

# Unicode Cree Syllabics for Windows and Macintosh

37th Algonquian Conference, Ottawa 2005

**Bill Jancewicz**

**SIL International and Naskapi Development Corporation**

## ABSTRACT

Submitted as an update to a presentation made at the 34th Algonquian Conference (Kingston). The ongoing development of the operating systems has included increased support for cross-platform use of Unicode syllabic script. Key improvements that were included in Macintosh's OS X.3 (Panther) operating system now allow direct keyboarding of Unicode characters by means of a user-defined input method. A summary and comparison of the available tools for handling Unicode on both Windows and Macintosh will be discussed.

## INTRODUCTION

Since 1988 the author has been working in the Naskapi language at Kawawachikamach with the primary purpose of linguistic analysis and Bible translation, sponsored by Wycliffe Bible Translators and SIL International. Related work includes mother tongue translator training, Naskapi literature and curriculum development.

Along with the language work the author also developed methods of production for Naskapi language materials in syllabic script by means of the computer. With the advent of high quality publishing capabilities in newer computers, procedures were updated to keep pace with the improving technology. With resources available from SIL International computer services department, a very satisfactory system of keyboarding syllabic texts in Naskapi was developed.

At the urging of colleagues working in related languages, the system was expanded to include the wider inventory of standard Eastern and Western Cree syllabics. While this has pushed the limit of what is possible with the current technology, Unicode makes this practical. Note that the system originally developed for keyboarding Naskapi syllabics is similar but not identical to the Cree system, because of the unique local orthography in use at Kawawachikamach. The Cree system described here is intended to be useful for most Cree communities.

An important feature of this system is the method of keyboarding. There are currently two basic philosophies with regard to keyboarding syllabics. The first has its roots in the early Cree typewriters, where each symbol had it's own key on the keyboard. A special keyboard layout had to be learned, but once learned a typist could keyboard syllabics rather quickly. In addition, the keyboard layout could be arranged somewhat mnemonically, so that one horizontal row corresponds to the /e/ syllabic series, and another corresponds to the /i/ series, and so on. However, this system also requires the practice of holding down a control key (like Shift or Alt or Both) while pressing the key to get all the various orientations of the characters or features such as vowel-length.

The other system described below, and in use at Kawawachikamach, follows the other philosophy of keyboarding.

Taking into account that a growing number of First Nations people have had some education in one of the majority languages, more and more aboriginal language speakers are increasingly computer literate. Further, the standard QWERTY keyboard is familiar to them. Also, many in the current generation of Naskapi and Cree people are beginning to have a strong desire to be able to read and write in their own language using the syllabic orthography. Syllabic script has a high traditional value, but it often seems out of reach. For the Cree and Naskapi speakers that have been trained, a system that combines what they already *know* (the English alphabet and standard computer keyboard) with what they *want to know* (how to read syllabics) has a very high success rate and a quick learning curve.

The system is essentially this: the user types in roman letters on a standard keyboard the way the word "sounds" in Cree. What he sees on the screen while he types is the word written in syllabics. With the advent of scaleable "TrueType" fonts and sophisticated word processors, anyone can now produce clear, publication-quality syllabic documents with a minimum of training.

In the Naskapi and Cree communities, some dozens of speakers have been trained; some elders but mostly people in their 20's and 30's who thought that they would never read syllabics, persons who are now producing their own syllabic literature in a quality that is accepted and read by others in the community. Note that the best success has been in training people who have never written in syllabics before, but we have also several people who *could* read and write but never "typed" before: and they all caught on quickly as well. A benefit to this system is that the hardware, the keyboard, does not need to be modified in any way with special key-caps or stickers, a new keyboard

does not have to be learned and memorized, and that syllabics and non-syllabics (roman) words can be mixed in at any time in any place in the document. Still, any users who already learned to type using the older “typewriter” layout may keep using that layout if they desire: users are able to choose whichever keyboard layout they prefer whenever they like.

