

Proposal to encode 116B8 TAKRI LETTER ARCHAIC KHA

Shriramana Sharma, samjnaa-at-gmail-dot-com, India

2017-Aug-01

§1. Introduction

This is a proposal to encode one character in the Takri block:

𑂔

116B8 TAKRI LETTER ARCHAIC KHA

§2. Discussion

Takri was proposed by L2/09-111. The proposal has attested the shape 𑂔 for *kha* on p 21*:

kh | 𑂔

However, the proposal shows that the other shape 𑂕 was also used for *kha* as on pp 36 and 40:

𑂕 | 𑂕, 𑂔 𑂕 | 𑂕 𑂔

Finally the proposal went on to encode 𑂕 alone as 116B8 TAKRI LETTER KHA, omitting 𑂔.

In retrospect, it would seem that the proposal should have encoded 𑂕 as TAKRI LETTER SSA, because it has provided copious attestations of 𑂕 being consistently used for *ṣa* in earlier periods and only in later periods has it been re-used for *kha* as seen on pp 21 and 23:

kh	𑂔		𑂔	𑂔		𑂔			𑂕	𑂕
ṣ	𑂕	𑂕	𑂕	𑂕	𑂕	𑂕		𑂕	𑂕	𑂕

Certainly the transition of 𑂔 to 𑂕 for *kha* cannot be a natural script evolution especially seeing as 𑂕 has already been consistently used for *ṣa*. Thus there are other linguistic factors at work here. (For more discussion see my L2/17-231.)

In any case, the proposal seems to have been led by a large number of other (later?) listings showing 𑂕 for *kha* to instead encode it as TAKRI LETTER KHA and not encode a TAKRI LETTER SSA at all. Re this the proposal wrote on p 10:

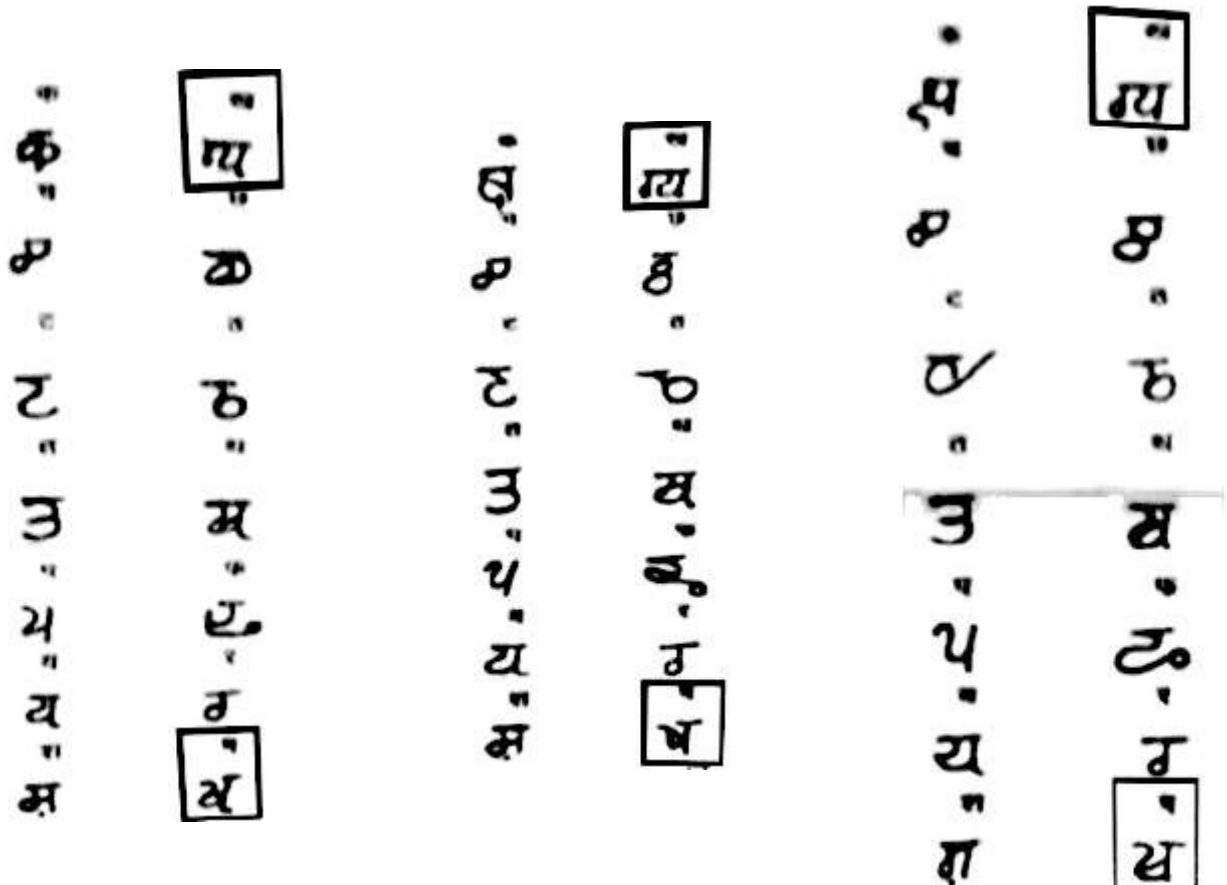
* All page numbers provided in this document are of the PDFs and not the internal numbering if any.

