Principles for an Ethiopic Text Editor

Daniel Yacob, Indiana University, IN
Yonas Fisseha, Software Engineer, Motorola Co.,
(yacob@apollo.aoe.vt.edu, fisseha@cig.mot.com)

Recently, when consulting on an editor for Ethiopic script, to be used from the standard Latin keyboards, it became of utility to write down and prioritize the features and principles we wished to implement. The document here was originally produced to aid GNU workers in their implementation of Ethiopic into the Multilingual version of Emacs. It is represented here as an itemized list of the features and concepts we thought valuable with only brief description of the numbered items. A discussion and defense of the list items would be the subject of another paper. Elements of the document are strongly biased towards SERA which was the Latin I/O system we were working with for IM at the time.

Notes:

The ASCII representation of Ethiopic characters in this document adhere to the SERA standards as defined in the papers of Yacob and Firdyiwek. SERA specific characters within quotes mean the actual Latin string whereas those within parentheses are the actual Ethiopic characters corresponding to the Latin string.

*Items marked with * may be considered advanced features and as such are nonessential.*

Input Method

1) Key Mapping

Key mappings of Ethiopic characters onto Latin or other keyboards be done in a logical and consistent manner in agreement with the common sounds that the user will have learned for both character sets.

2) Syllabic Entry

A minimal number of keystrokes be required to enter any character. The Latin characters “e”, “u”, “i”, ”a”, “E”, “o” and “W” are recommended for changing the syllabic form of some base consonant.

3) Dynamical Character Composition

The character being entered is updated at each phase of entry (at each new keystroke).

Example:

> “h” is struck and ሀ appears, “2” or “W” is struck and ኀ or ኋ appears, “E” is struck and finally ኄ or ኌ appears. Dynamic updating also applies for numbers.

4) Default Word Separator is Latin-Blank

The blank space, “ ”, is the default word separator for entry. The Ethiopic separator, “”, may be chosen with an arbitrary toggle key.

5) Default Punctuation is Ethiopic

Statistically, Ethiopic punctuation is in greater use than English punctuation in modern writing. A toggle key is used to enter English punctuation.

6) Default Numeral System is Arabic

Statistically, Arabic (0-9) numbers are in more regular use than Ethiopic. A toggle key is used to enter the Ethiopic numerals.

7) Quotation Entry

Ethiopic quotation marks be associated by glyph similarity to the math symbols < and > . The quotes may be entered with either single or double keystrokes. The math symbols should remain available with either single or double strike, in compliment with the quotation entry.

8) Availability of Native Punctuation

All punctuation provided on the keyboard of entry should be available while in an Ethiopic input mode. This may usually be accomplished by a double strike method whereby the mapped Ethiopic punctuation is exchanged for the native at the 2nd strike. Example : ";" is keyed and ፤ appears, and when ";" is struck a 2nd time ";" replaces ፤ and the entry terminates.

9)* Syllabic Form Editing (Diacritical Mark Edit Only)

A character entered previously may have its syllabic form updated with a single keystroke of the vowel for the desired new syllabic state. Example : ቻ has been entered previously, the character ቻ is selected [by highlighting or by "operate left" principle] and the key "o" is struck, the ቻ glyph becomes ቿ, ቿ remains chosen for additional changes.

10)* Syllabic Form Backspace (UnEnter Diacritical Mark)

A character just entered remains at the top of the edit queue until some new character is keyed. A "backspace" or "delete" key would have the effect of undoing the last vowel entry and returns the glyph to its previous form. Item (6) and (7) may be combined. Example : ሀ has just been keyed in, "backspace" is hit and ሀ → ሀ. An "a" is now keyed and ሀ → ሃ, the edit is complete.

11)* Selective Word Separator On/Off Toggle

The need may arise for the user to exchange the Latin word separator, " ", with the Ethiopic, "፡" or the Ethiopic for Latin. A toggle key or menu option may provide this to a section of selected ("highlighted") text, where all word separators are exchanged within the bounds of the selection.

12)* Julian Date Stamp

A menu item that offers conversion between the Ethio- Julian and Gregorian calendar systems will be an asset of high value to the user.

13)* Old Style Typewriter Entry

Support for the Italian devised system for keyed entry is desirable as the input method remains in popular use.

File I/O

1) Software Native I/O

Considered arbitrary.

2) Latin Transliteration.

Transliteration into a common 7-bit system of mixed Ethiopic and a foreign script is desirable for simple import/exportation between the software of different vendors. Transliteration is also required for encoded file transfer on 7-bit data networks.

a) The system of transliteration be along logical and phonetic guidelines such that the transliterated document may be read readily in encoded form.

b) The transliteration system permits output for the phonetical logics of other Ethiopian languages like Tigrigna, Amharic etc. without causing conflict for input.