

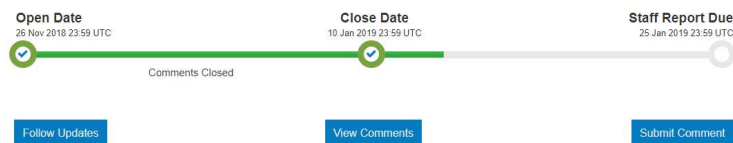
Dear all,
Greetings from Bangladesh. I would like to share my addendum and opinion on minutes of Bangla NBGP meeting at Delhi IHC dated 21-22 December 2018.
Regards
Mamun Or Rashid
Language Technology Specialist, Bangladesh Computer Council (EBLICT Project)
& Assistant Professor, Department of Bangla, Jahangirnagar University, Dhaka.

Day 2: point 4: Is MSR barrier for Nukta deletion?

Mr Mamun Or Rashid suggested the exclusion of *nukta* from the Bangla repertoire in the LGR document. He elaborated that when single stroke typing is already provided by the Unicode then there why the use of ঠ + ঙ and ঠ + ঙ. Mr. Akshat elaborately described **how nukta is relevant to the Bangla** repertoire keeping in mind the **limitations given by MSR**.

My comment: MSR is not constant like Bible or religious text, It is human generated rules. You know MSR 4 is going to be published. It is high time to includes our proposal (deletion of Nukta and include 3 characters) in MSR version 4. The MSR 4 is open now for staff reporting. [It's a matter of happiness that Mr Sarmad Hussain is key person/ staff contact of MSR documents as per website. Hope he will guide us and take necessary action to consider our proposal and stands.]

Maximal Starting Repertoire Version 4 (MSR-4) for Root Zone Label
Generation Rules (RZ-LGR)



Here I am enclosing a screenshot from MSR version 4 XML file which contains Nukta for Bangla. I would like to request you to change the rules regarding Bangla script as per meeting minutes. Here 09BC is stands for Nukta.

```
<range first-cp="097E" last-cp="097F" tag="sc:Deva" ref="8"/>
<range first-cp="0981" last-cp="0983" tag="sc:Beng" ref="0"/>
<range first-cp="0985" last-cp="098C" tag="sc:Beng" ref="0"/>
<range first-cp="098F" last-cp="0990" tag="sc:Beng" ref="0"/>
<range first-cp="0993" last-cp="09A8" tag="sc:Beng" ref="0"/>
<range first-cp="09AA" last-cp="09B0" tag="sc:Beng" ref="0"/>
<char cp="09B2" tag="sc:Beng" ref="0"/>
<range first-cp="09B6" last-cp="09B9" tag="sc:Beng" ref="0"/>
<char cp="09BC" tag="sc:Beng" ref="0"/>
<range first-cp="09BE" last-cp="09C4" tag="sc:Beng" ref="0"/>
<range first-cp="09C7" last-cp="09C8" tag="sc:Beng" ref="0"/>
<range first-cp="09CB" last-cp="09CD" tag="sc:Beng" ref="0"/>
<char cp="09CE" tag="sc:Beng" ref="7"/>
<char cp="09D7" tag="sc:Beng" ref="0"/>
<range first-cp="09F0" last-cp="09F1" tag="sc:Beng" ref="0"/>
<char cp="0A02" tag="sc:Guru" ref="0"/>
```

Day 2: point 4: Note of dissent issue

Dr Ajay Data emphasized that the Bangla LGR must go and there was no need or question of holding or delaying the task. Efforts may be made to make a note of this disagreements and attach with the proposal.

My comment: I think giving a note of dissent is not a better solution. But Bangladesh (also West Bengal even all Bangla speakers) always believe in inclusion of three characters ড, ঢ and ষ as atomic character in MSR as well as IDN. We cannot go forward with a broken orthographic pattern.

Day 2: History Part

[Not mentioned in minutes, though I propose in the meeting. Professor U N Singh Sir agreed with me. Please consider these points.]

3.1. Written Bangla

The 'Bangla alphabet' (বাংলালিপি - Bānglā lipi, ISO 15924) is derived from the Brāhmī writing system, which is related to the Devanāgarī script [108] as well as to Tirhutā writing system [106]. Considered to be fifth most widely used writing system in the world, this combined Bangla-Asamiyā-Manipuri Script (showing some variations for Asamiyā and Meitei or Bishnupriya Manipuri) (130), was used in the eastern Indian Sanskrit manuscripts too. It was once used also for Bodo and Santali as well both of which officially use Devanāgarī now. For Chakma in India and Bangladesh and for Kokborok in Tripura, it was and still is one of the scripts used. A close variant, called *Tirhutā* (123; now available also in UNICODE 10.0 as 11480 114DF; See 110) or *Mithilākṣara* was used for Maithili from the 14th Century until the early-20th century [106]. Some varieties of Bānglā lipi were also written in a system that derived from 'Nāgarī' lipi but showed a difference from both Devanāgarī and Bangla-Asamiyā-Manipuriscripts.

Comment [HN1]: This information are quoted form RC Majumder though he is historian not linguist.

Most of paleographer and epigraphist think both Bangla and Devanagari are derived from Brhami.

More, Bangla is not derived from Devanagari. They could be same generation script. They are sister in relation, not mother-daughter.

Better if we delete this divisive line.

Comment [HN2]: Not important information for Bangla language history.

Comment [HN3]: Controversial opinion. Should delete or rewrite this.

Day 2: Additional Explanation on Nukta

3.3.6 Nukta (◌̣ - U+09BC)

The nukta sign is placed below a certain number of consonants to represent sounds found only in words borrowed from Perso-Arabic. It is predominantly used in this manner in Bodo, Hindi, Kashmiri, Maithili, Santali and Sindhi. In Bangla, its use is further restricted. It can be optionally adjoined to ক KA (U+0995), খ KHA (U+0996), গ GA (U+0997), জ JA (U+099C) and ফ PHA (U+09AB) to show that words having these consonants with a nukta are to be pronounced in the Perso-Arabic style. e.g. ফিরোজ/firoz/ (U+09AB U+09BC U+09BF U+09B0 U+09CB U+090C U+09BC).

Comment [HN4]: First of all nukta is not a traditional character of Bangla script, alphabet or language. No grammarian and pundits include nukta as a letter or diacritics, even nobody declare the existence and uses of Nukta.

No manuscript/ puthi where we found with.

Pronunciation and phonetic feature of Loan word/ foreign words are very low frequent word. We could ignore these in domain name/IDN.

The domain name should follow common Bangla word like ফিরোজ, not ফিরোজ/

Comment [HN5]: This three characters have separate, single and independent (not conjunct) code block as 09DC for ড, 09DD for ঢ, 09DF for ষ.

Nukta based system is an alternative (as per Unicode chart). It create redundancy and variant problems.

We propose to decline nukta based writing in Bangla Domain.

As there are provisions made in the UNICODE character table for Bangla, it is strongly recommended that ষ YYA (U+09DF), ড RRA (U+09DC) and ঢ RRHA (U+09DD) are to be used in Bangla LGR instead of "ষ" YA (U+09AF)+ "◌̣" Nukta (U+09BC), "ড" DDA (U+09A1)+ "◌̣" Nukta (U+09BC) and "ঢ" DDHA (U+09A2)+ "◌̣" Nukta (U+09BC).