Proposal to encode ADLAM NASALIZATION MARK
for ADLaM script

Abdoulaye Barry, Ibrahima Barry, Peter Constable, Andrew Glass

September 17, 2018

Summary

ADLaM script was added to Unicode in version 9.0 (2016) based on Michael Everson's proposal documents submitted in 2013 and 2014 (L2/13-191 & L2/14-219R).The present document addresses a gap in the ADLaM encoding by proposing a new mark character as an addition to the existing ADLaM block:

Glyph	Code	Character name
'	1E94B	ADLAM NASALIZATION MARK

Usage requirements

ADLaM script uses an apostrophe-like sign called Nyondal to nasalize a succeeding consonant. For example:

ᠣᠥᠯᠡᠥᠥ	/hin.du/	'ancient'
ᠣᠥᠯᠡᠥᠥ	/hi.^{n}du/	'it'

This sign was known at the time the original proposals for ADLaM were submitted, but at that time, the use of European punctuation signs at their standard code points was expected to suffice (Everson 2013: 2, Everson 2014: 2). While implementing and using the ADLaM Unicode encoding, it has become clear that using ' (U+0027) or similar signs to represent ADLaM Nyondal does not work. Specific typographical features of Nyondal make it necessary to encode a distinct sign. These requirements are:

1. ADLaM Nyondal is a letter modifier not a punctuation sign and occurs in word internal position

2. As a letter, ADLaM Nyondal should have strong right-to-left directionality

3. Typographically, ADLaM Nyondal is straight, not hooked

4. ADLaM Nyondal is transparent to joining behavior, i.e., adjacent letters connect under it, but it doesn't cause joining by itself.

Comparable signs

Alternative characters already encoded in Unicode have been considered and rejected as follows:

- U+0027 APOSTROPHE

This character is functionally a punctuation character. It is subject to being replaced by a smart quote (U+2018 and U+2019) in many applications resulting in inconsistent spelling across applications. Not all

applications have the option to disable smart quotes and many users are not aware that the feature is enabled. Many applications give the apostrophe special logic in text-selection, etc. When used with joining scripts, it blocks joining behavior between adjacent signs.

- U+02BC MODIFIER LETTER APOSTROPHE

The default form is curved to work with LTR scripts. This character is used in several Latin-based orthography for African languages. As a result, modifying the design for ADLaM specifically will cause inconsistent appearance in mixed, script/language fonts. It blocks joining between adjacent signs. It has a strong-left directionality.

- U+A78B LATIN CAPITAL LETTER SALTILLO

- U+A78C LATIN SMALL LETTER SALTILLO

The default form of these characters are straight; but they have a strong-left directionality and will create the same inconsistent text ordering issues as U+02BC.

- U+05F3 HEBREW PUNCTUATION GERESH

The default form is not vertical, and it blocks joining behavior. Though it is strong RTL, it's association with Hebrew script will cause complications with text itemization and font fallback behaviors.

Therefore, the right approach to encoding this sign is to encode as a dedicated character, ADLAM NASALIZATION MARK. The suggestion here is to encode it at the first unassigned code point in the ADLAM block.

Joining behavior

ADLAM Nyondal is spacing and transparent to joining. That is, it has the following characteristics:

Position	With Nyondal	Without Nyondal
Initial, before ADLAM letter		
Medial position, between ADLAM letters		
Final position, after ADLAM letter		
Isolated position, not adjacent to ADLAM letters		

Properties

It should have the following properties:

Character properties In the format of UnicodeData.txt:

```
1E94B;ADLAM NASALIZATION MARK;Lm;0;R;;;;;N;;;;;
```

Linebreaking In the format of LineBreak.txt:

```
1E94B;AL  # Lm   [1] ADLAM NASALIZATION MARK
```

Joining categories In the format of ArabicShaping.txt:

```
1E94B; ADLAM NASALIZATION MARK; T; No_Joining_Group
```

References

Everson, Michael, 2013 "Preliminary proposal for encoding the Adlam script in the SMP of the UCS". N4488 L2/13-191. http://www.unicode.org/L2/L2013/13191-n4488-adlam.pdf

