

**Title:** Proposal to Encode the Sign ANJI for Bengali  
**Source:** Anshuman Pandey (pandey@umich.edu)  
**Status:** Individual Contribution  
**Action:** For consideration by UTC and WG2  
**Date:** 2012-04-23

## 1 Introduction

This is a proposal to encode a new character in the Bengali block of the Universal Character Set (ISO/IEC 10646). This document is a minor revision of L2/11-359 and it replaces the earlier proposal. Properties of the proposed character are:

GLYPH	CODE	CHARACTER NAME	GC	CCC	BIDI	MIRRORED	LINEBREAK
৯	0980	BENGALI ANJI	Lo	0	L	N	AL

The BENGALI ANJI is proposed for encoding at U+0980. This is the unassigned code-point at the beginning of the Bengali block, which is an appropriate allocation because the character is traditionally written at the beginning of texts as an invocation and at the head of the script.

## 2 Description

The Bengali sign ৯ ANJI (আঞ্জী *āñjī* < Skt. *añji* ‘ornament’) is written as an invocation at the beginning of texts. It is also called সিদ্ধম্ *siddham*. The character represents the Sanskrit phrase সিদ্ধিরস্তু *siddhirastu* “may there be success”. In some interpretations, it is considered a symbol of the Hindu deity Gaṇeśa in his aspect as patron of knowledge, represented specifically with his *ankuśa* “goad”, which symbolizes the notion of control or direction. For instance, in his *Bengali Grammar*, W. Yates, somewhat incorrectly, writes that the ৯ is a “contraction for the word *Ganesh*” (see figure 1). The sign is attested in inscriptions, manuscripts, and printed records.

D. C. Sircar offers the following details in *Indian Epigraphy* on the origins of the ANJI sign:

In inscriptions of the Gupta age, the auspicious word *siddham* at the beginning of epigraphs was sometimes represented by a crescent-like curved stroke opening towards the left. This mark resembles the symbol or mark of interpunctuation at the end of many of Aśoka’s edicts at Kalsī. It has been pointed out that some scholars regard it as standing for the *praṇava*, i.e. the auspicious syllable *om* [...] there is some evidence to show that the symbol was read as *siddham* or *siddhir=astu* [...] The symbol was gradually modified to the shape of a stroke slanting towards lower right with its top curved or looped inwards to the right or slanting towards lower left with its top curved or looped inwards to the left. Sometimes the lower end of the stroke exhibits an ornamental downward flourish or upward bend. (1965: 94–95)

He also describes the emergence of the form used in Bengali and related scripts:

In the medieval inscriptions of Eastern India, the symbol resembles the Bengali numerical figure 7 (a vertical stroke with its top curved towards the left and ending in an inward bend) and it was used in the central parts of Bengal with the addition of a sign called *candra-bindu* or *ānunāsika* above. But, in the eastern-most areas of Bengal, the same symbol was written with the lower end curved considerably towards the right and was called *āñjī*, and the Bengali alphabet with this symbol at the head was called *āñjī-ka-kha*. (1965: 96)

Amitabha Bhattacharya describes the ANJI as “a symbol which looks like either a *khaṇḍa* ta or an elephant’s trunk hanging down” (1994: 107).

Sircar’s description of the regional styles of writing ANJI are attested in the scripts related to Bengali. The Bengali ANJI corresponds to characters such as 𑒧 U+11480 TIRHUTA ANJI, 𑒧 U+AAF2 MEITEI MAYEK ANJI, and the Newar sign *siddhi* 𑒧. There may be a corresponding character in Devanagari that is used in the transliteration of Bengali (see figure 2); however, additional research is required in order to ascertain its usage.

### 3 Character Data

**Variants** The ANJI was written differently throughout the Bengal region. In some cases a *candrabindu* was placed above the sign: 𑒧. This form 𑒧 may be represented using the sequence <𑒧 ANJI, 𑒧 BENGALI SIGN CANDRABINDU>. Other variants are to be managed through fonts.

