MICROSOFT AMHARIC LANGUAGE PACK: AN EVALUATION OF WINDOWS INTEGRATED KEYBOARD AND USER EXPERIENCE

Girma A. Demeke

Abstract

This paper focused on Microsoft language interface pack for Amharic. It has examined Windows integrated keyboard and collation for Amharic. It involved a thorough analysis of Amharic characters and punctuation marks, culminating in the suggestion of a standardized solution. The paper has identified various challenges encountered when using the Windows integrated keyboard, particularly with the utilization of repeated and sequence of keys to encode a single character and the unreliability of the current sorting order, which often fails to adhere to its own guidelines. Furthermore, the analysis has delved into shortcomings within the Microsoft-prepared Style Guide for Amharic, including instances of misinformation, typos, and grammatical errors. Based on these findings, the paper has suggested ways to revise the style guide and the keyboard assignment.

Keywords: Microsoft Language Pack, Amharic Characters, Windows 10, Windows 11, Amharic Keyboard, Sorting Order

1. Introduction

Microsoft introduced Amharic language support in its language pack several years ago, allowing Amharic speakers to type in Amharic without

relying on third-party applications.[1] The display language can also be set to Amharic, enhancing the Windows experience for Amharic-speaking users. However, challenges exist in effectively utilizing the integrated Amharic keyboard within the Windows environment. This paper investigates the challenges encountered when using the Microsoft integrated language pack for Amharic input. It specifically examines the typing features/keyboards and collation/sorting order. Furthermore, it analyzes the Microsoft-recommended guideline and proposes potential solutions.

The primary focus of this paper is on the utilization of Microsoft language pack for the Amharic language in the Windows environment. Nevertheless, the recommendations presented may also be applicable or readily extendable to other languages that employ the Ethiopic writing system.

Section 2 introduces Amharic characters and their phonetic/ phonemic values. It also examines the most common symbols and punctuation marks currently in use with some historical background. Section 3 delves into the Unicode encoding of Amharic characters. Section 4 analyzes the implementation of Amharic in Windows integrated Amharic language keyboard and shows the challenges. This section also illustrates the problems with the sorting order and points out areas for potential improvement.

2. Amharic Alphabet and Symbols

Amharic does not make a distinction between uppercase and lowercase letters. Its script evolved from Ge'ez, with additions to accommodate unique Amharic sounds. Initially, Amharic was written alongside Ge'ez, primarily by church scholars familiar with Ge'ez script, who also adopted its punctuation marks. Gradually, Amharic superseded Ge'ez in written form, expanding its usage across diverse populations and knowledge levels. This evolution, combined with technological advancements and the increasing role of Amharic in various fields, led to significant differences between the Amharic and Ge'ez writing systems. Some Ge'ez punctuation marks were falling out of use in Amharic writing system, while others were adopted from English.

---

1   The Amharic keyboard and Nyala, the now-standard publishing font, predate the language interface pack. They were released to the public on January 30, 2007 with Windows Vista (Daniel Yacob, pc).

The Amharic alphabet is detailed in Section 2.1, and punctuation marks are discussed in Section 2.2.

2.1 Amharic Alphabet

The Amharic alphabet referred to ፊደል ገበታ fidäl gäbäta hence forth *Fidel Gebeta* in Amharic often presents the characters into two tables. The first table contains the "basic" characters and the second contains five labiovelars with five forms each. In addition, most of the consonants can be rounded/labialized and exclusively followed by the mid low vowel /a/. Characters representing those complex consonants plus the vowel /a/ are listed often separately below the two charts. This section discusses the characters with their phenetic/phonemic values.

2.1.1 Amharic Basic Characters

The Amharic "basic" alphabet/chart contains 34 characters each having seven forms. The following table contains the characters along with their current phonetic/phonemic values.

Table 1: Basic Amharic characters (as used in MA)

1st order C+ä	2nd order C+u	3rd order C+i	4th order C+a	5th order C+e	6th order C(+ɨ)[2]	7th order C+o	Phonetic/phon emic values of the consonants[3]
ሀ/ha/	ሁ	ሂ	ሃ	ሄ	ህ	ሆ	h
ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ	l
ሐ/ha/	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ	h
መ	ሙ	ሚ	ማ	ሜ	ም	ሞ	m
ሠ	ሡ	ሢ	ሣ	ሤ	ሥ	ሦ	s
ረ	ሩ	ሪ	ራ	ሬ	ር	ሮ	r
ሰ	ሱ	ሲ	ሳ	ሴ	ስ	ሶ	s
ሸ	ሹ	ሺ	ሻ	ሼ	ሽ	ሾ	š

---

2   Optional

3   The phonetic representation here and the transcription of Amharic words in this paper in general is a little bit modified version of IPA in line with the customary representation of Ethiopian languages in linguistic studies.

