Proposal to encode ADLAM LETTER APOSTROPHE
for ADLaM script

Abdoulaye Barry, Ibrahima Barry, Peter Constable, Andrew Glass

August 26, 2018

ADLaM script was added to Unicode in version 9.0 (2016) on the basis of 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 an apostrophe sign as an addition to the existing ADLaM block:

Glyph	Code	Character name
'	1E94B	ADLAM MODIFIER LETTER APOSTROPHE

ADLaM script uses the apostrophe sign to nasalize a succeeding consonant when used in conjunction with the ADLaM letter Nun. For example:

ᰀᰤᰀᰨᰵ	hindu	'ancient'
ᰀᰤᰀ'ᰨᰵ	hin'du	'it'

The original proposals called out that ADLaM uses European punctuation signs at their standard code points (Everson 2013: 2, Everson 2014: 2). The earlier proposals do not mention apostrophe explicitly, but examples of ADLaM apostrophe do occur in the figures (see below fig. 1). The expectation that ADLAM apostrophe could be represented with ' (U+0027) or similar signs has turned out to false. Specific typographical features of this sign make it necessary to encode it as a distinct sign in the ADLAM block. These requirements are:

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

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

3. Typographically, ADLaM apostrophe is straight, not hooked

4. ADLaM apostrophe is transparent to joining behavior, i.e., adjacent letters connect under it

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 RTL 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 MODIFIER LETTER APOSTROPHE. The suggestion here is to encode it at the first unassigned code point in the ADLAM block. It should have the following properties:

Character properties In the format of UnicodeData.txt:

1E94B;ADLAM MODIFIER LETTER APOSTROPHE;Lm;0;R;;;;;N;;;;;

Linebreaking In the format of LineBreak.txt:

1E94B;AL   # Lm   [1] ADLAM MODIFIER LETTER APOSTROPHE

Joining categories In the format of ArabicShaping.txt:

1E94B; ADLAM MODIFIER LETTER APOSTROPHE; 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	ხ 1E900	a 1E910	Ɗ 1E920	ⲧ 1E930	ꭹ 1E940	Ọ 1E950
1	Ч 1E901	Ꮧ 1E911	Ꮡ 1E921	ꞁ 1E931	ꭹ 1E941	ı 1E951
2	ꓳ 1E902	Ӡ 1E912	Ꮙ 1E922	ꞵ 1E932	ℚ 1E942	ꭰ 1E952
3	ꝯ 1E903	Ø 1E913	Ꮒ 1E923	Ⱶ 1E933	ꮃ 1E943	Ꮵ 1E953
4	Ꮒ 1E904	Ꮊ 1E914	ꓳ 1E924	з 1E934	◌̃ 1E944	ꮯ 1E954
5	Ш 1E905	ꮳ 1E915	ꙅ 1E925	⦰ 1E935	◌̆ 1E945	ꮝ 1E955
6	Ꭰ 1E906	Ꮄ 1E916	ꞃ 1E926	ꙃ 1E936	◌̄ 1E946	Ꞃ 1E956
7	Я 1E907	Ꮬ 1E917	ꙍ 1E927	ꝷ 1E937	◌̒ 1E947	Ꮞ 1E957
8	Ꮷ 1E908	ꭹ 1E918	ꝺ 1E928	Ɔ 1E938	◌̌ 1E948	ꭺ 1E958
9	Ꞵ 1E909	ꝸ 1E919	ꞡ 1E929	Ꝺ 1E939	◌̋ 1E949	ꝑ 1E959
A	Ꞝ 1E90A	ꟽ 1E91A	ꭼ 1E92A	ꭹ 1E93A	◌́ 1E94A	
B	∪ 1E90B	Ꮲ 1E91B	ꞵ 1E92B	ꞩ 1E93B	◌̓ 1E94B	
C	Ⲥ 1E90C	Ꮇ 1E91C	ꞓ 1E92C	ꭵ 1E93C		
D	Ꮵ 1E90D	Ꭾ 1E91D	∪ 1E92D	Ꮢ 1E93D		
E	ꝣ 1E90E	ꭹ 1E91E	ꝯ 1E92E	Ꮧ 1E93E		Ꝥ 1E95E
F	Ꮱ 1E90F	Ꮒ 1E91F	ꝺ 1E92F	ꭹ 1E93F		Ꮉ 1E95F

