L2/01-291

BILL JANCEWICZ

asamich
(snowshoes)

to: The Unicode Consortium
from: Bill Jancewicz
date: 7 August, 2001
re: Unified Canadian Aboriginal Syllabics

cc: file

Dear friends:

I present the following document for your consideration regarding the glyph shapes for a series of characters in The Unicode Standard 3.0.

Please feel free to contact me if you have any questions or concerns.

Sincerely,

Bill Jancewicz

Bill Jancewicz
BJ:ns
Enclosure: Proposal, Unified Canadian Aboriginal Syllabics

Bill Jancewicz - Naskapi Language studies
Naskapi Development Corporation, Box 5153, Kawawachikamach, Quebec G0G 2Z0
Phone (418) 585-2612 Fax (418) 585-3953 E-Mail <bill.jancewicz@sil.org>

PROPOSAL

SH Series, Unified Canadian Aboriginal Syllabics, U+1510 - U+1525

It has come to my attention that the shape and orientation of certain Canadian Syllabic glyphs as shown in The Unicode Standard 3.0 are in question. By means of this document I would like to recommend certain changes to the representation of these glyphs in the Standard.

Canadian Syllabics are glyphs that normally represent a consonant-vowel sequence. Simply put, the shape of the character defines the consonant, and the orientation of the character defines the vowel. “Orientation” means not only the rotational orientation, but also “mirror” orientation, as in English [b] may be considered a “mirror image” of [d], while [p] may be considered a 180 degree rotation of [d].

In Canadian Syllabics, we normally speak of four orientations, corresponding to four vowels, [e], [i], [o] and [a]. Further, for the Eastern orthographies, a final consonant is a smaller, usually raised version of the [a] orientation.

Vowel length is commonly indicated by an overdot, but this has no effect on the shape of the glyph: The shape of “SHA” and “SHAA” is identical, except that “SHAA” has an overdot.

The Unicode glyphs in question are found in the U+1510 to U+1525 range, comprising all the dotted and vowel combinations for the “SH” consonant series of Canadian Syllabics. Thus, we are actually only concerned with four basic glyphs, the four vowel combinations with SH, namely SHE, SHI, SHO, SHA. The “final” SH will be a superscript (half size, raised) version of SHA. For the moment, we will not discuss the dots.

Given these constraints, the five Unicode positions in question are U+1510 (SHE); U+1511 (SHI); U+1513 (SHO) and U+1515 (SHA), plus the final U+1525 (SH). Please refer to The Unicode Standard 3.0, Unified Canadian Aboriginal Syllabics, page 459.

It can be seen that all the other glyph shapes and orientations may be generated from these five glyphs.

I am providing facsimile evidence from various published sources for your consideration.

Table 1 is a summary of the various glyph shapes for these characters in the published material.

I am proposing that the glyphs shown in the Unicode Standard be changed so that they are similar in appearance to those in row 6.

U+1510 "SHE" Essentially, this is a mirror image of an English "S" shape. In most of the facsimile examples, a nearly straight line separates the two semi-circular curved sections. This straight line between the curves is an important feature, which assists the reader in recognizing the orientation.

In the glyph shown in the Unicode Standard 3.0, this straight line is nearly non-existent, and the glyph appears in the Standard as two connected semi-circles.

I recommend modifying the shape of the glyph in the Unicode Standard so that the left hand semicircle is "raised" such that the top of the "hump" is equal in height to the top of the other "tall" characters, up to the height of U+1489 "CHE". With this modification, the "shape" of this series of characters would be correct.

With regard to orientation, "straight line" part of the glyph for U+1510 (between the semi-circles) should be nearly vertical.

U+1511 "SHI" This glyph should be a mirror image of U+1510. As to orientation, the "straight line" part of the glyph (between the semi-circles) should be nearly vertical.

U+1512 "SHII" Same shape and orientation as U+1511. I will discuss dots later.

U+1513 "SHO" Same shape as U+1510, different orientation: The "straight line" part of this glyph should be represented as nearly horizontal, but rotated clockwise slightly. In rows 6, 7 and 8, this glyph is represented rotated to the "limit" clockwise, with both semi-circles resting full on the baseline. This is the maximum extreme rotation for this glyph.

I recommend that the glyph be represented in the Unicode Standard oriented like in row 6 (Table 1)

U+1514 "SHOO" Same shape and orientation as U+1513. Dots later.

U+1515 "SHA" This glyph should be a mirror image of U+1513. As to orientation, the "straight line" part of the glyph (between the semi-circles) should be nearly horizontal, but rotated counter-clockwise slightly. In my font, this glyph is represented rotated to the "limit" counter-clockwise, with both semi-circles resting full on the baseline. This is the maximum extreme rotation for this glyph.

I recommend that the glyph be represented in the Unicode Standard oriented like in row 6 (Table 1).

U+1516 "SHAA" Same shape and orientation as U+1515. Dots later.