## 8-BIT LEGACY ENCODINGS AND UNICODE

Computer character encoding systems evolved through the 1980s and 1990s. The hardware and software systems in use at that time limited the practical number of characters that could be used by any one font to only 256 8-bit characters. Practically speaking, when one eliminated the characters reserved for system control, a character set could have an inventory of just over 200 characters and symbols.

Whenever a developer wanted to include non-standard characters (like syllabics) in a character set, the common strategy was to replace existing standard character slots (A, B, é, ñ etc.) with the non-standard character. As can be imagined, this practice resulted in many conflicting standards. Further confusing the matter, Windows and Macintosh platforms use completely different encodings for the accented characters found in the upper half of the 256 available characters. If users kept their custom font encodings on one computing platform, they performed adequately for the most part. However, if one wanted to share projects across platforms or even between various applications, there the trouble began. The difficulties become most acute when material is intended for publication by commercial printers. Indeed it remains no small task to convert these custom font encodings without losing data.

The Unicode font encoding was developed in the 1990s in order to address the many requirements for multilingual and multi-platform computing. Unicode provides unique slots for over 65,000 characters, using 16-bit encoding. Beginning with the Windows 2000 and Macintosh OS X operating systems, a limited level of Unicode support was included for multi-lingual computing. With the release of Windows XP and later Macintosh OS X.3 (Panther), more adequate Unicode support was provided.

Of course, Unicode-compliant application programs must also be used in order to take advantage of the Unicode support built into the operating systems. The Microsoft Office suite of programs was one of the first to provide implementation of Unicode natively for both Windows (Office 2000 version 9 and beyond) and Macintosh (Office 2004, version 11 and beyond). While other application programs are beginning to be developed that include Unicode support, care must still be taken by the user to determine whether the applications used fully support Unicode. None of the applications that run on Macintosh “Classic” (OS 9) operating system support Unicode, and Windows 98 applications have little or no Unicode support. The popular lexical database and text-processing program “Shoobox”, by SIL International does not include Unicode support. However, SIL International has released a successor that does support Unicode, re-named “Toolbox”. Shoobox 8-bit data is forwardly compatible with Toolbox, but Toolbox Unicode data is not backwardly compatible with Shoobox.

The main advantage to using Unicode is that data and documents that are produced on one platform (Windows) can be shared and used on the other platform (Macintosh), because both the data and the font containing the character shapes is identical. Once the data is encoded in Unicode, there is no need for data conversion in order to use it on Windows or Macintosh. Of course, to be backwardly compatible with older systems or applications that cannot use Unicode, data conversion would still be necessary.

## CREE SYLLABICS IN UNICODE APPLICATIONS

Even with operating systems and applications that fully support Unicode, two other conditions are necessary to type and view Cree syllabics in Unicode: A font containing all the characters required for the language in question, and an “input method”, that interprets keystrokes on a keyboard that correspond with the syllabic characters desired.

### Unicode Cree Syllabic Fonts

Because of their unwieldy size and consumption of memory, it is not practical to use fonts that contain all 65,000 16-bit Unicode characters, nor is it desirable. Certain “blocks” of characters in the Unicode Standard are designated to cover various ranges of languages. The “Basic Latin” block of Unicode contains all the characters on a standard English keyboard, and the “Latin-1 Supplement” contains most of the common accented characters used in French and Spanish, along with numerical operators and normal punctuation. All of these should normally be included in a Cree Syllabic font as well, for convenience sake.

The “Canadian Aboriginal Syllabics” block (1400 - 167F) contains 640 character slots for Cree, Naskapi, Inuktitut, Blackfoot, and Carrier syllabics. Most Cree fonts include a subset of the Canadian Aboriginal Syllabics block, only

using character glyphs that are required for the language group for whom the font was developed. Although every Unicode font has slots for all 65,000 16-bit glyphs, it is unlikely that a Cree user would have any need for Arabic or Chinese characters, and so most of the slots are left empty.

There are currently several good Unicode Cree syllabic fonts available. These are listed in the appendix at the end of the paper.

