

Title: Proposal to Encode the Sign EKAM for Sharada in ISO/IEC 10646
Source: Script Encoding Initiative (SEI)
Author: Anshuman Pandey (pandey@umich.edu)
Status: Liaison Contribution
Action: For consideration by UTC and WG2
Date: 2011-11-02

1 Introduction

This is a proposal to encode a new character in the Sharada block of the Universal Character Set (ISO/IEC 10646). Sharada will be included in a future version of the UCS at the range U+11180..U+111DF (see N3595 L2/09-074R2). Properties of the proposed character are:

GLYPH	CODE	CHARACTER NAME	GC	CCC	BIDI	MIRRORED	LINEBREAK
ॐ	111C9	SHARADA EKAM	Lo	0	L	N	AL

Sharada EKAM should be sorted at the top of the script, before ॐ U+11183 SHARADA LETTER A.

2 Description

The Sharada ॐ EKAM is an invocation sign used at the beginning of texts. It represents the Sanskrit word ॐ *ekam* ‘one’ and is used in place of *ekam* in salutary phrases, such as *om svasti ekam siddham* ‘om hail! one, established!’. A detailed description of the character’s origins, meaning, and usage is provided by George A. Grierson in *The Journal of the Asiatic Society of Great Britain and Ireland* (1916: 678–680), which is reproduced here in figures 1–3.

3 References

Grierson, George A. 1916. “On the Sharada Alphabet.” *The Journal of the Asiatic Society of Great Britain and Ireland*, pp. 677–708.

Pandey, Anshuman. 2009. “Proposal to Encode the Sharada Script in ISO/IEC 10646”. N3595 L2/09-074R2. August 5, 2009. <http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3595.pdf>

4 Acknowledgments

This project was made possible in part by a grant from the United States National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at the University of California, Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.



to the Gurmukhī alphabet, but, unlike them, and like Nāgarī, it puts the letters *sa* and *ha* at the end of the alphabet, and not after the vowels. Kashmir is called the *Śāradā-kṣētra*, or holy land of the goddess Śāradā, and this is no doubt the origin of the name of the alphabet, although Elmslie, in his *Kāshmirī Vocabulary* (London, 1872), s.v. *Shāradā*, mentions a tradition that it is so called in honour of one Śāradānandana, who is said to have first reduced the Kāshmirī language to writing.

In India proper, when the alphabet is written down, it is usually preceded by the invocation *Ōm namaḥ siddham*, Ōm, reverence, established.¹ In Kashmir a slightly different formula of invocation is employed, viz. *Ōm svasti ēkaṁ siddham*, Ōm, hail! one, established. As regards the word *ēkaṁ*, one, it is a curious fact that, while, in writing the invocation, the words *ōm*, *svasti*, and *siddham* are fully written out—thus, **ॐ, स्वस्ति**, and **सिद्धं**—the word *ēkaṁ* is not written. Instead we have the mystic sign **ॐ**, which is named in Kāshmirī *okṣ sam gōr*, and is read as *ēkaṁ*. So that what is written in the Śāradā character is **ॐ स्वस्ति ॐ सिद्धं** read as *ōm svasti ēkaṁ siddham*. The traditional explanation of this is as follows: In order to master the theory of *mantras* in Kāshmirī Śaivism, it is necessary to learn the meaning or power of each letter composing a *mantra*, or the *mātrkā-cakra*. Each letter of the alphabet represents some mystic object. The vowels represent the various *śaktis*, the twenty-five consonants from *ka* to *ma* represent the twenty-five lower *tattvas*, and the other letters the higher *tattvas*, while *kṣa* represents the *prāṇa-bīja* or Life-seed.² In this way the


¹ Cf. Bühler, *On the Origin of the Indian Brāhma Alphabet*, p. 29 (Vienna, 1895), and Hoernle, on “The ‘Unknown Languages’ of Eastern Turkistān”, JRAS. 1911, p. 450. Bühler translates *siddham*, success.

² A full account of the *Mātrkā-cakra* will be found in Kṣēmarāja’s *Śivasūtravimarsinī*, ii, 7, translated in the *Indian Thought Series*, No. II.

