

Face to Face Neo-Brahmi Generation Panel (GP) Meeting Notes from meeting on 23-25 April 2018

Meeting Attendees (in alphabetical order)

GP members:

1. Ajay Data
2. Akshat Joshi
3. Chitrita Chatterjee
4. Dhanalakshmi Karanahalli Thyagaraj
5. Gangadhar Panday
6. Gurpreet Singh Lehal
7. Jay Paudyal
8. Khan Atiur Rahman
9. Kuldeep Patnaik
10. Rajiv Kumar
11. Neha Gupta
12. Shanmugam Rajabhathar
13. Shanmugaraja
14. Udaya Narayana Singh
15. U.B Pavanaja
16. Uma Maheshwar G
17. Veena Solomon
18. S.Maniam

Staff:

19. Akanksha
20. Pitinan Kooarmornpatana
21. Samiran Gupta
22. Sarmad Hussain

Meeting Notes (23 April 2018)

1. Oriya Script

- a. 0B57 ୱ 'ORIYA AU LENGTH MARK' (U+0B57) need to be reconsidered as it is not used but it is there for the rendering propose. The common used is already covered by “ୱ” U+0B4C Oriya Vowel Sign Au
- b. Cross-script variant Oriya-Bengali

Bengali	Oriya	Type	NBGP decision
ঐ U+0993	ୱ U+0B13	-	Yellow
ঐ U+0998	ୱ U+0B38	-	Green

c. Cross-script variant Oriya-Gujarati

Oriya	Gujarati	Type	NBGP decision
ଃ (0B03)	: (0A83)	-	Green
ଘ (0B2A)	ઘ (0A98)	-	Green
ଙ (0B25)	ઙ (0AA5)	-	Green

d. Oriya Telugu Kannada

Oriya	Telugu	Kannada	NBGP decision
ଠ (0B20)	ఠ (0C30)	ఠ (0CB0)	0B20 is not variant to these two.
ଠ (0B20)	ఠ (0C20)	ఠ (0CA0)	

However, it should be analyzed with Telugu ୠ (0C02), Kannada ୠ (0C82), Malayalam ୠ (0D02), Sinhala ୠ (0D82)

e. Dr. Rao will draft a sentence regards usage of Anusavara, to be added in the Oriya, Telugu Kannada, Malayalam LGR. So that all the scripts harmonized.

f. Rule#3 will be revised to:

Rule3: X (ୠ) must be preceded by C, N, M or V

g. Next version of .DOCX, .XML, .TXT will be ready to submit to the IP on 4 May 2018

2. Bengali Script

a. U+098C ୠ BENGALI LETTER VOCALIC L not in everyday use, will be excluded.

b. U+09D7 ୠ BENGALI AU LENGTH MARK, might be remove. Depending on the feedback from Assamese and Manipuri experts.

c. Cross-script variant Bengali – Devanagari

Bengali	Devanagari	Type	NBGP decision
ঔ U+0993	उ U+0909	-	Yellow
ঞ U+0998	ष U+0918	-	Yellow
ঠ U+09AE	म U+092E	blocked	Red
ড U+09BF	ढि U+093F	blocked	Red
ণ U+0981	ँ U+0945	-	Yellow

d. Cross-script variant Bengali – Gurmukhi

Bengali	Gurmukhi	Type	NBGP decision
ঔ U+0993	उ U+0A24	-	Green
ঞ U+0998	ष U+0A2C		Yellow
ঠ, ড U+09B6	ਮ U+0A05	-	Green
ঠ U+09AE	म U+0A38	blocked	Red

ঞ U+09AE	𑌎 U+0A2E	-	Green
ঠ U+09BF	𑌐 U+0A3F	blocked	Red
ঁ U+0981	ঁ U+0A71	-	Yellow
ঠ U+09AC and U+09BE	𑌑 U+0A17	-	Green

- e. Cross-script variant Bengali – Oriya, see 1.b.
- f. The name of the script will be “Bengali” as registered with UNICODE
- g. Next version of .DOCX, .XML, .TXT will be ready to submit to the IP on 6 May 2018

3. Tamil Script

- a. Additional Variant for “Sri” was defined. And the Type is allocatable. Rational of allocatable is required in the proposal.
- b. Rules which are derived from the input form Sri Lanka were discussed that it might be too strict. A call with expert from Sri Lanka will be set up to reconfirm.
- c. Next version of .DOCX, .XML, .TXT will be ready to submit to the IP on 4 May 2018

4. Malayalam Script

- a. Cross-script variant Malayalam – Sinhala

Malayalam	Sinhala	Type	NBGP decision
◌ 0D82	◌ 0D02	-	Not enough code variant code point
◌ 0D83	◌ 0D03	-	
◌ 0DC6	◌ 0D28	-	Green

- b. Cross-script variant Malayalam – Oriya

Malayalam	Oriya	Type	NBGP decision
◌ 0B02	◌ 0D02	blocked	Red
◌ 0B03	◌ 0D03	blocked	Red
◌ 0B20	◌ 0D20	blocked	Red

- c. WLE rules will be revised.
 - i. Edit rule#5 to:
Rule 5: CH must **not** be preceded by **H,B,X**
 - ii. Edit rule #6 to:
Rule 6: Label does not begin with CH
 - iii. Remove rule#8

However, this is not finalized yet. Will need more feedback from community.