## Unicode Input Methods

Depending upon the operating system, there are two different methods for using customized keyboards to insert Unicode characters. In both operating systems, there are utilities provided that allow the user to manually select the desired Unicode characters one-by-one. For Windows, the *Character Map* utility may be used (Start, Programs, Accessories, Character Map). For Macintosh, the *Character Palette* utility can be shown from the System Preferences, International menu. These utilities provide access to every Unicode character available in every font, but they must be selected one-by-one from a list or chart.

For practical input of Unicode characters, a customized keyboard is used.



## *Windows Unicode Keyboarding*

Keyman is a program developed by Tavultesoft <ENDNOTE> that allows the installation of multiple customized keyboards for access to Unicode characters in the Windows operating system. A developer's kit, TIKE is available that facilitates the creation of these keyboards, as well as user-installable font and keyboard packages. The author has created a package that contains all that is required to keyboard Unicode Cree syllabics in Windows, named *CreeKeysPRO*.

This system includes not only the Unicode TrueType font(s) but also the keyboard drivers for typing syllabics in roman.

### **CreeKeysPRO Installation**

Simply double click on the **CreeKeysPRO.EXE** file in your computer. The self-extracting installer will install the Keyman program, the keyboards and supporting documents, and the fonts. Accepting the default installation folders is recommended. After the installation, you will find the program files here:

```
C:\Program Files\Tavultesoft\Keyman\
```

And the user files (keyboards, copies of the fonts, a syllabic chart, etc.) are found here:

```
C:\Documents and Settings\\Application Data\Tavultesoft\Keyman\
```

As usual, administrator privileges are required for installation of fonts and programs, and user settings may need to be adjusted in order to view files in hidden folders.

### **Using the Font(s)**

The keyboards and fonts will work in any Windows application that supports Unicode. As of this writing an increasing number of commercial software applications that support Unicode are available.

1. Launch Keyman (this is the Keyboard driver)
2. Click START | PROGRAMS | Tavultesoft Keyman | Keyman  
"Tavultesoft Keyboard Manager" will briefly flash across the screen as the program loads. This message will disappear when one small icon appears in the taskbar Notification Area (formerly called the "Tray", normally on the lower right-hand side of the screen) like this:



This is the Keyman Button.

3. Open any Windows word processor, or other application.
4. Change the active font to **BJCree UNI**.

## To type in Roman and get Cree:

Now click on the Keyman button  in the Notification Area. Now click on the  **BJCree (east)** button. The icon in the tray changes to let you know which keyboard is active.

Now type in Roman the way Cree “sounds”. You should get syllabics. Double the vowels to get the diacritic (that is, type “paa” to get the long vowel dot over “pa”). “Ch” is “c”. or “ch” Back vowel is “u”, or “o”

To change back to English, just click on the  button in the Notification Area. Choose “No Keyman Keyboard”. You can keep typing in roman in the BJCree UNI font to get normal roman characters.

A complete syllabic chart may be found in:

C:\Documents and Settings\\Application Data\Tavultesoft\Keyman\

## To type using the Cree Keyboard Layout (similar to some Legacy keyboards):

Now click on the Keyman button  in the Notification Area again. This time, click on the  **Creeboard** button. The icon in the Notification Area changes to let you know which keyboard is active.

Now type following the Cree Keyboard layout. Press Shift+Alt+L to get ᓄ “suu”. You can find a copy of this keyboard layout to refer to in:

C:\Documents and Settings\\Application Data\Tavultesoft\Keyman\

To change back to English, just click on the  button in the Notification Area. Choose “No Keyman Keyboard”. You can keep typing in roman in the BJ Cree font to get normal roman characters.

## CreeKeysPRO Notes

To open the Keyman Configuration window (Help, Options, About, etc.) just Right Click on the Keyman button  in the Notification Area. Among other settings, shortcut keys may be programmed in this dialog in order to quickly turn the desired keyboard on or off. It is recommended to carefully choose an Alt+Function key combination that does not conflict with shortcut keys that are already assigned.

The author has designed other members of the font family for both Eastern and Western variations of Cree, with both Bold and Regular typefaces.

There are other versions of the Windows Roman keyboard available. The one for East Cree (described above) has been set up according to the preferences of the Cree School Board. This places the w-dot to the LEFT and the final u = small circle (°). The keyboard filename is `bjcruni.kmx`.

