Character encoding positions. In N4487 it was suggested that new Cherokee characters should be added to the SMP because there was little space left in the BMP (which had no free, unallocated blocks), and in order to preserve a constant offset for Cherokee case-pairs. While in general variable-length case-folding would not be problematic for many applications, it appears that for some others, such as certain database applications, it would be, at a deep enough level that it might not be possible to expect such folding to work (either in the short or long term, or ever), since a lesser-used script like Cherokee might not be reason enough to inspire corrections to that software, or to get such updates rolled out significantly. For this reason, the suggestion made recently by the U.S. National Body on the ballot for DIS ISO/IEC 10646 should be accepted. The boundaries of the Latin Extended-E block should be changed from 9 columns (U+AB30..U+ABBF) to 4 columns (U+AB30..U+AB6F), resulting in a block of 5 contiguous columns, which would enable a set of 80 characters to be encoded in the BMP.

We recommend retaining constant offset nevertheless, for the convenience of users of the code charts. Accordingly, CHEROKEE LETTER A through CHEROKEE LETTER YA will be encoded at U+13A0..U+13EF, and CHEROKEE SMALL LETTER A through CHEROKEE SMALL LETTER YA will be encoded at U+AB70..U+ABBF. The remaining letters CHEROKEE LETTER YE through CHEROKEE LETTER MV will be encoded at U+13F0..U+13F5, and CHEROKEE SMALL LETTER YE through CHEROKEE SMALL LETTER MV will be encoded at U+13F8..U+13D5.

5. Ordering. The letter *G Mv* should be inserted in its proper place in the syllabary, after *Y Mu*; Capital letters precede small letters as shown below.

6. Glyph representation of Cherokee upper- and lower-case. In at least some of the 19th-century material lower-case letters are represented by the “long primer” size and upper-case letters are represented by the “English” size of the lead-type fonts cut for Cherokee at the time. These are approximately 10 and 14 points respectively. Setting such text in with lead-type was simple enough, but on computers today unless the leading has been explicitly specified (not all applications can do this) mixing 14-point and 10-point type in the same paragraph can make the line-spacing look a little uneven.

[illegible]

¹ሊከሰው ሀገሩን ሲገባ፣ ወደ ሙሉ ሀገሩን ይመለስ ይቻላል፣ ወደ ሙሉ ሀገሩን ይመለስ ይቻላል። ²ሊከሰው ሀገሩን ሲገባ፣ ወደ ሙሉ ሀገሩን ይመለስ ይቻላል። ³ከሰ ለገሰ ሙሉ ሀገሩን ይመለስ ይቻላል። ⁴ሙሉ [ሀገሩን] ይመለስ ይቻላል። ወደ ሙሉ ይመለስ ይቻላል። ⁵ወደ ሙሉ ሀገሩን ይመለስ ይቻላል። ⁶ሀገሩን ይመለስ ይቻላል። ⁷ሙሉ ሀገሩን ይመለስ ይቻላል። ⁸ሀገሩን ይመለስ ይቻላል። ⁹ሙሉ ሀገሩን ይመለስ ይቻላል። ¹⁰ሙሉ ሀገሩን ይመለስ ይቻላል።

Note that a number of characters in this example have ascenders or descenders. Discussions with Cherokee users indicates that the use of such variation is found quite commonly in 19th-century handwriting, and is likely to be felt as natural to the script, although seeing it in type might be a surprise to some readers at first. Over the next period while, a certain amount of experimentation in this regard is to be expected. In the font used in the code chart, ascenders have been given to *o* *u*, *he*, *h ni*, *o que*, *o s*, *w ta*, *k tso*, and *o ya* (of which *h*, *w*, and *k* have not been attested per se in handwriting), and descenders have been given to *y gi*, *j gu*, *t hna*, *g nah*, *β ye*, *g yu* (of which only *j* has not been attested per se in handwriting).

7. Font style. It must be said that Cherokee users would like to have their fonts function with *italic* and **bold** and ***bold-italic*** styles just as Latin fonts do. To date specifically italic styles have never been designed (though 19th-century handwriting offers many opportunities for study); at a minimum an obliques Cherokee should be offered by font vendors.

8. Diacritical marks in Cherokee. It should be noted that a variety of combining diacritical marks are now used in Cherokee text. For ordinary readers and especially for children, U+0323 COMBINING DOT BELOW indicates shifts in consonant readings—such as voiced to voiceless, voiceless to voiced—and U+0331 COMBINING MACRON BELOW indicates the dropping of a vowel. (When a consonant is both shifted and has its vowel dropped, U+0324 COMBINING DIAERESIS BELOW is used.) In quite exceptional cases, nasalization can be indicated by U+03 combining tilde below. (Nasalization is common; marking it is not.) Diacritical marks above a letter indicate tones; these are typically used only by linguists and in lexicography.

[g] to [k]		ᵿ ke	ᵿ ki	ᵿ ko	ᵿ ku	ᵿ kv
[n] to [n]~[ᵿn]		ᵿ hne	ᵿ hni	ᵿ hno	ᵿ hnu	ᵿ hnv
[s] to [ᵿ]	ᵿ sha	ᵿ she	ᵿ shi	ᵿ sho	ᵿ shu	ᵿ shv
[d] to [t]				ᵿ to	ᵿ tu	ᵿ tv
[tᵿ] to [dᵿ]		ᵿ dle	ᵿ dli	ᵿ dlo	ᵿ dlu	ᵿ dlv
[ts]~[dᵿ] to [tᵿ]	ᵿ cha	ᵿ che	ᵿ chi	ᵿ cho	ᵿ chu	ᵿ chv
[w] to [ᵿ]	ᵿ wha	ᵿ whe	ᵿ whi	ᵿ who	ᵿ whu	ᵿ whv
[j] to [ᵿj]	ᵿ hya	ᵿ hye	ᵿ hyi	ᵿ hyo	ᵿ hyu	ᵿ hyv

Some additional modifiers have been discussed, for use in transliteration of names in fiction. For example, the non-Cherokee sound [r]~[ᵿ] might be represented by a dotted *l*-series:

[l] to [r]~[ᵿ]	ᵿ ra	ᵿ re	ᵿ ri	ᵿ ro	ᵿ ru	ᵿ rv
----------------	------	------	------	------	------	------

An example of these marks in use is in a spelling which is planned to be used in the forthcoming translation of *ᵿᵿᵿᵿ* (*Haquiti, The Hobbit*) in Cherokee, where a character's Sindarin name will be written not just Rᵿᵿᵿᵿ (*Elilonidi*) but Rᵿᵿᵿᵿᵿ (*Elrond*).

Tone marks above a letter, used to indicate pitch in technical transcription, follows the IPA conventions, and uses U+0304 COMBINING MACRON for “mid” (this is the default and need not be written), U+0302 COMBINING CIRCUMFLEX ACCENT for “falling”, U+030C COMBINING CARON for “rising”, U+0300 COMBINING GRAVE ACCENT for “low”, U+0301 COMBINING ACUTE ACCENT for “high”, and U+030B COMBINING DOUBLE ACUTE ACCENT for “super high”.

9. Timeliness. Since the identities of the characters in this proposal are not disputed by anyone (they are, after all, simply casing pairs of existing characters (plus *Mv/mv*), and since early implementation would be advantageous to the user community, we favour the placement of the Cherokee additions onto the earliest possible ballot.

10. Unicode Character Properties.

```

13A0;CHEROKEE LETTER A;Lu;0;L;;;;;N;;;;AB70;
13A1;CHEROKEE LETTER E;Lu;0;L;;;;;N;;;;AB71;
13A2;CHEROKEE LETTER I;Lu;0;L;;;;;N;;;;AB72;
13A3;CHEROKEE LETTER O;Lu;0;L;;;;;N;;;;AB73;
13A4;CHEROKEE LETTER U;Lu;0;L;;;;;N;;;;AB74;
13A5;CHEROKEE LETTER V;Lu;0;L;;;;;N;;;;AB75;
13A6;CHEROKEE LETTER GA;Lu;0;L;;;;;N;;;;AB76;
13A7;CHEROKEE LETTER KA;Lu;0;L;;;;;N;;;;AB77;
13A8;CHEROKEE LETTER GE;Lu;0;L;;;;;N;;;;AB78;
13A9;CHEROKEE LETTER GI;Lu;0;L;;;;;N;;;;AB79;
13AA;CHEROKEE LETTER GO;Lu;0;L;;;;;N;;;;AB7A;
13AB;CHEROKEE LETTER GU;Lu;0;L;;;;;N;;;;AB7B;
13AC;CHEROKEE LETTER GV;Lu;0;L;;;;;N;;;;AB7C;
13AD;CHEROKEE LETTER HA;Lu;0;L;;;;;N;;;;AB7D;
13AE;CHEROKEE LETTER HE;Lu;0;L;;;;;N;;;;AB7E;
13AF;CHEROKEE LETTER HI;Lu;0;L;;;;;N;;;;AB7F;
13B0;CHEROKEE LETTER HO;Lu;0;L;;;;;N;;;;AB80;
13B1;CHEROKEE LETTER HU;Lu;0;L;;;;;N;;;;AB81;
13B2;CHEROKEE LETTER HV;Lu;0;L;;;;;N;;;;AB82;
13B3;CHEROKEE LETTER LA;Lu;0;L;;;;;N;;;;AB83;
13B4;CHEROKEE LETTER LE;Lu;0;L;;;;;N;;;;AB84;
13B5;CHEROKEE LETTER LI;Lu;0;L;;;;;N;;;;AB85;
13B6;CHEROKEE LETTER LO;Lu;0;L;;;;;N;;;;AB86;
13B7;CHEROKEE LETTER LU;Lu;0;L;;;;;N;;;;AB87;
13B8;CHEROKEE LETTER LV;Lu;0;L;;;;;N;;;;AB88;
13B9;CHEROKEE LETTER MA;Lu;0;L;;;;;N;;;;AB89;
13BA;CHEROKEE LETTER ME;Lu;0;L;;;;;N;;;;AB8A;
13BB;CHEROKEE LETTER MI;Lu;0;L;;;;;N;;;;AB8B;
13BC;CHEROKEE LETTER MO;Lu;0;L;;;;;N;;;;AB8C;
13BD;CHEROKEE LETTER MU;Lu;0;L;;;;;N;;;;AB8D;
13BE;CHEROKEE LETTER NA;Lu;0;L;;;;;N;;;;AB8E;
13BF;CHEROKEE LETTER HNA;Lu;0;L;;;;;N;;;;AB8F;
13C0;CHEROKEE LETTER NAH;Lu;0;L;;;;;N;;;;AB90;
13C1;CHEROKEE LETTER NE;Lu;0;L;;;;;N;;;;AB91;
13C2;CHEROKEE LETTER NI;Lu;0;L;;;;;N;;;;AB92;
13C3;CHEROKEE LETTER NO;Lu;0;L;;;;;N;;;;AB93;
13C4;CHEROKEE LETTER NU;Lu;0;L;;;;;N;;;;AB94;
13C5;CHEROKEE LETTER NV;Lu;0;L;;;;;N;;;;AB95;
13C6;CHEROKEE LETTER QUA;Lu;0;L;;;;;N;;;;AB96;
13C7;CHEROKEE LETTER QUE;Lu;0;L;;;;;N;;;;AB97;
13C8;CHEROKEE LETTER QUI;Lu;0;L;;;;;N;;;;AB98;
13C9;CHEROKEE LETTER QUO;Lu;0;L;;;;;N;;;;AB99;
13CA;CHEROKEE LETTER QUU;Lu;0;L;;;;;N;;;;AB9A;
13CB;CHEROKEE LETTER QUV;Lu;0;L;;;;;N;;;;AB9B;
13CC;CHEROKEE LETTER SA;Lu;0;L;;;;;N;;;;AB9C;
13CD;CHEROKEE LETTER S;Lu;0;L;;;;;N;;;;AB9D;
13CE;CHEROKEE LETTER SE;Lu;0;L;;;;;N;;;;AB9E;
13CF;CHEROKEE LETTER SI;Lu;0;L;;;;;N;;;;AB9F;
13D0;CHEROKEE LETTER SO;Lu;0;L;;;;;N;;;;ABA0;
13D1;CHEROKEE LETTER SU;Lu;0;L;;;;;N;;;;ABA1;
13D2;CHEROKEE LETTER SV;Lu;0;L;;;;;N;;;;ABA2;
13D3;CHEROKEE LETTER DA;Lu;0;L;;;;;N;;;;ABA3;
13D4;CHEROKEE LETTER TA;Lu;0;L;;;;;N;;;;ABA4;
13D5;CHEROKEE LETTER DE;Lu;0;L;;;;;N;;;;ABA5;
13D6;CHEROKEE LETTER TE;Lu;0;L;;;;;N;;;;ABA6;
13D7;CHEROKEE LETTER DI;Lu;0;L;;;;;N;;;;ABA7;
13D8;CHEROKEE LETTER TI;Lu;0;L;;;;;N;;;;ABA8;
13D9;CHEROKEE LETTER DO;Lu;0;L;;;;;N;;;;ABA9;
13DA;CHEROKEE LETTER DU;Lu;0;L;;;;;N;;;;ABAA;
13DB;CHEROKEE LETTER DV;Lu;0;L;;;;;N;;;;ABAB;
13DC;CHEROKEE LETTER DLA;Lu;0;L;;;;;N;;;;ABAC;
13DD;CHEROKEE LETTER TLA;Lu;0;L;;;;;N;;;;ABAD;
13DE;CHEROKEE LETTER TLE;Lu;0;L;;;;;N;;;;ABAE;
13DF;CHEROKEE LETTER TLI;Lu;0;L;;;;;N;;;;ABAF;
13E0;CHEROKEE LETTER TLO;Lu;0;L;;;;;N;;;;ABB0;
13E1;CHEROKEE LETTER TLU;Lu;0;L;;;;;N;;;;ABB1;
13E2;CHEROKEE LETTER TLV;Lu;0;L;;;;;N;;;;ABB2;
13E3;CHEROKEE LETTER TSA;Lu;0;L;;;;;N;;;;ABB3;
13E4;CHEROKEE LETTER TSE;Lu;0;L;;;;;N;;;;ABB4;
13E5;CHEROKEE LETTER TSI;Lu;0;L;;;;;N;;;;ABB5;
13E6;CHEROKEE LETTER TSO;Lu;0;L;;;;;N;;;;ABB6;
13E7;CHEROKEE LETTER TSU;Lu;0;L;;;;;N;;;;ABB7;
13E8;CHEROKEE LETTER TSV;Lu;0;L;;;;;N;;;;ABB8;
13E9;CHEROKEE LETTER WA;Lu;0;L;;;;;N;;;;ABB9;
13EA;CHEROKEE LETTER WE;Lu;0;L;;;;;N;;;;ABBA;

```

13EB;CHEROKEE LETTER WI;Lu;0;L;;;;N;;;ABBB;
 13EC;CHEROKEE LETTER WO;Lu;0;L;;;;N;;;ABBC;
 13ED;CHEROKEE LETTER WU;Lu;0;L;;;;N;;;ABBD;
 13EE;CHEROKEE LETTER WV;Lu;0;L;;;;N;;;ABBE;
 13EF;CHEROKEE LETTER YA;Lu;0;L;;;;N;;;ABBF;
 13F0;CHEROKEE LETTER YE;Lu;0;L;;;;N;;;13F8;
 13F1;CHEROKEE LETTER YI;Lu;0;L;;;;N;;;13F9;
 13F2;CHEROKEE LETTER YO;Lu;0;L;;;;N;;;13FA;
 13F3;CHEROKEE LETTER YU;Lu;0;L;;;;N;;;13FB;
 13F4;CHEROKEE LETTER YV;Lu;0;L;;;;N;;;13FC;
 13F5;CHEROKEE LETTER MV;Lu;0;L;;;;N;;;13FD;
 13F8;CHEROKEE SMALL LETTER YE;Ll;0;L;;;;N;;;13F0;13F0
 13F9;CHEROKEE SMALL LETTER YI;Ll;0;L;;;;N;;;13F1;13F1
 13FA;CHEROKEE SMALL LETTER YO;Ll;0;L;;;;N;;;13F2;13F2
 13FB;CHEROKEE SMALL LETTER YU;Ll;0;L;;;;N;;;13F3;13F3
 13FC;CHEROKEE SMALL LETTER YV;Ll;0;L;;;;N;;;13F4;13F4
 13FD;CHEROKEE SMALL LETTER MV;Ll;0;L;;;;N;;;13F5;13F5
 AB70;CHEROKEE SMALL LETTER A;Ll;0;L;;;;N;;;13A0;13A0
 AB71;CHEROKEE SMALL LETTER E;Ll;0;L;;;;N;;;13A1;13A1
 AB72;CHEROKEE SMALL LETTER I;Ll;0;L;;;;N;;;13A2;13A2
 AB73;CHEROKEE SMALL LETTER O;Ll;0;L;;;;N;;;13A3;13A3
 AB74;CHEROKEE SMALL LETTER U;Ll;0;L;;;;N;;;13A4;13A4
 AB75;CHEROKEE SMALL LETTER V;Ll;0;L;;;;N;;;13A5;13A5
 AB76;CHEROKEE SMALL LETTER GA;Ll;0;L;;;;N;;;13A6;13A6
 AB77;CHEROKEE SMALL LETTER KA;Ll;0;L;;;;N;;;13A7;13A7
 AB78;CHEROKEE SMALL LETTER GE;Ll;0;L;;;;N;;;13A8;13A8
 AB79;CHEROKEE SMALL LETTER GI;Ll;0;L;;;;N;;;13A9;13A9
 AB7A;CHEROKEE SMALL LETTER GO;Ll;0;L;;;;N;;;13AA;13AA
 AB7B;CHEROKEE SMALL LETTER GU;Ll;0;L;;;;N;;;13AB;13AB
 AB7C;CHEROKEE SMALL LETTER GV;Ll;0;L;;;;N;;;13AC;13AC
 AB7D;CHEROKEE SMALL LETTER HA;Ll;0;L;;;;N;;;13AD;13AD
 AB7E;CHEROKEE SMALL LETTER HE;Ll;0;L;;;;N;;;13AE;13AE
 AB7F;CHEROKEE SMALL LETTER HI;Ll;0;L;;;;N;;;13AF;13AF
 AB80;CHEROKEE SMALL LETTER HO;Ll;0;L;;;;N;;;13B0;13B0
 AB81;CHEROKEE SMALL LETTER HU;Ll;0;L;;;;N;;;13B1;13B1
 AB82;CHEROKEE SMALL LETTER HV;Ll;0;L;;;;N;;;13B2;13B2
 AB83;CHEROKEE SMALL LETTER LA;Ll;0;L;;;;N;;;13B3;13B3
 AB84;CHEROKEE SMALL LETTER LE;Ll;0;L;;;;N;;;13B4;13B4
 AB85;CHEROKEE SMALL LETTER LI;Ll;0;L;;;;N;;;13B5;13B5
 AB86;CHEROKEE SMALL LETTER LO;Ll;0;L;;;;N;;;13B6;13B6
 AB87;CHEROKEE SMALL LETTER LL;Ll;0;L;;;;N;;;13B7;13B7
 AB88;CHEROKEE SMALL LETTER LV;Ll;0;L;;;;N;;;13B8;13B8
 AB89;CHEROKEE SMALL LETTER MA;Ll;0;L;;;;N;;;13B9;13B9
 AB8A;CHEROKEE SMALL LETTER ME;Ll;0;L;;;;N;;;13BA;13BA
 AB8B;CHEROKEE SMALL LETTER MI;Ll;0;L;;;;N;;;13BB;13BB
 AB8C;CHEROKEE SMALL LETTER MO;Ll;0;L;;;;N;;;13BC;13BC
 AB8D;CHEROKEE SMALL LETTER MU;Ll;0;L;;;;N;;;13BD;13BD
 AB8E;CHEROKEE SMALL LETTER NA;Ll;0;L;;;;N;;;13BE;13BE
 AB8F;CHEROKEE SMALL LETTER HNA;Ll;0;L;;;;N;;;13BF;13BF
 AB90;CHEROKEE SMALL LETTER NAH;Ll;0;L;;;;N;;;13C0;13C0
 AB91;CHEROKEE SMALL LETTER NE;Ll;0;L;;;;N;;;13C1;13C1
 AB92;CHEROKEE SMALL LETTER NI;Ll;0;L;;;;N;;;13C2;13C2
 AB93;CHEROKEE SMALL LETTER NO;Ll;0;L;;;;N;;;13C3;13C3
 AB94;CHEROKEE SMALL LETTER NU;Ll;0;L;;;;N;;;13C4;13C4
 AB95;CHEROKEE SMALL LETTER NV;Ll;0;L;;;;N;;;13C5;13C5
 AB96;CHEROKEE SMALL LETTER QUA;Ll;0;L;;;;N;;;13C6;13C6
 AB97;CHEROKEE SMALL LETTER QUE;Ll;0;L;;;;N;;;13C7;13C7
 AB98;CHEROKEE SMALL LETTER QUI;Ll;0;L;;;;N;;;13C8;13C8
 AB99;CHEROKEE SMALL LETTER QUO;Ll;0;L;;;;N;;;13C9;13C9
 AB9A;CHEROKEE SMALL LETTER QUU;Ll;0;L;;;;N;;;13CA;13CA
 AB9B;CHEROKEE SMALL LETTER QUV;Ll;0;L;;;;N;;;13CB;13CB
 AB9C;CHEROKEE SMALL LETTER SA;Ll;0;L;;;;N;;;13CC;13CC
 AB9D;CHEROKEE SMALL LETTER S;Ll;0;L;;;;N;;;13CD;13CD
 AB9E;CHEROKEE SMALL LETTER SE;Ll;0;L;;;;N;;;13CE;13CE
 AB9F;CHEROKEE SMALL LETTER SI;Ll;0;L;;;;N;;;13CF;13CF
 ABA0;CHEROKEE SMALL LETTER SO;Ll;0;L;;;;N;;;13D0;13D0
 ABA1;CHEROKEE SMALL LETTER SU;Ll;0;L;;;;N;;;13D1;13D1
 ABA2;CHEROKEE SMALL LETTER SV;Ll;0;L;;;;N;;;13D2;13D2
 ABA3;CHEROKEE SMALL LETTER DA;Ll;0;L;;;;N;;;13D3;13D3
 ABA4;CHEROKEE SMALL LETTER TA;Ll;0;L;;;;N;;;13D4;13D4
 ABA5;CHEROKEE SMALL LETTER DE;Ll;0;L;;;;N;;;13D5;13D5
 ABA6;CHEROKEE SMALL LETTER TE;Ll;0;L;;;;N;;;13D6;13D6
 ABA7;CHEROKEE SMALL LETTER DI;Ll;0;L;;;;N;;;13D7;13D7
 ABA8;CHEROKEE SMALL LETTER TI;Ll;0;L;;;;N;;;13D8;13D8
 ABA9;CHEROKEE SMALL LETTER DO;Ll;0;L;;;;N;;;13D9;13D9
 ABA;CHEROKEE SMALL LETTER DU;Ll;0;L;;;;N;;;13DA;13DA
 ABAB;CHEROKEE SMALL LETTER DV;Ll;0;L;;;;N;;;13DB;13DB
 ABAC;CHEROKEE SMALL LETTER DLA;Ll;0;L;;;;N;;;13DC;13DC
 ABAD;CHEROKEE SMALL LETTER TLA;Ll;0;L;;;;N;;;13DD;13DD
 ABAE;CHEROKEE SMALL LETTER TLE;Ll;0;L;;;;N;;;13DE;13DE
 ABAF;CHEROKEE SMALL LETTER TLI;Ll;0;L;;;;N;;;13DF;13DF
 ABB0;CHEROKEE SMALL LETTER TLO;Ll;0;L;;;;N;;;13E0;13E0
 ABB1;CHEROKEE SMALL LETTER TLL;Ll;0;L;;;;N;;;13E1;13E1
 ABB2;CHEROKEE SMALL LETTER TLV;Ll;0;L;;;;N;;;13E2;13E2
 ABB3;CHEROKEE SMALL LETTER TSA;Ll;0;L;;;;N;;;13E3;13E3
 ABB4;CHEROKEE SMALL LETTER TSE;Ll;0;L;;;;N;;;13E4;13E4

ABB5;CHEROKEE SMALL LETTER TSI;L1;0;L;;;N;;;13E5;13E5
 ABB6;CHEROKEE SMALL LETTER TSO;L1;0;L;;;N;;;13E6;13E6
 ABB7;CHEROKEE SMALL LETTER TSU;L1;0;L;;;N;;;13E7;13E7
 ABB8;CHEROKEE SMALL LETTER TSV;L1;0;L;;;N;;;13E8;13E8
 ABB9;CHEROKEE SMALL LETTER WA;L1;0;L;;;N;;;13E9;13E9
 ABBA;CHEROKEE SMALL LETTER WE;L1;0;L;;;N;;;13EA;13EA
 ABBB;CHEROKEE SMALL LETTER WI;L1;0;L;;;N;;;13EB;13EB
 ABBC;CHEROKEE SMALL LETTER WO;L1;0;L;;;N;;;13EC;13EC
 ABBD;CHEROKEE SMALL LETTER WU;L1;0;L;;;N;;;13ED;13ED
 ABBE;CHEROKEE SMALL LETTER WV;L1;0;L;;;N;;;13EE;13EE
 ABBF;CHEROKEE SMALL LETTER YA;L1;0;L;;;N;;;13EF;13EF

11. Acknowledgements. This project was made possible in part by support from the Cherokee Nation Language Program. Any views, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect those of the Cherokee Nation.

12. Figures.

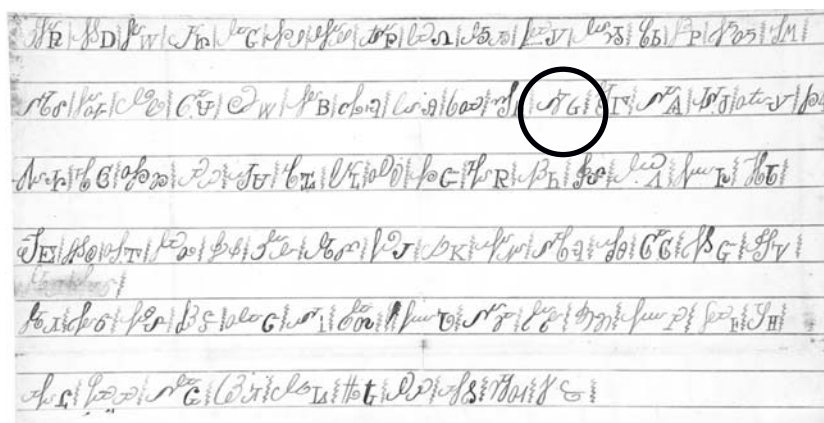


Figure 1. Sequoyah's original arrangement of the syllabary, showing *mv* in the original design and with a type-based version hand-drawn next to it.

Question 2. "What mean the three letters all three pronounced *un* (French) *u* & *i* of Guess, and what is become of them? What is become also of Clough *r*, Cleeh *e*, Clah *z*, Clegh *l*, Cloh *u*, and the nasal Gnaugh *z*?"

Ans. Only one of the three letters in question, viz. *z*, has simply the sound of *un* French. The other two are aspirated, and differ from each other only in that the sound of *u* is more open than that of *u*. This distinction has been regarded as of so little consequence, particularly by Maj. Lowrey, who has been the oracle on this subject, that the character *u* has been omitted, and, as no type has been cast for it, your printers can only represent it, as I perceive they have done, by breaking the English capital *G*. *Cherokee Phoenix Aug. 6. 1828*

Figure 2. Description of the problematic typography and use of the syllable *mv* in the *Cherokee Phoenix*, 1828-08-06.

ᏍᏍᏉ ᏚᏐᏐ, ᏐᏁ ᏅᏚᏚᏚᏚᏚ.

Dəvət 1

1 **I**ᏚᏐᏐᏐ ᏚᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ ᏚᏐᏐ ᏅᏚᏚᏚᏚᏚ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ ᏚᏐᏐ ᏅᏚᏚᏚᏚᏚ ᏐᏐᏐ.

2 **I**ᏚᏐᏐᏐ ᏐᏐᏐ ᏅᏚᏚᏚᏚᏚ ᏐᏐᏐ ᏐᏐᏐ.

3 **H**ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

4 **ᏐᏐᏐ** [ᏚᏐᏐ] ᏐᏐᏐ ᏐᏐᏐ; ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

5 **D**Ꮠ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

6 **Y**Ꮠ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

7 **ᏐᏐᏐ** ᏐᏐᏐ ᏚᏐᏐᏐ ᏐᏐᏐ, ᏐᏐᏐ **T**ᏐᏐ ᏚᏐᏐᏐ, ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

8 **i**Ꮠ ᏐᏐᏐ Ꮠ **T**ᏐᏐ ᏐᏐᏐ, ᏚᏐᏐᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ **T**ᏐᏐ.

9 **ᏐᏐᏐᏐᏐ** Ꮠ ᏚᏐᏐ ᏐᏐᏐ **T**ᏐᏐ ᏐᏐᏐ; ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

10 **ᏐᏐᏐ** ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

11 **ᏐᏐᏐ** ᏐᏐᏐ, ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

12 **ᏐᏐᏐ** ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

13 **ᏐᏐᏐ** ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

14 **D**Ꮠ ᏚᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐᏐ ᏐᏐᏐ ᏐᏐ ᏐᏐᏐ; ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

15 **G**Ꮠ ᏚᏐᏐᏐ ᏐᏐᏐ [ᏚᏐᏐ], ᏐᏐ ᏐᏐᏐ ᏐᏐ ᏐᏐᏐ; **A**Ꮠ ᏐᏐᏐ ᏐᏐᏐᏐᏐ, ᏐᏐ ᏐᏐᏐᏐᏐ, ᏐᏐ ᏐᏐᏐ **T**ᏐᏐ ᏐᏐᏐ, ᏐᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

16 **D**Ꮠ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐᏐᏐᏐ ᏐᏐᏐ ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

17 **I**ᏐᏐᏐᏐᏐ **Ꮠ**Ꮠ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, ᏐᏐᏐᏐᏐ ᏐᏐᏐ ᏐᏐ ᏐᏐᏐ **H**Ꮠ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

18 **i**Ꮠ ᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏅᏚᏚᏚᏚᏚ ᏅᏚᏚᏚᏚᏚ; ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ, **D**ᏐᏐᏐ ᏐᏐᏐ ᏐᏐ, ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

19 **D**Ꮠ ᏐᏐᏐ ᏐᏐ **G**Ꮠ ᏚᏐᏐᏐ, ᏐᏐᏐ **D**ᏐᏐ **H**ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ ᏐᏐᏐ.

20 **D**Ꮠ ᏐᏐᏐ, ᏐᏐ ᏐᏐ ᏐᏐᏐ, ᏐᏐᏐ ᏐᏐᏐ ᏐᏐ ᏐᏐᏐ; ᏐᏐ ᏐᏐ ᏐᏐᏐ ᏐᏐ.

Figure 3. The beginning of the Gospel of John (Gh Tsani) published in 1860 in Cherokee, showing capital letters in use throughout at the beginnings of sentences and personal and place names. (Red highlighting has been added.)