ቀ	ቁ	ቂ	ቃ	ቄ	ቅ	ቆ	k'
በ	ቡ	ቢ	ባ	ቤ	ብ	ቦ	b
ተ	ቱ	ቲ	ታ	ቴ	ት	ቶ	t
ቸ	ቹ	ቺ	ቻ	ቼ	ች	ቾ	č
ኀ/ha/	ኁ	ኂ	ኃ	ኄ	ኅ	ኆ	h
ነ	ኑ	ኒ	ና	ኔ	ን	ኖ	n
ኘ	ኙ	ኚ	ኛ	ኜ	ኝ	ኞ	ñ
አ/a/	ኡ/u/	ኢ/i/	ኣ/a/	ኤ/e/	እ/ɨ/	ኦ/o/	ʔ or just vowels4
ከ	ኩ	ኪ	ካ	ኬ	ክ	ኮ	k
ኸ	ኹ	ኺ	ኻ	ኼ	ኽ	ኾ	x/h
ወ	ዉ	ዊ	ዋ	ዌ	ው	ዎ	w
ዐ/ʔa/	ዑ	ዒ	ዓ	ዔ	ዕ	ዖ	ʔ5
ዘ	ዙ	ዚ	ዛ	ዜ	ዝ	ዞ	z
ዠ	ዡ	ዢ	ዣ	ዤ	ዥ	ዦ	ž
የ	ዩ	ዪ	ያ	ዬ	ይ	ዮ	y
ደ	ዱ	ዲ	ዳ	ዴ	ድ	ዶ	d
ጀ	ጁ	ጂ	ጃ	ጄ	ጅ	ጆ	ğ
ገ	ጉ	ጊ	ጋ	ጌ	ግ	ጎ	g
ጠ	ጡ	ጢ	ጣ	ጤ	ጥ	ጦ	t'

---

4 In limited contexts, this character is used to represent the glottal stop ʔ as in አእምሮ *ʔaʔmɨro* 'mind'. In word initial position the glottal stop may also optionally appear as in *(ʔ)antä 'you.m.s.',* and *(ʔ)arat* 'four'. In Tigrinya and Ge'ez ኸ represents the glottal stop ʔ and ዐ the pharyngeal voiced fricative ʕ.

5 The distinction between ኸ and ዐ does not exist in Modern Amharic. Both used alternatively as in, for instance ስእል vs. ስዕል both *siʔil* 'picture, painting'. ዐ represents in Old Amharic the pharyngeal voiced fricative ʕ as in Ge'ez, North Argobba, Tigrinya and Tigre.

ጨ	ጩ	ጪ	ጫ	ጬ	ጭ	ጮ	č'
ጰ	ጱ	ጲ	ጳ	ጴ	ጵ	ጶ	p'
ጸ	ጹ	ጺ	ጻ	ጼ	ጽ	ጾ	s'
ፀ	ፁ	ፂ	ፃ	ፄ	ፅ	ፆ	s'
ፈ	ፉ	ፊ	ፋ	ፌ	ፍ	ፎ	f
ፐ	ፑ	ፒ	ፓ	ፔ	ፕ	ፖ	p
ቨ	ቩ	ቪ	ቫ	ቬ	ቭ	ቮ	v

As can be seen in Table 1, in Modern Amharic there are some symbols that share the same sound. ሰ and ሠ stand for the alveolar fricative *s*, ጸ and ፀ for the ejective *s'*, አ and ዐ for the glottal stop *ʔ* or just for the vowels, and ሀ, ሐ, ኀ and ኸ for *h*. In Ge'ez ሰ, ሠ, አ, ዐ, ሀ, ሐ, and ኀ are assumed to have distinct phonetic values. There could also be phonetic distinctions between ጸ and ፀ in this language. Tigrinya and the other modern Ethio-Semitic languages do not make a phonetic distinction between ሰ and ሠ and between ጸ and ፀ. In Tigrinya, ኀ is just an alternative representation of the voiceless pharyngeal fricative ሐ *ḥ*. In the history of Amharic there is no evidence that ሰ and ሠ represent distinct phonemes. Nor does Amharic make phonetic distinction between ጸ and ፀ. In Old Amharic ሀ, ሐ, and ኸ as well as አ and ዐ have their own distinct phonetic values. We do not know, however, if ኀ has a distinct phonetic value at any period of the history of Amharic. The sound/consonant that ኀ represents in Ge'ez—which is a velar fricative *x*—is represented in Amharic by ኸ. See Demeke, (2014/2017) for more discussion on this.