U+1517 "SHWE" Same shape and orientation as U+1510. The "W" dot or "mid dot" diacritic should be the same height and weight as the dots in the other "W" combinations, like U+14F6 or U+14F7 say. This is true for U+1519 through U+1524.

Overdots: The dots signifying "length" (as in U+1512, U+1514, U+1516 above, and in all the "W" dot combinations): There are really two schools of thought on this matter.

One opinion is that the length dot (overdot) should be centered above the *highest point* of the character. This is the system that was followed by most of the overdots in the Standard: For example, U+1491 "CAA" has the dot over the stem (on the left) and U+148C "CII" has the dot centered, over the highest point of the curve.

On glyphs where it there is no "highest point", the dot is centered, as over U+14A6 "MII".

The other school of thought is to have the overdot ALWAYS centered. This scheme was employed in the Standard with U+140B "AA". My recommendation follows the former scheme, with the dot over the highest point (see dot note, below).

My recommendation is that *all overdots* for the entire character set (with the basic meaning of "vowel length") be positioned over the highest point of all the glyphs.

With regard to the "SH" series of characters, the former scheme of placing the dot over the highest point of the glyph is recommended.

Note that for my recommended glyph shapes, U+1514 "SHOO", U+1516 "SHAA", U+151F "SHWOO", U+1520 "WEST CREE SHWOO" U+1523 "SHWAA", and U+1524 "WEST CREE SHWAA", the overdot is centered. This is because at this rotation, (with both curves resting on the baseline) the glyph has *two* "highest points", and thus the dot must be centered.

However, if you choose to adopt the orientations shown in rows 2-5 for the representation of the glyphs in the Standard, you will not have such an extreme rotation of these "SHO" and "SHA" series glyphs. In this case, the glyph will have a "highest point" and the overdots should be over the highest point of the glyph, not centered on the width of the glyph.

In summary, I am recommending that the Unicode Standard be updated for characters in the range U+1510 - U+1525 according to the recommendations above as indicated on the last sheet in the Appendix, Unicode - Canadian Aboriginal Syllabics.

Please feel free to contact me again if you need further information about these or other glyphs in the Cree inventory.

Best regards, Bill Jancewicz

-Dot note: The overdots over the triangular characters actually are centered over the vertical stroke of the triangle, and not exactly over the "highest" point, which would be the outer edge of the vertical stroke. Visually, this placement aligned with the vertical stroke is correct in these cases.

Sometimes, dots can be two different sizes on the same character. The "W" dot (to the right or left) may be slightly heavier (up to 50% more in diameter) than the overdot (length dot).

Also, in many Cree encodings, the "W" dot (to the right or left) is rendered as a distinct glyph (U+1427).

Appendix

Table 1 row identification:

1) The Unicode Standard 3.0 page 459

2) CCALS: Computer Coding for Aboriginal Language Syllabics, Project overview prepared by Canada Federal Department of Communications (DOC) Phase III, 1993.

3) CBS The Canadian Bible Society. New Testament in Moose Cree. 1991. (ISBN 0-88834-424-4) Toronto. The Childhood of Jesus, Scripture Selections in Moose Cree. 1991. (ISBN 0-88834-680-8) Toronto.

4) Old Testament, as used in the Liturgy. Jules Leguerrier, O.M.I. Bishop of Moosonee. 1987. Diocese of Moosonee.

5) Common Prayer and Hymns in the Cree Indian Language (James and Hudson Bays, E.) Translated by the Rev. and Mrs. W.G. Walton (C.M.S.) 1923.

An Exposition of the Ten Commandments in the Language of the Cree Indians who inhabit the territory East of Hudson Bay, Canada. By the Rev. and Mrs. W.G. Walton, C.M.S. Missionaries. 1921.

The Gospel of John. Attributed to Rev. W. G. Walton. 1930.

The Book of Common Prayer, Translated into the language of the Cree Indians in the Diocese of Moosonee. By the Right Rev. John Horden, D.D. Bishop of Moosonee. 1890.

Proper Lessons from the Old Testament, in the Cree language. By the Right Rev. John Horden, D.D. Bishop of Moosonee. 1878.

6) BJCree font, developed for Cree Programs, Cree School Board, Chisassibi, Quebec. Bill Jancewicz. 1999.

7) Cree Lexicon Eastern James Bay Dialects. Marguerite MacKenzie et al. 1987. Cree School Board (ISBN 0-920245-80-3). Typesetting by Nortext Information Design Ltd.

8) James Bay Cree New Testament, Mistissini Quebec. 2001. Typeset by Canadian Bible Society, SILECree Font.

	SHE	SHI	SHO	SHA	SH
1) Unicode 3.0	1510	1511	1513	1515	1525
2) CCALS					
3) CBS Moose Cree					
4) Leguerrier Old Testament					
5) Walton / Horden					
6) BJCree font, Cree School Board					
7) Cree Lexicon					
8) James Bay Cree New Testament					

Table 1