This character resembles the consonant conjunct mha shown in Figure 10. Further investigation is needed in order to determine the existence of a TAKRI SSA.

It is curious and somewhat unfortunate that the proposal has considered only the shape in the last column (perhaps because it is labeled “Modern Takri”) for determining the linguistic identity of the character. It is true that the shape given for *ṣa* in the last column resembles the shape given for *mha* on p 25, but examining the samples on pp 21~23 one finds that the last column isn’t really representative of a true evolution of the earlier columns.

Be that as it may, the current situation is that **ਖ** is encoded as 1168B TAKRI LETTER KHA, and **ਘ** is not encoded. This is now proposed for encoding. While later Takri writings may not use **ਘ**, earlier documents use it in contrast with **ਖ** as seen from the earlier attestations.

More attestations of **ਘ** in contrast with **ਖ** from various styles of Takri are in L2/17-209 pp 7~9:



In fact L2/17-209 has provided these attestations in favour of encoding a new character for the written form **ਖ** already encoded at 1168B and changing the glyph of that 1168B to **ਘ** on the presumption that that is the “correct form of KHA”. However, since Unicode encodes written forms and names are only secondary, I do not believe it is appropriate to thus effectively relocate **ਖ** and hence propose that it is the unrepresented **ਘ** which should instead be encoded newly.

§3. Character properties

The properties including line-breaking etc are as in other Takri/Sharada consonants:

116B8;TAKRI LETTER ARCHAIC KHA;Lo;0;L;;;;N;;;;;

Collation will however need to be tailored because if this new character is used, then it will represent *kha* and the existing 𑖑 1168B TAKRI LETTER KHA will represent *ṣa*, whereas in styles which do not use this new character, 𑖑 1168B will represent both *kha* and *ṣa* as per context.

The new character 116B8 should be annotated as “used in earlier writings to denote kha” and the existing character 1168B as “also used to denote ssa”.

§4. Official Proposal Summary Form

(Based on N4502-F)

A. Administrative

1. Title

Proposal to encode 116B8 TAKRI LETTER ARCHAIC KHA

2. Requester's name

Shriramana Sharma

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution

4. Submission date

2017-Aug-01

5. Requester's reference (if applicable)

6. Choose one of the following: This is a complete proposal (or) More information will be provided later

This is a complete proposal.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters), Proposed name of script

No

1b. The proposal is for addition of character(s) to an existing block, Name of the existing block

Yes, Takri

2. Number of characters in proposal

1 (one)

3. Proposed category

Category C, major extinct

4. Is a repertoire including character names provided?

Yes

4a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes

4b. Are the character shapes attached in a legible form suitable for review?

Yes

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Not necessary. The glyph may be copied from 11192 SHARADA LETTER KHA.

b. Identify the party granting a license for use of the font by the editors (include address, e-mail etc.)

Anshuman Pandey

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes

7. 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)?

Yes

8. 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.

See detailed proposal.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

No

2b. If YES, with whom?

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Those who wish to store documents in Takri script as digital text

4a. The context of use for the proposed characters (type of use; common or rare)

Rare

4b. Reference

See detailed proposal.

5a. Are the proposed characters in current use by the user community?

Only academic use

5b. If YES, where?

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

No

6b. If YES, is a rationale provided?

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Only one character is proposed. It is proposed to be kept contiguous with other linguistic characters of Takri.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?

The character is similar to 11192 SHARADA LETTER KHA.

10b. If YES, is a rationale for its inclusion provided?

The script property would be different, script=takri.

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences?

No

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

13c. If YES, reference: