Proposal for Ethiopic Script Root Zone LGR

LGR Version 2

Date: 2017-05-17

Document version: 5.2

Authors: Ethiopic Script Generation Panel

Contents

1 General Information/ Overview/ Abstract

This document describes the proposed LGR for Ethiopic script. It describes the salient features of the script and the communities or ethnic groups at home and abroad using the script in the writing systems. It also gives an overview of the various decisions taken by the Ethiopic script generation panel to come up with the LGR proposal. It also describes the methodology, encompassing the procedure as well as the criteria the panel used to select the languages for considerations in the LGR proposal.

The formal specification of the LGR in XML format is available in the file:

- Proposed-LGR-EthiopicScript-20170517.xml

After generating the necessary tests cases, the panel has also carried out label generation tests on the eight languages targeted by the LGR proposal. Sample labels are given in the test file:

- Labels-EthiopicScript-20170517.txt

2 Script for which the LGR is proposed

Ethiopic script has the following specification:

ISO 15924 Code: Ethi

ISO 15924 N°: 430

English Name: Ethiopic (Geʻ ez)

The complete set of characters in the Ethiopic script fall in the following Unicode ranges:

Ethiopic (Geʻ ez): U+1200 – U137F

Ethiopic Supplement: U+1380 – U+139F

Ethiopic Extended: U+2D80 – U+2DDF

Ethiopic Extended-A: U+AB00 – U+AB2F

For the current version of the LGR only code points shortlisted in the Maximal Starting Repertoire (MSR) version 2 [MSR-2] are considered.

The Unicode range 1380-139F is one of the whole block exclusions in the MSR-2 document, hence it is preemptively out of consideration in the LGR formulation. The entire Unicode range U+1200 – U+137F along with some parts of the Unicode points range in the Ethiopic Supplement block are the focal point of considerations in the LGR formulation. The GP has considered the following Unicode points range as its starting repertoire:

- U+1200 – U+137F,

- U+2D80 – U+2D96,
- U+AB00 – U+AB0F.

3   Background on Script and Principal Languages Using It

Ethiopia has more than 80 languages with up to 200 different dialects spoken. The largest ethnic and linguistic groups are the Oromos, Amharas and Tigrayans. As Oromos right now use Latin script or alphabet called Qubee in their writing system they are not relevant to the Ethiopic Script LGR proposal which the Ethiopic Generation Panel has developed. In general, currently, the languages that are spoken in the country can be grouped into four major groups:

- **Semitic**: The Semitic languages are spoken in northern, central and eastern Ethiopia (mainly in Amhara, Tigray, Gurage Zone, Harar and Argobba). These languages use the Ethiopic (Ge'ez) script in their writing systems.
- **Cushitic**: The Cushitic languages are mostly spoken in central, southern and eastern Ethiopia (mainly in the Afar, Oromia and Somali regions). Most of the Cushitic languages use the Latin alphabet in their writing systems. There are languages that use the Ethiopic script, which fall under this category; they are Saho, Awingi, Hamt'agna and Kebena
- **Omotic**: The Omotic language is predominantly spoken between the lakes of the southern Rift Valley and the Omo River. Basketo is one of the Omotic languages which use the Ethiopic script. It is spoken in a special woreda[1] of the Southern Nations, Nationalities Peoples' Region (SNNPR) of Ethiopia named after its inhabitants (the Basketo people). It is known to be the homeland of Omotic-speaking ethnic group.
- **Nilo-Saharan**: The Nilo-Saharan languages are largely spoken in the western part of the country along the border with Sudan (mainly in Gambella and Benshangul regions).

Ethiopia is the homeland of a remarkable variety of communities. Their cultural and linguistic diversity is believed to be a result of complex historical background, as well as geographical and social differences. In Ethiopia, a script of considerable age known as Ethiopic (Ge'ez) and several writing systems based on it are used. The Ethiopic script is an abugida (alpha-syllabary). It is used in the writing systems for several languages spoken in Ethiopia and the neighboring country, Eritrea. It is the most commonly used script in the writing systems of the majority of languages spoken in the country. It originated as an abjad (consonant-only alphabet) and was first used to write Ge'ez, now the liturgical language of the Ethiopian and the Eritrean Orthodox Tewahedo Churches. The earliest known inscriptions in the Ethiopic script date back to the 8th century BC.

3.1   Local Languages Using the Script

There are several languages, which use or are in the process of adopting the Ethiopic script in their writing systems. However, with a repertoire constrained by Unicode version 6.3 [Unicode 6.3] as well as [MSR-2], only eight of them qualify for considerations in the LGR proposal at the time of writing. Thus, the GP confirms that the cut-offs by MSR-2 and Unicode 6.3, and the

---

[1] A *woreda* is the lowest administrative unit in the Ethiopian government's administrative power echelon

Unicode points range which the GP used as a starting point repertoire in general are accommodative of the code points available in the various languages which use a version of the Ethiopic script. The table below, table 3.1 lists the languages which use Ethiopic script in their writing systems along with their EGIDS score obtained from Ethnologue [EGIDS].

Table 3.1 Writing Systems which use Ethiopic Script and along with their EGIDS score

No.	Language	EGIDS
1	Amharic	1
2	Argobba	4
3	Awingi	4
4	Harari	4
5	Hamt'agna(Xamtanga)	4
6	Silt'e	4
7	Tigrè(in use in Eritrea)	4
8	Tigrigna	2

Some code points exhibit special properties as they are used in the writing system of the Amharic language, and the GP has considered their effects critically in building the code point repertoire for the script. The following are pertinent code point properties that the GP has investigated and addressed during analysis of the code points and creation of the code point repertoire:

- **Gemination**: Some Amharic words convey different meanings or exhibit twinning because of gemination (a spoken consonant is pronounced for an audibly longer period of time). However, gemination is generally not represented in the orthography (not relevant in labels), despite being distinctive in speech. As a result gemination in the Amharic language does not require any special treatment in the LGR proposal.

- **Code Point Redundancy or Phonemic Decay:** While being orthographically different, some of the code points in Ethiopic script have same sound and meaning for writing in Amharic language, and as a result they can be interchangeably used. Much of Amharic's vocabulary has its origins in the Ge'ez language. Amharic maintains the entirety of Ge'ez alphabets or code points while preserving most, but not all of, Ge'ez's phonology (the way sounds are organized and used). As a consequence, the phonemes of a number of symbols in the orthography have decayed into the nearest Amharic equivalent (alternatively used code points or alphabets). The resulting redundancy in symbols in the Amharic language creates observable influence that makes the DNS system open to

spoofing or other related security risks [DY]. To circumvent the security risks, the GP has treated the redundant code points as variants of each other.

Below, table 3.2, lists out the redundant code points in the writing system of the Amharic language.

Table 3.2 Phonemic Decay or Redundant Code Points in the Amharic Writing System

Code Point	Code Point Name	Code Point	Code Point Name	Code Point	Code Point Name
(1200) ሀ	HA	(1210) ሐ	HHA	(1280) ኀ	XA
(1201) ሁ	HU	(1211) ሑ	HHU	(1281) ኁ	XU
(1202) ሂ	HI	(1212) ሒ	HHI	(1282) ኂ	XI
(1203) ሃ	HAA	(1213) ሓ	HHAA	(1283) ኃ	XAA
(1204) ሄ	HEE	(1214) ሔ	HHEE	(1284) ኄ	XEE
(1205) ህ	HE	(1215) ሕ	HHE	(1285) ኅ	XE
(1206) ሆ	HO	(1216) ሖ	HHO	(1286) ኆ	XO
-	-	(1217) ሗ	HWA	(1288) ኈ	XWA
(1338) ጸ	TSA	(1340) ፀ	TZA	-	-
(1339) ጹ	TSU	(1341) ፁ	TZU	-	-
(133A) ጺ	TSI	(1342) ፂ	TZI	-	-
(133B) ጻ	TSAA	(1343) ፃ	TZAA	-	-
(133C) ጼ	TSEE	(1344) ፄ	TZEE	-	-
(133D) ጽ	TSE	(1345) ፅ	TZE	-	-
(133E) ጾ	TSO	(1346) ፆ	TZO	-	-
(1230) ሰ	SA	(1220) ሠ	SZA	-	-
(1231) ሱ	SU	(1221) ሡ	SZU	-	-
(1232) ሲ	SI	(1222) ሢ	SZI	-	-
(1233) ሳ	SAA	(1223) ሣ	SZAA	-	-
(1234) ሴ	SEE	(1224) ሤ	SZEE	-	-
(1235) ስ	SE	(1225) ሥ	SZE	-	-
(1236) ሶ	SO	(1226) ሦ	SZO	-	-
(1237) ሷ	SWA	(1227) ሧ	SZWA	-	-
(12A0) አ	GLOTTAL A	(12D0) ዐ	PHARYNGEAL A	-	-
(12A1) ኡ	GLOTTAL U	(12D1) ዑ	PHARYNGEAL U	-	-
(12A2) ኢ	GLOTTAL I	(12D2) ዒ	PHARYNGEAL I	-	-
(12A3) ኣ	GLOTTAL AA	(12D3) ዓ	PHARYNGEAL AA	-	-
(12A4) ኤ	GLOTTAL EE	(12D4) ዔ	PHARYNGEAL EE	-	-
(12A5) እ	GLOTTAL E	(12D5) ዕ	PHARYNGEAL E	-	-
(12A6) ኦ	GLOTTAL O	(12D6) ዖ	PHARYNGEAL O	-	-

(12AE)ኮ	KO	(12B0)ኰ	KWA	-	-
(12A0)አ	GLOTTAL A	(12D3) ዓ	GLOTTAL AA		
(12D0)ዐ	PHARYNGEAL A	(12D3)ዓ	PHARYNGEAL AA		

The code points listed in each row in the above table can be interchangeably used in the writing system of the Amharic language- the historical background about how some of the code points are redundantly or interchangeably used, along with a broader list of Amharic words that are orthographically different but mean and sound alike, is available in [HomonymList, HomonymsPart1, HomonymsPart2]. Amharic is the second-most spoken Semitic language in the world, after Arabic, and the official working language of the Federal Democratic Republic of Ethiopia. It is also the official or working language of several of the states within the federal system [Wikipedia]. The GP has considered security related issues posed by the presence of redundant code points in the LGR proposal at large, and hence the code points described in table 3.2 are treated as script wide variants (see section 4.3- analysis of code point variants).

3.2   Geographic Territories of the Language or the Language Map of Ethiopia

A significant number of Diaspora who use Ethiopic script are living in United States and some parts of Europe. Owing to historical ties, the Ethiopic script is also used in neighboring Eritrea (at large) and in Djibouti (to a certain degree). The following are list of countries outside of Ethiopia where significant user communities are believed to exist who are capable of using the Ethiopic script:

- Eritrea
- Djibouti
- Egypt
- United States
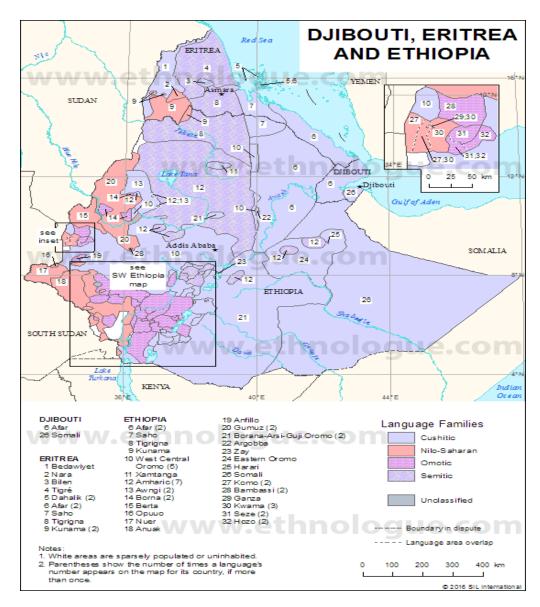- Canada
- Israel
- Europe
- Australia

Figure 1- The Language Map of Djibouti, Eritrea and Ethiopia [LanguageMap]

4   Overall Development Process and Methodology

The GP started its LGR development work by exploring the "Guidelines for Developing Script-Specific Label Generation Rules for Integration into the Root Zone LGR" [Guidelines] document that explains the various pertinent issues with regards to IDN and LGR proposal development procedures and processes with the aim to gain a better and unanimous clarity on what the panel need to do to produce the proposal for the Ethiopic Script LGR for the Root Zone.

4.1   Sources Consulted to Determine the Repertoire

In order to determine the code point included as well as excluded from the repertoire, in addition to the panel members knowledge and expertise, the GP consulted several available sources

including completed and ongoing research works of scholars working on the digital usability of the languages considered in the LGR proposal.

Moreover, the GP used corpus analysis for the two relatively well-resourced languages, Amharic and Tigrigna, as auxiliary evidences for basing its decision on including as well as excluding a code point from the repertoire. To this effect, a total of 14,850 html files with Amharic contents, and 598 html and 40 PDF files with Tigrigna contents were crawled.

Cognizant of the limitations of using corpus particularly for those code points being absent from the corpora used, the GP to base its code point exclusion decision used additional sources such as public comments and the online encyclopedia of Writing Systems and Languages [Omniglot], which published Ethiopic script versions for the eight languages considered in the LGR proposal.

4.2 Team Composition and Diversity

The Generation Panel is composed of members from different disciplines with complementary expertise. Members' professional backgrounds include: technical community with thorough knowledge and skills in DNS, its internal makings and operations, linguists, academics, Information Technology and Computer Science professionals working in the industry, Legal researchers, Cyber-law experts, and researchers on natural language processing.

Even where the panel was not able to include additional members for the eight languages considered for the LGR formulation, the linguists on the panel include members who have also carried out extensive studies in these languages.

The panel started the LGR development from its establishment in January 2016. Following its establishment, the panel has had three face-to-face meetings in January 2016, March 2016 and May 2016. To gain a better understanding of pertinent issues on the ICANN policy, IDN and the notion of LGR for the Root Zone, the panel in collaboration with the Department of Computer Science of the Addis Ababa University and ICANN organized a two-day training workshop on the 22nd and 23rd of July 2016. In the training workshop several invited participants drawn from academia, ISP and the media have taken part, in addition to panel members.

The activities the panel has carried out are organized into six major steps or phases:

- Defining the code point principles

- Conducting corpus analyses for Amharic and Tigrigna languages using contents published online

- Defining the frequency cut of a character or code point from the corpus data

- Building the code point repertoire

- Defining the code point variant principles

- Establishing the code point variant sets

- Analyses of allocatable vs. blocked variants

- Producing a documentation for the process

- Creating an XML for Ethiopic Script LGR proposal

- Submitting the Final LGR proposal for Ethiopic Script along with its XML specification

4.3   Analysis of Code Point Repertoire

The panel held a lengthy discussion on setting criteria or principles to be used to assess the inclusion, exclusion or deferral of code points in the LGR. The panel started the analysis of code point repertoire based on the requirements set forth in the [Guidelines] for the eight languages selected, which comply with the minimum status level (4) on the Expanded Graded Intergenerational Disruption Scale (EGIDS) [EGIDS]. The discussion resulted in the formulation and consideration of the criteria or principles listed below which the panel has used to build the Ethiopic script code point repertoire.

Inclusion criteria or principles:

- A code point representing a letter in established contemporary use in the writing systems of any of the eight languages targeted by the LGR proposal.

- Any code point with a single occurrence in the corpus data analysed is included: no matter what the frequency value is, a single occurrence is considered as a testament of the character or code point's contemporary use.

Exclusion criteria or principles:

- A code point not supported by the IDNA 2008 protocol

- Any code point representing a security or stability issue, which cannot be addressed at any of the stages the analysis process (i.e., stage of determining code points, variants or whole label rules).

- Any code point either deprecated or not recommended for use in the Unicode Standard.

- Any code point with historic significance and with little or no use in the contemporary writing system of the languages selected.

- Any code point with limited or declining use.

The following are additional reasons taken into account for the exclusion of code points:

- Lack of sufficient information or the availability of limited information about a code point use in the contemporary writing system of the eight languages selected.

- Languages with a numeric status of 4 or less are selected based on the EGIDS scale; a particular code point is associated only with a language whose status level is 5 or more is excluded.

- Code points which are in contemporary use but not part of the MSR-2.

In the event that the GP produced evidence demonstrating the use of code-points in contemporary writing systems of the eight languages but could not include them because of insufficient clarity of the evidence with respect to the rules and scope defined by the [Guidelines], the GP opted to defer the code-points (and they are therefore excluded from the current version of the LGR proposal).

4.4   Analysis of Code Point Variants

There are significant numbers of code points that can be interchangeably used in the writing system of the Amharic language. These code points pose serious security problems and the GP has identified the code point pairs or triples with this property as variants. Because of the presence of code point redundancy in the Amharic writing system, two orthographically different lexical words or non-canonical spellings (labels) can be alike both in sound and meaning.   Due to the security problems these variants pose and the fact that the Amharic language is the second-most spoken Semitic language in the world, after Arabic, and the official working language of the Federal Democratic Republic of Ethiopia and many of the states within the federal system [Wikipedia], the GP treat the code point variants in the most restrictive approach in that they have been considered as script wide variants beyond the Amharic language. Thus, in the LGR for Ethiopic script for the root zone, lexical terms or non-lexical labels that involve the code points listed in Table 1.1 have been treated as code point variants across the entire Ethiopic script. To this effect, formulating the following code point variant rules has become pertinent:

1. The labels that can be generated from the code point repertoire can be of either of the two forms:

    a. Canonical Label: A label which has a dictionary meaning.

    b. Non- Canonical Label: A label which has no dictionary meaning.

2. Two code points are variants if they can be interchangeably used within the writing systems of the eight languages considered in the LGR proposal and such use:

    a. either yields canonical labels that are based on equivalent(alternatively spelt) lexical terms

    b. or yields non-canonical labels that can be spelt alternatively.

3. Variant sets are symmetric and transitive

    a. If A is a variant of B then the vice versa also holds true

b. If A is a variant of B and B is a variant of C then A is also a variant of C.

The following are additional principles the GP has used to treat code point variants:

4. By default, all code point variant pairs have a blocked type.

5. Any pair of code points, which are interchangeably used within the writing system of the eight languages have been treated as blocked variants of each other.

5 Repertoire

In this section, the code points included and excluded on the basis of the principles outlined above are presented.

5.1 Summary of Code Points Included and Excluded

The Ethiopic Script Repertoire is presented in tables that show the code points included in the repertoire using selected background colors. The general presentation is similar to that used for character code charts in the Unicode Standard. The background color indicates the status of the code point:

- Green: code points that are part of the LGR.
- White: code points that are not PVALID in IDNA 2008 [RFC5892][IDNAREG].
- Pink: code points that are excluded from the Root Zone in a generic fashion (digits, hyphen), or by being excluded from the Maximal Starting Repertoire [MSR-2].
- Lavender: code points not included in the LGR as result of decisions by the Ethiopic Generation Panel.

	120	121	122	123	124	125	126	127	128	129	12A	12B
0	ሀ 1200	ሐ 1210	ሠ 1220	ሰ 1230	ቀ 1240	ቐ 1250	በ 1260	ተ 1270	ኀ 1280	ነ 1290	አ 12A0	ኰ 12B0
1	ሁ 1201	ሑ 1211	ሡ 1221	ሱ 1231	ቁ 1241	ቑ 1251	ቡ 1261	ቱ 1271	ኁ 1281	ኑ 1291	ኡ 12A1	▨
2	ሂ 1202	ሒ 1212	ሢ 1222	ሲ 1232	ቂ 1242	ቒ 1252	ቢ 1262	ቲ 1272	ኂ 1282	ኒ 1292	ኢ 12A2	ኲ 12B2
3	ሃ 1203	ሓ 1213	ሣ 1223	ሳ 1233	ቃ 1243	ቓ 1253	ባ 1263	ታ 1273	ኃ 1283	ና 1293	ኣ 12A3	ኳ 12B3
4	ሄ 1204	ሔ 1214	ሤ 1224	ሴ 1234	ቄ 1244	ቔ 1254	ቤ 1264	ቴ 1274	ኄ 1284	ኔ 1294	ኤ 12A4	ኴ 12B4
5	ህ 1205	ሕ 1215	ሥ 1225	ስ 1235	ቅ 1245	ቕ 1255	ብ 1265	ት 1275	ኅ 1285	ን 1295	እ 12A5	ኵ 12B5
6	ሆ 1206	ሖ 1216	ሦ 1226	ሶ 1236	ቆ 1246	ቖ 1256	ቦ 1266	ቶ 1276	ኆ 1286	ኖ 1296	ኦ 12A6	▨
7	ሇ 1207	ሗ 1217	ሧ 1227	ሷ 1237	ቇ 1247	▨	ቧ 1267	ቷ 1277	ኇ 1287	ኗ 1297	ኧ 12A7	▨
8	ለ 1208	መ 1218	ረ 1228	ሸ 1238	ቈ 1248	ቘ 1258	ቨ 1268	ቸ 1278	ኈ 1288	ኘ 1298	ከ 12A8	ኸ 12B8
9	ሉ 1209	ሙ 1219	ሩ 1229	ሹ 1239	▨	▨	ቩ 1269	ቹ 1279	▨	ኙ 1299	ኩ 12A9	ኹ 12B9
A	ሊ 120A	ሚ 121A	ሪ 122A	ሺ 123A	ቊ 124A	ቚ 125A	ቪ 126A	ቺ 127A	ኊ 128A	ኚ 129A	ኪ 12AA	ኺ 12BA
B	ላ 120B	ማ 121B	ራ 122B	ሻ 123B	ቋ 124B	ቛ 125B	ቫ 126B	ቻ 127B	ኋ 128B	ኛ 129B	ካ 12AB	ኻ 12BB
C	ሌ 120C	ሜ 121C	ሬ 122C	ሼ 123C	ቌ 124C	ቜ 125C	ቬ 126C	ቼ 127C	ኌ 128C	ኜ 129C	ኬ 12AC	ኼ 12BC
D	ል 120D	ም 121D	ር 122D	ሽ 123D	ቍ 124D	ቝ 125D	ቭ 126D	ች 127D	ኍ 128D	ኝ 129D	ክ 12AD	ኽ 12BD
E	ሎ 120E	ሞ 121E	ሮ 122E	ሾ 123E	▨	▨	ቮ 126E	ቾ 127E	▨	ኞ 129E	ኮ 12AE	ኾ 12BE
F	ሏ 120F	ሟ 121F	ሯ 122F	ሿ 123F	▨	▨	ቯ 126F	ቿ 127F	▨	ኟ 129F	ኯ 12AF	▨

	12C	12D	12E	12F	130	131	132	133	134	135	136	137
0	12C0	12D0	12E0	12F0	1300	1310	1320	1330	1340	1350	1360	1370
1		12D1	12E1	12F1	1301		1321	1331	1341	1351	1361	1371
2	12C2	12D2	12E2	12F2	1302	1312	1322	1332	1342	1352	1362	1372
3	12C3	12D3	12E3	12F3	1303	1313	1323	1333	1343	1353	1363	1373
4	12C4	12D4	12E4	12F4	1304	1314	1324	1334	1344	1354	1364	1374
5	12C5	12D5	12E5	12F5	1305	1315	1325	1335	1345	1355	1365	1375
6		12D6	12E6	12F6	1306		1326	1336	1346	1356	1366	1376
7			12E7	12F7	1307		1327	1337	1347	1357	1367	1377
8	12C8	12D8	12E8	12F8	1308	1318	1328	1338	1348	1358	1368	1378
9	12C9	12D9	12E9	12F9	1309	1319	1329	1339	1349	1359	1369	1379
A	12CA	12DA	12EA	12FA	130A	131A	132A	133A	134A	135A	136A	137A
B	12CB	12DB	12EB	12FB	130B	131B	132B	133B	134B		136B	137B
C	12CC	12DC	12EC	12FC	130C	131C	132C	133C	134C		136C	137C
D	12CD	12DD	12ED	12FD	130D	131D	132D	133D	134D	135D	136D	
E	12CE	12DE	12EE	12FE	130E	131E	132E	133E	134E	135E	136E	
F	12CF	12DF	12EF	12FF	130F	131F	132F	133F	134F	135F	136F	

	138	139
0	1380	1390
1	1381	1391
2	1382	1392
3	1383	1393
4	1384	1394
5	1385	1395
6	1386	1396
7	1387	1397
8	1388	1398
9	1389	1399
A	138A	
B	138B	
C	138C	
D	138D	
E	138E	
F	138F	

	2D8	2D9	2DA	2DB	2DC	2DD
0	2D80	2D90	2DA0	2DB0	2DC0	2DD0
1	2D81	2D91	2DA1	2DB1	2DC1	2DD1
2	2D82	2D92	2DA2	2DB2	2DC2	2DD2
3	2D83	2D93	2DA3	2DB3	2DC3	2DD3
4	2D84	2D94	2DA4	2DB4	2DC4	2DD4
5	2D85	2D95	2DA5	2DB5	2DC5	2DD5
6	2D86	2D96	2DA6	2DB6	2DC6	2DD6
7	2D87					
8	2D88		2DA8	2DB8	2DC8	2DD8
9	2D89		2DA9	2DB9	2DC9	2DD9
A	2D8A		2DAA	2DBA	2DCA	2DDA
B	2D8B		2DAB	2DBB	2DCB	2DDB
C	2D8C		2DAC	2DBC	2DCC	2DDC
D	2D8D		2DAD	2DBD	2DCD	2DDD
E	2D8E		2DAE	2DBE	2DCE	2DDE
F	2D8F					

15

5.2 Code Point Repertoire Included

The GP has carried out corpus analysis for the two relatively well-resourced languages, Amharic and Tigrigna. The remaining six languages as their EGIDS level suggest (all are EGIDS level of 4) are under-resourced (their use is restricted to educational for the local community and primarily contents are available in paper forms and contents of digital nature are non-existent) and as a result the GP couldn't carry out similar analysis. The corpus analysis result has been used as auxiliary evidence of some of the code points use in the modern writing systems of the two languages.