- d. Next version of .DOCX, .XML, .TXT will be ready to submit to the IP on 4 May 2018

Meeting Notes (24 April 2018)

5. Malayalam Script (cont.)

- a. From the revised rules yesterday, there are still rejecting 17K labels out of 60K corpus. Further analysis found that:
 - (1) The rule #6 must remain but only CH cannot be at the beginning of a label.
 - (2) In corpus there are a lot of ZWJ which is not allow in IDNA2008, therefore the failed rate is high.Action: Sarmad will update the XML and filter the test corpus and retest. The test result will be share again.

6. Devanagari Script

- a. Going through the LGR document as a sample of other scripts. All proposal should complete this check list before then next submission.

Check List for all Proposals

#	Action Item	Yes?
1	First page. <i>LGR Version:</i> 3.0 → LGRs are likely to be integrated into LGR3.0 <i>Date:</i> 2018-04-22 → Please use format yyyy-mm-dd <i>Document version:</i> 3.0 →Up to author <i>Authors:</i> Neo-Brahmi Generation Panel [NBGP]	
2	File name of the XML and Test Labels File must be updated at every submission The naming convention for xml = “Proposed-LGR- Lang - yyyymmdd .xml” → Lang = script e.g. Deva The naming convention for test file = “LGR- Lang -test-labels- yyyymmdd .xml” → Lang = script e.g. Deva	
3	The paragraph about the EGIDS should be the same for all scripts. Please copy-paste from Devanagari	
4	The section 4 should be the same for all scripts, except the examples from each script. Please see Devanagari	
5	All tables must have numbers and names	
6	Add the MSR table, the same way as Devanagari	
7	Use the terms, “candidate variant” for the analyzing text, once finalized it should be either: ‘variant’ or ‘confusable’ or ‘distinguishable’	
8	The variant type explanation section should be copied over from Devanagari	

Check List for Malayalam

#	Action Item	Yes?
1	Malayalam – add the eye lash (the same as Devanagari)	

b. Check list on Devanagari Proposal

#	Action Item	Yes?
1	The paragraph about the EGIDS should be the same for all - Add 'significant' usage - Level 5, not in wide spread usage	
2	3.2.1 Sanskrit Addingon social media as well..."and lot of creative writing in Sanskrit is published every year." Add a reference for this Sanskrit section	

c. Devanagari, Gurmukhi and Bangla will be sent the Next version of .DOCX, .XML, .TXT to the IP right after the meeting.

7. Gurmukhi Script

- Cross-script variant between Gurmukhi and Devanagari was agreed as reflected in Devanagari proposal.
- Devanagari, Gurmukhi and Bangla will be sent the Next version of .DOCX, .XML, .TXT to the IP right after the meeting.

8. Conjunct characters variant analysis

- It was agreed that each script will create the possible combination list for all other script to review and feedback if they found any possible variants.

9. Gujarati Script

- Cross-script variant with Oriya is all distinguishable
- 093A DEVANAGARI VOWEL SIGN OE and 0A82 \$ GUJARATI SIGN ANUSVARA were discussed and concluded that they are similar but not enough similar character to form variant labels.

10. Telugu Script

- For 0C55 ̣ TELUGU LENGTH MARK in the excluded code points table, the reason should indicate that it is not available in the general keyboard.
- The in-script variant case of Character seq. [Ca+e+u] and [Ca+o], the analysis table will be included in the proposal, but this case is not defined as variants as they WLE is not allowed.

Cross-script variant Telugu-Kannada has been discussed. The total number of cross script variant between Telugu and Kannada is 34 sets. NBGP reconfirm the following pairs:

Telugu	Kannada	Type	NBGP decision
ఝ (0C1E)	ಝ (0C9E)	blocked	Red
ఞ (0C1F)	ಞ (0C9F)	blocked	Red
ఞ (0C21)	ఞ (0CA1)	blocked	Red
ఞ (0C22)	ఞ (0CA2)	blocked	Red

- c. Cross-script variant Telugu-Sinhala has been discussed and reconfirm the following pairs:

Telugu	Sinhala	Type	NBGP decision
ఞ (0C02)	ඞ (0D82)	blocked	Red
ఞ (0C03)	ඞ (0D83)	blocked	Red
ఞ (0C30)	ඞ (0DBB)	blocked	Red

11. Kannada Script

- a. Cross-script variant Telugu – Kannada: As discussed above in 6.b.
b. Cross-script variant Telugu-Sinhala has been discussed and reconfirm the following pairs:

Kannada	Sinhala	Type	NBGP decision
ఞ (0C83)	(0D83)	blocked	Red
ఞ (0C82)	(0D82)	blocked	Red
ఞ (0CB0)	ඞ (0DBB)	blocked	Red

- c. WLE Rule1 revised to be:
H must be preceded by C

12. Next online meeting:

- a. Tuesday 1 May 2018 – Updated document submission deadline.
b. Friday 4 May 2018 – NBGP online meeting to finalize the document.