**General Category** The character is assigned ‘gc=Lo’ because it may be combined with marks such as CANDRABINDU.

**Collation** The ANJI should be sorted at the head of the script before 𑒧 U+0985 BENGALI LETTER A.

**‘Confusable’ Data** The 𑒧 ANJI resembles 𑒧 U+09ED BENGALI DIGIT SEVEN.

**Annotation to the Names List** The following annotation is to be added to the names list:

```
@      Sign
0980   BENGALI ANJI
       = siddham, siddhirastu
       * used at the beginning of texts as an invocation
```

### 4 References

Bhattacharyya, Amitabha. 1994. “Some Aspects of the Development of the Bengali Script”. *Journal of Ancient Indian History*, vol. XIX., parts 1–2, 1989–90, p. 97–110.

Sircar, D. C. 1965. *Indian Epigraphy*. Delhi: Motilal Banarsidass.

———. 1979. *Some Epigraphical Records of the Medieval Period from Eastern India*. New Delhi: Abhinav Publications.

Yates, W. 1849. *A Bengali Grammar*. Reprinted, with improvements, from his introduction to the Bengali language. Edited by J. Wenger. Calcutta: Baptist Mission Press.

## COMPOUND LETTERS.

9

Some consonants, when combined with this form of র, assume a peculiar shape ; as,

ক্র *kra*, ত্র *tra*, ত্ত্র *ttra*, ত্ত্র *ntra*, ত্ত্র *ndra*, ত্ত্র *stra*.

The vowel উ combined with র, is thus expressed : রু *ru*.

The vowel উ combined with র, is thus expressed : রু *ru*.

So ক্র *bhru*, শ্র *shru*, দ্র *dru*, দ্রু *drú*.

Sometimes the vowel উ has this form ( ু ) ; as,

গু *gu*, শু *shu*, ন্তু *ntu*, স্তু *stu*.

Remark also হু *hu*, and হ্র *hri*, and ে (for ত্) *t*.

In addition to the above, which are very common, a few other compound letters are worthy of notice, as they are frequently used by the natives, although generally avoided in good printing.

ত্ for তু *tu*.

ক্ for কু *ku*.

ম্ for মু *mu*.

ক্র for হু *hra*.

( ̣ ) over a letter signifies that it is nasal.

গ is a contraction for the word *Ganesh*.

৳ is a contraction for the name of God.

SECTION VII.—Of the Permutation of the Letters,  
(Sandhi.)\*

The changing of the letters ই and ঐ to এ ; of উ and ঊ to ও ; of ঞ and ঞ্ to অর, is called *gun* (pron. *goon*.)

The changing of the letters অ to আ ; ই ঐ and এ to ঐ ; উ ঊ and ও to ঔ ; and ঞ ঞ্ to আর, is called *briddhi*.

\* This section may be entirely omitted by the learner on going through the grammar the first time, and studied afterwards, in connexion with the chapter on compound words. As it strictly belongs to orthography, order requires its insertion in this place.

Figure 1: Description of গ ANJI as “a contraction for the word *Ganesh*” (from Yates 1849: 9).

Bihar

39

## TEXT IN NĀGARĪ SCRIPT

(१) q सिद्धिरस्तु । विक्रमादित्य-नृपतेः । संवत् १२[६]७ ज्येष्ठ-वदि १५ [गु]रो ॥ प्रव(ब)-  
 लतरानेकाहि(२)त-तृण-व(ब)हल-कक्ष-दावानले । श्री-सुरताण-मौजदीने राज्णं कुर्वारो । क्षत्रि-  
 (३) य-प्रवरकुल-सोमथान्वयेनाजिहिल(न)-गोत्रिणा । जगत्पाल-प्रपितामहेन । उत्त(४)र-पितामहेन ।  
 अजयपाल-गित्केन(ण) । श्रीमंत्रेश्वर-कामदेवेन गया कृ(५)ता । सपत्नीकेन । तथा भ्रात्रा राजदेवेन  
 जैत्रपाल-जनकेन गया कृत्ता(ता) । (६) प्रपितामहः साक्षी ॥ श्रुभं भवतु । जगत ॥ गया उद्धृता गता-  
 (ताः) ॥ ॥