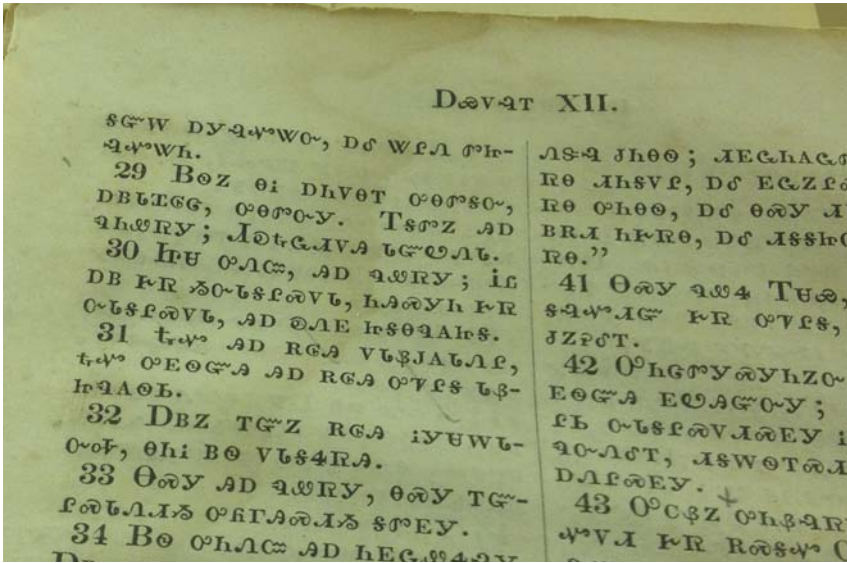


Figure 4. Photo of part of the 1860 Cherokee New Testament.

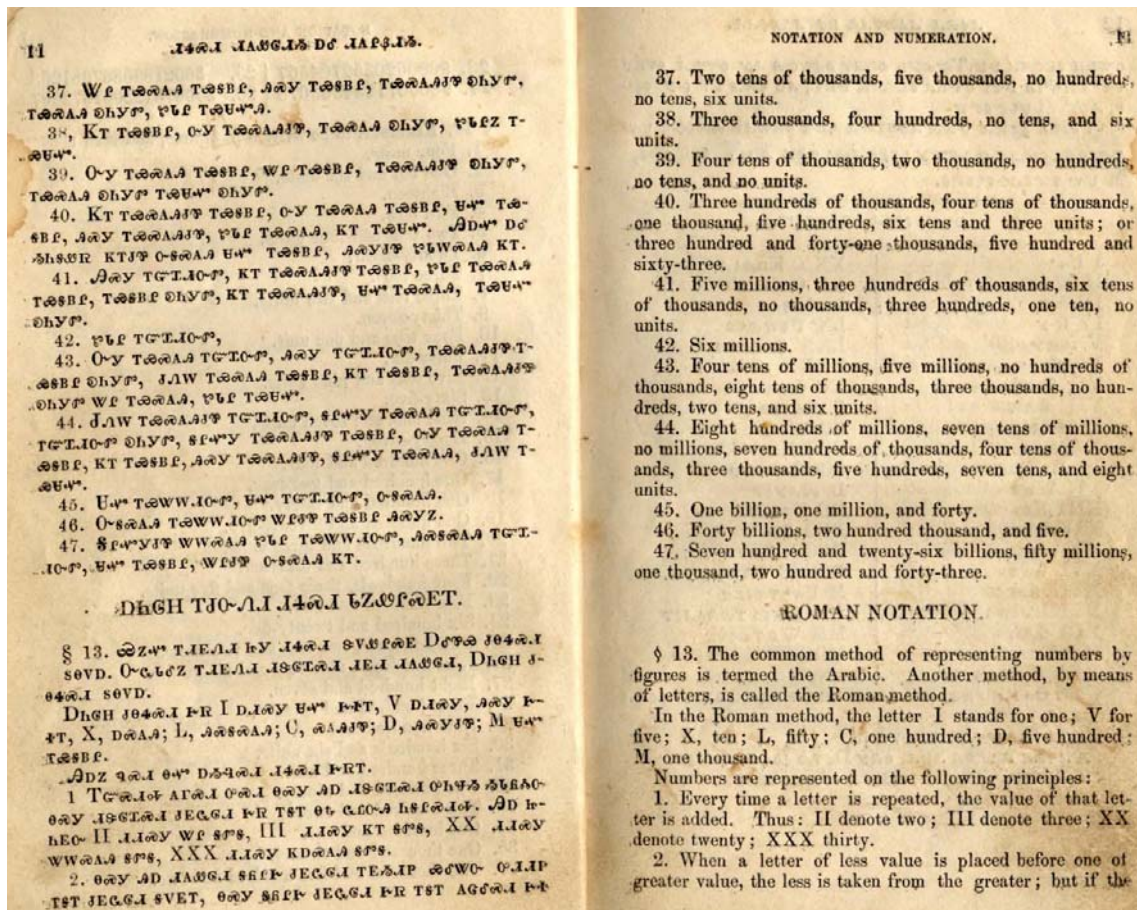


Figure 5. From John B. Jones' *Elementary Arithmetic in Cherokee and English*, 1870, showing capital letters in use throughout at the beginnings of sentences. In the section on Roman Notation, The Roman method is called *DhGH Anilomi* ('people of Lomi' = Rome) and the Arabic *Dṛṣṣ Aleqwiya* (which transliterates *Arabia*).

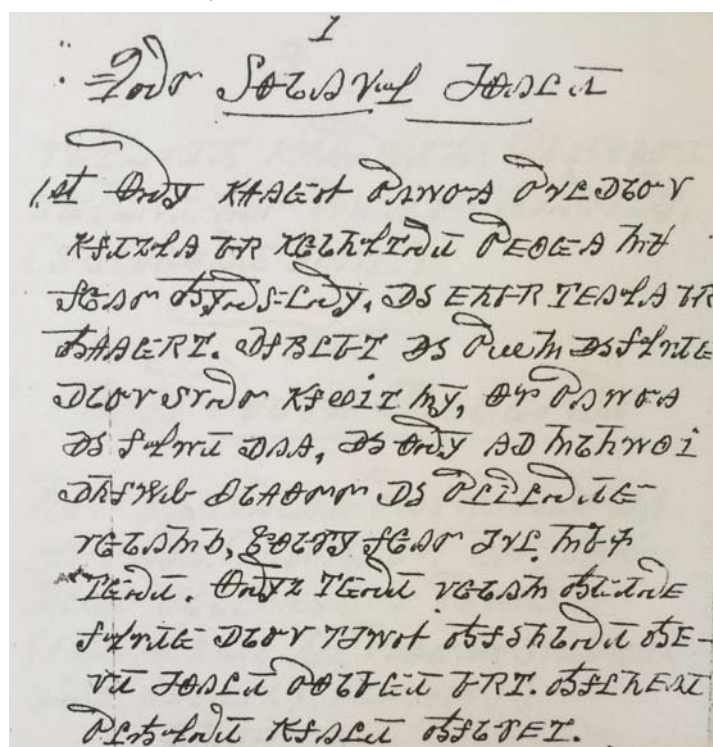


Figure 6. Sample of 19th-century handwriting in which ascenders and a few descenders can be seen.

16 ოქლიც აქაჲსოჲ ოქლიც დოღჳსოჲ იოჲ 10, 1881.

ოქლიც ტეჲ DJBA ტელიჲ ჯაიაჲ აჲფ DS-ჲ ოჲჲოჲ.

ოქლიც Dჲf Kfwoy, ოჲყ აჲფ DS-ჲ ოჲჲოჲ 271 აჲფ
§65.75, ოჲჲოჲ DJBA ოქლიც იოჲ 8,
1879, აჲფ ოჲჲოჲ იოჲ 9, 1879, ტეჲ DJBA ტელიჲ ოჲჲოჲ.

ოჲ ოჲჲოჲ, ოჲ-VST 16, 1881.

ოჲჲოჲ,

DYEOT.

ოქლიც DS-ჲ DJBA ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ.