The GP decided to include a code point with a single occurrence or a cut-off value of one based on the idea that any code point warrants inclusion if it has been used in the modern-day writing system (i.e. it appears in any of the texts published by the sources used for the corpus analysis). The GP's firmly believes that a single appearance is enough to warrant a code point inclusion- it is impossible to set a cut-off point on the frequency of use of a code point for inclusion: some code points by their nature are very infrequently used while some are heavily frequently used. However, in doing so the GP has exercised maximum caution to make sure that the actual incidence of low-frequency instances in the corpus are non-typographic errors.

The table below, table 5.1, lists the code points included for the proposed LGR. Relevant references have been cited as testament to the code point use in contemporary writing systems of the eight languages. The EGIDS level and additional documentation have also been used as supporting evidence for active use of code points in the writing systems. The list of languages given is representative but not exhaustive, as an exhaustive list would require significantly more analysis, which is not needed because using a single example is a sufficient criterion for inclusion. In the table below the column named "code point name" is extracted from the file named "UnicodeData.txt" [2]

In order to reduce clutter as well as making the LGR proposal document more readable the following conventions have been formulated for the "Languages Using the Code Point" column of the code points repertoire table, Table 5.1, presented below.

Code Point Convention:
1- **Common- to-all:** Code points common to all the eight languages or writing systems considered by the LGR proposal are labeled as common-to-all. In the proposed LGR for Ethiopic script the code point repertoire has been built from the following eight languages or writing systems:
   - Amharic, Argobba, Awngi, Hamt'agna, Harari, Silt'e, Tigrè and Tigrigna
2- **Common-to-all-except:** Code points common to any of the 6 or 7 languages or writing systems considered by the LGR proposal are labeled as common-to-all-except followed by the name of the language where the code point is not being used.
3- When four or fewer of the languages use a code point in their writing system the languages are listed individually.

---

[2] The Unicode Consortium, Unicode Character Database, available online at
http://www.unicode.org/Public/UCD/latest/

Table 5.1 Code Point Repertoire

No.	Code Point	Glyph	Code Point Name	Languages Using the Code Point	EGIDS Level	Reference/ Comments
1	1200	ሀ	ETHIOPIC SYLLABLE HA	Common-to-all	1	[CA, CT, Omniglot]
2	1201	ሁ	ETHIOPIC SYLLABLE HU	Common-to-all	1	[CA,CT, Omniglot]
3	1202	ሂ	ETHIOPIC SYLLABLE HI	Common-to-all	1	[CA, CT, Omniglot]
4	1203	ሃ	ETHIOPIC SYLLABLE HAA	Common-to-all	1	[CA, CT, Omniglot]
5	1204	ሄ	ETHIOPIC SYLLABLE HEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
6	1205	ህ	ETHIOPIC SYLLABLE HE	Common-to-all	1	[CA, CT, Omniglot]
7	1206	ሆ	ETHIOPIC SYLLABLE HO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
8	1208	ለ	ETHIOPIC SYLLABLE LA	Common-to-all	1	[CA, CT, Omniglot]
9	1209	ሉ	ETHIOPIC SYLLABLE LU	Common-to-all	1	[CA, CT, Omniglot]
10	120A	ሊ	ETHIOPIC SYLLABLE LI	Common-to-all	1	[CA, CT, Omniglot]
11	120B	ላ	ETHIOPIC SYLLABLE LAA	Common-to-all	1	[CA, CT, Omniglot]
12	120C	ሌ	ETHIOPIC SYLLABLE LEE	Common-to-all	1	[CA, CT, Omniglot]
13	120D	ል	ETHIOPIC SYLLABLE LE	Common-to-all-except: Harari	1	[CA, CT, Omniglot]
14	120E	ሎ	ETHIOPIC SYLLABLE LO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
15	120F	ሏ	ETHIOPIC SYLLABLE LWA	Amharic	1	[CA, CT, Omniglot]
16	1210	ሐ	ETHIOPIC SYLLABLE HHA	Amharic, Argobba, Tigrigna	1	[CA, CT, Omniglot]
17	1211	ሑ	ETHIOPIC SYLLABLE HHU	Amharic, Argobba, Tigrigna	1	[CA, CT, Omniglot]
18	1212	ሒ	ETHIOPIC SYLLABLE HHI	Amharic, Argobba, Tigrigna	1	[CA, CT, Omniglot]
19	1213	ሓ	ETHIOPIC SYLLABLE HHAA	Amharic, Argobba, Tigrigna	1	[CA, CT, Omniglot]
20	1214	ሔ	ETHIOPIC SYLLABLE HHEE	Amharic, Argobba, Tigrigna	1	[CA, CT, Omniglot]
21	1215	ሕ	ETHIOPIC SYLLABLE HHE	Amharic, Argobba, Tigrigna	1	[CA, CT, Omniglot]
22	1216	ሖ	ETHIOPIC SYLLABLE HHO	Amharic, Argobba, Tigrigna	1	[CA, CT, Omniglot]
23	1217	ሗ	ETHIOPIC SYLLABLE HHWA	Amharic	1	[CA]
24	1218	መ	ETHIOPIC SYLLABLE MA	Common-to-all	1	[CA, CT, Omniglot]

25	1219	ሙ	ETHIOPIC SYLLABLE MU	Common-to-all	1	[CA, CT, Omniglot]
26	121A	ሚ	ETHIOPIC SYLLABLE MI	Common-to-all	1	[CA, CT, Omniglot]
27	121B	ማ	ETHIOPIC SYLLABLE MAA	Common-to-all	1	[CA, CT, Omniglot]
28	121C	ሜ	ETHIOPIC SYLLABLE MEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
29	121D	ም	ETHIOPIC SYLLABLE ME	Common-to-all	1	[CA, CT, Omniglot]
30	121E	ሞ	ETHIOPIC SYLLABLE MO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
31	121F	ሟ	ETHIOPIC SYLLABLE MWA	Amharic, Tigrigna	1	[CA, CT, Omniglot]
32	1220	ሠ	ETHIOPIC SYLLABLE SZA	Amharic	1	[CA, Omniglot]
33	1221	ሡ	ETHIOPIC SYLLABLE SZU	Amharic	1	[CA, Omniglot]
34	1222	ሢ	ETHIOPIC SYLLABLE SZI	Amharic	1	[CA, Omniglot]
35	1223	ሣ	ETHIOPIC SYLLABLE SZAA	Amharic	1	[CA, Omniglot]
36	1224	ሤ	ETHIOPIC SYLLABLE SZEE	Amharic	1	[CA, Omniglot]
37	1225	ሥ	ETHIOPIC SYLLABLE SZE	Amharic	1	[CA, Omniglot]
38	1226	ሦ	ETHIOPIC SYLLABLE SZO	Amharic	1	[CA, Omniglot]
39	1227	ሧ	ETHIOPIC SYLLABLE SZWA	Amharic	1	[CA]
40	1228	ረ	ETHIOPIC SYLLABLE RA	Common-to-all	1	[CA, CT, Omniglot]
41	1229	ሩ	ETHIOPIC SYLLABLE RU	Common-to-all	1	[CA, CT, Omniglot]
42	122A	ሪ	ETHIOPIC SYLLABLE RI	Common-to-all	1	[CA, CT, Omniglot]
43	122B	ራ	ETHIOPIC SYLLABLE RAA	Common-to-all	1	[CA, CT, Omniglot]
44	122C	ሬ	ETHIOPIC SYLLABLE REE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
45	122D	ር	ETHIOPIC SYLLABLE RE	Common-to-all	1	[CA, CT, Omniglot]
46	122E	ሮ	ETHIOPIC SYLLABLE RO	Common-to-all-except: 47Hamt'agna	1	[CA, CT, Omniglot]
47	122F	ሯ	ETHIOPIC SYLLABLE RWA	Amharic, Tigrigna	1	[CT, Omniglot]
48	1230	ሰ	ETHIOPIC SYLLABLE SA	Common-to-all	1	[CA, CT, Omniglot]
49	1231	ሱ	ETHIOPIC SYLLABLE SU	Common-to-all	1	[CA, CT, Omniglot]
50	1232	ሲ	ETHIOPIC SYLLABLE SI	Common-to-all	1	[CA, CT, Omniglot]
51	1233	ሳ	ETHIOPIC SYLLABLE SAA	Common-to-all	1	[CA, CT, Omniglot]
52	1234	ሴ	ETHIOPIC SYLLABLE SEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
53	1235	ስ	ETHIOPIC SYLLABLE SE	Common-to-all	1	[CA, CT, Omniglot]

54	1236	ሶ	ETHIOPIC SYLLABLE SO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
55	1237	ሷ	ETHIOPIC SYLLABLE SWA	Amharic, Tigrigna	1	[CA, CT, Omniglot]
56	1238	ሸ	ETHIOPIC SYLLABLE SHA	Common-to-all	1	[CA, CT, Omniglot]
57	1239	ሹ	ETHIOPIC SYLLABLE SHU	Common-to-all	1	[CA, CT, Omniglot]
58	123A	ሺ	ETHIOPIC SYLLABLE SHI	Common-to-all	1	[CA, CT, Omniglot]
59	123B	ሻ	ETHIOPIC SYLLABLE SHAA	Common-to-all	1	[CA, CT, Omniglot]
60	123C	ሼ	ETHIOPIC SYLLABLE SHEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
61	123D	ሽ	ETHIOPIC SYLLABLE SHE	Common-to-all	1	[CA, CT, Omniglot]
62	123E	ሾ	ETHIOPIC SYLLABLE SHO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
63	123F	ሿ	ETHIOPIC SYLLABLE SHWA	Amharic, Tigrigna	1	[CT, Omniglot]
64	1240	ቀ	ETHIOPIC SYLLABLE QA	Common-to-all-except: Awngi	1	[CA, Omniglot]
65	1241	ቁ	ETHIOPIC SYLLABLE QU	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
66	1242	ቂ	ETHIOPIC SYLLABLE QI	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
67	1243	ቃ	ETHIOPIC SYLLABLE QAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
68	1244	ቄ	ETHIOPIC SYLLABLE QEE	Common-to-all-except: Awngi, Hamt'agna	1	[CA, CT, Omniglot]
69	1245	ቅ	ETHIOPIC SYLLABLE QE	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
70	1246	ቆ	ETHIOPIC SYLLABLE QO	Common-to-all-except: Awngi, Hamt'agna	1	[CA, CT, Omniglot]
71	1247	ቇ	ETHIOPIC SYLLABLE QOA	Amharic	1	[CA]
72	1248	ቈ	ETHIOPIC SYLLABLE QWA	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
73	124A	ቊ	ETHIOPIC SYLLABLE QWI	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
74	124B	ቋ	ETHIOPIC SYLLABLE QWAA	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
75	124C	ቌ	ETHIOPIC SYLLABLE QWEE	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
76	124D	ቍ	ETHIOPIC SYLLABLE QWE	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]

77	1250	ቐ	ETHIOPIC SYLLABLE QHA	Amharic, Awngi, Tigrigna	1	[CT, Omniglot]
78	1251	ቑ	ETHIOPIC SYLLABLE QHU	Amharic, Awngi, Tigrigna	1	[Omniglot]
79	1252	ቒ	ETHIOPIC SYLLABLE QHI	Amharic, Awngi, Tigrigna	1	[CT, Omniglot]
80	1253	ቓ	ETHIOPIC SYLLABLE QHAA	Amharic, Tigrigna	1	[Omniglot]
81	1254	ቔ	ETHIOPIC SYLLABLE QHEE	Amharic, Awngi, Tigrigna	1	[CT, Omniglot]
82	1255	ቕ	ETHIOPIC SYLLABLE QHE	Amharic, Awngi, Tigrigna	1	[Omniglot]
83	1256	ቖ	ETHIOPIC SYLLABLE QHO	Amharic, Awngi, Tigrigna	1	[Omniglot]
84	1258	ቘ	ETHIOPIC SYLLABLE QHWA	Awngi, Tigrigna	2	[Omniglot]
85	125A	ቚ	ETHIOPIC SYLLABLE QHWI	Awngi, Tigrigna	2	[Omniglot]
86	125B	ቛ	ETHIOPIC SYLLABLE QHWAA	Tigrigna	2	[Omniglot]
87	125C	ቜ	ETHIOPIC SYLLABLE QHWEE	Awngi, Tigrigna	2	[Omniglot]
88	125D	ቝ	ETHIOPIC SYLLABLE QHWE	Tigrigna	2	[CT, Omniglot]
89	1260	በ	ETHIOPIC SYLLABLE BA	Common-to-all-except: Argobba	1	[CA, CT, Omniglot]
90	1261	ቡ	ETHIOPIC SYLLABLE BU	Common-to-all-except: Argobba	1	[CA, CT, Omniglot]
91	1262	ቢ	ETHIOPIC SYLLABLE BI	Common-to-all-except: Argobba	1	[CA, CT, Omniglot]
92	1263	ባ	ETHIOPIC SYLLABLE BAA	Common-to-all-except: Argobba, Awngi	1	[CA, CT, Omniglot]
93	1264	ቤ	ETHIOPIC SYLLABLE BEE	Common-to-all-except: Argobba, Hamt'agna	1	[CA, CT, Omniglot]
94	1265	ብ	ETHIOPIC SYLLABLE BE	Common-to-all-except: Argobba	1	[CA, CT, Omniglot]
95	1266	ቦ	ETHIOPIC SYLLABLE BO	Common-to-all-except: Argobba, Hamt'agna	1	[CA, Omniglot]
96	1267	ቧ	ETHIOPIC SYLLABLE BWA	Amharic, Tigrigna	1	[CT, Omniglot]
97	1268	ቨ	ETHIOPIC SYLLABLE VA	Amharic, Tigrè, Tigrigna	1	[Omniglot]
98	1269	ቩ	ETHIOPIC SYLLABLE VU	Amharic, Tigrè, Tigrigna	1	[Omniglot]

99	126A	ቪ	ETHIOPIC SYLLABLE VI	Amharic, Tigrè, Tigrigna	1	[Omniglot]
100	126B	ቫ	ETHIOPIC SYLLABLE VAA	Amharic, Tigrè, Tigrigna	1	[CA, Omniglot]
101	126C	ቬ	ETHIOPIC SYLLABLE VEE	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
102	126D	ቭ	ETHIOPIC SYLLABLE VE	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
103	126E	ቮ	ETHIOPIC SYLLABLE VO	Amharic, Tigrè, Tigrigna	1	[CA, Omniglot]
104	126F	ቯ	ETHIOPIC SYLLABLE VWA	Amharic	1	[CA]
105	1270	ተ	ETHIOPIC SYLLABLE TA	Common-to-all	1	[CA, CT, Omniglot]
106	1271	ቱ	ETHIOPIC SYLLABLE TU	Common-to-all	1	[CA, CT, Omniglot]
107	1272	ቲ	ETHIOPIC SYLLABLE TI	Common-to-all	1	[CA, CT, Omniglot]
108	1273	ታ	ETHIOPIC SYLLABLE TAA	Common-to-all-except: Hamt'agna	1	[CA, Ct, Omniglot]
109	1274	ቴ	ETHIOPIC SYLLABLE TEE	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
110	1275	ት	ETHIOPIC SYLLABLE TE	Common-to-all	1	[CA, CT, Omniglot]
111	1276	ቶ	ETHIOPIC SYLLABLE TO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
112	1277	ቷ	ETHIOPIC SYLLABLE TWA	Amharic, Tigrigna	1	[CA, CT, Omniglot]
113	1278	ቸ	ETHIOPIC SYLLABLE CA	Common-to-all	1	[CA, CT, Omniglot]
114	1279	ቹ	ETHIOPIC SYLLABLE CU	Common-to-all	1	[CA, CT, Omniglot]
115	127A	ቺ	ETHIOPIC SYLLABLE CI	Common-to-all	1	[CA, CT, Omniglot]
116	127B	ቻ	ETHIOPIC SYLLABLE CAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
117	127C	ቼ	ETHIOPIC SYLLABLE CEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
118	127D	ች	ETHIOPIC SYLLABLE CE	Common-to-all	1	[CA, CT, Omniglot]
119	127E	ቾ	ETHIOPIC SYLLABLE CO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
120	127F	ቿ	ETHIOPIC SYLLABLE CWA	Amharic, Tigrigna	1	[CA, CT, Omniglot]
121	1280	ኀ	ETHIOPIC SYLLABLE XA	Amharic, Hamt'agna	1	[CA, Omniglot]
122	1281	ኁ	ETHIOPIC SYLLABLE XU	Amharic, Hamt'agna	1	[CA, Omniglot]
123	1282	ኂ	ETHIOPIC SYLLABLE XI	Amharic, Hamt'agna	1	[CA, Omniglot]

124	1283	ኃ	ETHIOPIC SYLLABLE XAA	Amharic, Hamt'agna	1	[CA, Omniglot]
125	1284	ኄ	ETHIOPIC SYLLABLE XEE	Amharic	1	[CA, Omniglot]
126	1285	ኅ	ETHIOPIC SYLLABLE XE	Amharic, Hamt'agna	1	[CA, Omniglot]
127	1286	ኆ	ETHIOPIC SYLLABLE XO	Amharic	1	[CA, Omniglot]
128	1288	ኈ	ETHIOPIC SYLLABLE XWA	Hamt'agna, Tigrè	4	[Omniglot]
129	128A	ኊ	ETHIOPIC SYLLABLE XWI	Amharic, Hamt'agna, Tigrè	1	[CA, Omniglot]
130	128B	ኋ	ETHIOPIC SYLLABLE XWAA	Amharic, Hamt'agna, Tigrè	1	[CA, Omniglot]
131	128C	ኌ	ETHIOPIC SYLLABLE XWEE	Amharic, Tigrè	1	[CA, Omniglot]
132	128D	ኍ	ETHIOPIC SYLLABLE XWE	Amharic, Hamt'agna, Tigrè	1	[CA, Omniglot]
133	1290	ነ	ETHIOPIC SYLLABLE NA	Common-to-all	1	[CA, CT, Omniglot]
134	1291	ኑ	ETHIOPIC SYLLABLE NU	Common-to-all	1	[CA, CT, Omniglot]
135	1292	ኒ	ETHIOPIC SYLLABLE NI	Common-to-all	1	[CA, CT, Omniglot]
136	1293	ና	ETHIOPIC SYLLABLE NAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
137	1294	ኔ	ETHIOPIC SYLLABLE NEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
138	1295	ን	ETHIOPIC SYLLABLE NE	Common-to-all	1	[CA, CT, Omniglot]
139	1296	ኖ	ETHIOPIC SYLLABLE NO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
140	1297	ኗ	ETHIOPIC SYLLABLE NWA	Amharic	1	[CA, Omniglot]
141	1298	ኘ	ETHIOPIC SYLLABLE NYA	Amharic, Harari, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
142	1299	ኙ	ETHIOPIC SYLLABLE NYU	Amharic, Harari, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
143	129A	ኚ	ETHIOPIC SYLLABLE NYI	Amharic, Harari, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
144	129B	ኛ	ETHIOPIC SYLLABLE NYAA	Amharic, Harari, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
145	129C	ኜ	ETHIOPIC SYLLABLE NYEE	Amharic, Harari, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
146	129D	ኝ	ETHIOPIC SYLLABLE NYE	Amharic, Harari, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
147	129E	ኞ	ETHIOPIC SYLLABLE NYO	Amharic, Harari, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
148	129F	ኟ	ETHIOPIC SYLLABLE NYWA	Amharic	1	[CA, Omniglot]

149	12A0	አ	ETHIOPIC SYLLABLE GLOTTAL A	Common-to-all	1	[CA, CT, Omniglot]
150	12A1	ኡ	ETHIOPIC SYLLABLE GLOTTAL U	Common-to-all	1	[CA, CT, Omniglot]
151	12A2	ኢ	ETHIOPIC SYLLABLE GLOTTAL I	Common-to-all	1	[CA, CT, Omniglot]
152	12A3	ኣ	ETHIOPIC SYLLABLE GLOTTAL AA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
153	12A4	ኤ	ETHIOPIC SYLLABLE GLOTTAL EE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
154	12A5	እ	ETHIOPIC SYLLABLE GLOTTAL E	Common-to-all	1	[CA, CT, Omniglot]
155	12A6	ኦ	ETHIOPIC SYLLABLE GLOTTAL O	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
156	12A7	ኧ	ETHIOPIC SYLLABLE GLOTTAL WA	Amharic	1	[CA, Omniglot]
157	12A8	ከ	ETHIOPIC SYLLABLE KA	Common-to-all	1	[CA, CT, Omniglot]
158	12A9	ኩ	ETHIOPIC SYLLABLE KU	Common-to-all	1	[CA, CT, Omniglot]
159	12AA	ኪ	ETHIOPIC SYLLABLE KI	Common-to-all	1	[CA, CT, Omniglot]
160	12AB	ካ	ETHIOPIC SYLLABLE KAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
161	12AC	ኬ	ETHIOPIC SYLLABLE KEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
162	12AD	ክ	ETHIOPIC SYLLABLE KE	Common-to-all	1	[CA, CT, Omniglot]
163	12AE	ኮ	ETHIOPIC SYLLABLE KO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
164	12B0	ኰ	ETHIOPIC SYLLABLE KWA	Common-to-all-except: Argobba, Harari	1	[CA, CT, Omniglot]
165	12B2	ኲ	ETHIOPIC SYLLABLE KWI	Common-to-all-except: Argobba, Harari	1	[CA, CT, Omniglot]
166	12B3	ኳ	ETHIOPIC SYLLABLE KWAA	Common-to-all-except: Argobba, Harari	1	[CA, CT, Omniglot]
167	12B4	ኴ	ETHIOPIC SYLLABLE KWEE	Amharic, Awngi, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
168	12B5	ኵ	ETHIOPIC SYLLABLE KWE	Amharic, Awngi, Tigrigna	1	[CA, CT, Omniglot]
169	12B8	ኸ	ETHIOPIC SYLLABLE KXA	Amharic, Awngi, Hamt'agna, Tigrigna	1	[CA, CT, Omniglot]

170	12B9	ኹ	ETHIOPIC SYLLABLE KXU	Amharic, Awngi, Hamt'agna, Tigrigna	1	[CA, CT, Omniglot]
171	12BA	ኺ	ETHIOPIC SYLLABLE KXI	Amharic, Awngi, Hamt'agna, Tigrigna	1	[CA, CT, Omniglot]
172	12BB	ኻ	ETHIOPIC SYLLABLE KXAA	Amharic, Hamt'agna, Tigrigna	1	[CA, CT, Omniglot]
173	12BC	ኼ	ETHIOPIC SYLLABLE KXEE	Amharic, Awngi, Tigrigna	1	[CA, CT, Omniglot]
174	12BD	ኽ	ETHIOPIC SYLLABLE KXE	Amharic, Awngi, Hamt'agna, Tigrigna	1	[CA, CT, Omniglot]
175	12BE	ኾ	ETHIOPIC SYLLABLE KXO	Amharic, Awngi, Tigrigna	1	[CA, CT, Omniglot]
176	12C0	ዀ	ETHIOPIC SYLLABLE KXWA	Amharic	1	[CA]
177	12C2	ዂ	ETHIOPIC SYLLABLE KXWI	Amharic, Tigrigna	1	[CA, CT, Omniglot]
178	12C3	ዃ	ETHIOPIC SYLLABLE KXWAA	Amharic, Tigrigna	1	[CA, CT ,Omnilgot]
179	12C4	ዄ	ETHIOPIC SYLLABLE KXWEE	Awngi, Tigrigna,	2	[Omnilgot]
180	12C5	ዅ	ETHIOPIC SYLLABLE KXWE	Amharic, Awngi, Tigrigna	1	[CA, CT, Omniglot]
181	12C8	ወ	ETHIOPIC SYLLABLE WA	Common-to-all	1	[CA, CT, Omniglot]
182	12C9	ዉ	ETHIOPIC SYLLABLE WU	Common-to-all	1	[CA, CT, Omniglot]
183	12CA	ዊ	ETHIOPIC SYLLABLE WI	Common-to-all	1	[CA, CT, Omniglot]
184	12CB	ዋ	ETHIOPIC SYLLABLE WAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
185	12CC	ዌ	ETHIOPIC SYLLABLE WEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
186	12CD	ው	ETHIOPIC SYLLABLE WE	Common-to-all	1	[CA, CT, Omniglot]
187	12CE	ዎ	ETHIOPIC SYLLABLE WO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
188	12CF	ዏ	ETHIOPIC SYLLABLE WOA	Tigrigna	2	[CT]
189	12D0	ዐ	ETHIOPIC SYLLABLE PHARYNGEAL A	Common-to-all-except: Awngi, Hamt'agna, Tigrè	1	[CA, CT, Omniglot]
190	12D1	ዑ	ETHIOPIC SYLLABLE PHARYNGEAL U	Common-to-all-except: Awngi, Hamt'agna, Tigrè	1	[CA, CT, Omniglot]

191	12D2	᎐	ETHIOPIC SYLLABLE PHARYNGEAL I	Common-to-all-except: Awngi, Hamt'agna, Tigrè	1	[CA, CT, Omniglot]
192	12D3	᎐	ETHIOPIC SYLLABLE PHARYNGEAL AA	Common-to-all-except: Awngi, Hamt'agna, Tigrè	1	[CA, CT, Omniglot]
193	12D4	᎐	ETHIOPIC SYLLABLE PHARYNGEAL EE	Common-to-all-except: Awngi, Hamt'agna, Tigrè	1	[CA, CT, Omniglot]
194	12D5	᎐	ETHIOPIC SYLLABLE PHARYNGEAL E	Common-to-all-except: Awngi, Hamt'agna, Tigrè	1	[CA, CT, Omniglot]
195	12D6	᎐	ETHIOPIC SYLLABLE PHARYNGEAL O	Common-to-all-except: Awngi, Hamt'agna, Tigrè	1	[CA, CT, Omniglot]
196	12D8	ዘ	ETHIOPIC SYLLABLE ZA	Common-to-all	1	[CA, CT, Omniglot]
197	12D9	ዙ	ETHIOPIC SYLLABLE ZU	Common-to-all	1	[CA, CT, Omniglot]
198	12DA	ዚ	ETHIOPIC SYLLABLE ZI	Common-to-all	1	[CA, CT, Omniglot]
199	12DB	ዛ	ETHIOPIC SYLLABLE ZAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
200	12DC	ዜ	ETHIOPIC SYLLABLE ZEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
201	12DD	ዝ	ETHIOPIC SYLLABLE ZE	Common-to-all	1	[CA, CT, Omniglot]
202	12DE	ዞ	ETHIOPIC SYLLABLE ZO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
203	12DF	ዟ	ETHIOPIC SYLLABLE ZWA	Amharic	1	[CA, Omniglot]
204	12E0	ዠ	ETHIOPIC SYLLABLE ZHA	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
205	12E1	ዡ	ETHIOPIC SYLLABLE ZHU	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
206	12E2	ዢ	ETHIOPIC SYLLABLE ZHI	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
207	12E3	ዣ	ETHIOPIC SYLLABLE ZHAA	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
208	12E4	ዤ	ETHIOPIC SYLLABLE ZHEE	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
209	12E5	ዥ	ETHIOPIC SYLLABLE ZHE	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]

210	12E6	ዦ	ETHIOPIC SYLLABLE ZHO	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
211	12E7	ዧ	ETHIOPIC SYLLABLE ZHWA	Amharic	1	[CA, Omniglot]
212	12E8	የ	ETHIOPIC SYLLABLE YA	Common-to-all	1	[CA, CT, Omniglot]
213	12E9	ዩ	ETHIOPIC SYLLABLE YU	Common-to-all	1	[CA, CT, Omniglot]
214	12EA	ዪ	ETHIOPIC SYLLABLE YI	Common-to-all	1	[CA, CT, Omniglot]
215	12EB	ያ	ETHIOPIC SYLLABLE YAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
216	12EC	ዬ	ETHIOPIC SYLLABLE YEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
217	12ED	ይ	ETHIOPIC SYLLABLE YE	Common-to-all	1	[CA, CT, Omniglot]
218	12EE	ዮ	ETHIOPIC SYLLABLE YO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
219	12EF	ዯ	ETHIOPIC SYLLABLE YOA	Tigrigna	2	[CT]
220	12F0	ደ	ETHIOPIC SYLLABLE DA	Common-to-all	1	[CA, CT, Omniglot]
221	12F1	ዱ	ETHIOPIC SYLLABLE DU	Common-to-all	1	[CA, CT, Omniglot]
222	12F2	ዲ	ETHIOPIC SYLLABLE DI	Common-to-all	1	[CA, CT, Omniglot]
223	12F3	ዳ	ETHIOPIC SYLLABLE DAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
224	12F4	ዴ	ETHIOPIC SYLLABLE DEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
225	12F5	ድ	ETHIOPIC SYLLABLE DE	Common-to-all	1	[CA, CT, Omniglot]
226	12F6	ዶ	ETHIOPIC SYLLABLE DO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
227	12F7	ዷ	ETHIOPIC SYLLABLE DWA	Amharic, Tigrigna	1	[CA, CT, Omniglot]
228	1300	ጀ	ETHIOPIC SYLLABLE JA	Common-to-all	1	[CA, CT, Omniglot]
229	1301	ጁ	ETHIOPIC SYLLABLE JU	Common-to-all	1	[CA, CT, Omniglot]
230	1302	ጂ	ETHIOPIC SYLLABLE JI	Common-to-all	1	[CA, CT, Omniglot]
231	1303	ጃ	ETHIOPIC SYLLABLE JAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
232	1304	ጄ	ETHIOPIC SYLLABLE JEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
233	1305	ጅ	ETHIOPIC SYLLABLE JE	Common-to-all	1	[CA, CT, Omniglot]
234	1306	ጆ	ETHIOPIC SYLLABLE JO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
235	1307	ጇ	ETHIOPIC SYLLABLE JWA	Amharic, Tigrigna	1	[CA, CT, Omniglot]
236	1308	ገ	ETHIOPIC SYLLABLE GA	Common-to-all	1	[CA, CT, Omniglot]

237	1309	ጉ	ETHIOPIC SYLLABLE GU	Common-to-all	1	[CA, CT, Omniglot]
238	130A	ጊ	ETHIOPIC SYLLABLE GI	Common-to-all	1	[CA, CT, Omniglot]
239	130B	ጋ	ETHIOPIC SYLLABLE GAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
240	130C	ጌ	ETHIOPIC SYLLABLE GEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
241	130D	ግ	ETHIOPIC SYLLABLE GE	Common-to-all	1	[CA, CT, Omniglot]
242	130E	ጎ	ETHIOPIC SYLLABLE GO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
243	1310	ጐ	ETHIOPIC SYLLABLE GWA	Common-to-all-except: Argobba, Harari, Silt'e	1	[CA, CT, Omniglot]
244	1312	ጒ	ETHIOPIC SYLLABLE GWI	Common-to-all-except: Argobba, Harari, Silt'e	1	[CT, Omniglot]
245	1313	ጓ	ETHIOPIC SYLLABLE GWAA	Common-to-all-except: Argobba, Awngi, Harari, Silt'e	1	[CA, CT, Omniglot]
246	1314	ጔ	ETHIOPIC SYLLABLE GWEE	Common-to-all-except: Argobba, Hamt'agna, Harari, Silt'e	1	[CA, Omniglot]
247	1315	ጕ	ETHIOPIC SYLLABLE GWE	Common-to-all-except: Argobba, Harari, Silt'e	1	[CA, CT, Omniglot]
248	1318	ጘ	ETHIOPIC SYLLABLE GGA	Amharic, Awngi, Hamt'agna	1	[CA, Omniglot]
249	1319	ጙ	ETHIOPIC SYLLABLE GGU	Awngi, Hamt'agna	4	[Omniglot]
250	131A	ጚ	ETHIOPIC SYLLABLE GGI	Awngi, Hamt'agna	4	[Omniglot]
251	131B	ጛ	ETHIOPIC SYLLABLE GGAA	Hamt'agna	4	[Omniglot]
252	131C	ጜ	ETHIOPIC SYLLABLE GGEE	Awngi	4	[Omniglot]
253	131D	ጝ	ETHIOPIC SYLLABLE GGE	Awngi, Hamt'agna	4	[Omniglot]
254	131E	ጞ	ETHIOPIC SYLLABLE GGO	Awngi	4	[Omniglot]
255	1320	ጠ	ETHIOPIC SYLLABLE THA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
256	1321	ጡ	ETHIOPIC SYLLABLE THU	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
257	1322	ጢ	ETHIOPIC SYLLABLE THI	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
258	1323	ጣ	ETHIOPIC SYLLABLE THAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]

259	1324	ጤ	ETHIOPIC SYLLABLE THEE	Common-to-all-except: Awngi, Hamt'agna	1	[CA, CT, Omniglot]
260	1325	ጥ	ETHIOPIC SYLLABLE THE	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
261	1326	ጦ	ETHIOPIC SYLLABLE THO	Common-to-all-except: Awngi, Hamt'agna	1	[CA, CT, Omniglot]
262	1327	ጧ	ETHIOPIC SYLLABLE THWA	Amharic	1	[CA, Omniglot]
263	1328	ጨ	ETHIOPIC SYLLABLE CHA	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
264	1329	ጩ	ETHIOPIC SYLLABLE CHU	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
265	132A	ጪ	ETHIOPIC SYLLABLE CHI	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
266	132B	ጫ	ETHIOPIC SYLLABLE CHAA	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
267	132C	ጬ	ETHIOPIC SYLLABLE CHEE	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
268	132D	ጭ	ETHIOPIC SYLLABLE CHE	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
269	132E	ጮ	ETHIOPIC SYLLABLE CHO	Common-to-all-except: Argobba, Awngi, Hamt'agna	1	[CA, CT, Omniglot]
270	132F	ጯ	ETHIOPIC SYLLABLE CHWA	Amharic	1	[CA, Omniglot]
271	1330	ጰ	ETHIOPIC SYLLABLE PHA	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
272	1331	ጱ	ETHIOPIC SYLLABLE PHU	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
273	1332	ጲ	ETHIOPIC SYLLABLE PHI	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
274	1333	ጳ	ETHIOPIC SYLLABLE PHAA	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
275	1334	ጴ	ETHIOPIC SYLLABLE PHEE	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]

276	1335	ጵ	ETHIOPIC SYLLABLE PHE	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
277	1336	ጶ	ETHIOPIC SYLLABLE PHO	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
278	1337	ጷ	ETHIOPIC SYLLABLE PHWA	Amharic	1	[CA]
279	1338	ጸ	ETHIOPIC SYLLABLE TSA	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
280	1339	ጹ	ETHIOPIC SYLLABLE TSU	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
281	133A	ጺ	ETHIOPIC SYLLABLE TSI	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
282	133B	ጻ	ETHIOPIC SYLLABLE TSAA	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
283	133C	ጼ	ETHIOPIC SYLLABLE TSEE	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
284	133D	ጽ	ETHIOPIC SYLLABLE TSE	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
285	133E	ጾ	ETHIOPIC SYLLABLE TSO	Amharic, Hamt'agna, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
286	133F	ጿ	ETHIOPIC SYLLABLE TSWA	Amharic, Tigrè, Tigrigna	1	[CA, CT, Omniglot]
287	1340	ፀ	ETHIOPIC SYLLABLE TZA	Amharic, Awngi	1	[CA, Omniglot]
288	1341	ፁ	ETHIOPIC SYLLABLE TZU	Amharic, Awngi	1	[CA, Omniglot]
289	1342	ፂ	ETHIOPIC SYLLABLE TZI	Amharic, Awngi	1	[CA, Omniglot]
290	1343	ፃ	ETHIOPIC SYLLABLE TZAA	Amharic	1	[CA, Omniglot]
291	1344	ፄ	ETHIOPIC SYLLABLE TZEE	Amharic, Awngi	1	[CA, Omniglot]
292	1345	ፅ	ETHIOPIC SYLLABLE TZE	Amharic, Awngi	1	[CA, Omniglot]
293	1346	ፆ	ETHIOPIC SYLLABLE TZO	Amharic, Awngi	1	[CA, Omniglot]
294	1348	ፈ	ETHIOPIC SYLLABLE FA	Common-to-all	1	[CA, CT, Omniglot]
295	1349	ፉ	ETHIOPIC SYLLABLE FU	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
296	134A	ፊ	ETHIOPIC SYLLABLE FI	Common-to-all	1	[CA, CT, Omniglot]

