Proposal to Encode the Khojki Script in ISO/IEC 10646

Anshuman Pandey
University of Michigan
Ann Arbor, Michigan, U.S.A.
pandey@umich.edu

March 4, 2009

Contents

List of Tables

List of Figures

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]

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 Khojki Script in ISO/IEC 10646**
2. Requester's name: **Anshuman Pandey (pandey@umich.edu)**
3. Requester type (Member Body/Liaison/Individual contribution): **Individual contribution**
4. Submission date: **March 4, 2009**
5. Requester's reference (if applicable): **N/A**
6. Choose one of the following:
   (a) This is a complete proposal: **No**
   (b) or, More information will be provided later: **Yes**

B. Technical - General

1. Choose one of the following:
   (a) This proposal is for a new script (set of characters): **Yes**
       i. Proposed name of script: **Khojki**
   (b) The proposal is for addition of character(s) to an existing block: **No**
       i. Name of the existing block: **N/A**
2. Number of characters in proposal: **66**
3. Proposed category: **B.1 - Specialized (small collections of characters)**
4. Is a repertoire including character names provided?: **Yes**
   (a) If Yes, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?:
       Yes
   (b) Are the character shapes attached in a legible form suitable for review?: **Yes**
5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for
   publishing the standard?: **Anshuman Pandey; True Type format**
   (a) If available now, identify source(s) for the font and indicate the tools used: **The characters of the digitized
       Khojki font are based on normalized forms of printed Khojki characters. The font includes characters
       from a digitized Khojki font called KhojkiJiwa.ttf (provided by Shafique Virani), as well as several new
       characters not found in KhojkiJiwa and revisions to several glyphs. The font is a work-in-progress
       and the final font will be a new work that will be produced by Anshuman Pandey from printed Khojki
       sources.**
6. References:
   (a) Are references (to other character sets, dictionaries, descriptive texts etc.) provided?: **Yes**
   (b) Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed
       characters attached?: **Yes**
7. Special encoding issues:
   (a) 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; see proposal for
       additional details.**
8. Additional Information: Submitters are invited to provide any additional information about Properties of the pro-
   posed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the pro-
   posed 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 be-
   haviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equiv-
   alence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org
   for such information on other scripts. Also see http://www.unicode.org/Public/UNIDATA/UCD.html and associ-
   ated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for
   inclusion in the Unicode Standard. **Character properties and numeric information are included.**

---

[1] Form number: N3102-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)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?: **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.)? **Yes**
   (a) If Yes, with whom?:
      - **Shafique Virani (shafique.virani@utoronto.ca), University of Toronto - Mississauga**
         i. If Yes, available relevant documents: **N/A**
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? **Yes**
   (a) Reference: **See text of proposal**
4. The context of use for the proposed characters (type of use; common or rare): **Common**
   (a) Reference: **The script was used primarily by the Ismaili religious community in South Asia.**
5. Are the proposed characters in current use by the user community?: **Yes. Members of the Ismaili community and specialists studying Ismaili literature require the script.**
   (a) If Yes, where? Reference: **In the United Kingdom, Canada, United States, and Pakistan.**
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?: **No**
   (a) If Yes, is a rationale provided?: **N/A**
      i. If Yes, reference: **N/A**
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? **Yes**
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? **No**
   (a) If Yes, is a rationale for its inclusion provided?: **N/A**
      i. If Yes, reference: **N/A**
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? **No**
   (a) If Yes, is a rationale provided?: **N/A**
      i. If Yes, reference: **N/A**
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? **Yes**
    (a) If Yes, is a rationale for its inclusion provided? **Yes**
       i. If Yes, reference: **See text of proposal**
11. Does the proposal include use of combining characters and/or use of composite sequences? **Yes**
    (a) If Yes, is a rationale for such use provided? **Yes**
       i. If Yes, reference: **See text of proposal**
    (b) Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? **N/A**
       i. If Yes, reference: **N/A**
12. Does the proposal contain characters with any special properties such as control function or similar semantics? **Yes**
    (a) If Yes, describe in detail (include attachment if necessary): **Virama**
13. Does the proposal contain any Ideographic compatibility character(s)? **No**
    (a) If Yes, is the equivalent corresponding unified ideographic character(s) identified? **N/A**
       i. If Yes, reference: **N/A**

1   Introduction

Purpose   This is a formal proposal to encode the Khojki script in the Universal Character Set (ISO/IEC 10646). It replaces the following document:

- L2/08-201: "Towards an Encoding for the Khojki Script in ISO/IEC 10646", May 2008

2   Background

Name   The Khojki (𑈈𑈬𑈐𑈈𑈭) script takes its name from the Sindhi form *khojā* of the Persian خواجہ *k̲h̲wājah* "master, lord". Literally, it means "of the master" (*khojā* + *kī*). The name appears in transliteration as Khojakī and Khojkī. The common form 'Khojki' is recommended as the formal name of the script in the UCS.