ოქლიც Dჲf Kfwoy, ოჲყ WWBA ოჲჲოჲ DS-ჲ KYOTJZ
(\$27.75) ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ DS-ჲ ოჲჲოჲ ოჲჲოჲ
ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ
ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ
DS-ჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ ოჲჲოჲ.

ოჲ ოჲჲოჲ, ოჲ-VST 22, 1881.

ოჲჲოჲ,

DYEOT.

18

ACT AMENDING AN ACT.

AN ACT to make a certain warrant preferred.

Be it enacted by the National Council, That war-
rant No. 271, issued to Dr. L. M. Cravens for \$56.75,
under an appropriation dated December 8th, 1879, is-
sued December 9th, 1879, be and the same is hereby
made a preferred warrant.

Approved November 16th, 1881,

D. W. BUSHYHEAD,

Principal Chief.

AN ACT making appropriation for the Auditor's
Office.

Be it enacted by the National Council, That the
sum of twenty-seven dollars and seventy-five cents
(\$27.75), be and the same is hereby appropriated out of
any money in the Treasury belonging to the general
fund, not otherwise appropriated, for the use and ben-
efit of the Auditor's Office, and the Chief is hereby
authorized to draw a preferred warrant for the same.

Approved November 22d, 1881,

D. W. BUSHYHEAD,

Principal Chief.

Figure 7. Cherokee and English acts, from *Laws and Joint Resolutions of the Cherokee Nation*, enacted during the Regular and Special Sessions of the years 1881–2–3 (Tahlequah: E. C. Boudinot, 1884; Wilmington: Scholarly Resources, 1975). Capital letters can be seen in the Cherokee text. (Red highlighting has been added.)

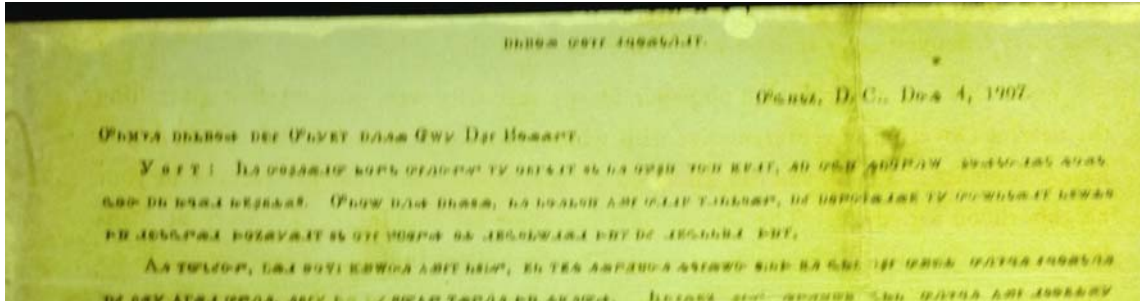


Figure 8. A warning from the Commissioner of Indian Affairs (Washington, D.C., 1907-03-04), concerning land fraud, showing the use of capital letters.

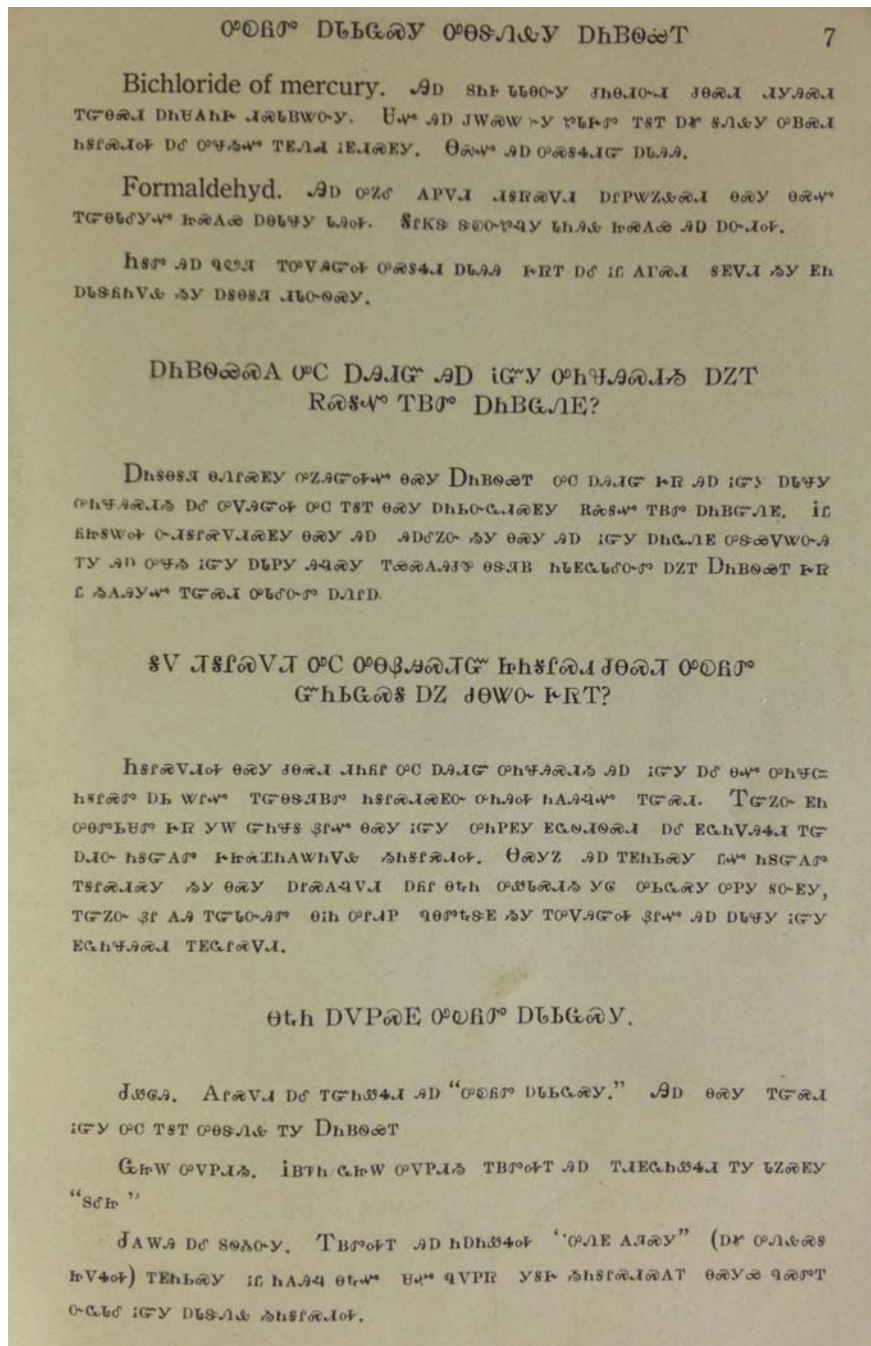


Figure 9. From a Cherokee health pamphlet, dated 1918, in which capital letters are used.



ᑕᑭᑭᑦᑕᑦ ᐱᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ

ᐱᑦᑕᑦ ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ", ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ.

ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ", ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ", ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦ", ᐱᑦᑕᑦ ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦ", ᐱᑦᑕᑦ ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ.

ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ", ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ.

ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ", ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ.

ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ" ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦᑕᑦ ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ" ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. "ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ", ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ, ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ.

ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ. ᐱᑦᑕᑦ ᐱᑦᑕᑦ ᐱᑦᑕᑦ.

44

Figure 10. A story from the Second-Grade Cherokee school curriculum (*Language Arts and Reading Comprehension*, 2008), showing capital letters in use throughout at the beginnings of sentences and personal and place names.