297	134B	ፋ	ETHIOPIC SYLLABLE FAA	Common-to-all-except: Awngi	1	[CA, CT, Omniglot]
298	134C	ፌ	ETHIOPIC SYLLABLE FEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
299	134D	ፍ	ETHIOPIC SYLLABLE FE	Common-to-all	1	[CA, CT, Omniglot]
300	134E	ፎ	ETHIOPIC SYLLABLE FO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
301	134F	ፏ	ETHIOPIC SYLLABLE FWA	Amharic	1	[CA, Omniglot]
302	1350	ፐ	ETHIOPIC SYLLABLE PA	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
303	1351	ፑ	ETHIOPIC SYLLABLE PU	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
304	1352	ፒ	ETHIOPIC SYLLABLE PI	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
305	1353	ፓ	ETHIOPIC SYLLABLE PAA	Common-to-all-except: Awngi, Hamt'agna	1	[CA, CT, Omniglot]
306	1354	ፔ	ETHIOPIC SYLLABLE PEE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
307	1355	ፕ	ETHIOPIC SYLLABLE PE	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
308	1356	ፖ	ETHIOPIC SYLLABLE PO	Common-to-all-except: Hamt'agna	1	[CA, CT, Omniglot]
309	1357	ፗ	ETHIOPIC SYLLABLE PWA	Amharic	1	[CA]
310	1358	ፘ	ETHIOPIC SYLLABLE RYA	Amharic	1	[Omniglot]
311	1359	ፙ	ETHIOPIC SYLLABLE MYA	Amharic	1	[Omniglot]

5.3 Code Points Excluded from the Repertoire

The two blocks listed below are whole block exclusions in the MSR-2 document, hence they are preemptively out of consideration in the LGR proposal:

- 2D80..2DDF; Ethiopic Extended
- AB00..AB2F; Ethiopic Extended-A

As described in section 4.1, the GP used the corpus as auxiliary evidence for the Amharic and Tigrigna languages, and for these two languages the decision to exclude a code point from the repertoire was made based on the non-appearance of a code point both from the script version published by the online encyclopedia and the frequency list of the corpus data analysed. The

Ethiopic GP has used the panel members' knowledge and expertise, public comments and several available resources to base its decision on excluding code points. For the six languages, since the languages are not well-resourced (as their EGIDS score confirms) a corpus data was not used, and the decision to exclude a code point was made based on the non-appearance of the code point from the script version published by the online encyclopedia of Writing Systems and Languages.

Accordingly, the following code points have been excluded from the code point repertoire of the Ethiopic script LGR as there has not been any occurrence of the code point in the corpus data analysed for online contents published in Amharic and Tigrigna languages, as well as the eight Ethiopic script versions published for the eight languages in the online encyclopedia of Writing Systems and Languages.

Table 5.2 Code Points Excluded from the Repertoire

No.	Unicode	Glyph	Code Point Name and Properties
1	1207	ሇ	ETHIOPIC SYLLABLE HOA
2	1287	ዀ	ETHIOPIC SYLLABLE XOA
3	12AF	ኯ	ETHIOPIC SYLLABLE KOA
4	12FF	ዿ	ETHIOPIC SYLLABLE DDWA
5	130F	ጏ	ETHIOPIC SYLLABLE GOA
6	131F	ጟ	ETHIOPIC SYLLABLE GGWAA
7	1347	ፇ	ETHIOPIC SYLLABLE TZOA
8	135A	ፚ	ETHIOPIC SYLLABLE FYA
9	12F8	□	ETHIOPIC SYLLABLE DDA
10	12F9	□	ETHIOPIC SYLLABLE DDU
11	12FA	□	ETHIOPIC SYLLABLE DDI
12	12FB	□	ETHIOPIC SYLLABLE DDAA
13	12FC	□	ETHIOPIC SYLLABLE DDEE
14	12FD	□	ETHIOPIC SYLLABLE DDE
15	12FE	□	ETHIOPIC SYLLABLE DDO

6   Variants

There are no cross-script variants in Ethiopic script. However there are within-script variants that have been addressed based on the principles outlined in section 4.3.

6.1 Description of Variants Assigned

As it did for the Amharic writing system, the GP has reviewed the writing systems of the seven languages considered in the LG proposal for the presence of variant code points and it appears that the seven languages considered in the LGR do not have code point variants. As explained in

section 3.1, variant code points is the phenomenon of the Amharic writing system, and the remaining seven writing systems do not have this property.

In this section, the GP decisions as well as treatment of variants on Ethiopic script in respect to the principles presented in the "Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels" [Procedure] have been described. The GP considered script wide variants as opposed to language or writing-system-specific variants.

The variant analysis described below for the Ethiopic script is a result of the presence of code point redundancy or phonemic decay (see Section 3.1 for the notions description) on some code points in the Amharic language or writing system [HomonymsPart1, HomonymsPart2, HomonymList].

Table 6.1 Code Points Proposed as Variants by the GP

Code Point	Code Point Name	Code Point	Code Point Name	Code Point	Code Point Name	Type (Blocked/ Allocatable)	Applicable Principle No.
(1200) ሀ	HA	(1210) ሐ	HHA	(1280) ኀ	XA	B	1,2
(1201) ሁ	HU	(1211) ሑ	HHU	(1281) ኁ	XU	B	1,2
(1202) ሂ	HI	(1212) ሒ	HHI	(1282) ኂ	XI	B	1,2
(1203) ሃ	HAA	(1213) ሓ	HHAA	(1283) ኃ	XAA	B	1,2
(1204) ሄ	HEE	(1214) ሔ	HHEE	(1284) ኄ	XEE	B	1,2
(1205) ህ	HE	(1215) ሕ	HHE	(1285) ኅ	XE	B	1,2
(1206) ሆ	HO	(1216) ሖ	HHO	(1286) ኆ	XO	B	1,2
-	-	(1217) ሗ	HWA	(1288) ኈ	XWA	B	1,2
(1230) ሰ	SA	(1220) ሠ	SZA	-	-	B	1
(1231) ሱ	SU	(1221) ሡ	SZU	-	-	B	1
(1232) ሲ	SI	(1222) ሢ	SZI	-	-	B	1
(1233) ሳ	SAA	(1223) ሣ	SZAA	-	-	B	1
(1234) ሴ	SEE	(1224) ሤ	SZEE	-	-	B	1
(1235) ስ	SE	(1225) ሥ	SZE	-	-	B	1
(1236) ሶ	SO	(1226) ሦ	SZO	-	-	B	1
(1237) ሷ	SWA	(1227) ሧ	SZWA	-	-	B	1
(12A0) አ	GLOTTAL A	(12D0) ዐ	PHARYNGEAL A	-	-	B	1
(12A1) ኡ	GLOTTAL U	(12D1) ዑ	PHARYNGEAL U	-	-	B	1
(12A2) ኢ	GLOTTAL I	(12D2) ዒ	PHARYNGEAL I	-	-	B	1

(12A3)አ	GLOTTAL AA	(12D3)ዓ	PHARYNGEAL AA	-	-	B	1
(12A4)ኄ	GLOTTAL EE	(12D4)ዔ	PHARYNGEAL EE	-	-	B	1
(12A5)እ	GLOTTAL E	(12D5)ዕ	PHARYNGEAL E	-	-	B	1
(12A6)ኆ	GLOTTAL O	(12D6)ዖ	PHARYNGEAL O	-	-	B	1
(12AE)ከ	KO	(12B0)ኰ	KWA	-	-	B	1
(1338) ጸ	TSA	(1340)ፀ	TZA	-	-	B	1
(1339)ጹ	TSU	(1341)ፁ	TZU	-	-	B	1
(133A)ጺ	TSI	(1342)ፂ	TZI	-	-	B	1
(133B)ጻ	TSAA	(1343)ፃ	TZAA	-	-	B	1
(133C)ጼ	TSEE	(1344)ፄ	TZEE	-	-	B	1
(133D)ጽ	TSE	(1345)ፅ	TZE	-	-	B	1
(133E)ጾ	TSO	(1346)ፆ	TZO	-	-	B	1
(12A0)አ	GLOTTAL A	(12D3) ዓ	GLOTTAL AA	-	-	B	1,2
(12D0)ዐ	PHARYNGEAL A	(12D3)ዓ	PHARYNGEAL AA	-	-	B	1,2

6.2 Why Variants

Due to the security problems that allocatable variants pose, the variants have been classified as blocked variants; yet this classification is necessary, for without this classification or restriction the system will be open to spoofing and other related security risks. With this intent, the GP treated the code point variants in the most restrictive way possible, in that they have been considered as script-wide variants, beyond the Amharic language or writing system. Thus, any pair of code points considered variants in the writing system of the Amharic language will be taken as a variant pair across the entire set of languages using the script.

6.3 The Impact of Variants on Other Languages

In the following table, Table 6.2, sample words which use code points interchangeably are presented in order to explain what code point redundancy or phonemic decay in the Amharic writing system means. The sample words are taken from a much broader list available in [HomonymList]: its first page, with column-headers translated, is annexed in Appendix C.

Table 6.2 – Sample Words with Interchangeably Used Code Points

Amharic Words With Pair of Interchangeably Code Points used	Meaning in the Amharic Language	Code Points Interchangeably Used	Language Affected Due to Collision Caused by Variants	Remark
ሕብር ጎብር	Both "ሕብር" and "ጎብር" mean: ቀለም/Color	(1215)ሕ and (1285)ጎ	None	The word "ሕብር" is a loan word for both Amharic and Tigrigna languages(loaned from Geʼez) and the Tigrigna language is affected by same collision
ምሁር ምሑር	Both "ምሁር" and "ምሑር" mean: ሊቅ/ Educated	(1201)ሁ and (1211)ሑ	None	The word "ምሑር" is a loan word for both Amharic and Tigrigna languages(loaned from Geʼez) and the Tigrigna language is affected by the same collision
ሣለ ሳለ	Both "ሣለ" and "ሳለ" mean: ምሥልን ሠራ/He painted a Picture	(1223)ሣ and (1233)ሳ	None	
ሣይ ሳይ	Both "ሣይ" and "ሳይ" mean: ሠዓሊ./ Painter or ስመለከት/Seeing	(1223)ሣ and (1233)ሳ	None	
መእት, መዓት	Both "መእት" and "መዓት" mean: እጅግ ብዙ/ Too much(used to describe quantity) or ቁጣ /Heavy punishment	(12A0)አ and (12D3)ዓ	None	
አደገ ዐደገ	Both "አደገ" and "ዐደገ" mean: ተለቀ/Getting Grown	(12A0)አ and (12D0)ዐ	None	

The Amharic words described above are sample words taken to explain why the Ethiopic GP treated the code points described in Table 6.1 as variants, and the list is not meant to be exhaustive.

Besides the Amharic language, the GP has also analyzed the impact of the proposed variant code points identified in Table 6.1 on the other seven languages using the script, and considered in the LGR proposal. There are 98 code points identified as variants in Ethiopic script. Some of the code points treated as variants in Amharic are not used by other seven languages. In view of differences in the use of the script among the respective writing systems of the languages, the chance of collision between labels from one language based on certain code points being treated as variants due to some other language is negligible because of the differences in the use of the script among the languages. (In other words, most pairs of real-world labels would contain some code point unique to that language, and therefore unaffected by variant definitions in another language). This is true whether canonical or non-canonical forms are concerned.