Versions of the keyboards suitable for Plains Cree, and Woods Cree and Swampy Cree are in development.



## *Macintosh Unicode Keyboarding*

While Apple implemented only limited support in Macintosh OS X for Unicode, subsequent upgrades to the operating system provided increased Unicode support, including (by OS X.3 “Panther”) support for user-definable Unicode keyboards via an XML text file. <ENDNOTE> While writing an XML keyboard definition program by hand is possible, a few utilities have been developed to facilitate the creation of these keyboard layout files. *Ukelele* <ENDNOTE> provides assistance with building a keyboard layout that incorporates modifier keys (Option, Ctrl, Command+key), and *KeyLayoutMaker* <ENDNOTE> provides for situations where many multi-key sequences are required to generate single Unicode characters, as required when typing the roman-spelled sequences to get syllabic characters.

When keyboard layouts created this way are compiled together with all the necessary fonts using Apple’s Developer Tools applications “PackageMaker” and “Installer”, a one-file distribution can be used to install the Unicode fonts and keyboards on a Macintosh computer. The author has created a package that contains all that is required to keyboard Unicode Cree syllabics in Macintosh, named *McCree3EastOSX*.

## McCree3EastOSX Installation

Simply double click on the **McCree3EastOSX.pkg** file in your computer. The self-extracting installer will install the keyboard layouts, and supporting documents, and the fonts. Accepting the default installation folders is recommended. After the installation, you will find the keyboard layout files here:

Macintosh HD\Library\Keyboard Layouts\

And the fonts are installed here:

Macintosh HD\Library\Fonts

As usual, administrator privileges are required for installation of fonts and programs, and user settings may need to be adjusted in order to view files in system folders.

Your system will require you to log out and log on again to make the fonts and keyboards available to the user. Sometimes this occurs automatically during the installation. If it does not, log out manually.

Before using the keyboard for the first time, you may have to activate the International Settings, found in System Preferences.

1. Launch System Preferences (under the Apple Menu)
2. Under "Personal", click the International icon.
3. Under the "Input Menu" tab, turn on the keyboard layouts and input methods to appear in the Input Menu.  **U.S.** may be selected by default. Scroll through the list and find  **BJECree Syllabics** and (if desired)  **Creebrd Syllabics**. You can also select any other input methods desired at this time.
4. Close System Preferences.

## Using the Font(s)

The keyboards and fonts will work in any Macintosh OS X application that supports Unicode. As of this writing an increasing number of commercial software applications that support Unicode are available.

1. Look for the Input Menu icon on the right side of the Menu Bar, near the status icons and the clock, at the top of the screen. The Input Menu icon normally looks like this:



2. Open any Macintosh word processor, or other application.
3. Change the active font to **BJCree UNI**.

To type in Roman and get Cree: 



Now click on the Input Menu icon  in the Menu Bar. Now click on the  **BJECree Syllabics** icon. The icon in the Menu Bar changes to let you know which keyboard is active.

Now type in Roman the way Cree "sounds". You should get syllabics. Double the vowels to get the diacritic (that is, type "paa" to get the long vowel dot over "pa"). "Ch" is "c". or "ch" Back vowel is "u", or "o"

To change back to English, just click on the  icon in the Menu Bar. Choose  **U.S.** You can keep typing in roman using the BJCree UNI font to get normal roman characters.

A complete syllabic chart may be found in:

Macintosh HD\Library\Documentation\McCree 3 East OSX\bjcreeUNIeastmac.pdf



develop some creative font designs that every user of Unicode syllabics should have installed in their fonts folder for variety.

Ron Ogawa's "Ballymun RO" Unicode font is an interesting variation that uses an unorthodox treatment of the glyph stems. (varying stroke width, as with a quill pen) <http://nexus.brocku.ca/rogawa/ucas/>. This site also has some good samples of instructions about "how to configure your browser" for viewing syllabics on the Web.

Another font that contains a huge number of glyphs is "Code2000", used to test browser compatibility with Unicode. But the glyph shapes (for Canadian Syllabics, anyway) tend to be crudely designed. <http://home.att.net/~jameskass/> However, this site contains many resources for people who are working with Unicode fonts on the web, and it is worth checking out.

The Canadian Bible Society's Jeff Klassen has produced CBSUCAS, another full set. It is better looking than some but the version that I have has some proportional, and spacing issues with the "dots" (<http://www.ubs-translations.org/~jklassen/>)

Michael Everson created Everson Mono Unicode. He says on his website: "Everson Mono is a simple, elegant, monowidth font." (<http://www.evertype.com/emono/>)

## Other Syllabic Unicode fonts

Tiro Typeworks, in its work for the Nunavut government, has developed some fine Inuktitut syllabic fonts. Unfortunately, most of their work (as of this writing) does not contain many common Cree syllabic glyphs.

However, a recent visit to the Tiro Typeworks website noted that they are currently developing a new Unicode syllabic font Canadian Syllabics & Latin Scripts named "Euphemia":

"Euphemia covers most languages which use the Canadian Syllabic script including various Cree orthographies, Inuktitut and the historical Carrier script. Three fonts are available with free end-user licenses in TrueType-OpenType format. (Retrieved 15 Oct 2005 [http://www.tiro.com/syllabics/resources/syllabic\\_resources.html](http://www.tiro.com/syllabics/resources/syllabic_resources.html))"

Euphemia now comes with Mac OS X.4, and it is expected to also be released with later Windows OS's. As of this writing, it is not yet available for individual download from the Tiro Typeworks website, but seems to indicate that it will be available "soon".

[http://www.tiro.com/syllabics/resources/syllabic\\_resources.html](http://www.tiro.com/syllabics/resources/syllabic_resources.html)

[http://www.alanwood.net/unicode/unified\\_canadian\\_aboriginal\\_syllabics.html](http://www.alanwood.net/unicode/unified_canadian_aboriginal_syllabics.html)

David McCreedy's Gallery of Unicode Fonts ([http://www.travelphrases.info/gallery/Fonts\\_Cree.html](http://www.travelphrases.info/gallery/Fonts_Cree.html)) appears to be a suitably up-to-date inventory of Unicode fonts available on-line.

## Notes and Acknowledgments:

The syllabic fonts named BJCree UNI were designed by the author. They were originally developed to meet local (Naskapi) needs, and later expanded as needed to suit East Cree, West (Plains) Cree, and TH- (Woods) Cree. It also covers Moose Cree, Severn Cree, Oji-Cree and Ojibwe. The font was developed with extensive reader testing to arrive at the basic shapes, stroke weight and spacing. As noted above there are other Unicode fonts available that have a more complete inventory. The font name is "BJCree UNI" (BJCRUS.TTF) and it contains (in the UCAS block) 351 syllabic glyphs. The glyph shapes were originally laid out using a font editing application named "Fontmonger", and later refined using both "Fontographer" and "FontLab" They conform to the TrueType outline standard.

Tavultesoft Keyboard Manager was written by Marc Durdin. See "About Keyman" in the Keyman configuration menu.

Ukelele was written by John Brownie, SIL International.

KeyLayoutMaker was written by Jonathan Kew, SIL International Non-Roman Script Initiative.

Cree Programs and the Cree Regional Authority provided some of the funding for the development of the BJCree (east) and Creeboard keyboard layouts, originally for the Windows operating system and then for the Macintosh OS. The packages may be freely downloaded for non-governmental and non-commercial use. This document should be included with acknowledgments.

<http://developer.apple.com/technotes/tn2002/pdf/tn2056.pdf>

[http://scripts.sil.org/cms/scripts/page.php?site\\_id=nrsi&item\\_id=ukelele](http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&item_id=ukelele)  
[http://scripts.sil.org/cms/scripts/page.php?site\\_id=nrsi&item\\_id=keylayoutmaker](http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&item_id=keylayoutmaker)  
Bill Jancewicz,  
Preston, Connecticut 2005  
860-889-1798 <bill\_jancewicz@sil.org>