Publisher Survey of Ethiopic Layout & Formatting Practices

A Survey of the   
W3C Ethiopic Layout Requirements Task Force

# Letter Shape and Appearance

## English Font Weights

The historic “weight” (the letter thickness) of Ethiopic is heavier than found in European writing. When European words appear in a document, they are sometimes made heavier to match the Ethiopic letters. For Example:

ካገላበጥኳቸው ስራዎች ውስጥ ግን ታዬ አሰፋ እና ሽፈራው በቀለ የተባሉ ምሁራን በጋራ ያሳተሙት The Study of Amharic Literature: An Overview የሚለው ጥናት አጥጋቢ መልስ የሚሰጥ ሆኖ አግኝቼዋሉ።

English words have their usual weight.

ካገላበጥኳቸው ስራዎች ውስጥ ግን ታዬ አሰፋ እና ሽፈራው በቀለ የተባሉ ምሁራን በጋራ ያሳተሙት The Study of Amharic Literature: An Overview የሚለው ጥናት አጥጋቢ መልስ የሚሰጥ ሆኖ አግኝቼዋሉ።

English words are made heavier.

1. When publishing in the classical Ethiopic weight, how should English words appear? in their default English weight?

☐ English default weight ☐ Matching Ethiopic weight

Other: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. If heavier Latin letters are desirable, how much so?

☐ Equal to Ethiopic letter thickness  
Other: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Quotation

In Ethiopic literature two styles of citation marks can be found. They are the guillemets « » and the quote marks “ ”. Both styles have compliments in singular mark forms: ‹ › and ‘ ’.

1. Which style should be the default for Ethiopic literature: ☐ « » ☐ “ ”
2. Are both styles appropriate for Ethiopic literature? ☐ No ☐ Yes (Explain)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. What form is appropriate for outer quotation: ☐ Single ☐ Double
2. What form is appropriate for inner quotation: ☐ Single ☐ Double

## Ethiopicized Punctuation

Modern Ethiopic borrows a number of foreign symbols to show currency, express emphasis, time, mathematics, or communicate with Internet protocols (such as "@" , "://"). The following are believed to be in common use:

1234567890 ? ! ¡ . / () [] {} < = > \ # % & $ : , € @ … \_ - + ± × ÷ ‘ ’ “ ” ‹ › « »

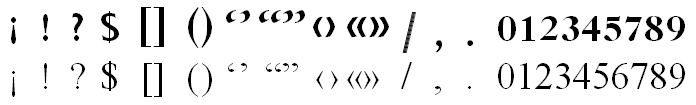
1. Are there any additional western symbols required for Ethiopic writing? If so which:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Are any symbols from other foreign script (Chinese, Arabic, etc.) that should be included? If so which:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

A number of fonts will provide the western punctuation with a change of shape to visually match better with Ge’ez letters. For example:



1. What symbols benefit from shape change?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. What symbols benefit from weight change (but not weight)?

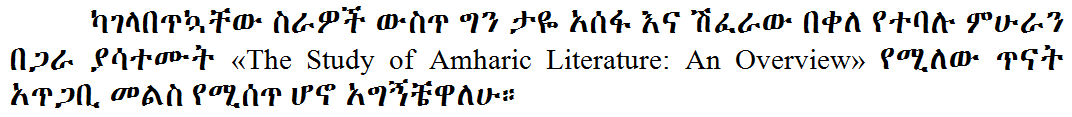
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Quotation of Foreign Phrases

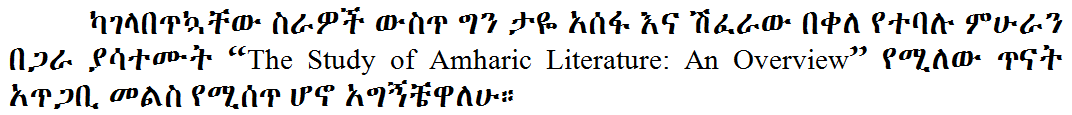
When Ethiopicized punctuation is in use and a foreign phrase is to be quoted, an issue appears regarding the proper weight and style of the quotation marks. Should the quotation marks use the Ethiopic style or those of the foreign phrase? Some examples:

## Macintosh HD:Users:dmekonnen:git:elreq:images:Punctuation-At-Latin-Boundary-Ethiopic-Guillemets.png

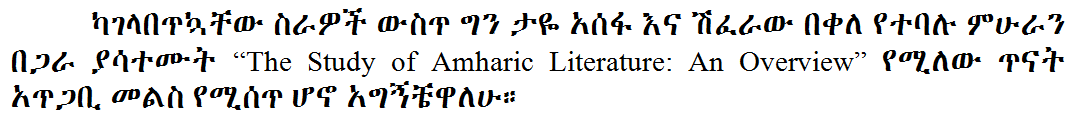
A: Latin text enclosed with European weight guillemets.



B: Latin text enclosed with Ethiopic weight guillemets.



C: Latin text enclosed with Ethiopic weight quotes.

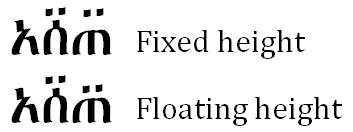


D: Latin text enclosed with European weight quotes.

1. What quotation weight should be used around European phrases:   
    ☐ European ☐ Ethiopic

## Positioning of the Gemination Mark (ጥበቅ)

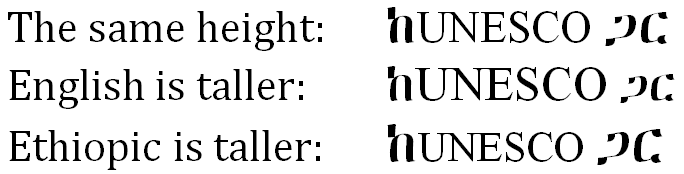
The Ethiopic germination mark has at least a century of use yet is not commonly found outside of linguistic literature. The mark, known as ጥበቅ, helps indicate stress on a consonant (e.g. ገና compared to ገና፟) and may appear in a single dot form. Outside of linguistic works, the germination mark appears most famously in Haddis Alemayehu’s used a *Fikre Eske Makber* where the single dot mark is used. The choice of one or two dots, as well as its position, might be different when written by hand versus a printing technology. For example:



1. The best style for modern publishing is: ☐ fixed ☐ floating

## Height of English Used with Ethiopic

In many Ethiopic documents with English (or European) text the English letters might appear shorter or taller than the Ethiopic. For example:



1. When used together in a sentence, how tall should English letters be?

☐ The same height

☐ English is taller

☐ Ethiopic is taller

# Lines and Paragraph Formatting

1. *How much should a paragraph be indented?*
2. *Should the first paragraph of a section be indented?*
3. *Should a new paragraph be indented following a list?*
4. *Are there any other special rules for when to, or not to, indent a paragraph?*

# Document Layout

Discuss: line composition rules, Punctuations, Mixed script text, Paragraph adjustment, Directionality and bidi, Tab setting, Alignment,[*Justification*](http://w3c.github.io/elreq/#def_justify), Hyphenation, Word/Sentence boundaries, Special cases (e.g. poetry, math, vertical, etc.)

## First Paragraph Rule

Paragraph indentation is a modern Ethiopic practice. The initial paragraph of a section is sometimes not indented. This practice may be idiosyncratic to an author but may represent a convention in use by a publishing house. A best practice supported by stakeholders should be established here.