The default vowel associated with the first characters is the mid central vowel. However, the first order of አ, ዐ, ሀ, ሐ, and ኀ and the corresponding fourth order characters have the same phonetic values. That is why in Table 1 above the phonetic/phonemic transcription in the first order of these characters is given. In the case of the characters referred to as Alfa, there is a symbol ኧ representing the mid-central vowel *ä* or the glottal sound followed by the mid-central vowel *ʔä*. This character is rarely used in Amharic as such combination is found as far as we know only in one interjection word ኧረ /ʔärä/ somewhat similar to 'Oh'.

2.1.2 Amharic Characters Representing Labiovelars

As pointed out above, the Amharic *Fidel Gebeta* has also five labiovelars with five orders. These labiovelars are often listed in the Amharic alphabet following the main chart separately.

Table 2: Labiovelars

Cʷ+ä	Cʷ+i	Cʷ+a	Cʷ+e	Cʷ+i	IPA representation of the consonants
ቈ	ቊ	ቋ	ቌ	ቍ	kʼʷ
ኈ	ኊ	ኋ	ኌ	ኍ	xʷ
ኰ	ኲ	ኳ	ኴ	ኵ	kʷ
ዀ	ዂ	ዃ	ዄ	ዅ	xʷ
ጐ	ጒ	ጓ	ጔ	ጕ	gʷ

As can be seen in table 2, ኈ and ዀ do not have distinct phonetic values. Both letters represent the labiovelar fricative x^w. In fact, in today's Amharic except the third order of the first three letters in Table 2, i.e. ቋ, ኋ, and ኳ, and the last, i.e. ጓ, all the others are hardly in use as they do not represent any distinct phonemes. These include the symbol ዀ and all its forms.

In Old Amharic the labiovelars are assumed to have phonemic value where words in Modern Amharic like kʼorrätʼ-ä decide/cut$_{\text{perf}}$-3mss,[6] koläkkol-ä tickle/line up$_{\text{perf}}$-3mss, and gommäd-ä cut$_{\text{perf}}$-3mss were written and pronounced as ቈረጠ kʷʼärrätʼ-ä decide/cut$_{\text{perf}}$-3mss, ኰለኰለ kʷäläkʷkʷäl-ä tickle/line up$_{\text{perf}}$-3mss, and ጐመደ gʷämmäd-ä cut$_{\text{perf}}$-3mss respectively. These labiovelars have contrastive meaning which is distinct from their counterpart velar sounds as can be seen from the following minimal pairs: ቀረጠ kʼärrätʼ-ä tax$_{\text{perf}}$-3mss, ከለከለ käläkkäl-ä forbid$_{\text{perf}}$-3mss, and ገመደ gämmäd-ä twist$_{\text{perf}}$-3mss[7].

In addition to the characters in tables one and two above, the fourth orders of the basic characters, except a few characters, can add the

---

6   Perf stands for the perfective aspect whereas 3mss stands for third person
     masculine singular subject agreement.
7   Twist for a rope.

horizontal line (_) on the foot or on the top to bring a labialized, i.e. rounded, reading as in Cʷ+a, where Cʷ here stands for a rounded consonant. The former applies to most of the characters. We have thus letters such as ሏ *lʷa*, ቧ *bʷa*, ሧ *sʷa*, ኧ *šʷa*, ቷ *tʷa*, etc. The latter type applies for a few cases such as ፘ *rʷa* and ᎀ *mʷa*. Even the latter characters are often replaced by letters having the lower horizontal line (_) in uniformity with the majority as in ሟ mʷa and ፉ fʷa. Currently symbols with top line are often used to represent palatalized sounds in other languages and found in the Unicode chart as ፘ *ṛ̌i* (1358), and ፚ *f̣i* (135A). ᎀ *mʷa* is found in the Unicode chart 1359.

2.2 Amharic Punctuation Marks

As Amharic adopted the characters from Ge'ez it also adopted punctuation marks. The punctuation marks used in Ge'ez listed in the Unicode Standard, Version 16.0 are the following with their description:

Table 3: Ge'ez Punctuation Marks

Symbol	Unicode	Description
፠	U+1360	Ethiopic Section Mark
፡	U+1361	Ethiopic Wordspace
።	U+1362	Ethiopic Full Stop
፣	U+1363	Ethiopic Comma
፤	U+1364	Ethiopic Semicolon
፥	U+1365	Ethiopic Colon
፦	U+1366	Ethiopic Preface Colon
፧	U+1367	Ethiopic Question Mark
፨	U+1368	Ethiopic Paragraph Separator

Amharic has undergone a significant transformation in its written form over the past 150 years. This evolution is closely linked to the introduction of Western-style education and government systems, which increased the demand for Amharic to fulfill a wider range of written communication needs. Traditional punctuation marks, inherited from Ge'ez, proved inadequate for these new demands. The rise of printing and the increased availability of paper further necessitated changes. Some traditional punctuation marks, particularly those designed to save space on expensive parchments—such as paragraph marks, the double-dot

word separator, and section separators—were phased out and replaced with English punctuation marks or simply by using spaces.