Capital letters

1E900	ꦧ	ADLAM CAPITAL LETTER ALIF
1E901	Ꙋ	ADLAM CAPITAL LETTER DAALI
1E902	Ꙅ	ADLAM CAPITAL LETTER LAAM
1E903	Ꙃ	ADLAM CAPITAL LETTER MIIM
1E904	Ꝥ	ADLAM CAPITAL LETTER BA
1E905	Ш	ADLAM CAPITAL LETTER SINNYIIYHE
1E906	Ꝑ	ADLAM CAPITAL LETTER PE
1E907	Я	ADLAM CAPITAL LETTER BHE
1E908	Ꝯ	ADLAM CAPITAL LETTER RA
1E909	ß	ADLAM CAPITAL LETTER E
1E90A	Ꮆ	ADLAM CAPITAL LETTER FA
1E90B	U	ADLAM CAPITAL LETTER I
1E90C	Ꙅ	ADLAM CAPITAL LETTER O
1E90D	Ꮟ	ADLAM CAPITAL LETTER DHA
1E90E	Ꮜ	ADLAM CAPITAL LETTER YHE
1E90F	Ꝑ	ADLAM CAPITAL LETTER WAW
1E910	Ꙉ	ADLAM CAPITAL LETTER NUN
1E911	ꟷ	ADLAM CAPITAL LETTER KAF
1E912	Ꝫ	ADLAM CAPITAL LETTER YA
1E913	Ꙇ	ADLAM CAPITAL LETTER U
1E914	Ꙍ	ADLAM CAPITAL LETTER JIIM
1E915	Ꙃ	ADLAM CAPITAL LETTER CHI
1E916	Ꙏ	ADLAM CAPITAL LETTER HA
1E917	Ꙛ	ADLAM CAPITAL LETTER QAAF
1E918	Ꙋ	ADLAM CAPITAL LETTER GA
1E919	Ꙅ	ADLAM CAPITAL LETTER NYA
1E91A	Ꙇ	ADLAM CAPITAL LETTER TU
1E91B	�	ADLAM CAPITAL LETTER NHA

Supplementary capital letters

1E91C	Ꙙ	ADLAM CAPITAL LETTER VA
1E91D	Ꙝ	ADLAM CAPITAL LETTER KHA
1E91E	Ꙅ	ADLAM CAPITAL LETTER GBE
1E91F	Ꙋ	ADLAM CAPITAL LETTER ZAL
1E920	Ꙝ	ADLAM CAPITAL LETTER KPO
1E921	Ꝯ	ADLAM CAPITAL LETTER SHA

Small letters

1E922	℧	ADLAM SMALL LETTER ALIF
1E923	ᴉ	ADLAM SMALL LETTER DAALI
1E924	ɔ	ADLAM SMALL LETTER LAAM
1E925	ƨ	ADLAM SMALL LETTER MIIM
1E926	ɪ	ADLAM SMALL LETTER BA
1E927	ɰ	ADLAM SMALL LETTER SINNYIIYHE
1E928	ℓ	ADLAM SMALL LETTER PE
1E929	ꙅ	ADLAM SMALL LETTER BHE
1E92A	ʁ	ADLAM SMALL LETTER RA
1E92B	ɞ	ADLAM SMALL LETTER E
1E92C	ɞ	ADLAM SMALL LETTER FA
1E92D	℧	ADLAM SMALL LETTER I
1E92E	℧	ADLAM SMALL LETTER O
1E92F	℧	ADLAM SMALL LETTER DHA
1E930	℧	ADLAM SMALL LETTER YHE
1E931	ℓ	ADLAM SMALL LETTER WAW
1E932	℧	ADLAM SMALL LETTER NUN
1E933	ꟷ	ADLAM SMALL LETTER KAF
1E934	℥	ADLAM SMALL LETTER YA
1E935	℧	ADLAM SMALL LETTER U
1E936	℧	ADLAM SMALL LETTER JIIM
1E937	℧	ADLAM SMALL LETTER CHI
1E938	℧	ADLAM SMALL LETTER HA

(continued)

1E939	ℊ	ADLAM SMALL LETTER QAAF
1E93A	℧	ADLAM SMALL LETTER GA
1E93B	℧	ADLAM SMALL LETTER NYA
1E93C	℧	ADLAM SMALL LETTER TU
1E93D	℧	ADLAM SMALL LETTER NHA

Supplementary small letters

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

Diacritical marks

1E944	◌	ADLAM ALIF LENGTHENER
1E945	◌	ADLAM VOWEL LENGTHENER
1E946	◌	ADLAM GEMINATION MARK
1E947	◌	ADLAM HAMZA
1E948	◌	ADLAM CONSONANT MODIFIER
1E949	◌	ADLAM GEMINATE CONSONANT MODIFIER
1E94A	◌	ADLAM NUKTA

• called hoortobbhere when drawn above and lestobbhere when drawn below the base character

1E94A	◌	ADLAM NUKTA
1E94B	'	ADLAM MODIFIER LETTER APOSTROPHE

Digits

1E950	𝟶	ADLAM DIGIT ZERO
1E951	𝟷	ADLAM DIGIT ONE
1E952	𝟸	ADLAM DIGIT TWO
1E953	𝟹	ADLAM DIGIT THREE
1E954	𝟺	ADLAM DIGIT FOUR
1E955	𝟻	ADLAM DIGIT FIVE
1E956	𝟼	ADLAM DIGIT SIX
1E957	𝟽	ADLAM DIGIT SEVEN
1E958	𝟾	ADLAM DIGIT EIGHT
1E959	𝟿	ADLAM DIGIT NINE

Punctuation

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

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

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

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