Issues:

1. *How much should a paragraph be indented?*
2. *Should the first paragraph of a section be indented?*
3. *Should a new paragraph be indented following a list?*
4. *Are there any other special rules for when to, or not to, indent a paragraph?*

## Hyphenation

* A word is split at a location where the user would not be able to infer from context whether one or two words appears across lines. I.E. a split should not occur at the boundary of a compound word.
* When whitespace is the word delimiter, the hyphen symbol should be applied when a word is split over lines.
* Numbers are not split over lines.
* Words of a single syllable are not split over lines.
* Words of two letters are not split over lines.
* Only letters, numbers, or opening punctuation may start a new line.
* Words are split along a syllable boundary.
* A consonant cluster may not be split.

## Word Boundaries

Regular rules apply? One consideration for Ethiopic would be that a newline is not a dependable word boundary in the common case where words are split across lines without a hyphen symbol and white space is the default wordspace. Word boundaries are known here only by context.

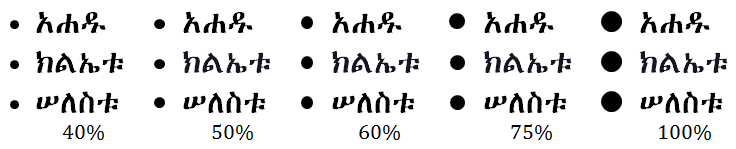
## Line Breaking

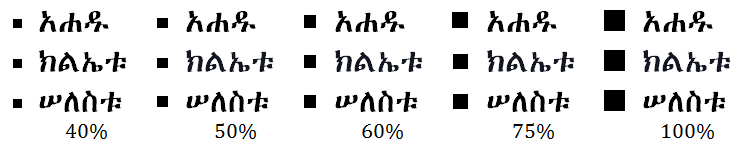
Are western rules ok for Modern Ethiopic? Cover breaking rules with Ethiopic Wordspace in Classical Ethiopic

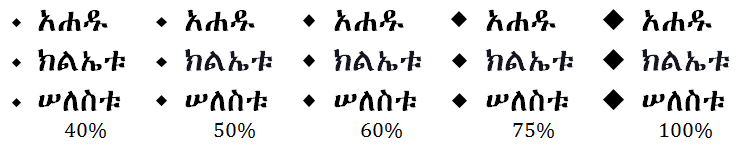
# List Styles

## Bullet Lists

Bullet lists are utilized regularly in Ethiopic literature. Authors using a computer or typewriter will work with the list marker symbols made available by their software or machine. Many marker, or “bullet”, symbols are accepted for Ethiopic literature though not all will be considered optimal.

Fig. 23Circle bullet size relative to letter height.

Fig. 24Square bullet size relative to letter height.

Fig. 25 Diamond bullet size relative to letter height.

Issues:

1. *What are the preferred bullet shapes?*
2. *What are the preferred shape sequences in nested lists (e.g. bullet hierarchy)?*
3. *What is bullet height and width relative to letter dimensions?*
4. *What is the default distance of the list item text from the bullet?*
5. *Are lists indented (left margin)? If so, how much?*
6. *Examples are needed for pre-digital, classical, literature.*
7. Diamond shaped bullets are found commonly and may represent a desired default. The Unicode standard provides a number of similar diamond shaped symbols. From this set of diamond symbols, those suitable for adorning Ethiopic bullet lists must be determined.

⬩ (U+2B29)  
◆ (U+25C6)  
⬥ (U+2B25)  
❖ (U+2756)  
♦ (U+2666)

1. Do copy editors or publishers set a policy for list symbols, or leave it up to the author to decide?
2. Of interest: when are bullet lists first found in Ethiopic writing and in what form?

## Ordered List Item Marker

In Ethiopic ordered lists a number of symbols are used for the list marker (suffix). For example "፦" , "." , ")" and even "፡" (Ethiopic Wordspace).

Issues:

1. *What are the preferred list marker symbols? What is the best default?*
2. *How much, if any, white space follows the suffix?*
3. *Do copy editors or publishers set a policy for list item markers, or leave it up to the author to decide?*
4. *Of interest: when are ordered lists first found in Ethiopic writing and in what form?*

## List Item Marker Alignment

Ethiopic corpus will present lists with two styles of alignment. These are a left side alignment at the list counter, or alignment along the list marker (suffix). Layout software will align a list at the marker in keeping with the later style. The former style (left justified at counter) may reflect a limitation of the layout technology employed and not a preference of the author, copy editor or typesetter. A depiction of these two alignment styles is presented in the following figures:

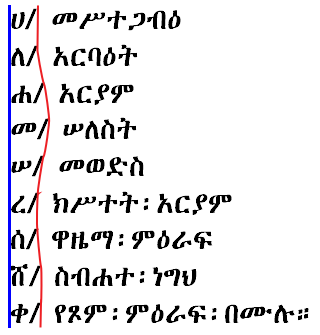


Fig. 26 Alphabetical list aligned justified on counter left.

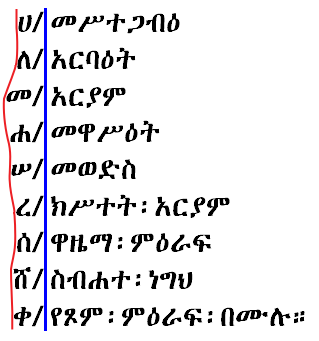


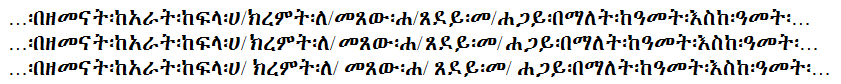
Fig. 27 Alphabetical list aligned on marker.

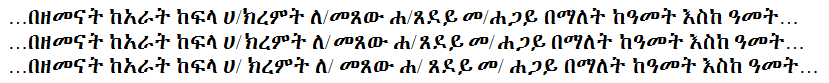
Issues:

1. *What alignment style(s) do stakeholders desire?*

## Inline Lists

How much space after marker? Is the space after marker the different when Ethiopic wordspace is in use?

Fig. 28 Ethiopic inline list sample using wordspaces with 3 spacings (none, hair and thin) following list marker.

Fig. 29 Ethiopic inline list sample using spaces with 3 spacings (none, hair and thin) following list marker.

## Numbered Lists

Ethiopic literature will apply ordered numbered list using both Ethiopic and Western numeral systems. Ethiopic numeral lists are addressed in the [Ethiopic Numeric Counter Style](https://www.w3.org/TR/css-counter-styles-3/#valdef-counter-style-name-ethiopic-numeric) section of the [CSS Counter Styles Level 3](https://www.w3.org/TR/css-counter-styles-3/)specification.