Everson, Michael, 2014. "Revised proposal for encoding the Adlam script in the SMP of the UCS". N4628R L2/14-219R. http://www.unicode.org/L2/L2014/14219r-n4628-adlam.pdf

	1E90	1E91	1E92	1E93	1E94	1E95
0	ᤐ 1E900	ᤑ 1E910	ᤒ 1E920	ᤓ 1E930	ᤔ 1E940	ᤕ 1E950
1	1E901	1E911	1E921	1E931	1E941	1E951
2	1E902	1E912	1E922	1E932	1E942	1E952
3	1E903	1E913	1E923	1E933	1E943	1E953
4	1E904	1E914	1E924	1E934	1E944	1E954
5	1E905	1E915	1E925	1E935	1E945	1E955
6	1E906	1E916	1E926	1E936	1E946	1E956
7	1E907	1E917	1E927	1E937	1E947	1E957
8	1E908	1E918	1E928	1E938	1E948	1E958
9	1E909	1E919	1E929	1E939	1E949	1E959
A	1E90A	1E91A	1E92A	1E93A	1E94A	
B	1E90B	1E91B	1E92B	1E93B	1E94B	
C	1E90C	1E91C	1E92C	1E93C		
D	1E90D	1E91D	1E92D	1E93D		
E	1E90E	1E91E	1E92E	1E93E		1E95E
F	1E90F	1E91F	1E92F	1E93F		1E95F

Capital letters

1E900 ხ ADLAM CAPITAL LETTER ALIF
1E901 �映 ADLAM CAPITAL LETTER DAALI
1E902 ᑕ ADLAM CAPITAL LETTER LAAM
1E903 ᑫ ADLAM CAPITAL LETTER MIIM
1E904 Ᏸ ADLAM CAPITAL LETTER BA
1E905 ɰ ADLAM CAPITAL LETTER SINNYIIYHE
1E906 ᑲ ADLAM CAPITAL LETTER PE
1E907 Я ADLAM CAPITAL LETTER BHE
1E908 Ⴑ ADLAM CAPITAL LETTER RA
1E909 ᘔ ADLAM CAPITAL LETTER E
1E90A ᕲ ADLAM CAPITAL LETTER FA
1E90B U ADLAM CAPITAL LETTER I
1E90C ᑎ ADLAM CAPITAL LETTER O
1E90D ᕼ ADLAM CAPITAL LETTER DHA
1E90E ᖾ ADLAM CAPITAL LETTER YHE
1E90F ᖴ ADLAM CAPITAL LETTER WAW
1E910 ᖳ ADLAM CAPITAL LETTER NUN
1E911 Ꮞ ADLAM CAPITAL LETTER KAF
1E912 З ADLAM CAPITAL LETTER YA
1E913 ᘓ ADLAM CAPITAL LETTER U
1E914 ᗯ ADLAM CAPITAL LETTER JIIM
1E915 ᖲ ADLAM CAPITAL LETTER CHI
1E916 ᑐ ADLAM CAPITAL LETTER HA
1E917 ᖱ ADLAM CAPITAL LETTER QAAF
1E918 ᖴ ADLAM CAPITAL LETTER GA
1E919 Ꮥ ADLAM CAPITAL LETTER NYA
1E91A ᕰ ADLAM CAPITAL LETTER TU
1E91B Ꮳ ADLAM CAPITAL LETTER NHA

Supplementary capital letters

1E91C Ᏼ ADLAM CAPITAL LETTER VA
1E91D ᖷ ADLAM CAPITAL LETTER KHA
1E91E Ꮞ ADLAM CAPITAL LETTER GBE
1E91F Ꮞ ADLAM CAPITAL LETTER ZAL
1E920 ᗤ ADLAM CAPITAL LETTER KPO
1E921 ᖵ ADLAM CAPITAL LETTER SHA

