ENCODING THE W-DOT IN SYLLABICS

CHRIS HARVEY © 2010

CONTENTS

WHAT IS IT?

In the syllabics orthographies of the Algonquian languages: Cree, Ojibwe, and Oji-Cree, there are two categories of symbols: the full syllabic character and the final. A *final* is simply a consonant symbol which is written when that sound is at the end of a syllable: the **n** in **giin** (ᑭᐣ) is a final and is written ᐣ or ᓐ (depending on the dialect). A *full syllabic character* is a single symbol which incorporates an initial consonant (or no consonant at all) plus one vowel: ᓇ is pronounced na – the one symbol contains both the **n** sound as well as the vowel **a**.

With finals and full syllabic characters, it is possible to write any syllable which fits the syllable pattern: V, VC, CV or CVC (where C is any consonant and V is any vowel), for example: ᐊ ᐋᐣ ᐸ ᐸᐣ (a, an, pa, pan). In addition to these syllable types, the Algonquian languages mentioned above commonly put a **w**-sound in between a consonant and a vowel: as in **ēkwa**. The syllabification of this word is *ē-kwa,* not *ēk-wa,* so using a k-final here (ᐁᑊᐊ·) would be incorrect (as indicated by the red colour) and clunky. Instead, a mid-level dot is written beside the full syllabic character to indicate the inserted w-sound: ᐁᑲ·. The w-dot is a marker attached to a full syllabic character indicating that a w-sound should be pronounced before the vowel sound.

*For foreign words written in Syllabics, some syllable formations cannot be written with the standard Consonant + Vowel. In these situations, finals can be used in non-standard positions. For example, Chris could be written ᑊᕒᐃᣟ (k-r-i-s). Typically, once a word has been fully adopted into the language it takes a more Algonquian phonetic pattern. Some sounds, like **h**, only have a final form.*

Languages and dialects position the w-dot differently. Generally speaking, the further east you travel, the more likely the w-dot will occur *before* the full syllabic character, the further west, the w-dot will probably occur *after* the full syllabic character. Thus **miinwa** could be ᒪ̇·ᓇ or ᒪ̇ᓇ: depending on the dialect.

SIMILAR SYMBOLS

Syllabics uses a dot-like symbo in different contexts to mean different things.

1. The w-dot (mid-level) indicates a w-sound is pronounced between the consonant and vowel in a full syllabic character. The position of the w-dot – before or after the full syllabic character – depends on dialect. Both ·ᑲ and ᑲ· are pronounced **kwa**.
2. A dot directly above a full syllabic character marks a long vowel: ᑲ has a short vowel **ka,** ᑳ has a long vowel **kā.** Many writers choose not to write the long-vowel dot as very few words are differentiated only by vowel length: **maci** ᒪᒋ *bad* vs **māci** ᒫᒋ *begin to do*. Texts which contain the long-vowel dot are called *pointed* to differentiate them from texts with no vowel-length marking.
3. A high-level dot on its own is used as a y-final in some dialects of Algonquian

*The **e** vowel in Cree, Oji-Cree, and Ojibwe is always long, consequently syllabics writing does not need to put the long-vowel dot atop a full syllabics character with the vowel **e**, ᐁ.*

languages. The y-dot can be seen at the end of the unpointed word **āsay** ᐊᓴ□ *already.*

A word with all of the first three types of dot is ***kīkwāy*** *ᑭᑲ·□ what*

4. Some writers prefer a Latin-style dot-period over the syllabics x-shaped period x
5. In an archaic form of Moose Cree writing, a mid-level dot is used for a w-sound at the *end* of a syllable, in other words, a w-final. Because Moose Cree places its mid-syllable w-dot before the full syllabic character, there is no chance in confusing the mid-syllable w-dot with the final w-dot. For example: **kīsiswēw** ᑭᓯ·ᔦ· *s/he is cooking that animate thing.* The modern w-final is a ring ∘, ᑭᓯ·ᔦ∘.

ENCODING THE W-DOT

Unicode is the current, universal system by which languages of the world encode their letters and characters on computers. See the FAQ section of this web-site for more details on Unicode. In Unicode, syllabics characters are arranged under the acronym UCAS (Unified Canadian Aboriginal Syllabics).