Issues:

1. *What is the list item marker preference with Ethiopic numbers?*
2. *What is the list item marker preference with western numbers?*
3. *Are there any sublist (nested list) considerations?*

## Alphabetical (ሀለሐመ) Lists

The Unicode standard encodes Ethiopic syllables for many languages using Ethiopic script past and present. Alphabetic lists are commonplace in Ethiopic literature, but will conform to the letter inventory of the language of the surrounding content. The alphabetical style lists specified here do not encompass all languages using Ethiopic, but only those with a demonstrated requirement as found in corpus or stakeholder input.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ge'ez** | **Amharic** | **Blin** | **Tigrinya (Eritrean)** | **Tigrinya (Ethiopian)** |
| * ሀ/ * ለ/ * ሐ/ * መ/ * ሠ/ * ረ/ * ሰ/ * ቀ/ * በ/ * ተ/ * ኀ/ * ነ/ * አ/ * ከ/ * ወ/ * ዐ/ * ዘ/ * የ/ * ደ/ * ገ/ * ጠ/ * ጰ/ * ጸ/ * ፀ/ * ፈ/ * ፐ/ | * ሀ/ * ለ/ * ሐ/ * መ/ * ሠ/ * ረ/ * ሰ/ * ሸ/ * ቀ/ * በ/ * ተ/ * ቸ/ * ኀ/ * ነ/ * ኘ/ * አ/ * ከ/ * ኸ/ * ወ/ * ዐ/ * ዘ/ * ዠ/ * የ/ * ደ/ * ጀ/ * ገ/ * ጠ/ * ጨ/ * ጰ/ * ጸ/ * ፀ/ * ፈ/ * ፐ/ | * ሀ/ * ለ/ * ሐ/ * መ/ * ረ/ * ሰ/ * ሸ/ * ቀ/ * ቐ/ * በ/ * ተ/ * ቸ/ * ኀ/ * ነ/ * ኘ/ * አ/ * ከ/ * ኸ/ * ወ/ * ዐ/ * ዘ/ * ዠ/ * የ/ * ደ/ * ጀ/ * ገ/ * ጘ/ * ጠ/ * ጨ/ * ጰ/ * ጸ/ * ፀ/ * ፈ/ * ፐ/ | * ሀ/ * ለ/ * ሐ/ * መ/ * ረ/ * ሰ/ * ሸ/ * ቀ/ * ቐ/ * በ/ * ተ/ * ቸ/ * ነ/ * ኘ/ * አ/ * ከ/ * ኸ/ * ወ/ * ዐ/ * ዘ/ * ዠ/ * የ/ * ደ/ * ጀ/ * ገ/ * ጠ/ * ጨ/ * ጰ/ * ጸ/ * ፈ/ * ፐ/ | * ሀ/ * ለ/ * ሐ/ * መ/ * ሠ/ * ረ/ * ሰ/ * ሸ/ * ቀ/ * ቐ/ * በ/ * ተ/ * ቸ/ * ኀ/ * ነ/ * ኘ/ * አ/ * ከ/ * ኸ/ * ወ/ * ዐ/ * ዘ/ * ዠ/ * የ/ * ደ/ * ጀ/ * ገ/ * ጠ/ * ጨ/ * ጰ/ * ጸ/ * ፀ/ * ፈ/ * ፐ/ |

These tables are expressed in the following as per the [CSS Counter Styles Level 3](https://www.w3.org/TR/css-counter-styles-3/) recommendation.

@**counter**-**style** **ethiopic-halehame-gez** { **system**: **alphabetic**; **speak-as**: **upper-latin**; **symbols**: ሀ ለ ሐ መ ሠ ረ ሰ ቀ በ ተ ኀ ነ አ ከ ወ ዐ ዘ የ ደ ገ ጠ ጰ ጸ ፀ ፈ ፐ ; **suffix**: "፦&**#x2009**;"; }

@counter-style **ethiopic-halehame-am-et** { system: alphabetic; speak-**as**: upper-latin; symbols: ሀ ለ ሐ መ ሠ ረ ሰ ሸ ቀ በ ተ ቸ ኀ ነ ኘ አ ከ ኸ ወ ዐ ዘ ዠ የ ደ ጀ ገ ጠ ጨ ጰ ጸ ፀ ፈ ፐ ; suffix: "/&#x2009;"; }

@counter-style **ethiopic-halehame-ti-et** { system: alphabetic; speak-**as**: upper-latin; symbols: ሀ ለ ሐ መ ሠ ረ ሰ ሸ ቀ ቐ በ ተ ቸ ኀ ነ ኘ አ ከ ኸ ወ ዐ ዘ ዠ የ ደ ጀ ገ ጠ ጨ ጰ ጸ ፀ ፈ ፐ ; suffix: "/&#x2009;"; }

@counter-style **ethiopic-halehame-ti-er** { system: alphabetic; speak-**as**: upper-latin; symbols: ሀ ለ ሐ መ ረ ሰ ሸ ቀ ቐ በ ተ ቸ ነ ኘ አ ከ ኸ ወ ዐ ዘ ዠ የ ደ ጀ ገ ጠ ጨ ጰ ጸ ፈ ፐ ; suffix: "/&#x2009;"; }

@counter-style **ethiopic-halehame-byn** { system: alphabetic; speak-**as**: upper-latin; symbols: ሀ ለ ሐ መ ረ ሰ ሸ ቀ ቐ በ ተ ቸ ነ ኘ አ ከ ኸ ወ ዐ ዘ ዠ የ ደ ጀ ገ ጘ ጠ ጨ ጰ ጸ ፈ ፐ ; suffix: "/&#x2009;"; }

Issues:

1. *ቨ appears to be equally included and omitted in lists.*
2. *How is end of sequence handled?*
   1. *Sequence as radix, e.g.: ... ፈ ... ፐ ... ሀሀ ... ሀለ ...*
   2. *Columnar radix, e.g.: ... ፈ ... ፐ ... ሀሁ ... ሀሉ ...*
   3. *Wrap to next column, e.g.: ... ፈ ... ፐ ... ሁ ... ሉ ...*
   4. *With any approach, how is the end of matrix handled?*

## አበገደ Lists

The አበገደ ordering of the Ethiopic syllabary is an alignment with the Coptic and Greek alphabets possibly to facilitate interdenominational communication or for the transfer of gematria practices. The ordering is used today largely for pedagogical purposes and has been used by some authors for the collation of entire works such as dictionaries. More often authors will apply the ordering for list orders.

The አበገደ ordering is potentially desirable to any language using the Ethiopic syllabary. The ordering is less likely to be found in the writing practices of languages that have a written tradition of under a hundred years. The language specific orders shown here are only those found utilized in corpus.