	13A	13B	13C	13D	13E	13F
0	D 13A0	Ꭱ 13B0	G 13C0	Ꭲ 13D0	Ꭳ 13E0	Ꭴ 13F0
1	R 13A1	Ꭶ 13B1	Ꭰ 13C1	Ꭺ 13D1	Ꭵ 13E1	Ꭷ 13F1
2	T 13A2	Ꭺ 13B2	Ꭱ 13C2	R 13D2	P 13E2	Ꭶ 13F2
3	Ꭰ 13A3	W 13B3	Z 13C3	L 13D3	G 13E3	G ^w 13F3
4	Ꭱ 13A4	Ꭲ 13B4	Ꭳ 13C4	Ꭴ 13D4	Ꭵ 13E4	B 13F4
5	i 13A5	Ꭲ 13B5	Ꭳ 13C5	Ꭴ 13D5	Ꭵ 13E5	G 13F5
6	S 13A6	G 13B6	T 13C6	Ꭵ 13D6	K 13E6	
7	Ꭱ 13A7	M 13B7	Ꭲ 13C7	Ꭳ 13D7	Ꭴ 13E7	
8	Ꭲ 13A8	Ꭳ 13B8	Ꭴ 13C8	Ꭵ 13D8	C ^w 13E8	Ꭶ 13F8
9	Y 13A9	Ꭲ 13B9	Ꭳ 13C9	V 13D9	G 13E9	Ꭷ 13F9
A	A 13AA	Ꭱ 13BA	Ꭲ 13CA	S 13DA	Ꭵ 13EA	Ꭶ 13FA
B	J 13AB	H 13BB	E 13CB	Ꭲ 13DB	Ꭳ 13EB	G ^w 13FB
C	E 13AC	Ꭲ 13BC	Ꭳ 13CC	Ꭴ 13DC	Ꭵ 13EC	B 13FC
D	Ꭲ 13AD	Y 13BD	Ꭳ 13CD	L 13DD	Ꭵ 13ED	G 13FD
E	Ꭲ 13AE	Ꭳ 13BE	Ꭴ 13CE	L 13DE	Ꭵ 13EE	
F	Ꭲ 13AF	Ꭳ 13BF	Ꭴ 13CF	C 13DF	Ꭵ 13EF	

Most lower-case Cherokee syllables are encoded at U+AB70-ABBF.

Upper-case syllables

13A0	D	CHEROKEE LETTER A
13A1	R	CHEROKEE LETTER E
13A2	T	CHEROKEE LETTER I
13A3	Ꮘ	CHEROKEE LETTER O
13A4	Ꮚ	CHEROKEE LETTER U
13A5	i	CHEROKEE LETTER V
13A6	Ꮝ	CHEROKEE LETTER GA
13A7	Ꮜ	CHEROKEE LETTER KA
13A8	Ꮞ	CHEROKEE LETTER GE
13A9	Ꮟ	CHEROKEE LETTER GI
13AA	A	CHEROKEE LETTER GO
13AB	J	CHEROKEE LETTER GU
13AC	E	CHEROKEE LETTER GV
13AD	Ꮵ	CHEROKEE LETTER HA
13AE	Ꮶ	CHEROKEE LETTER HE
13AF	Ꮷ	CHEROKEE LETTER HI
13B0	Ꮹ	CHEROKEE LETTER HO
13B1	Ꮺ	CHEROKEE LETTER HU
13B2	Ꮻ	CHEROKEE LETTER HV
13B3	W	CHEROKEE LETTER LA
13B4	Ꮼ	CHEROKEE LETTER LE
13B5	Ꮽ	CHEROKEE LETTER LI
13B6	G	CHEROKEE LETTER LO
13B7	M	CHEROKEE LETTER LU
13B8	Ꮾ	CHEROKEE LETTER LV
13B9	Ꮿ	CHEROKEE LETTER MA
13BA	Ᏸ	CHEROKEE LETTER ME
13BB	H	CHEROKEE LETTER MI
13BC	Ᏺ	CHEROKEE LETTER MO
13BD	Ᏻ	CHEROKEE LETTER MU
13BE	Ᏼ	CHEROKEE LETTER NA
13BF	Ᏽ	CHEROKEE LETTER HNA
13C0	G	CHEROKEE LETTER NAH
13C1	᏷	CHEROKEE LETTER NE
13C2	ᏸ	CHEROKEE LETTER NI
13C3	Z	CHEROKEE LETTER NO
13C4	ᏺ	CHEROKEE LETTER NU
13C5	ᏻ	CHEROKEE LETTER NV
13C6	ᏼ	CHEROKEE LETTER QUA
13C7	᏾	CHEROKEE LETTER QUE
13C8	᏿	CHEROKEE LETTER QUI
13C9	᐀	CHEROKEE LETTER QUO
13CA	ᐁ	CHEROKEE LETTER QUU
13CB	ᐂ	CHEROKEE LETTER QUV
13CC	ᐃ	CHEROKEE LETTER SA
13CD	ᐄ	CHEROKEE LETTER S
13CE	ᐅ	CHEROKEE LETTER SE
13CF	ᐆ	CHEROKEE LETTER SI
13D0	ᐇ	CHEROKEE LETTER SO
13D1	ᐈ	CHEROKEE LETTER SU
13D2	R	CHEROKEE LETTER SV
13D3	ᐊ	CHEROKEE LETTER DA
13D4	W	CHEROKEE LETTER TA
13D5	S	CHEROKEE LETTER DE
13D6	ᐌ	CHEROKEE LETTER TE
13D7	ᐍ	CHEROKEE LETTER DI
13D8	ᐎ	CHEROKEE LETTER TI
13D9	V	CHEROKEE LETTER DO
13DA	S	CHEROKEE LETTER DU
13DB	ᐐ	CHEROKEE LETTER DV
13DC	ᐑ	CHEROKEE LETTER DLA
13DD	ᐒ	CHEROKEE LETTER TLA
13DE	L	CHEROKEE LETTER TLE
13DF	C	CHEROKEE LETTER TLI

13E0	ᐃ	CHEROKEE LETTER TLO
13E1	ᐄ	CHEROKEE LETTER TLU
13E2	P	CHEROKEE LETTER TLV
13E3	G	CHEROKEE LETTER TSA
13E4	ᐇ	CHEROKEE LETTER TSE
13E5	ᐈ	CHEROKEE LETTER TSI
13E6	K	CHEROKEE LETTER TSO
13E7	ᐊ	CHEROKEE LETTER TSU
13E8	ᐋ	CHEROKEE LETTER TSV
13E9	G	CHEROKEE LETTER WA
13EA	ᐌ	CHEROKEE LETTER WE
13EB	ᐍ	CHEROKEE LETTER WI
13EC	ᐎ	CHEROKEE LETTER WO
13ED	ᐏ	CHEROKEE LETTER WU
13EE	ᐐ	CHEROKEE LETTER WV
13EF	ᐑ	CHEROKEE LETTER YA
13F0	ᐒ	CHEROKEE LETTER YE
13F1	ᐓ	CHEROKEE LETTER YI
13F2	ᐔ	CHEROKEE LETTER YO
13F3	ᐕ	CHEROKEE LETTER YU
13F4	B	CHEROKEE LETTER YV

Archaic upper-case syllable

13F5	G	CHEROKEE LETTER MV
------	---	--------------------

Lower-case syllables

13F8	ᐁ	CHEROKEE SMALL LETTER YE
13F9	ᐂ	CHEROKEE SMALL LETTER YI
13FA	ᐃ	CHEROKEE SMALL LETTER YO
13FB	ᐄ	CHEROKEE SMALL LETTER YU
13FC	B	CHEROKEE SMALL LETTER YV