## TRANSLATION

[Symbol meaning ‘Let there be success.’] Let there be success. On Thursday, the 15th lunar day of the dark half of Jyeshṭha in the year 1297 of king Vikramāditya, when *Suratāṇa* Maujadīna, who is the forest-fire to the grass and dry wood that are the numerous powerful enemies, was ruling the kingdom, the illustrious Mantreśvara-Kāmadeva who is the son of Ajayapāla, grandson of Uttara and great-grandson of Jagatpāla and belongs to the Ajihila (Ajihina)-gotra and to the Sopatha family that is the foremost clan of the Kṣatriya community, performed Gayā (i.e. the *śrāddha* of his ancestors at Gayā together with his wife. Likewise, his brother Rājadeva, father of Jaitrapāla, also performed Gayā (i.e. the *śrāddha* probably of Jaitrapāla, at Gayā). The witness [to the performance of the ceremonies] is Prapitāmaha (i.e. the god Prapitāmaheśvara of Gayā). Let there be welfare for the world. [The people whose *śrāddha* has been performed] at Gayā have obtained their deliverance [from their sins or the effects of their *karman*].

## ANSARI'S TRANSCRIPT AND TRANSLATION OF THE RECORD

## TEXT

- 1 Om siddhi rastu / vikramadityanṛpateh / samvat 1257 jyeṣṭha vadi 15 ravau / Pravalatarani Kṛ-
- 2 ta ṛṇa dviṣat kakṣa dāvānalo Śrī Suratāṇa Mojdina rājyaṁ kurvāṇe Nṛ-
- 3 pa pravara kulsoya dhanārthaya Nājahila gotriṇe Śrī Jagatapāla Prapitāmahena /
- 4 . . . ? rapitāmahena / Ajayapāla Pitrakena / Śrī Mantreśvara Kāmadevena Gaja-
- 5 tā sapatnīkena / tadā (yā?) trātrā (tathā bhrātrā?) Bājadevena Jaitrapālaja labdhena Gayākṛta
- 6 Prapitāmahakṣayāhni / sā janjava (śrutāmbha?) śrī (vato?) jagatagho (jagatotha?) Gayāsu hatā (Gayā uddhṛta?) / matā (gatā?) /

Figure 2: Excerpt showing the use of U+0071 LATIN SMALL LETTER Q for representing १ ANJI in Devanagari text (from Sircar 1979: 39). The sign appears at top in the Devanagari as “q” in the phrase “q सिद्धिरस्तु” and is glossed as “[Symbol meaning ‘Let there be success.’]”.

**ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>**

**Please fill all the sections A, B and C below.**

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

**A. Administrative**

1. Title:	<b>Proposal to Encode the Sign ANJI for Bengali</b>
2. Requester's name:	<i>Anshuman Pandey (pandey@umich.edu)</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>
4. Submission date:	<i>2012-04-23</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<input checked="" type="checkbox"/> <i>Yes</i>
(or) More information will be provided later:	<input type="checkbox"/>

**B. Technical – General**

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):		
Proposed name of script:		
b. The proposal is for addition of character(s) to an existing block:		
Name of the existing block:	<i>Bengali</i>	
2. Number of characters in proposal:	<i>1</i>	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>	B.2-Specialized (large collection) <input type="checkbox"/>
C-Major extinct <input type="checkbox"/>	D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<input checked="" type="checkbox"/> <i>Yes</i>	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/> <i>Yes</i>	
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/> <i>Yes</i>	
5. Fonts related:		
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Anshuman Pandey</i>	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Anshuman Pandey (pandey@umich.edu)</i>	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input checked="" type="checkbox"/> <i>Yes</i>	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/> <i>Yes</i>	
7. Special encoding issues:		
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?	<input checked="" type="checkbox"/> <i>Yes</i>	

**8. Additional Information:**

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database ( <http://www.unicode.org/reports/tr44/> ) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

<sup>1</sup> Form number: N3902-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03)

### C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain	No
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom? If YES, available relevant documents:	No
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	Yes <i>Size of user community is unknown.</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	Common <i>Character used as invocation symbol in inscriptions, manuscripts, and other records</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes <i>Character is historical, but may be in current use</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	Yes Yes <i>Character proposed for encoding in the Bengali block of BMP</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	N/A
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:	No
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	No
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	Yes Yes <i>Corresponds to characters used for similar purposes in other scripts</i>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:	No
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	No
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:	No