|  |  |  |  |
| --- | --- | --- | --- |
| **Ge'ez** | **Amharic** | **Tigrinya (Eritrean)** | **Tigrinya (Ethiopian)** |
| * አ/ * በ/ * ገ/ * ደ/ * ሀ/ * ወ/ * ዘ/ * ሐ/ * ጠ/ * የ/ * ከ/ * ለ/ * መ/ * ነ/ * ሠ/ * ዐ/ * ፈ/ * ጸ/ * ቀ/ * ረ/ * ሰ/ * ተ/ * ኀ/ * ፀ/ * ጰ/ * ፐ/ | * አ/ * በ/ * ገ/ * ደ/ * ጀ/ * ሀ/ * ወ/ * ዘ/ * ዠ/ * ሐ/ * ጠ/ * ጨ/ * የ/ * ከ/ * ኸ/ * ለ/ * መ/ * ነ/ * ኘ/ * ሠ/ * ዐ/ * ፈ/ * ጸ/ * ቀ/ * ረ/ * ሰ/ * ሸ/ * ተ/ * ቸ/ * ኀ/ * ፀ/ * ጰ/ * ፐ/ | * አ/ * በ/ * ገ/ * ደ/ * ጀ/ * ሀ/ * ወ/ * ዘ/ * ዠ/ * ሐ/ * ጠ/ * ጨ/ * የ/ * ከ/ * ኸ/ * ለ/ * መ/ * ነ/ * ኘ/ * ዐ/ * ፈ/ * ጸ/ * ቀ/ * ቐ/ * ረ/ * ሰ/ * ሸ/ * ተ/ * ቸ/ * ጰ/ * ፐ/ | * አ/ * በ/ * ገ/ * ደ/ * ጀ/ * ሀ/ * ወ/ * ዘ/ * ዠ/ * ሐ/ * ጠ/ * ጨ/ * የ/ * ከ/ * ኸ/ * ለ/ * መ/ * ነ/ * ኘ/ * ሠ/ * ዐ/ * ፈ/ * ጸ/ * ቀ/ * ቐ/ * ረ/ * ሰ/ * ሸ/ * ተ/ * ቸ/ * ኀ/ * ፀ/ * ጰ/ * ፐ/ |

These tables are expressed in the following as per the [CSS Counter Styles Level 3](https://www.w3.org/TR/css-counter-styles-3/) recommendation.

@**counter**-**style** **ethiopic-abegede-gez** { **system**: **alphabetic**; **speak-as**: **upper-latin**; **symbols**: አ በ ገ ደ ሀ ወ ዘ ሐ ጠ የ ከ ለ መ ነ ሠ ዐ ፈ ጸ ቀ ረ ሰ ተ ኀ ፀ ጰ ፐ ; **suffix**: "፦&**#x2009**;"; }

@counter-style **ethiopic-abegede-am-et** { system: alphabetic; speak-**as**: upper-latin; symbols: አ በ ገ ደ ጀ ሀ ወ ዘ ዠ ሐ ጠ ጨ የ ከ ኸ ለ መ ነ ኘ ሠ ዐ ፈ ጸ ቀ ረ ሰ ሸ ተ ቸ ኀ ፀ ጰ ፐ ; suffix: "/&#x2009;"; }

@counter-style **ethiopic-abegede-ti-et** { system: alphabetic; speak-**as**: upper-latin; symbols: አ በ ገ ደ ጀ ሀ ወ ዘ ዠ ሐ ጠ ጨ የ ከ ኸ ለ መ ነ ኘ ሠ ዐ ፈ ጸ ቀ ቐ ረ ሰ ሸ ተ ቸ ኀ ፀ ጰ ፐ ; suffix: "/&#x2009;"; }

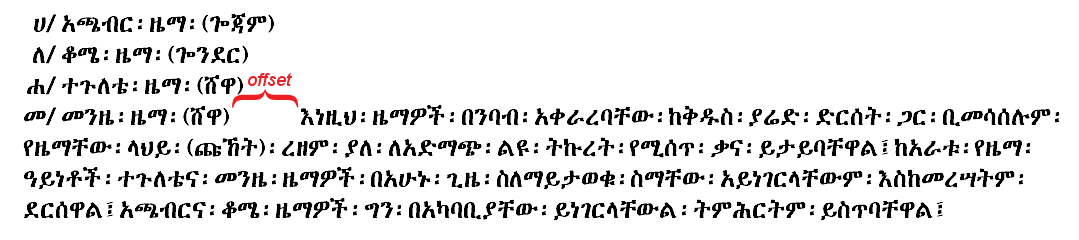
@counter-style **ethiopic-abegede-ti-er** { system: alphabetic; speak-**as**: upper-latin; symbols: አ በ ገ ደ ጀ ሀ ወ ዘ ዠ ሐ ጠ ጨ የ ከ ኸ ለ መ ነ ኘ ሠ ዐ ፈ ጸ ቀ ቐ ረ ሰ ሸ ተ ቸ ኀ ጰ ፐ ; suffix: "/&#x2009;"; }

Issues:

1. *ቨ appears to be equally included and omitted in lists.*
2. *How is end of sequence handled?*
   1. *Sequence as radix, e.g.: ... ጰ ... ፐ ... አአ ... አበ ...*
   2. *Columnar radix, e.g.: ... ጰ ... ፐ ... አቡ ... አጉ ...*
   3. *Wrap to next column, e.g.: ... ጰ ... ፐ ... ቡ ... ጉ ...*
   4. *With any approach, how is the end of matrix handled?*

## Inline End of List Continuation

An observed formatting practice is the start of a following paragraph inline with the last list item. The paragraph may flow immediately from the last item, or some indentation may be applied. This practice is illustrated in the following figure:

Fig. 30Sample for inlined paragraph continuation at end of list.

Issues:

1. *Is this an older style? Idiosyncratic to a few authors? The desired norm?*
2. *When appropriate, what is the width of the paragraph offset (if any) from the final list item?*

# Ethiopic Numeral Formatting

## Ge’ez Numeral Bars

The Ge’ez numbers can be found written in two styles where the upper and lower bars (lines) can be either separated or connected.

|  |  |
| --- | --- |
|  | Isolated style – the numeral bars do not connect. |
| Joining style – the numeral bars are connected. |

**Survey Questions:**

1. Which style should be the default? ☐ Isolated ☐ Joined
2. Are both styles desirable? ☐ Yes ☐ No
3. If both styles are desired:
   1. When would the isolated style be used? Please explain:  
      \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
      \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
   2. When would the joined style be used? Please explain:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Ge’ez Numeral Vertical Alignment

The Ge’ez numbers use a system where an amount greater than 10 can be written with a single number. For example 50 uses the single number ፶ and 100 uses the single number ፻. This difference leads to different ways the numbers can be aligned vertically. Three examples are:

|  |  |  |
| --- | --- | --- |
| **Left Justify** | **Decimal Place Justify** | **Right Justify** |
| ............... ፻፶፭  ............... ፶፭  ............... ፭  ............... ፻  ............... ፲  ............... ፩ | ............... ፻፶፭  ............... ፶፭  ............... ፭  ............... ፻  ............... ፲  ............... ፩ | ............... ፻፶፭  ............... ፶፭  ............... ፭  ............... ፻  ............... ፲  ............... ፩ |

1. Which alignment style makes the best default?

☐ Left ☐ Right ☐ Decimal ☐ Other (Explain Below)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Are the vertical alignment rules applied for Roman numerals applicable for Ethiopic [TBD: determine the Roman numeral rules]?

☐ Yes ☐ No

# Ethiopic Wordspace Formatting

## Ethiopic Wordspace at Latin Boundaries

Classical era Ethiopic documents do not exhibit a very high degree of consistency with respect to the presence of Ethiopic wordspace before or following Latin punctuation. Exceptions to an apparent convention are usually found within an individual document, making internal consistency problematic for a work of any significant length beyond 20 pages. It is likely that the lack of a well established convention here would have kept copy editors and publishers view formatting conformance here as critical. Observed practices are:

**Always**

An Ethiopic word ends with either a wordspace or punctuation. This rule allows a word to be split across lines without the need for a hyphenation symbol. Ethiopic wordspace does not precede or follow another punctuation, with the possible exception when preceding a right-side enclosing symbol (brackets and quotation).