Archaic lower-case syllable

13FD	G	CHEROKEE SMALL LETTER MV
------	---	--------------------------

	AB7	AB8	AB9	ABA	ABB
0	D AB70	Ꮟ AB80	Ꮠ AB90	Ꮡ ABA0	Ꮢ ABB0
1	R AB71	Ꮣ AB81	Ꮤ AB91	Ꮥ ABA1	Ꮦ ABB1
2	T AB72	Ꮧ AB82	Ꮨ AB92	Ꮩ ABA2	Ꮪ ABB2
3	Ꮫ AB73	Ꮬ AB83	Ꮭ AB93	Ꮮ ABA3	Ꮯ ABB3
4	Ꮰ AB74	Ꮱ AB84	Ꮲ AB94	Ꮳ ABA4	Ꮴ ABB4
5	Ꮵ AB75	Ꮶ AB85	Ꮷ AB95	Ꮸ ABA5	Ꮹ ABB5
6	Ꮺ AB76	Ꮻ AB86	Ꮼ AB96	Ꮽ ABA6	Ꮾ ABB6
7	Ᏸ AB77	Ᏹ AB87	Ᏺ AB97	Ᏻ ABA7	Ᏼ ABB7
8	᏶ AB78	᏷ AB88	ᏸ AB98	ᏹ ABA8	ᏺ ABB8
9	ᏼ AB79	ᏽ AB89	᏾ AB99	᏿ ABA9	᐀ ABB9
A	ᐁ AB7A	ᐂ AB8A	ᐃ AB9A	ᐄ ABAA	ᐅ ABBA
B	ᐇ AB7B	ᐈ AB8B	ᐉ AB9B	ᐊ ABAB	ᐋ ABBB
C	ᐍ AB7C	ᐎ AB8C	ᐏ AB9C	ᐐ ABAC	ᐑ ABBC
D	ᐓ AB7D	ᐔ AB8D	ᐕ AB9D	ᐖ ABAD	ᐗ ABBD
E	ᐙ AB7E	ᐚ AB8E	ᐛ AB9E	ᐜ ABAE	ᐝ ABBE
F	ᐟ AB7F	ᐠ AB8F	ᐡ AB9F	ᐢ ABAF	ᐣ ABBF

The upper-case Cherokee syllables are encoded at U+13A0-13FF. Lower-case ye, yi, yo, yu, yv, and mv are encoded in that same block.

Lower-case syllables

AB70	Ꭰ	CHEROKEE SMALL LETTER A
AB71	Ꭱ	CHEROKEE SMALL LETTER E
AB72	Ꭲ	CHEROKEE SMALL LETTER I
AB73	Ꭳ	CHEROKEE SMALL LETTER O
AB74	Ꭴ	CHEROKEE SMALL LETTER U
AB75	Ꭵ	CHEROKEE SMALL LETTER V
AB76	Ꭶ	CHEROKEE SMALL LETTER GA
AB77	Ꭷ	CHEROKEE SMALL LETTER KA
AB78	Ꭸ	CHEROKEE SMALL LETTER GE
AB79	Ꭹ	CHEROKEE SMALL LETTER GI
AB7A	Ꭺ	CHEROKEE SMALL LETTER GO
AB7B	Ꭻ	CHEROKEE SMALL LETTER GU
AB7C	Ꭼ	CHEROKEE SMALL LETTER GV
AB7D	Ꭽ	CHEROKEE SMALL LETTER HA
AB7E	Ꭾ	CHEROKEE SMALL LETTER HE
AB7F	Ꭿ	CHEROKEE SMALL LETTER HI
AB80	Ꮀ	CHEROKEE SMALL LETTER HO
AB81	Ꮁ	CHEROKEE SMALL LETTER HU
AB82	Ꮂ	CHEROKEE SMALL LETTER HV
AB83	Ꮃ	CHEROKEE SMALL LETTER LA
AB84	Ꮄ	CHEROKEE SMALL LETTER LE
AB85	Ꮅ	CHEROKEE SMALL LETTER LI
AB86	Ꮆ	CHEROKEE SMALL LETTER LO
AB87	Ꮇ	CHEROKEE SMALL LETTER LU
AB88	Ꮈ	CHEROKEE SMALL LETTER LV
AB89	Ꮉ	CHEROKEE SMALL LETTER MA
AB8A	Ꮊ	CHEROKEE SMALL LETTER ME
AB8B	Ꮋ	CHEROKEE SMALL LETTER MI
AB8C	Ꮌ	CHEROKEE SMALL LETTER MO
AB8D	Ꮍ	CHEROKEE SMALL LETTER MU
AB8E	Ꮎ	CHEROKEE SMALL LETTER NA
AB8F	Ꮏ	CHEROKEE SMALL LETTER HNA
AB90	Ꮐ	CHEROKEE SMALL LETTER NAH
AB91	Ꮑ	CHEROKEE SMALL LETTER NE
AB92	Ꮒ	CHEROKEE SMALL LETTER NI
AB93	Ꮓ	CHEROKEE SMALL LETTER NO
AB94	Ꮔ	CHEROKEE SMALL LETTER NU
AB95	Ꮕ	CHEROKEE SMALL LETTER NV
AB96	Ꮖ	CHEROKEE SMALL LETTER QUA
AB97	Ꮗ	CHEROKEE SMALL LETTER QUE
AB98	Ꮘ	CHEROKEE SMALL LETTER QUI
AB99	Ꮙ	CHEROKEE SMALL LETTER QUO
AB9A	Ꮚ	CHEROKEE SMALL LETTER QUU
AB9B	Ꮛ	CHEROKEE SMALL LETTER QUV
AB9C	Ꮜ	CHEROKEE SMALL LETTER SA
AB9D	Ꮝ	CHEROKEE SMALL LETTER S
AB9E	Ꮞ	CHEROKEE SMALL LETTER SE
AB9F	Ꮟ	CHEROKEE SMALL LETTER SI
ABA0	Ꮠ	CHEROKEE SMALL LETTER SO
ABA1	Ꮡ	CHEROKEE SMALL LETTER SU
ABA2	Ꮢ	CHEROKEE SMALL LETTER SV
ABA3	Ꮣ	CHEROKEE SMALL LETTER DA
ABA4	Ꮤ	CHEROKEE SMALL LETTER TA
ABA5	Ꮥ	CHEROKEE SMALL LETTER DE
ABA6	Ꮦ	CHEROKEE SMALL LETTER TE
ABA7	Ꮧ	CHEROKEE SMALL LETTER DI
ABA8	Ꮨ	CHEROKEE SMALL LETTER TI
ABA9	Ꮩ	CHEROKEE SMALL LETTER DO
ABAA	Ꮪ	CHEROKEE SMALL LETTER DU
ABAB	Ꮫ	CHEROKEE SMALL LETTER DV
ABAC	Ꮬ	CHEROKEE SMALL LETTER DLA
ABAD	Ꮭ	CHEROKEE SMALL LETTER TLA
ABAE	Ꮮ	CHEROKEE SMALL LETTER TLE

ABAF	Ꮯ	CHEROKEE SMALL LETTER TLI
ABB0	᠀	CHEROKEE SMALL LETTER TLO
ABB1	᠁	CHEROKEE SMALL LETTER TLU
ABB2	᠂	CHEROKEE SMALL LETTER TLV
ABB3	᠃	CHEROKEE SMALL LETTER TSA
ABB4	᠄	CHEROKEE SMALL LETTER TSE
ABB5	᠅	CHEROKEE SMALL LETTER TSI
ABB6	᠆	CHEROKEE SMALL LETTER TSO
ABB7	᠇	CHEROKEE SMALL LETTER TSU
ABB8	᠈	CHEROKEE SMALL LETTER TSV
ABB9	᠉	CHEROKEE SMALL LETTER WA
ABBA	᠊	CHEROKEE SMALL LETTER WE
ABBB	᠋	CHEROKEE SMALL LETTER WI
ABBC	᠌	CHEROKEE SMALL LETTER WO
ABBD	᠍	CHEROKEE SMALL LETTER WU
ABBE	᠎	CHEROKEE SMALL LETTER WV
ABBF	᠏	CHEROKEE SMALL LETTER YA

A. Administrative

1. Title

Revised proposal for the addition of Cherokee characters to the UCS

2. Requester's name

Michael Everson, Durbin Feeling

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

Individual contribution.

4. Submission date

2014-02-25

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)

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

Cherokee, and a new Cherokee Supplement block

2. Number of characters in proposal

87 (7, 80).

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?

Roy Boney, David Crawler, Durbin Feeling, Sarah Holcomb, Sherry Holcomb, John Ross, Anna Sixkiller, Dennis Sixkiller, Pauline Teehee, Candessa Tehee.

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?

Cherokee speakers and writers.

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

Used historically and in modern editions.

4b. Reference

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

Yes.

5b. If YES, where?

Various publications, though the use has been limited by the lack of availability of capital letters.

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)?

No.

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?

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)?

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?