Small letters

1E922 ʊ ADLAM SMALL LETTER ALIF
1E923 ᴨ ADLAM SMALL LETTER DAALI
1E924 ᴐ ADLAM SMALL LETTER LAAM
1E925 ᴢ ADLAM SMALL LETTER MIIM
1E926 ʮ ADLAM SMALL LETTER BA
1E927 ɰ ADLAM SMALL LETTER SINNYIIYHE
1E928 ℮ ADLAM SMALL LETTER PE
1E929 ᴃ ADLAM SMALL LETTER BHE
1E92A ʮ ADLAM SMALL LETTER RA
1E92B ᴆ ADLAM SMALL LETTER E
1E92C ᕱ ADLAM SMALL LETTER FA
1E92D ʊ ADLAM SMALL LETTER I
1E92E ᴑ ADLAM SMALL LETTER O
1E92F ᴑ ADLAM SMALL LETTER DHA
1E930 ᴄ ADLAM SMALL LETTER YHE
1E931 ᴀ ADLAM SMALL LETTER WAW
1E932 ᴀ ADLAM SMALL LETTER NUN
1E933 ᴴ ADLAM SMALL LETTER KAF
1E934 ᴈ ADLAM SMALL LETTER YA
1E935 ʊ ADLAM SMALL LETTER U
1E936 ᴨ ADLAM SMALL LETTER JIIM
1E937 ᴜ ADLAM SMALL LETTER CHI
1E938 ᴐ ADLAM SMALL LETTER HA

1E939 ᴧ ADLAM SMALL LETTER QAAF
1E93A ᴤ ADLAM SMALL LETTER GA
1E93B ᴤ ADLAM SMALL LETTER NYA
1E93C ᴫ ADLAM SMALL LETTER TU
1E93D ᴂ ADLAM SMALL LETTER NHA

Supplementary small letters

1E93E ᴀ ADLAM SMALL LETTER VA
1E93F ℮ ADLAM SMALL LETTER KHA
1E940 ᴤ ADLAM SMALL LETTER GBE
1E941 ᴤ ADLAM SMALL LETTER ZAL
1E942 ℮ ADLAM SMALL LETTER KPO
1E943 ℮ ADLAM SMALL LETTER SHA

Diacritical marks

1E944 $\tilde{\circ}$ ADLAM ALIF LENGTHENER
1E945 $\circ$ ADLAM VOWEL LENGTHENER
1E946 $\circ$ ADLAM GEMINATION MARK
1E947 $\circ$ ADLAM HAMZA
1E948 $\circ$ ADLAM CONSONANT MODIFIER
1E949 $\circ$ ADLAM GEMINATE CONSONANT MODIFIER
1E94A $\circ$ ADLAM NUKTA
       • called hoortobbhere when drawn above and
         lestobbhere when drawn below the base
         character
1E94B $\acute{}$ **ADLAM NASALIZATION MARK**
       = nyondal

Digits

1E950 Ꝺ ADLAM DIGIT ZERO
1E951 Ι ADLAM DIGIT ONE
1E952 ᴤ ADLAM DIGIT TWO
1E953 ᴥ ADLAM DIGIT THREE
1E954 ᴌ ADLAM DIGIT FOUR
1E955 ᴄ ADLAM DIGIT FIVE
1E956 ᴃ ADLAM DIGIT SIX
1E957 Ꭱ ADLAM DIGIT SEVEN
1E958 Ꭰ ADLAM DIGIT EIGHT
1E959 Ꝑ ADLAM DIGIT NINE

Punctuation

1E95E ᾱ ADLAM INITIAL EXCLAMATION MARK
1E95F ᾱ ADLAM INITIAL QUESTION MARK

n'der

n'gal

n'gol

n'gol

Fig. 1. Excerpt from booklet in Fulani discussing democracy, entitled *Dɛmokarāku* (Everson 2014: 14). Highlighted examples show ADLaM Nyondal with correct joining behavior.

Fig. 2. Story in Fulani of "the laborer and his children". Examples of ADLaM Nyondal in red.

Fig. 3. Sample of ADLaM Nyondal with defective rendering. In this example, the expected joining behavior is not possible due to the use of standard apostrophe that blocks joining.