**Right-Side Enclosing Symbol Convention 1**

Ethiopic wordspace does not precede but does follow a right-side enclosing symbol. This appears to be the most commonly used convention during imperial times.

**Right-Side Enclosing Symbol Convention 2**

Ethiopic wordspace will precede a right-side enclosing symbol but not follow it. In these cases white space generally appears after the Latin punctuation, this whitespace is however observed as an artifact of justification spacing and does not demonstrate an inserted space symbol. The figure in this document illustrating [centered wordspace justification style](http://w3c.github.io/elreq/#fig_justification_centered) provides an example of white space after quotation. ፍቅር እስከ መቃብር by Haddis Alemayehu (Alemayehu, 1965 (1958 EC)) is also an easily found reference here.

**Right-Side Enclosing Symbol Convention 3**

Ethiopic wordspace does not precede or follow the right-side enclosing symbol. This style is observed but some width of white space will appear outside the enclosing symbols.

[TBD: somewhere in this document mention the mouse text selection rule that a following wordspace should be automatically selected with text, analagous to the rule applied to white space. MS Word does this.]

## Rules of Formatting the Ethiopic Wordspace-Punctuation Boundary

This section has become redundant and should be combined with[*Ethiopic Wordspace at Latin Boundaries*](http://w3c.github.io/elreq/#ethiopic_wordspace_at_latin_boundaries)

## Rules of Applying Whitespace in Ethiopic Text

The intent of this section is to discuss when white space (and what width) can be added to documents using Ethiopic wordspace. Examples of when this is found to occur (some of these topics are discussed elsewhere):

* *indentation*
* *following a list item marker*
* *within a numeric sequence, like a phone number*
* *within a date format (sometimes, not always)*
* *within a scripture reference (sometimes not always)*
* *before and following quotation marks*
* *following an ellipsis (?)*
* *within a block of embedded foreign script*

[image samples needed]

## Justification

### Justification When ETHIOPIC WORDSPACE is the Word Delimiter

Since the arrival of the printing press to Ethiopia in 1863 (Pankhurst, 1998), full justification of Ethiopic has been a common typesetting practice in Ethiopian, and later Eritrean, publishing houses. Earlier, Ethiopic justification rules are a feature of Hiob Ludolf’s Historia Æthiopica which is noted as the first use of movable type for Ethiopic script (Ludolf, 1681). Prior to letterpress typography, calligraphic manuscripts rendered on parchment also featured full, or approximately full, justification. Though the latter likely reflects the scribe’s desire not to waste a millimeter of available lateral writing space.

The placement of Ethiopic wordspace presents a complication to the justification of Ethiopic text. Two placement styles developed in typeset literature which will be referred to here as “word bounded” and “centered” styles. Additionally, the word spacing following an Ethiopic fullstop may (or may not) be governed by a special rule and in combination with the two wordspace spacing styles. These spacing rules are discussed in the following sections.

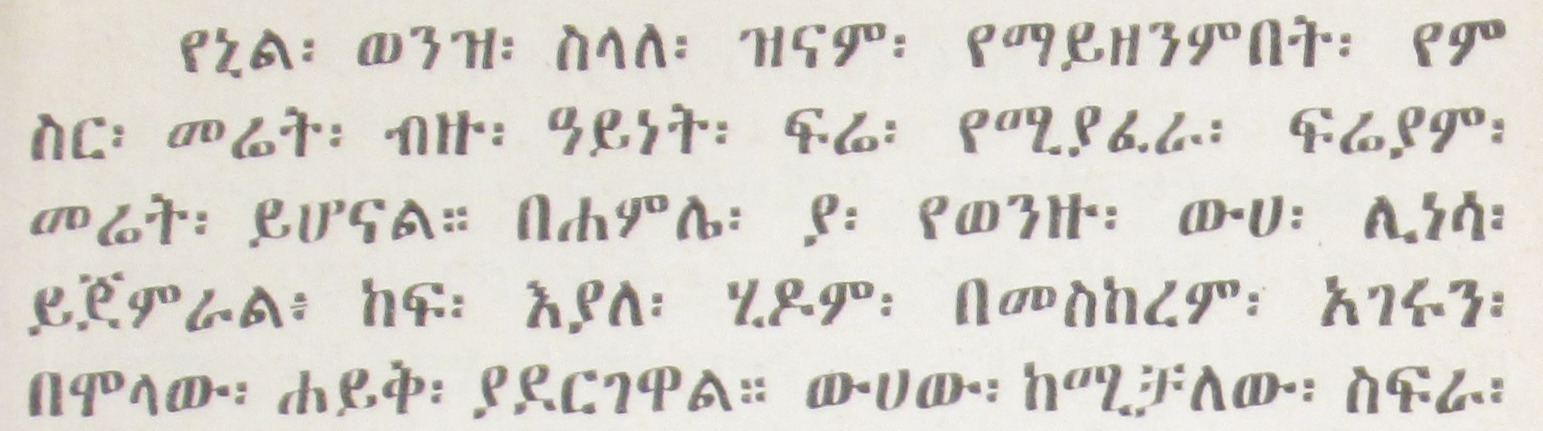


Fig. 15 Ethiopic Justification in Historia Æthiopica (Ludolf, 1681)

### Justification with Word Bounded Wordspace and Punctuation

In keeping with line justification for Latin script, the non-printed or “white space” between words is treated as stretchable. The width of the space symbol itself will be elongated to some aesthetic width value that may vary from space symbol to space symbol across a printed line. In Ethiopic justification, the white space between the Ethiopic word separator and the words it separates is likewise allowed to stretch. This stretching of white space may be either symmetrical (“centered”) or asymmetrical but in the latter case space stretching is always between the right side of the separator and the following word –referred to here as “word bounded”.

In “word bounded” justification the word separator, which may be either a punctuation symbol or U+1361 ETHIOPIC WORDSPACE [፡], appears to adhere to the word to the left as if it were its final character. Figures [Fig.16Ethiopic justification in word bounded style (Erikson, 1921 (1913 EC))](http://w3c.github.io/elreq/#justification_word_bounded) and [Fig. 15Ethiopic Justification inHistoria Æthiopica (Ludolf, 1681)](http://w3c.github.io/elreq/#justification_in_historia_aethiopica) both illustrate the word bounded style.

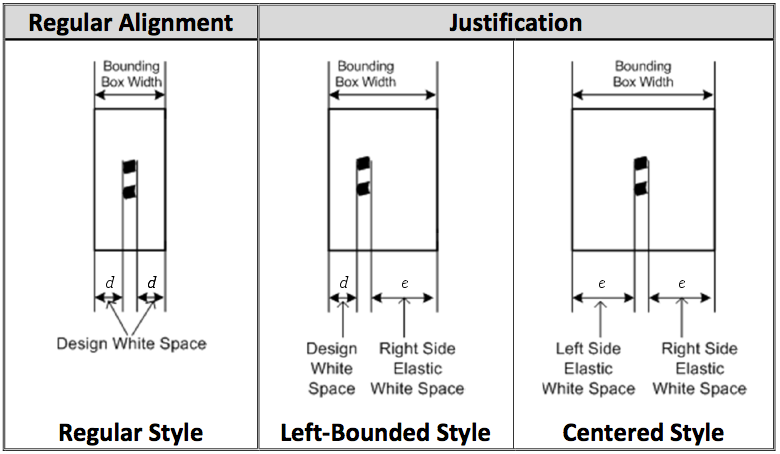
Fig. 16 Ethiopic justification in word bounded style (Erikson, 1921 (1913 EC))

### Justification with Centered Wordspace and Punctuation

In the second major form of Ethiopic justification the white space around word separators is stretched equally on both the left and right sides; giving the appearance of the separator being centered between the words it divides.[Fig. 17Ethiopic justification in centered style (Gubenya, 1973 (1966 EC)).](http://w3c.github.io/elreq/#fig_justification_centered) depicts the white space stretching in both forms of justification using Ethiopic Wordspace as an example, though the stretching rules apply equally to Ethiopic punctuation as well.