Figure 1: Description of **ॐ** EKAM (from Grierson 1916: 678).

letter *a* represents the *jñāna-śakti*. It also indicates the Supreme (*anuttara*) and Solitary (*akula = kulōttīrṇa*) Śiva. The sign  is composed of three parts. The horizontal line — represents the letter *a*, i.e. also Śiva; the two perpendicular strokes || represent the other vowels, and also the *śaktis*; and the two curved marks  represent a plough (*hala*), and hence all the consonants, which are called by grammarians “*hal*”. The whole sign therefore represents all the vowels plus the consonants, or, in other words, the entire alphabet. On the mystical side it also represents Śiva plus all the *śaktis* and *tattvas*, i.e. Śiva and all his developments in the way of so-called creation.

In the Kāshmirī name *ok^u sam gōr*, *ok^u* means “one”, “non-dual”; *sam* is a contraction of *saṁvittva*, or condition of *parā saṁvit*, the Supreme Experience; and *gōr* is for *gōr^u*, it has been inquired into (and therefore understood). With *siddham* added the whole means “the supreme monist experience has been mastered (for it has been established in the *āgamas*)”. *Ēkam siddham* has the same meaning.

A less mystical interpretation has been kindly given to me by Professor Barnett. He points out that the *siddham* is probably derived from the first *sūtra* of the Kātantra, which runs *siddhō varṇasamāmnāyah*, i.e. “the traditional order of the letters is established (as follows)”, and that this grammar was, over a thousand years ago, the most popular handbook in Northern India¹ and the Buddhist regions of Central Asia. The mark  is evidently one of the sacred symbols used at the commencement or end of any important writing, such as are referred to by Bühler on p. 85 of his *Indische Palaeographie*, and has practically the force of a sign of punctuation. A not

¹ When I was in India its use in Northern India seems to have been confined to Eastern Bengal, where I studied it with the local Paṇḍits. In the rest of Bengal the Mugdhabōdha was in general use.

Figure 2: Description of  EKAM (from Grierson 1916: 679).

very dissimilar sign will be found at the end of the plate facing p. 281 of vol. ii of Rājendra Lāla Mitra's *Notices of Sanskrit Manuscripts* (२८१). Taking this sign as one of punctuation, it would be natural to mark this first division-point by mentally interjecting *ēkaṁ*, as a distant echo of the numbering of the first *sūtra* of the Kātantra, and in course of time the word *ēkaṁ* would become petrified, the meaning would be forgotten, and a new mystical meaning given to it.

In those parts of Northern India with which I am acquainted there are, except in the Pañjāb, no special names for the various letters. अ *a* is called *a-kāra*, क *ka* is called *ka-kāra*, and so on. In the Gurmukhī alphabet, used in the Pañjāb, it is different. Here the vowels have each its own name. Thus, initial अ *a* is called *āirā*, non-initial ऀ *ā* is called *ā-kannā*, and so on. The consonants are also named by enunciating each twice and doubling the consonant mentioned the second time. Thus क *ka* is called *kakkā*, ख *kha* is called *khakkhā*, and so on.

In the schools of Kashmir this Pañjāb system is carried much further. Every vowel form and every consonant has its separate individual name. Most of these names have no definite meaning apart from this connotation, and, as names of letters, do not seem to have been invented on any regular system. Even each syllable of *ōm svastī*, and of *siddham*, and the sign ॐ for *ēkaṁ*, has its own name. I give these names in the following tables, written in the Nāgarī and Śāradā characters, with a transliteration into the Roman character. As these names are not Sanskrit, but are in the Kāshmirī language, the system of transliteration followed is that which is applied to Kāshmirī, and which differs slightly from the transliteration of the corresponding Nāgarī or Śāradā letters when used for Sanskrit.

Figure 3: Description of ॐ EKAM (from Grierson 1916: 680).

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

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:	Proposal to Encode the Sign EKAM for Sharada in ISO/IEC 10646
2. Requester's name:	<i>Script Encoding Initiative (SEI) / Anshuman Pandey (pandey@umich.edu)</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Liaison contribution</i>
4. Submission date:	<i>2011-11-02</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<input checked="" type="checkbox"/> Yes
(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>Sharada</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"/> Yes	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/> Yes	
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/> Yes	
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"/> Yes	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/> Yes	
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"/> Yes	

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.

¹ 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 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:	No
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:	No
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