- In terms of the Syllabics encoding, it was decided early on that there be distinct characters for each full syllabic and for each final. Thus ᑲ **ka** is one character, as is ᐢ **n**.
- The long-vowel dot was determined to be an essential part of a syllabics letter and would not be encoded separately. What this means is that ᑳ **kā** is not treated by the computer as ᑲ plus the long-vowel dot, ᑳ is a single character, an indivisible unit. If you try backspacing after the ᑳ character, you will see that it is a single character: the backspace does not delete the long-vowel dot separately.
- Similarly, the w-dot (either left-side or right-side) is understood to be part of a single character. Both ᑊᑲ and ᑲᑊ are single characters, *not* ᑲ plus a w-dot character. This was an important, and in my opinion a correct, decision. To use the specialized "archaic Moose-Cree w-final" as a regular w-dot is not best practice. Here is why:

TRACKING

Tracking is when extra space is added between each letter of a word or sentence to spread the text out slightly. It can also refer to reduced space between letters to crunch words together. Tracking is an automatied process – you don't have to type in the spaces manually, and it is very important in graphic design and typography. In English, adding tracking increases letter spacing one character at a time,

t h i s s e n t e n c e h a s l o o s e t r a c k i n g .

In tracking a syllabics text, we want the w-dot to stay close to the full syllabic character it is modifying.

ᐅ ᐦ ᐃ ᐃ ᑌᑊ ᐃᑊ ᓇ ᐊ ᕀ ᑲ ᐢ ᑳ ᐊᑊ

If the w-dot is processed as a separate character, the result after tracking incorrectly separates the w-dots.

ᐅ ᐦ ᐃ ᐃ ᑌ ᑊ ᐃ ᑊ ᓇ ᐊ ᕀ ᑲ ᐢ ᑳ ᐊ ᑊ

VERTICAL TEXT

When words in syllabics are written vertically from top-to-bottom, the w-dot should stay on the same line as the full syllabic character it is modifying.

Correct Incorrect

In each case, tracking and vertical text, it is important to keep the w-dot and the full syllabic character together as one character. Using the archaic Moose-Cree w-final as a convenient w-dot results in incorrect formatting.

The sequence ᑲ· (**ka** + archaic Moose-Cree w-final) is different from the single character single character ᑳ **kwa**. The former has the Unicode reference 1472 1427 while the latter is 147F. A computer will not recognize these two sequences as the same thing for searches or spell-checking.

IMPLICATIONS

The problem for Syllabics typists is, which dot to use? From a theoretical standpoint, the answer is simple: unless you are typing an obsolete orthography for Moose-Cree or Blackfoot syllabics, stay away from the mid-dot syllabics character U+1427. Use the suggested pre-combined characters which include both the full syllabic plus the w-dot.

From a practical standpoint, the decision is more complex.

1. Many people coming from an older non-Unicode syllabics font will almost certainly be accustomed to a single w-dot key. They would type **kwa** with two keystrokes: a 'ka' key then a 'w-dot' key. The result was two characters. In correct Unicode this should instead be a single characte combining the **ka** and **w**.
    i. For Eastern orthographies, where the w-dot comes before the full syllabic, there is no keyboard-design issue. The w-dot key is pressed first, which calls a *dead key*. A dead key makes the computer wait until the next keystroke before displaying the resulting character. It is way to create one character with two keystrokes.
    ii. For Western orthographies, there is a problem. A typist used to their old idiosyncratic font would expect to type the dot *after* the full syllabic character. There is no elegant way to make the second keystroke the dead key.
2. When editing a sentence, the typist may wish to add in a single w-dot to correct a spelling mistake. The best way to do so is to delete the un-w-dotted character and retype it in its entirety. For example, the word ᐃᑌᐊᐠ is misspelled, there should be a w-dot after the second syllabic character. The person editing would naturally want to position the cursor between the ᑌ and the ᐊ· and type the w-dot key. Instead, they should delete the entire ᑌ and type in ᑌ·.

The solution to problem 2 is simple: typists learn the new system and there should be no access to the archaic Moose Cree mid-dot.

Getting a dead-key-based keyboard as described in 1-i is definitely possible. The problem as outlined in 1-ii, however, is difficult indeed. Any solution would require making the syllabic character itself a dead-key, which causes more technical problems than it solves. The Languagegeek syllabics keyboards use separate keys for t-syllabics and tw-syllabics, k- and kw-, c- and cw- etc. ᑌ is typed *t e* while ᑌ· is typed *r e*.

For a keyboard which strives to mimic a pre-Unicode font layout I would suggest that a

Shift or Alt state be assigned to the w-dot characters. For example:

- ᵬ could be the ‘f’ key.
- ᵭ could be the ‘Shift + f’ key
- ᵬ· could be the ‘Right-Alt + f’ key
- ᵭ· could be the ‘Shift + Right-Alt + f’ key

In practice, this may be too akward to type smoothly.

Home Previous Page