Among the first introduced English punctuation marks to Amharic is a question mark. A three dot ellipsis (…) to indicate something is missing or to mean etc., brackets, hyphen, and both dashes, i.e. ndash and mdash, single and double quotation marks, exclamation mark, and slash especially front slash are used in a similar way and function to English. Period (single dot) which is referred in Amharic as ነቁጥ näk'ut' is used with abbreviations and with numbers especially in listing items and of course to separate fraction from a whole number.

Except for አራት ነጥብ arat nät'ib (።) which is equivalent to a full stop/period, ነጠላ ሰረዝ nät'äla säräz (፣) somewhat similar to coma in its usage in a text, and ድርብ ሰረዝ dirrib säräz (፤) somewhat similar to semi colon (which is also often used as colon), the others neither are uniformly nor often used.

Symbols used in mathematics or in general hard science are almost all adopted into Amharic with the usage of Arabic numerals. The traditional Ethiopic numerals are left often for a few roles. The usage of the English colon (:) which is identical in shape with the Amharic ሁለት ነጥብ hulät nät'ib (:) is used not on its traditional role but with the role that it has in English, especially when used with numerals, such as in the case of separating hour from minutes, year and page number in the case of citations within a text as in Girma (2014:45). The percent marker % is also adopted to Amharic and frequently used in texts as well. This is despite the claim made in the Microsoft's Amharic Style Guide that "[p]ercentages in Amharic are expressed in saying (+) 50 በመቶ. The symbol <%> is not used" (p. 17).

A good keyboard and collation for Amharic writing should consider the above facts to simplify ease of typing and sorting. However, as we will see in section 4, neither the guideline nor the Windows integrated Amharic keyboard makes typing and sorting an easy task for Ethiopic based languages.

3. Unicode Assignment of Amharic characters
The following Unicode assignment of Amharic characters is based on Version 16.0.[8] The Ethiopic characters range is 1200–137F. We will list

---

8   Ethiopic. https://www.unicode.org/charts/PDF/U1200.pdf Accessed on 10/20/2024.

here only the Amharic characters and punctuation marks currently in use as this will show the sorting order based on the ሀለሐመ HaLäHaMä/HaLäHaMä Fidel Gebeta.

120	121	122	123	124	126	127	128	129	12A	12B
ሀ 1200	ሐ 1210	ሠ 1220	ሰ 1230	ቀ 1240	በ 1260	ተ 1270	ኀ 1280	ነ 1290	አ 12A0	
ሁ 1201	ሑ 1211	ሡ 1221	ሱ 1231	ቁ 1241	ቡ 1261	ቱ 1271	ኁ 1281	ኑ 1291	ኡ 12A1	
ሂ 1202	ሒ 1212	ሢ 1222	ሲ 1232	ቂ 1242	ቢ 1262	ቲ 1272	ኂ 1282	ኒ 1292	ኢ 12A2	
ሃ 1203	ሓ 1213	ሣ 1223	ሳ 1233	ቃ 1243	ባ 1263	ታ 1273	ኃ 1283	ና 1293	ኣ 12A3	ኳ 12B3
ሄ 1204	ሔ 1214	ሤ 1224	ሴ 1234	ቄ 1244	ቤ 1264	ቴ 1274	ኄ 1284	ኔ 1294	ኤ 12A4	
ህ 1205	ሕ 1215	ሥ 1225	ስ 1235	ቅ 1245	ብ 1265	ት 1275	ኅ 1285	ን 1295	እ 12A5	
ሆ 1206	ሖ 1216	ሦ 1226	ሶ 1236	ቆ 1246	ቦ 1266	ቶ 1276		ኖ 1296	ኦ 12A6	
		ሧ 1227	ሷ 1237		ቧ 1267	ቷ 1277		ኗ 1297	ኧ 12A7	
ለ 1208	መ 1218	ረ 1228	ሸ 1238		ሸ 1268	ቸ 1278		ኘ 1298	ከ 12A8	ኸ 12B8
ሉ 1209	ሙ 1219	ሩ 1229	ሹ 1239		ሹ 1269	ቹ 1279		ኙ 1299	ኩ 12A9	ኹ 12B9
ሊ 120A	ሚ 121A	ሪ 122A	ሺ 123A		ሺ 126A	ቺ 127A		ኚ 129A	ኪ 12AA	ኺ 12BA
ላ 120B	ማ 121B	ራ 122B	ሻ 123B	ቧ 124B	ሻ 126B	ቻ 127B	ኌ 128B	ኛ 129B	ካ 12AB	ኻ 12BB
ሌ 120C	ሜ 121C	ሬ 122C	ሼ 123C		ሼ 126C	ቼ 127C		ኜ 129C	ኬ 12AC	ኼ 12BC
ል	ም	ር	ሽ		ሽ	ች		ኝ	ክ	ኽ

120D	121D	122D	123D		126D	127D		129D	12AD	12BD
120E	121E	122E	123E		126E	127E		129E	12AE	12BE
120F	121F	122F	123F			127F		129F		

12C	12D	12E	12F	130	131	132	133	134	135	136
	12D0	12E0	12F0	1300		1320	1330	1340	1350	
	12D1	12E1	12F1	1301		1321	1331	1341	1351	1361
	12D2	12E2	12F2	1302		1322	1332	1342	1352	1362
	12D3	12E3	12F3	1303	1313	1323	1333	1343	1353	1363
	12D4	12E4	12F4	1304		1324	1334	1344	1354	1364
	12D5	12E5	12F5	1305		1325	1335	1345	1355	1365
	12D6	12E6	12F6	1306		1326	1336	1346	1356	1366
		12E7	12F7	1307		1327	1337		1357	
12C8	12D8	12E8	12F8	1308		1328	1338	1348		
12C9	12D9	12E9	12F9	1309		1329	1339	1349		

ፊ 12CA	ዚ 12DA	ዪ 12EA	ዺ 12FA	ጊ 130A		ፚ 132A	ፚ 133A	ፚ 134A		
ፋ 12CB	ዛ 12DB	ያ 12EB	ዻ 12FB	ጋ 130B		ፚ 132B	ፚ 133B	ፚ 134B		
ፌ 12CC	ዜ 12DC	ዬ 12EC	ዼ 12FC	ጌ 130C		ፚ 132C	ፚ 133C	ፚ 134C		
ው 12CD	ዝ 12DD	ይ 12ED	ዽ 12FD	ግ 130D		ፚ 132D	ፚ 133D	ፚ 134D		
ፎ 12CE	ዞ 12DE	ዮ 12EE	ዾ 12FE	ጎ 130E		ፚ 132E	ፚ 133E	ፚ 134E		
	ዟ 12DF					ፚ 132F	ፚ 133F	ፚ 134F		

In the above table the shaded boxes are from the original document. The empty boxes in the Unicode document contain Ethiopic characters which are neither Amharic symbols nor currently in use by Amharic. As we will see in the following section, the sorting order proposed for Amharic in the Microsoft Style Guide is based on the above Unicode order.

4. Windows Keyboard for Amharic and Collation
4.1 Microsoft Style Guide for Amharic

Microsoft has developed a style guide for Amharic which last revised February 2011.[9] As stated on the introduction of the guideline, its aim is "to provide everybody involved in the localization of Amharic Microsoft products with Microsoft-specific linguistic guidelines and standard conventions that differ from or are more prescriptive than those found in language reference materials" (Amharic Style Guide, 2011: 5). Microsoft claims that it adopted the conventions "after considering

---

9   "This Style Guide went through major revision in February 2011 in order to remove outdated and unnecessary content. It contains information pertaining to all Microsoft products and services" (Amharic Style Guide, 2011: 5).

context based on various needs, but above all, they are easy to follow and applicable for all types of software to be localized" (Amharic Style Guide, 2011: 5).

The style guide indicates sorting order, use of punctuation marks, some basic grammatical usages, and formatting issues among others. In terms of scope it covers almost all basic information needed for the localization standards. The problem is it is poorly researched and full of errors. Some of the problems that we will see in a moment in using the integrated Amharic language pack drives from the problem found on the Style Guide.

Although not directly relevant the Amharic examples presented in the Microsoft Style Guide are full of typos and grammatical errors. To notice this, one does not even go further beyond the sample Amharic text given in the first pages of the Style Guide. In this sample text punctuation marks are random. Grammatical errors and typos are all over. The text is neither coherent. Both content and stylistic wise nothing good can be said about the sample text.

On the grammatical description, not enough information is given in the Style Guide. In some cases, wrong information is provided. For instance, on the section for inflection we find the note "[t]his section does not apply to Amharic." Unless the term inflection means something different for the Style Guide, Amharic is one of the most productive languages in using inflections. Although this is not directly affecting the usage of the language pack, the complete misunderstanding of punctuation marks currently in use in Amharic has major effect as we will see in the following section.

Besides typos, inconsistencies, and wrong information pointed out above, every point raised on the style guide needs careful revision. For instance, the long day format is given as dd MMMM yyyy with the following example: 17 መጋቢት 2011 (p. 10). Unless this is a stipulation for technical reasons, this form is nonexistent in Amharic be it in written or spoken form. On the long day format month name always comes first. The example with the currency is also misleading: 0.50 ሳንቲም / Santim (p. 16). There is no 0.50 ሳንቲም / Santim in the Ethiopian currency. The Ethiopian Birr like the US dollar has 100 cents and the smallest is 1 ሳንቲም 'cent.'

4.2 Windows Keyboard for Amharic

4.2.1 Amharic Characters

We couldn't find instructions or a preview of Amharic keyboard. The On-Screen Keyboard for **Amharic Input Method (version 1.0)** has no preview as can be seen below in the screenshot of on-screen keyboard of Windows 10.

Fig. 1: Amharic virtual keyboard, Windows 10
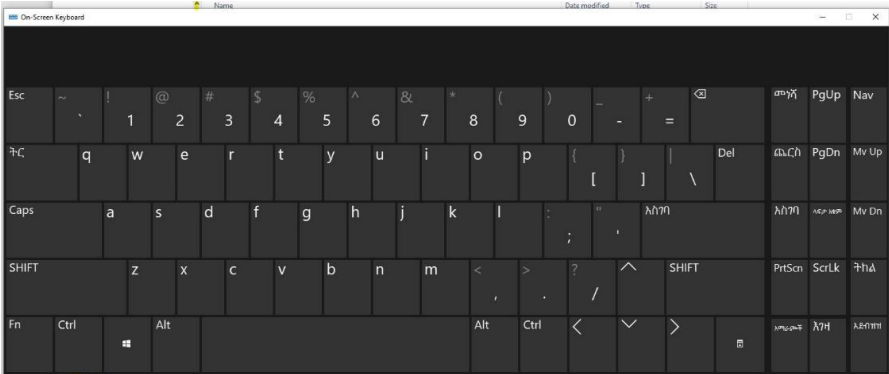
Amharic is written using Latin characters which automatically are converted to Amharic characters. The list below is based on actual typing using Windows 10 and Windows 11 imaged computers. As pointed out above, Amharic does not make a distinction between capital and small letters. Using a shift button may create a different character.

Table 4: Basic Amharic haracters (as used in MA)

1st order C+e	2nd order C+u	3rd order C+i	4th order C+a	5th order C+ie	6th order C	7th order C+o	Rounded C+ua[10]	Windows Keyboard Representation
ሀ	ሁ	ሂ	ሃ	ሄ	ህ	ሆ		h
ለ	ሉ	ሊ	ላ	ሌ	ል	ሎ	ሏ	l
ሐ	ሑ	ሒ	ሓ	ሔ	ሕ	ሖ		hh
መ	ሙ	ሚ	ማ	ሜ	ም	ሞ	ሟ	m

---

10   The letters with the asterisk below are rarely used in Amharic.

ሠ	ሡ	ሢ	ሣ	ሤ	ሥ	ሦ	ሧ	ss
ረ	ሩ	ሪ	ራ	ሬ	ር	ሮ	ሯ	r
ሰ	ሱ	ሲ	ሳ	ሴ	ስ	ሶ	ሷ	s
ሸ	ሹ	ሺ	ሻ	ሼ	ሽ	ሾ	ሿ	x or Shift + s
ቀ	ቁ	ቂ	ቃ	ቄ	ቅ	ቆ	ቋ	q
በ	ቡ	ቢ	ባ	ቤ	ብ	ቦ	ቧ	b
ተ	ቱ	ቲ	ታ	ቴ	ት	ቶ	ቷ	t
ቸ	ቹ	ቺ	ቻ	ቼ	ች	ቾ	ቿ	c
ሀ /ha/	ሁ	ሂ	ሃ	ሄ	ህ	ሆ	ኋ	hhh
ነ	ኑ	ኒ	ና	ኔ	ን	ኖ	ኗ	n
ኘ	ኙ	ኚ	ኛ	ኜ	ኝ	ኞ	ኟ	Shift + n, ny, or gn
አ/a/	ኡ /u/	ኢ /i/	ኣ /aa/	ኤ /ie/	እ/e/	ኦ/o/		
ከ	ኩ	ኪ	ካ	ኬ	ክ	ኮ	ኳ	k
ኸ	ኹ	ኺ	ኻ	ኼ	ኽ	ኾ	ዃ	x/h
ወ	ዉ	ዊ	ዋ	ዌ	ው	ዎ		w
ዐ/ Shift +a/	ዑ / Shift +u/	ዒ / Shift +i/	ዓ / Shift +aa/	ዔ / Shift +ie/	ዕ / Shift +e/	ዖ / Shift +o/		
ዘ	ዙ	ዚ	ዛ	ዜ	ዝ	ዞ	ዟ	z
ዠ	ዡ	ዢ	ዣ	ዤ	ዥ	ዦ	ዧ	Shift +z
የ	ዩ	ዪ	ያ	ዬ	ይ	ዮ		y
ደ	ዱ	ዲ	ዳ	ዴ	ድ	ዶ	ዷ	d
ጀ	ጁ	ጂ	ጃ	ጄ	ጅ	ጆ	ጇ	j
ገ	ጉ	ጊ	ጋ	ጌ	ግ	ጎ	ጓ	g
ጠ	ጡ	ጢ	ጣ	ጤ	ጥ	ጦ	ጧ	Shift + t or t.

ጨ	ጩ	ጪ	ጫ	ጬ	ጭ	ጮ	ጯ	Shift + c or c.
ጰ	ጱ	ጲ	ጳ	ጴ	ጵ	ጶ	ጷ*	Shift + p or p.
ጸ	ጹ	ጺ	ጻ	ጼ	ጽ	ጾ	ጿ	ts
θ	ፁ	ፂ	ፃ	ፄ	ፅ	ፆ		tz
ፈ	ፉ	ፊ	ፋ	ፌ	ፍ	ፎ	ፏ	f
ፐ	ፑ	ፒ	ፓ	ፔ	ፕ	ፖ	ፗ*	p
ቨ	ቩ	ቪ	ቫ	ቬ	ቭ	ቮ	ቯ*	v

As can be seen from the above table the problem is clear even without going into details. Repeated and combined keys to encode a single character pose significant challenges. For example, the palatal nasal has three different representations, making it incredibly difficult to type common words such as ግን 'but' and ምክንያት 'because' in Amharic which have a squence of ny and gn. The Windows keyboard usually turns them into the palatal nasal. The problem is worse with the combination of period with another key to represent ejectives. For instance, when a full stop is added to a sentence final ች, the palatal voiceless affricate represented with the Latin c in the keyboard, will be turned to ጭ, the palatal ejective affricate, encoded in the keyboard as c.[11] This is because the period is also used to encode the Amharic full stop, አራት ነጥብ arat nät'ib (።). We have no information why squence of characters are used at all to begin with. Since Amharic does not have capital letters, shift + C could have been utilized for characters exhaustevely to avoid such problems.

4.2.2 Punctuation Marks

On the section that discusses "Dashes and Hyphens" on the Microsoft Amharic guideline, it states that "[t]his section does not apply to Amharic." However, dashes and hyphens are commonly used in Amharic, similar to English. The key designated for hyphens on English keyboards typically adds two dots above the preceding character when

---

11   Sentence-final ች is frequent in Amharic as the agreement marker on the verb, , for third person feminine subjects is -ăčč. In the default word order in Amharic a verb takes sentence-final position as in ሳባ ዛሬ መጣች saba zare mät' t'ačč 'Saba came today'.

used on an Amharic keyboard. To use hyphens or dashes in Amharic with the current Windows integrated keyboard, users must switch the keyboard input to English or use search and insert from Symbols. This presents a practical challenge for users who need to seamlessly incorporate these punctuation marks within Amharic text.

In the list of Ethiopic Unicode characters, the top two-dots assigned as marking gemination. This is in fact based on the suggestion of the Ethiopic alphabet to encode gemination for a certain language. A top two-dots symbol has not been used to represent consonant gemination in Amharic writing system. However, we find this usage in a few grammatical works such as Cohen (1936) with citation of Amharic examples.[12] Haddis Alemayehu in the 60s also proposes in his work of fiction ፍቅር እስከ *መቃብር* 'Love onto Grave' to implement a similar symbol, a dot, for gemination. The author even did not use it in his work uniformly. Amharic does not use (and need) to indicate gemination in regular writings. It is very strange to include that symbol based on usages/transcription in earlier linguistics work and a suggestion which is mostly on paper.

4.3 **Sorting Order**

The Microsoft style guideline for Amharic states that "[a]lphabetical order is not necessarily indicative of sorting order."[13] The character sorting order from Microsoft guide is similar with the sorting order given in Unicode. However, Microsoft sorting order is a mess. Consider the following randomly selected words sorted with Windows integrated Amharic keyboard and the actual order.

Table 5: Sample Sorted Texts

Windows sorting	Correct sorting/order
ሐሌታ	ህሳብ
ህሳብ	ህገር
ሐብል	ሁለት
ኃብት	ሃይማኖት
ኃይል	ሆድ
ሃይማኖት	ሊባ

ሀገር	ሐሌታ
ሁለት	ሐብል
ሕዝብ	ሕዝብ
ሆድ	መስኮት
ሌባ	ማገር
መስኮት	ሠራ
ማገር	ረዳት
ሠራ	ሰዋ
ሰዋ	ሰፈር
ሰፈር	ሸዋ
ረዳት	ኌብት
ሸዋ	ኍይል
ፀሀይ	ጸለየ
ጸለየ	ጸጋ
ጸጋ	ፀሀይ
ያም	ያም

Although it seems to me that the above sorting order may be triggered by the phonetic transcription of Amharic, I will not speculate on that further. The point is that the sorting order produced by the Microsoft language pack does not even follow its own guideline prepared for the language.

> The sorting order follows the standard Unicode model of Letters-> Punctuations-> Numerals."[14] The guideline also claims that "[t]here is no accepted standard sorting order (Microsoft Style Guide for Amharic, 2011:18).

Amharic and Ge'ez use two types of sorting orders: the አቡጊዳ 'abugida' order and the ሀለሐመ HaLeHaMe order. Although the former order is used in previous time with dictionaries, currently sorting order is usually done with the latter order. The Amharic alphabet in formal schools do not include even the abugida chart as alternative, although it is common in the Ethiopian orthodox church-based schools. Hence, the sorting order if followed as presented in the Unicode chart should have been good.

While the sorting order within the Windows integrated keyboard might present challenges, its search functionality remains effective. Each

---

14 http://abyssiniagateway.net/fidel/unicode/new/recommendation.html

character is distinctly recognized. Although direct web search implementation may require adjustments, the current Amharic encoding within the Windows keyboard proves valuable for indexing, search-and-replace operations, and character/word manipulation within Office 365 applications.

5. Conclusion

Windows OS offers the commendable feature of enabling Amharic writing without requiring third-party software. However, as we have seen in this paper, the current implementation presents significant challenges. This paper has highlighted persistent issues with using the Windows integrated keyboard for Amharic input, issues that have been recognized for some time.

While we acknowledge that Amharic may not be a primary market focus for Microsoft, given their investment in including Amharic language support, minor revisions would be a valuable improvement. A crucial starting point for these revisions lies in the Style Guide.

The current Style Guide relies on poorly selected references, hindering accurate and consistent Amharic input. It omits crucial references for Amharic grammatical studies, notably works by Leslau (1995) and Baye Yimam (1986 EC). While textbooks like Dagnachew Worku (1985) exist, the Style Guide lacks any references specifically addressing Amharic writing style. Furthermore, the sample text should be derived from a published source. Guidance on punctuation usage should be informed by professional resources or, at the very least, an analysis of contemporary publications across different fields.

To fix the current typing problem, combination of keys should be avoided as much as possible. Using Shift should be utilized for any consonant which are not in the English keyboard. Rather than adding uncommon gemination mark, the hyphen key should be left as in English.

Although the examination in this paper focused on Microsoft language pack for Amharic in the Windows environment, the revision may benefit any system that uses Ethiopic writing system.

References

Baye Yimam. 2000 EC. *የአማርኛ ሰዋስው* *'Amharic Grammar'*, 2nd Edition. Addis Ababa: Birhanina Selam Printing Press.

Cohen, Marcel. 1936. *Traité de Langue Amharique.* Paris: Institute D'Ethnologie.

Dagnachew Worku. 1985. *የጽሑፍ ጥበብ መመሪያ* '*A Guide to Writing Skills'.* Addis Ababa: Kuraz Publishing Agency.

Girma A. Demeke. 2014/2017. *Grammatical Changes in Semitic: The Diachronic Grammar of Amharic.* Trenton: The Red Sea Press.

Haddis Alemayehu. 1968/1996. *ፍቅር እስከ መቃብር* '*Love onto Grave*,' 9th Printing, Addis Ababa: Mega Publishing.

Leslau, Wolf. 1995. *Reference Grammar of Amharic.* Wiesbaden: Harrassowitz Verlag.

Microsoft. Amharic Style Guide. February 2011. https://download. microsoft.com/download/6/2/a/62a993f1-d3b2-4f28-b3e7-22965eb91426/amh-eth-styleguide.pdf

Acknowledgment

Girma A. Demeke,
Lawrenceville, New Jersey