Fig. 17Ethiopic justification in centered style (Gubenya, 1973 (1966 EC)).

To further illustrate the justification spacing applied to both Ethiopic punctuation and wordspace, [Fig. 18Depiction of white space around Ethiopic wordspace for three modes of text justification.](http://w3c.github.io/elreq/#fig_white_space_round_wordspace) presents white space stretching from the point of view of the symbol’s typographic bounding box. Here the “design white space”, the space between the visible symbol and the box border, is itself stretched as needed to meet line justification:

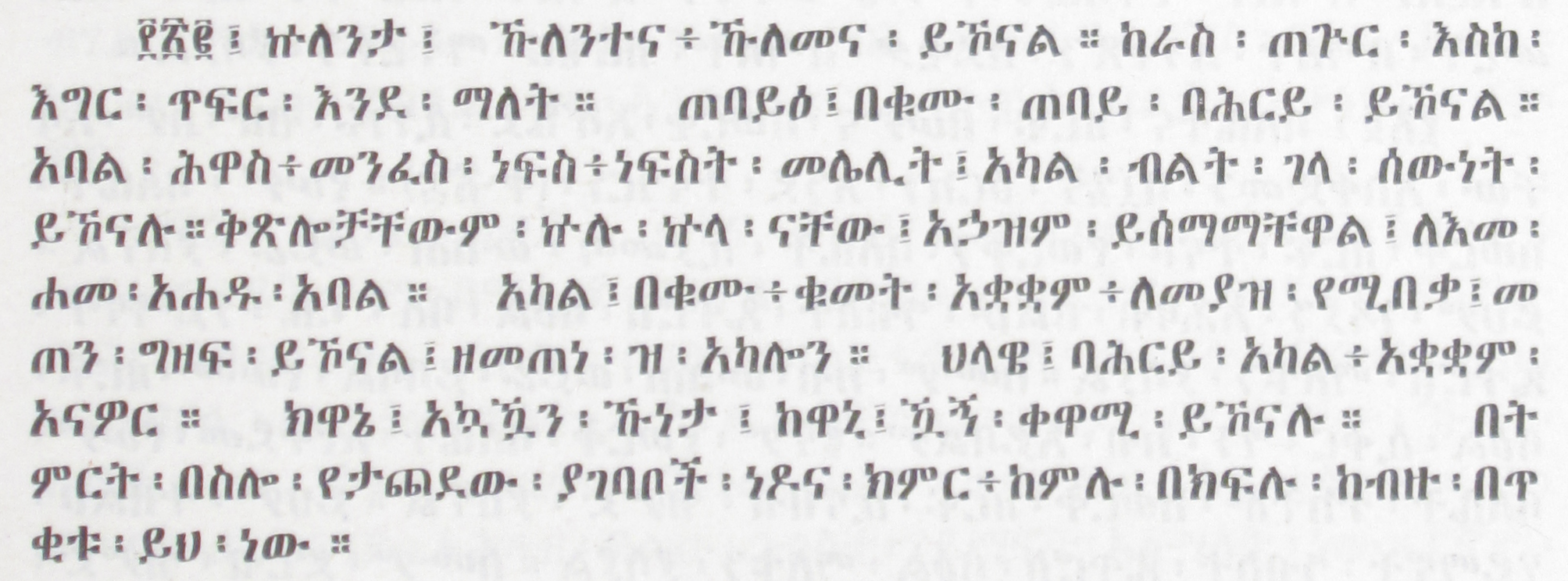
Fig. 18 Depiction of white space around Ethiopic wordspace for three modes of text justification.

Issues:

1. *Which spacing style should be the default used by applications?*
2. *Is the original term, “left bounded”, a better (more descriptive) term than “word bounded”?*

### Spacing Threshold Following Full Stop

In the regular mode of Ethiopic justification (both forms) U+1362 ETHIOPIC FULL STOP [።] will be treated equally with all other punctuation symbols. In a second mode, the Ethiopic full stop will have special spacing rules applied to it whereby more separation space is allowed following the symbol and the start of the next word. In a sense, the right side space of the full stop is “more elastic” than in the regular mode. The elasticity rule and the visual effect are similar to that of the final line of a fully justified paragraph in Western text. When the final line of a paragraph of Latin script crosses a certain horizontal threshold, the line will become fully justified. Below that threshold the line will appear left aligned. The same rule appears to be applied to the Ethiopic full stop but on any line of the paragraph. An illustration of this sub-mode is depicted in the following:

Fig. 19 Justification Sample from Maṣḥafa Sawāsew, Page 159 (Kifle, 1955 (1948 EC))