In an attempt to measure the effect of the Amharic language derived variants on the other languages considered in the LGR proposal objectively, two independent test runs were conducted to scan the occurrence of collisions and they yield compatible results. For the test about 300K unique Amharic words and 30K unique Tigrigna words have been used. Below is a summary of the test results:

a. **Summary of the Test Run:**

Table 6.3 Amharic-then-Tigrigna: Under the Assumption that 300k labels (Amharic Words) Already Registered

Language	Number of Words	Unique Labels	Successfully Registered	Errors	Collisions	Exact Collisions
Amharic	394214	394214	290308	91447	12459	0
Tigrigna	37593	37593	27839	5516	1367	2871

Table 6.4 Tigrigna-then-Amharic: Under the Assumption that 30k labels (Tigrigna Words) Already Registered

Language	Number of Words	Unique Labels	Successfully Registered	Errors	Collisions	Exact Collisions
Tigrigna	37593	37593	31437	5516	640	0
Amharic	394214	394214	286710	91447	12979	3078

As can be seen from the two tables described above (Table 6.3 and 6.4):

- 2% of labels (640) experience collision within the Tigrigna set due to variants defined for the Amharic language;
- about 4% of labels (1367) experience collisions with variants of Amharic words; and
- over 9% of labels(2871) match some Amharic word precisely.

It can be said that the effect may in fact be proportional to the percentages of code point overlap between languages (see a snapshot of the extent of overlaps in Table 6.5 below). The test results also show case that the effect of words simply being the same (both Amharic and Tigrigna writing systems use words loaned from Ge'ez) or being spelled the same has a proportionally bigger influence than the effect of the variants. In summary, the impact of Amharic-driven variants on other languages is acceptable and the GP's decision to trade the impact for security is reasonable.

We give also a summary of the extent of Amharic-driven variants impact on the other languages as described in the table below (identified in Table 6.5). These impacts on non-Amharic, given as a percentage of the relevant Amharic code points involved, are summarized in the table below, Table 6.5.

Table 6.5 The Extent of Code Point Overlap (Collisions) If and When the Seven

Languages use Amharic-Driven Variants

Language	Common variants	% of common variants	Remark
Tigrigna	36	36%	64% of code points treated as variants are not used by the Tigrigna language at all / They are not part of the Ethiopic script version used by the Tigrigna language.
Argobba	34	34%	66% of code points treated as variants are not used by the Argobba language at all / They are not part of the Ethiopic script version used by the Argobba language.
Hamt'agna	22	22%	78% of code points treated as variants are not used by the Hamt'agna language at all / They are not part of the Ethiopic script version used by the Hamt'agna language.
Tigrè	20	20%	80% of code points treated as variants are not used by the Tigrè language at all- in other words they are not part of the Ethiopic script version used by the Tigrè language.

Silt'e	18	18%	82% of code points treated as variants are not used by the Silt'e language at all / They are not part of the Ethiopic script version used by the Silt'e language.
Harari	16	16%	84% of code points treated as variants are not used by the Harari language at all / They are not part of the Ethiopic script version used by the Harari language.
Awngi	14	14%	86% of code points treated as variants are not used by the Argobba language at all / They are not part of the Ethiopic script version used in Argobba language.

7   Whole Label Evaluation Rules (WLE)

As it stands no WLE rules specific to the Ethiopic script are needed. As required, the LGR specification for this proposal copies the default rules and actions of MSR-2.

8 Contributors

The Ethiopic GP is composed of the following members:

No.	Name	Role in the Panel	Organization
1	Dessalegn Mequanint	**Chair**	Addis Ababa University, Department Computer Science.
2	Mulugeta Seyoum	**Secretary**	Addis Ababa University, Academy of Ethiopian Languages and Cultures
3	Kinfe Micheal Yilma	**Member**	Ethiopian Justice and Legal System Research Institute
4	Gezahegn Tadesse	**Member**	Awash International Bank
5	Balcha Reba	**Member**	Ministry of Communication and Information Technology
6	Assefa Kore	**Member**	Ethio-Telecom
7	Tigabu Dagne	**Member**	Addis Ababa Institute of Technology
8	Halefom Hailu	**Member**	INSA Ethiopia
9	Teshome Yehualashet	**Member**	Addis Ababa University, Academy of Ethiopian Languages and Cultures.
10	Kassahun Lemlemu	**Member**	Mahibere Kidusan
11	Destaye Alemayehu	**Member**	Addis Ababa University, Department Computer Science
12	Senait Kiros	**Member**	Addis Ababa University, Department Computer Science

9 References

[Guidelines]   ICANN, "Guidelines for Developing Script-Specific Label Generation Rules for Integration into the Root Zone",
https://www.icann.org/en/system/files/files/Guidelines-for-LGR-20150424.pdf

[Procedure]   ICANN," Procedure to Develop and Maintain the Label Generation Rules for the Root Zone in Respect of IDNA Labels.",
https://www.icann.org/en/system/files/files/lgr-procedure-20mar13-en.pdf

[MSR-2]   Integration Panel, "Maximal Starting Repertoire — MSR-2 Overview and Rationale", 14 April 2015,
https://www.icann.org/en/system/files/files/msr-2-overview-14apr15-en.pdf

[Unicode 6.3]   Unicode® 6.3.0, 30 September 2013,
http://unicode.org/versions/Unicode6.3.0/

[RFC6912]   Sullivan, A., Thaler, D., Klensin, J., and O. Kolkman, "Principles for Unicode Code Point Inclusion in Labels in the DNS", RFC 6912, April 2013, http://www.rfc-editor.org/info/rfc6912.

[RFC7940]   Davies, K. and A. Freytag, "Representing Label Generation Rulesets Using XML", RFC 7940, August 2016, http://www.rfc-editor.org/info/rfc7940.

[Requirements]   Integration Panel "Requirements for LGR Proposals from Generation Panels" available online ashttps://www.icann.org/en/system/files/files/Requirements-for-LGR-Proposals-20150424.pdf

[EGIDS]   Ethnologue- Languages of the World,
https://www.ethnologue.com/browse/names

[CA]   Corpus Analysis performed by crawling 14,850 html Amharic Contents of size 1.8 GB published online

[Omniglot]   The Online Encyclopedia of Writing Systems & Languages, Date Accessed- August, 2016:   http://www.omniglot.com/writing/amharic.htm, http://www.omniglot.com/writing/argobba.htm, http://www.omniglot.com/writing/awngi.htm, http://www.omniglot.com/writing/harari.htm, http://www.omniglot.com/writing/xamtanga.htm, http://www.omniglot.com/writing/silte.htm, http://www.omniglot.com/writing/tigre.htm, http://www.omniglot.com/writing/tigrinya.htm.

[CT]   Corpus Analysis performed by crawling 598 html and 40 PDF files with Tigrigna Contents published online.

[ANT]   Alphabet National du Tchad, Nat. standard for Chad; Fig 11 in [L2/10-288R] (pp. 19-20)

[L2/10-288R]   Lorna A. Priest, Martin Hosken, "Proposal to add Arabic script characters for African and Asian languages", http://www.unicode.org/L2/L2010/10288r-arabic-proposal.pdf

[HomonymsPart1]   Amharic Orthography and Homonyms Part 1(በዐማርኛ ሥርዐተ-ጽሕፈት ውስጥ የድምፀ-ሞክሼ ሆሄያት አጠቃቀም ክፍል አንድ), http://www.senamirmir.org/downloads/Homonyms-Note01.pdf

[HomonymsPart2]   Amharic Orthography and Homonyms Part 2(በዐማርኛ ሥርዐተ-ጽሕፈት ውስጥ የድምፀ-ሞክሼ ሆሄያት አጠቃቀም ክፍል ሁለት), http://www.senamirmir.org/downloads/Homonyms-List02.pdf

[HomonymList]   Usage of Sound alike Alphabets in Amharic Writing System, http://www.senamirmir.org/downloads/Homonyms-List02.pdf

[Wikipedia]   https://en.wikipedia.org/wiki/Amharic

[LanguageMap]   The Language Map of Ethiopia, Djibouti and Eritrea, http://archive.ethnologue.com/16/show_map.asp?name=ER&seq=10

[DY]   Daniel Yacob " Application of the Double Metaphone Algorithm to Amharic Orthography", International Conference of Ethiopian Studies XV, https://pdfs.semanticscholar.org/2f71/033d74d2f17a9502867e4a43dc4374500726.pdf

[IDNAREG]   IANA Registry: "IDNA Parameters". For Unicode 6.3 available at: http://www.iana.org/assignments/idna-tables-6.3.0/idna-tables-6.3.0.xml. Visited 2013-11-20.

[RFC5892]   Faltstrom, P., Ed., "The Unicode Code Points and Internationalized Domain Names for Applications (IDNA)", RFC 5892, August 2010.

Appendix A: Corpus Analysis of Amharic Language

Corpus analysis on online contents published in Amharic language has been performed by crawling 14,850 html files with estimated size of 1.8GB. The table below lists the frequency of occurrences for code points or Ethiopic characters on Amharic contents published online.

Character	Frequency	Character	Frequency	Character	Frequency
ን	2525998	ዮ	65312	፤	4142
ት	1878630	ጭ	65025	ጂ	3908
ው	1593696	ጊ	63852	ዢ	3743
የ	1560828	ሀ	60623	ፎ	3684
በ	1493784	ዱ	60319	ፉ	3580
ለ	1396708	ኤ	59123	ጹ	3514
እ	1332877	ፓ	58636	ዖ	3510
ር	1307398	ጽ	57051	ፐ	3479
፡	1306162	ዶ	56960	ፗ	3426
መ	1260038	ዜ	56661	ፅ	3201
ል	1256820	ፏ	56451	ፄ	3183
ም	1215335	ጎ	55347	ዒ	2848
ስ	1205601	ሽ	55306	ፔ	2615
ተ	992996	ሬ	55131	ኹ	2605
ነ	971586	ፌ	53773	፜	2573
ያ	941156	ጡ	53534	ጸ	2565
ኣ	931894	ቼ	53455	ዚ	2332
ይ	922067	ጆ	50723	ዣ	2216
ና	790777	ፕ	49860	ኡ	2070
ከ	702312	ቋ	49444	ዬ	2030
ብ	694590	ሕ	47428	ፚ	2016
ች	692254	ቀ	46933	ሒ	1913
ደ	680546	ፉ	46398	፝	1837
ግ	651892	ሠ	45593	ዑ	1812
ሏ	629937	ቄ	45428	ሓ	1755
ገ	624375	ሐ	45182	ሡ	1743
ማ	600978	ቲ	44365	ኦ	1705
ድ	592919	ኝ	43251	ዬ	1603
ታ	549058	ሶ	41525	ፍ	1262
ረ	548124	ዮ	39770	ኍ	1187
ሰ	514767	ሒ	39309	ጇ	1129
ሩ	492345	ኞ	38178	ዟ	1124
ወ	491291	ኋ	37947	ዔ	1065
ባ	490711	ጂ	37092	ጬ	985
ሚ	485566	θ	36896	ሽ	827

ቀ	434698	ஊ	34684	ኽ	702
ሀ	428588	ኦ	32170	ዖ	629
ፕ	378794	ጨ	30833	፦	602
ክ	352449	ፒ	30468	ኩ	600
ጠ	347746	ኂ	30369	ፗ	590
ሰ	318148	ዠ	29294	ፀ	528
ዳ	309331	ጀ	28706	ዊ	518
ቅ	293318	ሺ	27062	ኟ	517
ፍ	288159	ኑ	23885	ዦ	472
ፈ	284733	ጌ	23773	ጼ	438
ጣ	277429	ጐ	23261	ኸ	434
፣	275560	ጦ	22476	ኢ	412
ጋ	269353	ኬ	22142	ፁ	382
ሬ	267432	ሐ	22087	ஷ	374
ሆ	261545	ን	21613	ጂ	335
ካ	258339	ፆ	20474	ዩ	325
ቻ	253836	ቀ	20268	ዦ	321
ሁ	251483	ማ	18366	ፅ	318
ፖ	243566	ዉ	18353	ዒ	305
ቱ	224860	ሲ	17347	ዔ	280
ሉ	213984	ቼ	17214	ፗ	279
ዘ	208553	ጀ	16394	ፀ	264
ኛ	205007	ሊ	16042	ፅ	231
ሲ	180435	ጸ	15144	ፚ	198
ዲ	179602	ጻ	14554	ቐ	182
፤	178170	ዜ	13405	ፗ	179
ም	175126	ኾ	12888	ጔ	177
ቶ	174296	ዚ	12662	ኁ	141
ዝ	166857	ሣ	12255	ዪ	139
ዎ	160873	ጎ	11453	ዕ	124
ር	154232	ዷ	11241	ሑ	121
ቃ	154003	ጇ	10962	ዼ	118
ሩ	139747	ஹ	10358	ጐ	108
ቢ	136466	ሷ	10134	ፀ	107
ሊ	136224	ሽ	9956	ፗ	99
ዓ	132454	ኚ	9739	ፘ	87
ዕ	130379	ኗ	9733	ኹ	87
ዮ	129131	ቡ	9716	ቆ	80
ዜ	126001	ሹ	9351	ፗ	73
ዊ	125446	ቿ	9247	ቄ	73
ኝ	124385	ሾ	8577	ፗ	65

