IM Principles For An Ethiopic Text Editor

Daniel Yaqob, Indiana University, Bloomington, IN
yacob@padis.gn.apc.org
Yonas Fisseha, Software Engineer, Motorola Co.
fisseha@cig.mot.com

March 2, 1996

Experience with the **S**ystem for **E**thiopic **R**epresentation in **A**SCII (SERA) as an input method has lead to specialized adjustments in the SERA system to facilitate a more natural and rapid input system. This form of SERA, hereafter referred to as "SERA-IM", has been developed for the standard Roman 101 keyboards under principles that are extensible to any keyboard type. Though developed as an ASCII based system for the Ethiopic writing system, the methodologies applied in SERA and SERA-IM are extendible to any syllabic writing system.

This paper describes in brief the application of SERA-IM in a simple text editor. Additional guidelines are offered to describe the features believed to make up a good Ethiopic editor. Familiarity with Ethiopic script and the SERA system will benefit the reader. Multilingual Emacs demonstrates an idea implementation of SERA-IM. *Note:* 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 the base consonant.

3) Dynamical Character Composition

The character being entered is updated at each phase of composition (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[1].

4) Bang-Bang Composition

The phonetical and functional duplicate characters found in lesser use in the present day; 'se,'he,'Se,'a, '?, etc, may be input with a keyboard double strike of the representative phoneme. Example:

"ss" and "aa" would render ሠ and ዐ, "??" would render ፧, etc.

5) Default Word Separator is Latin-Blank

The blank space, " ", is the default word separator for primary entry. The Ethiopic separator, "፡" may be chosen with an arbitrary toggle key. A mode should be provided to exchange the default word separators. Shift-space bar is recommended to provide the current secondary word separator. Note: When space bar is used to enter Ethiopic word space a logic is desirable where space bar strikes following punctuation entry (including word space) will enter blank space in the editor.

6) 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.

7) 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.

8) Functional and Glyph Mapped Punctuation Entry

It is desirable to allow punctuation input following the mapping of the functional equivalence of the native punctuation on the keyboard. It is equally desirable for the typist to key-in punctuation following a glyph mapping of the Ethiopic punctuation onto keyboard's punctuation.

Functional Mapping:

The space bar is hit and ፡ is rendered.
Period is hit and ። is rendered.

[1] Note in this example that "hW" renders the more frequent ኋ and *not* the seldom needed ሗ of transcription SERA. The apostrophe terminator "hW' " is then required to compose ሗ; an equivalent form under transcription rules. Likewise for all consonant classes having 12 forms.

Glyph Mapping:

> Colon is hit and ፡ is rendered.
> Colon is hit again and ። is rendered.
> (Colon hit a third time finally gives Roman ":".)
> Colon and Hyphen are hit and ፦ is rendered.

9) 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.

10)* 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:

> ቸ has been entered previously, the character ቸ is selected [by highlighting or by "operate left" principle] and the key "o" is struck, the ቸ glyph becomes ቾ, ቾ remains chosen for additional changes.

11)* Syllabic Form Backspace (UnEnter Diacritical Mark)

A character just entered remains at the top of the edit stack 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 (2), (3), and (10) may be combined. Example:

> ሁ has just been keyed in, "backspace" is hit and ሁ → ሀ.
> An "a" is now keyed and ሀ → ሃ, the edit is complete.

12)* 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.

13)* 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.

14)* Old Style Typewriter Entry

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

File I/O

1) Software Native I/O

Unicode I/O should be provided; otherwise native I/O handling may be considered arbitrary.

2) Latin Transcription.

Transcription 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. Transcription is also required for encoded file transfer on 7-bit data networks. A good transcription system would provides at a minimum:

a) The system of transcription be along logical and phonetic guidelines such that the transliterated document remains human readable in encoded form.

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

SERA is matured transcription system satisfying these requirements.