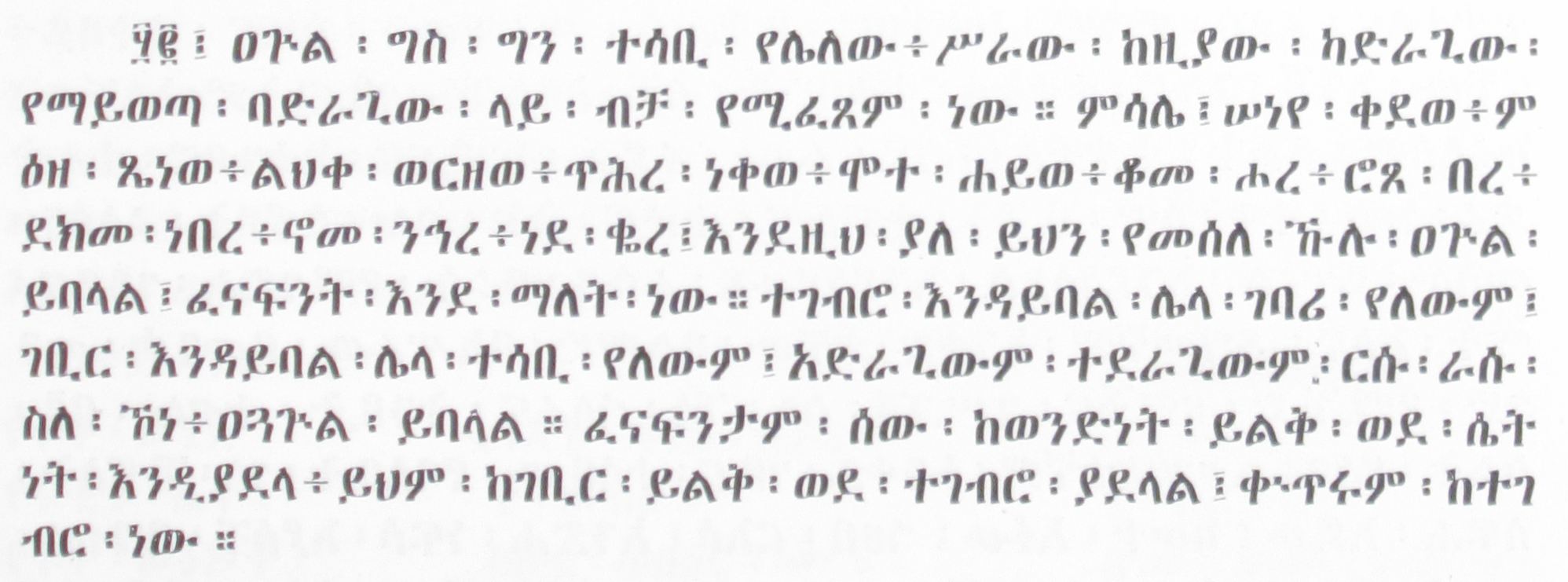
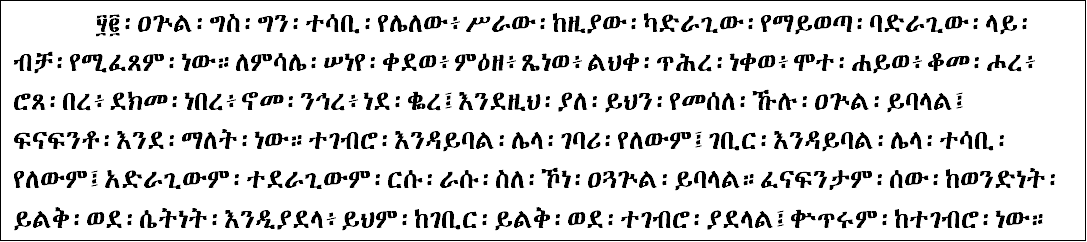
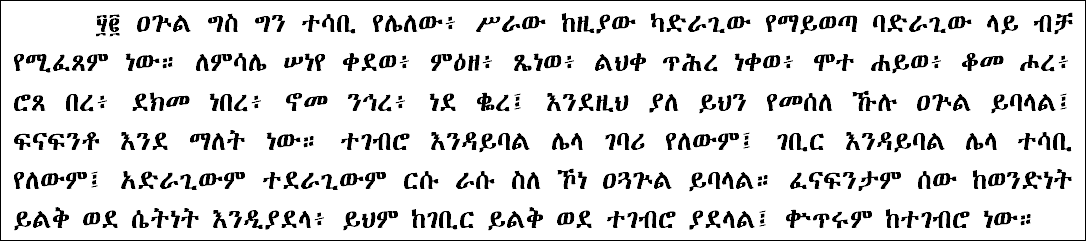
Issues:

1. *Does the additional spacing after fullstop apply to other Ethiopic punctuation? Additional spacing is also seen in the above sample following ETHIOPIC SEMICOLON.*
2. *Is additional spacing after punctuation desirable by current day publishers? If so, how much space and for which punctuation symbols?*

### Shortcomings Found in Electronic Typesetting Systems

To date, computer software that typesets text has applied justification rules for white space stretching that were designed to meet publishing requirements in the Western world. When the same rules are applied to Ethiopic text, the results are unsatisfactory as they do not meet user expectations. Largely responsible for the formatting dissonance when Western justification is applied to Ethiopic text, is the absence of a white space symbol in the writing system. There is no explicit white space symbol (in classic Ethiopic writing) to be “stretched”.

Formatting algorithms will then process U+1361 ETHIOPIC WORDSPACE [፡] as a punctuation symbol where word enclosing rules, rather than word spacing rules, will be applied. While still stretchable, “white space” in the Ethiopic wordspace is implicit rather than explicit. For a complete solution, software will ultimately need to be enhanced to stretch implicit space as required. Reclassifying the Ethiopic wordspace as a “Zs” symbol is expected to help alleviate justification issues and clears the way for software firms to implement comprehensive support for Ethiopic justification. Since the Ethiopic wordspace interferes with justification in present day software, users may opt not to use it or may “pad” wordspace and Ethiopic punctuation with explicit white space to obtain the formatting desired. The following samples depict formatting of Kidane Wolde Kifle’s seminal workMaṣḥafa Sawāsew with a popular word processor (Kifle, 1955 (1948 EC)) under the limitations of Western spacing rules justification.

Fig. 20Full Justification Sample from Maṣḥafa Sawāsew, Page 65 (Kifle, 1955 (1948 EC))Fig. 21The sample from the previous figure with a 12 point font and fully justified by Microsoft Word 2010 within a 6.5 inch margin with line-breaking rules applied. No spaces (U+0020) in sample.Fig. 22The sample from the earlier figure with space (U+0020) replacing Ethiopic wordspace. Space characters have additionally been added following Ethiopic punctuation.

In digital documents such as in web pages and eBooks, it is recommended that the appearance of either U+0020 SPACE [ ] or U+1361 ETHIOPIC WORDSPACE [፡] be configurable as a user preference. An easy to access “space” toggle button would enhance a viewing application’s usability.

# Writing Style Guidance

## Abbreviation formation

Abbreviations in Ethiopic languages will apply an abbreviation marker ("/" or ".") placed between the first letters of each word in a phrase. Single word abbreviation that applies a right side truncation (vs mid) will apply a "." as the abbreviation maker. In a multi-word abbreviations the last word may remain whole. Two letter words are only abbreviated in a multi-word abbreviation.

**Examples:**  
Single Word  
ሚኒስትር ⇒ ሚ/ር  
ሆስፒታል ⇒ ሆ/ል  
Multi Word  
ጠቅላይ ሚኒስትር ⇒ ጠ/ሚ/ር  
ኢትዮጵያ ኦርቶዶክስ ተዋሕዶ ቤተ ክርስቲያን ⇒ ኢ/ኦ/ተ/ቤ/ክ

Issues:

1. *When "." is used as the abbreviation marker, it is sometimes found at the end of the abbreviation.*
2. *Confirm single word abbreviation rule.*
3. *Are mixed abbreviation markers appropriate? Some well established abbreviations apply a "." as the marker, for example "አዲስ አበባ ዩኒቨርሲቲ" becomes "አ.አ.ዩ". This may reflect the chosen convention of the institute.*
4. *What symbol is applied for foreign words transcribed into an Ethiopic languages?*
5. *Do copy editors or publishers set a policy for abbreviation, or leave it up to the author to decide?*
6. *Of interest: when are abbreviations first found in Ethiopic writing and in what form?*

## Ordinal Notation

An ordinal is formed in Amharic when "ኛ", and in Tigrinya when "ይ", follows a cardinal number. The ordinal marker is often, but not always, rendered in superscript form. The superscript practice is most prevalent with ordinals in western numerals, but is also applied with Ethiopic numerals.