ሽ	124019		ጬ	8260		ፀ	60
ሎ	121334		0	7614		ኺ	60
ኢ	119673		ሠ	7567		ጊ	56
ጉ	119667		ፇ	7513		ኌ	47
ዚ	119163		ኸ	6854		ጬ	42
ሌ	118895		ጢ	6133		ኋ	40
ማ	117011		ሸ	5978		..	36
ኑ	111814		ፗ	5576		ቛ	36
ቡ	110726		ጧ	5275		ቡ	36
ሬ	110136		ኗ	5133		ወ	35
ቲ	109898		ፀ	5096		ኹ	34
ቤ	107252		ፒ	4753		ኌ	33
መ	105127		ሴ	4627		ኡ	28
ጀ	102846		ሑ	4579		ሽ	24
ዛ	101619		ኽ	4489		ኹ	24
ኮ	94829					ቍ	23
ፈ	93370					ዧ	23
ሥ	92633					ቓ	22
።	91436					ኟ	20
ጅ	90464					ቛ	18
ቻ	90252					ጛ	18
ፉ	89918					ቍ	17
ጨ	88727					ኾ	15
ሱ	86808					ኺ	13
ዥ	84801					ኳ	9
ዙ	83513					ቛ	9
ያ	83226					ኗ	9
ዬ	81110					ፒ	8
ኩ	79423					ዼ	7
ቀ	79050					ቛ	5
ሳ	78901					ኟ	4
ዐ	78076					ዧ	4
ቆ	77030					ኹ	4
ሜ	76810					ሒ	3
ዮ	71598					ፗ	2
ኪ	69065					ኍ	2
ሜ	67952					ዺ	2
በ	67767					ፇ	1
ፖ	65651					ፔ	1
						ፇ	1
						ዺ	1

Appendix B: Corpus Analysis of Tigrigna Language

Corpus analysis on online contents published in Tigrigna language has been performed by crawling 598 html and 40 PDF files. The table below lists the frequency of occurrences for code points or Ethiopic characters on contents published online in Tigrigna language.

Character	Frequency
ን	89629
ት	69181
ብ	59072
ም	57032
ይ	55251
እ	42824
ር	41387
ኣ	41056
ል	35958
ዝ	34841
ስ	31531
መ	31375
ተ	31180
ና	30920
ነ	26368
ላ	24768
በ	24108
ራ	22500
ው	22365
ግ	22105
ሓ	19370
ታ	19286
ክ	18792
ድ	18781
፡	18653
ቲ	16865
ረ	16693
ገ	16671
ሰ	16274
ዘ	16251
ሕ	15614
ወ	15496
ከ	14680
ሉ	14115

Character	Frequency
ጆ	1919
ኦ	1915
ሆ	1915
ቃ	1868
ሊ	1779
ሒ	1751
ሐ	1620
ጀ	1585
ሞ	1539
ሬ	1531
ቄ	1463
ዮ	1447
ሌ	1356
ጥ	1321
ፉ	1313
ሻ	1309
ዦ	1265
ሁ	1261
ቆ	1260
ቸ	1248
ቀ	1240
ኪ	1233
ጊ	1168
ዙ	1098
ሽ	1078
ዑ	998
ቤ	976
ዚ	974
ዉ	944
ጅ	870
ዣ	868
ጢ	861
ኺ	824
ዐ	815

Character	Frequency
ጸ	26
ፑ	25
ፒ	24
ዬ	24
ቹ	24
ጌ	24
ጧ	23
ሹ	23
ሏ	22
ጿ	21
ኆ	20
ፕ	19
ዒ	19
ዟ	16
ጄ	14
ፏ	14
ዢ	14
ቁ	13
ፌ	13
ዔ	13
ኸ	12
ኹ	11
ጉ	10
ፉ	10
ኀ	10
ዠ	9
ዷ	8
ዿ	8
ዼ	7
ኩ	7
ዯ	7
ዺ	6
ፖ	6
ፔ	6

ለ	14006	ሶ	803	ኺ	6
ማ	13953	ቴ	802	ዚ	6
ደ	13839	ጡ	786	ኊ	5
ዕ	13590	ፊ	786	ዿ	5
ህ	13509	ህ	777	ቍ	5
የ	13072	ጸ	750	ጒ	4
ካ	12720	ፉ	733	ኺ	3
ፍ	12561	ጇ	724	ኤ	3
ያ	12229	ሜ	723	ጂ	3
ጋ	11576	ቀ	714	ጒ	3
ሉ	10759	ቐ	713	ሻ	3
።	10376	ኔ	666	ሮ	3
ኢ	9531	ኸ	650	ዢ	3
ዮ	8966	ጨ	621	ዷ	2
ዳ	8690	ኝ	614	ዺ	2
ዊ	8122	ቻ	611	ሧ	2
ሬ	8051	ጠ	610	ዧ	2
አ	7668	ጄ	555	ሯ	2
ፕ	7642	ፓ	534	ዎ	2
ሃ	7568	ኺ	487	ጄ	1
ሬ	7340	ጌ	474	ኤ	1
ሎ	7261	ጓ	452	ዷ	1
ዕ	7238	ጸ	425	ዃ	1
ጋ	7230	ሐ	424	ዻ	1
ዋ	6780	ጊ	418	፤	1
ቀ	6774	ቖ	381	ዿ	1
ሰ	6622	ሴ	359	ዿ	1
ዜ	6600	ሥ	357	ቍ	1
ሊ	6187	ጨ	354	፟	1
ቢ	6130	ኼ	322	ዿ	1
ሩ	6103	፦	320	ሮ	1
፣	5974	ቔ	297		
ዮ	5946	ቒ	274		
ሚ	5929	ዖ	270		
ክ	5761	ጎ	262		
ቶ	5377	ቧ	261		
ዲ	5294	ዖ	254		
ጠ	5174	ሔ	221		
ዋ	5171	ቪ	209		
ኮ	5086	ኺ	207		
ፀ	4892	ሹ	207		

ሽ	4674	ዴ	206
ር	4560	ሷ	194
ክ	4526	ዱ	182
ቆ	4464	ፆ	152
ኑ	4302	ቄ	131
ቃ	4301	ኝ	131
ቅ	4265	ሠ	130
ኻ	4145	ዮ	127
ፀ	4003	ኩ	123
ሲ	3972	ኬ	113
ኮ	3930	ጤ	104
ኤ	3847	ፄ	100
ቱ	3740	ኙ	98
ጉ	3729	ኺ	84
ኡ	3508	ቲ	83
ጊ	3309	ቭ	78
አ	3265	ጪ	76
ዘ	3120	ጎ	74
ም	3103	ፒ	72
ች	3039	ፚ	70
ቡ	2863	ጸ	70
ኽ	2791	ኂ	68
፤	2737	ኹ	60
ጣ	2681	ዦ	59
ጌ	2584	ፏ	59
ጌ	2540	ዕ	58
ሱ	2505	፥	54
በ	2503	ፐ	47
ፉ	2473	ሽ	47
ሑ	2447	ጬ	45
ፎ	2432	ቸ	45
ዊ	2432	ፈ	40
ሙ	2178	ቫ	38
ሬ	2168	ጕ	36
ዙ	2141	ማ	34
ዱ	2108	ፙ	33
ኝ	1989	ፖ	32
ጎ	1969	ዬ	32
ኹ	1964	ሣ	29
ጽ	1951	ሴ	29
ዶ	1926	ሼ	27

Appendix C: Homonym List

Translation of [HomonymList] headings from its first page.

The pair of words which are sound alike are presented in a row with a space separating each row containing pair of sound alike words. In the first column sound alike pair of words are presented, in the second column each words meaning in Ge'ez is described

Usage of Sound alike Alphabets in Amharic Writing System					
Sound alike words	Meaning	Verb root in Ge'ez	Verb root in Amharic	Common Nouns	Adjective

በዐማርኛ ድምፀ-ሞክሼ ቃላት
ውስጥ የድምፀ-ሞክሼ ሆሄያት አጠቃቀም መደበል

ድምፀ-ሞክሼ ቃላት	ፍቺዎች	ግእዝ ስም ፤ ግስ/	በዐማርኛ /ግስ ፥ ስም	የተጸውኦ/የነገር ስሞች	ቅጽሎች እና ሌሎች ምሥርቾች
ሐልዮ	ዝማሬ	ሐልዬ ፤ ሐለየ/		ማሕሌት	መሐልል ፥ ማሕሊታይ?
ጎልዮ	ንድፈ-ሐሳብ ፤ ዕሳቤ	ጎለየ/		ጎልዮት	
ሕብር	ቀለም ፤ መልክ	ሕብር/		ሕብሬ ፤ ሕብረ-ሰማይ	ሕብራዊ ፥ ሕብራዊ
ጎብር	ትብብር	ጎቢር ፤ ጎበረ/		ጎብረት ፤ ማጎበር	
ጎብር	ጥምር ፍቺ ያለው ቃል (ሐረግ)	ጎቢር ፤ ጎበረ/		ጎብር ቅኔ	
ሐየ	አሐየ		/አሐየ		
ኸየ	ኹለት ጊዜ በሥር		/ኸየ		ኸያኛ
ምሁር	በዕውቀት የጎለበተ ፥ ሊቅ	ምሂር ፤ መሀረ/		መምሀር ፤ ትምህርት	ምሁራዊ
ምሑር	ይቅር የተባለ ፤ ምሕረት ያገኘ	መሐረ ፤ መሐረ/		መሐሪ ፤ ምሕረት	
መርሕብ	ስፊ ሜዳ ፤ አውላላ	ርሒብ ፤ ርሕብ/			ርኁብ ፤ ራብተኛ
መርጎብ	አሰራቢ ፤ የሚያስርብ	ርጊብ ፤ ርጎብ/	/ሪብ ፤ ራብ	ርጓብ (ረጓብ)	
መሣሣት	መቅጠን ፤ ዘርዝራ ወይም ሥሥ መኾን	ሠሥዓ ፤ ሠሥዐ/	/ሣሣ		ሥሥ
መሣሣት	በሥሥዓት ወይም በዋንቃቄ መያዝ	ሠሥዓ ፤ ሠሥዐ/	/ሣሣ	ሥሣት	ሥሡ ፥ ሠማች
መሳሳት	ትክክል አለመኾን ፤ አቅጣጫን አለማወቅ	ስሒት ፤ ስሕት/	/ሳተ	ስሕተት	ሳች ፤ አሳቻ
መንጎል	መንቀያ ፤ ከሥር መመንገያ	ነጊል ፤ ነገለ/	/ነገለ		ነገይ
መንጎል	ወፍራም ንፍጥ		/ነጐለ	ነጐል	ነጐሊ ፥ መንጐላም
መንጐል	የቆዳ ፥ ወረቀት ፥ ወዘተ. መብሻ		/መንጐል		
መአት	እጅግ ብዙ	ምእት ፤ መአተ/		ምእት ዓመት	
መዓት	ቁጣ	መዐተ ፥ መዓት/			መዓተኛ
ሣለ	ምሥልን (ሥራ) ነደፈ	ሥዔል ፤ ሠዐለ/		ሥዕል	ሠዓሊ ፤ ሣይ
ሳለ	ጠየቀ ፤ ለመነ	ስኢል ፤ ስአለ/		ስእለት	ስኣሊ ፥ ሳይ
ሳለ	ኡህ ኡህ አለ	ስዓል ፤ ሰዐለ/		ሳል	ሳይ
ሳለ	መቁረጫ ጠረዝን አተባ	በሒል ፤ በሐለ/		በለት	በል ፥ በሐለ