Issues:

1. *What letters should be superscripted, and it what contexts? For example, "ሳ" and "ብር" may also be superscripted in presentation of prices but is not common in print (found mostly in store front advertising).*
2. *Always superscript? Language or locale specific? As a setting?*
3. *An argument goes that superscript of "ኛ" should not follow an Ethiopic numerals since the counting system is put into a Ge'ez context and Ge'ez literature does not use a superscript. Yet "ኛ" is appropriate to Amharic language and is not a letter found in Ge'ez. The class of literature may be at issue here, such a religious materials.*

## Ethiopic Comma Usage

In Ethiopic writing practices three encoded symbols will be used in the context of comma, however they are generally not used together. Looked at another way, the Ethiopic comma may appear with three different glyphs. The western comma also has an important role in Ethiopic writing. Usage rules are as follows:

**፣ - ነጠላ ሠረዝ**

ETHIOPIC COMMA (አማ/ ነጠላ ሠረዝ, ትግ/ ንጽል ሰረዝ) is used as a comma in Ethiopian practices and as a semicolon in Eritrea. ፣ will also be used in Ge'ez writing as a semicolon (is this a modern usage or classic?). ETHIOPIC SEMICOLON, ፤, in turn is used as a colon in Eritrean writing (verify).

**፥ - ንዑስ ሠረዝ**

ETHIOPIC COLON in contrast to its Unicode name, is never(?) used as a colon in any known Ethiopic practices. Rather, it is a classical glyph variant of ETHIOPIC COMMA (፣). ፥ is the default comma glyph in Ge'ez writing and is the preferred comma glyph by some modern writers writing in other languages. ፥ is the default comma in most Amharic bibles, likely in keeping with the Ge'ez Bible that the Amharic would have been translated from. Some modern Bible translations into Amharic, as well as other religious literature, will use ፣ as the default comma in prose but maintain ፥ for passage references, e.g. ማቴ4፥23 , ዮሐ13፥16 , etc. Given this common usage in religious materials "ecclesiastical comma" is a practical English term for referring to the ፥ glyph variant. Kidane Wolde Kifle and Desta Tekle Wolde note the equivalence of the two commas.

**፡ - ንዑስ ነጥብ**

ETHIOPIC WORDSPACE (አማ/ ሁለት ነጥብ, ትግ/ ክልተ ነጥቢ) in keeping with Ge'ez heritage, is used as the word separator (i.e. space) in the classical writing practices of both Eritrean and Ethiopia. In modern Eritrean writing practices ETHIOPIC WORDSPACE will be used in the role of a comma (illustrated in Figure [Fig. 31Ecclesiastical emphasis sample from Maṣḥafa Ṣalot Mes Ser'ate Kiddase Betegreññā, Page 32 (Weldemariam, 1995 (1988 EC))](http://w3c.github.io/elreq/#ecclesiastical_emphasis_sample)). When used in modern Ethiopian writing, it retains its classic role as word separator (space). See [6.2.4 Ethiopic Wordspace in the Context of Comma](http://w3c.github.io/elreq/#wordspace_in_comma_context) for elaboration on this topic.

**, - comma**

COMMA is utilized in Eritrean and Ethiopian writing practices in the formatting of western numbers only.

## Parenthetical Expressions

Parenthetical expressions are found regularly in modern Ethiopic writing and will apply any of the enclosing symbol pairs: // , () and [].

Issues:

1. *What are the preferred parenthetical enclosing symbol pairs for modern writing?*
2. *Do special rules apply for inner expressions?*
3. *Do copy editors or publishers set a policy for parenthesis symbols, or leave it up to the author to decide?*
4. *Of interest: when are parenthesis first found in Ethiopic writing and in what form?*

### Parenthetical Expressions with Ethiopic Wordspace

When Ethiopic Wordspace is the space delimiter in a document, an issue for consideration arises with the closing parenthesis. As a general rule, an Ethiopic word will be terminated by either an Ethiopic Wordspace or a punctuation symbol. In practice, however, the wordspace is sometimes found both before and following the closing parenthesis. An example to illustrate the two forms:

1) With inner wordspace:

… ፅፋት፡(ፅፍዓት፡)፡ከመ፡ኪደተ፡ንስር፡(እንደ፡አሞራ፡ፈጣን፡)።

2) Without inner wordspace:

… ፅፋት፡(ፅፍዓት)፡ከመ፡ኪደተ፡ንስር፡(እንደ፡አሞራ፡ፈጣን)።

The identical scenarios are likewise found with Ethiopic wordspace in the presence of a closing quotation mark and must also be addressed.

Issues:

1. *Which of the above two forms is considered most appropriate?*
2. *Do special rules, or exceptions, apply with Ethiopic wordspace around closing symbols?*

## Quotation

Classical Ethiopic literature applying quotation marks will employ double guillemet (« ») in a primary style and single guillemets (‹ ›) in a secondary style. Single guillemets will be used for inner-quotation and single word quotation. Modern Ethiopic writing will additionally utilize Latin quotation marks similarly (“ ” ‘ ’, U+201C, U+201D, U+2018, U+2019). The choice of Latin script quotation may represent either an author preference or a software limitation that made guillemets unavailable or difficult to access.

Issues:

1. *What are the preferred quotation marks for modern Ethiopic literature?*
2. *Is the single quotation mark usage statement correct?*
3. *Do any other rules apply to quotation usage in Ethiopic writing?*
4. *Do copy editors or publishers set a policy for quotation mark choice, or leave it up to the author to decide?*
5. *Do the same rules apply for Ethiopic wordspace surrounding a closing quotation mark as with a closing parenthesis?*
6. *Of interest: When are quotation marks first found in Ethiopic writing and in what form?*

## Ellipsis

Both pre-composed and mid-line ellipsis are found in Ethiopic literature. The presence of one over the other may simply be an artifact of the publishing technology and not necessarily in line with the publisher's preference.

Issues:

1. *Which ellipsis form is the preferred as the default?*
2. *How much spacing between dots is desirable?*
3. *Are square dots preferable over circular?*
4. *What spacing (if any) before and after the dots?*

## Letter Spacing

Any number of kerning pairs and ligatures are possible for Ethiopic typography that would lead to better visual quality of printed litterature. While beyond the present scope of this task force raising the topic with stakeholders to gauge the level of interest would be beneficial for setting the direction of future work. Producing visual samples (common lone words and full paragraphs) would help in the discussion of this topic.

## Bibliographic Citation

A formally recognized standard for bibliographic citation of Ethiopic publications is not found in the Ethiopian publishing community, and bibliographic convention is left to the discretion of individual authors. Establishing a standard is recommended by the present authors and will aid in document consistency and in the machine processing of reference citations. To address a book citation convention, a strong starting point is available from the work of a recognized subject matter expert, [Dereje Gebre](http://www.aau.edu.et/chls/ato-dereje-gebre/) of the AAU Amharic Language Department, and past Vice President of the Ethiopian Writer's Association. Professor Dereje employs the following convention:

     <Citation> ::= <Author Full Name> "፤" <Publication Date> "፤" <i><Terminated Title>|(<Unterminated Title> "።")</i> <City> "፤" <Publisher> "።"

where

     <Title> ::= <Terminated Title> | <Unterminated Title>  
     <Terminated Title> ::= <Title> [፡፤።?]  
     <Unterminated Title> ::= <Text> [:Letter:]

An example:

     ደረጀ ገብሬ፤ ሚያዝያ 1996፤ *ተግባራዊ፡የጽህፈት፡መማሪያ።* ንግድ ማተሚያ ድርጅት፤ አዲስ አበባ።

Issues:

1. *Does the Ethiopian Writer's Association endorse any conventions for media citation?*
2. *What are formats for other media types?*