Overview   Khojki is the name of the writing system used specifically by the Nizari Ismaili community of South Asia for recording their religious literature. It is a Brahmi-based script of the Landa family, which is a class of mercantile scripts that is derived ultimately from Sharada. Khojki is a refined form of the Lohāṇakī variety of Landa used in Sind that was purposely developed as an ecclesiastical script. It is believed to have become distinct in the 16th century CE. Popular Ismaili tradition states that the script was invented by Pir Sadruddin, an Ismaili missionary actively working with the Hindu Lohana community of Sind. Khojki is one of two mercantile Landa scripts that were transformed into literary scripts in order to serve the purposes of religious communities; the other is Gurmukhi.

Developed in and used predominantly in Sind (in modern Pakistan), Khojki was used to represent Sindhi, Gujarati, Persian, and other regional languages. The majority of Khojki materials are manuscripts; however, hundreds of lithographs and books were published in the 20th century after Laljibhai Devraj produced metal types for Khojki in Germany for use at his Khoja Sindhi Printing Press in Bombay. Large collections of Khojki materials are housed at the Institute of Ismaili Studies in London and at Harvard University.

Although Khojki possesses a limited user base, manuscripts and books were used in Ismaili religious ceremonies not only in South Asia but in East and South Africa, where large diaspora communities were formed starting in the 19th century. While the position of Khojki as a living script is declining, it is still used by members of the Ismaili community and by specialists who study the history and literature of the the community. An encoding for Khojki in the Universal Character Set will provide a means for the Ismaili community and specialists to preserve literature in the script and to advance the study of Ismaili religious and literary culture.

	1128	1129	112A	112B	112C
0	◌̊ 11280	ধা 11290	દ 112A0	ঝ 112B0	⁝ 112C0
1	◌̇ 11281	૩ઃ 11291	২ 112A1	ડ 112B1	— 112C1
2	ꠀ 11282	ਜ 11292	૦ 112A2	◌) 112B2	= 112C2
3	ꠀ) 11283	২૯ 11293	પ 112A3	◌ſ 112B3	[囗] 112C3
4	꞉꞉ꞌ 11284	ꠇ 11294	৯ 112A4	◌ૂ 112B4	
5	૭૯ 11285	ꠃ 11295	৩ 112A5	◌ 112B5	
6	৩ 11286	ꠉ 11296	ম 112A6		
7	৩ 11287	꠱ঃ 11297	ꠎ 112A7	◌̀ 112B7	
8	૫ 11288	꠨ 11298	ꠎ 112A8	◌̂ 112B8	
9	ꠣ 11289	꠩ 11299	০ণ 112A9	◌) 112B9	
A	6 1128A	꠪ 1129A	૨ 112AA	◌) 112BA	
B	ꠀ) 1128B	২৭ 1129B	ন 112AB	◌/ 112BB	
C	꠪ 1128C	꠬ 1129C	৮ 112AC	◌̈ 112BC	
D	꠫ 1128D	꠭ 1129D	૧ 112AD	◌ᵚ 112BD	
E	꠯ 1128E	৫ 1129E	꠲ 112AE	୧ 112BE	
F	ꠣ 1128F	૧ 1129F		୧୧ 112BF	

Table 1: Glyph chart for Khojki

Various signs

11280 ँ KHOJKI SIGN CANDRABINDU
11281 ं KHOJKI SIGN ANUSVARA

Independent vowels

11282 ᴧ KHOJKI LETTER A
11283 ᴧ KHOJKI LETTER AA
11284 ᴂ KHOJKI LETTER I
11285 ᴂ KHOJKI LETTER II
11286 ᴇ KHOJKI LETTER U
11287 ᴇ KHOJKI LETTER UU
11288 ᴠ KHOJKI LETTER E
11289 ᴧ KHOJKI LETTER AI
1128A ᴇ KHOJKI LETTER O
1128B ᴧ KHOJKI LETTER AU

Consonants

1128C ᴈ KHOJKI LETTER KA
1128D ᴠ KHOJKI LETTER KHA
1128E ᴈ KHOJKI LETTER GA
1128F ᴈ KHOJKI LETTER GGA
11290 ᴦ KHOJKI LETTER GHA
11291 ᴊ KHOJKI LETTER NGA
11292 ᴊ KHOJKI LETTER CA
11293 ᴂ KHOJKI LETTER CHA
11294 ᴣ KHOJKI LETTER JA
11295 ᴣ KHOJKI LETTER JJA
11296 ᴣ KHOJKI LETTER JHA
11297 ᴣ KHOJKI LETTER NYA
11298 ᴈ KHOJKI LETTER TTA
11299 ᴈ KHOJKI LETTER TTHA
1129A ᴦ KHOJKI LETTER DDA
1129B ᴈ KHOJKI LETTER DDHA
1129C ᴊ KHOJKI LETTER NNA
1129D ᴦ KHOJKI LETTER TA
1129E ᴦ KHOJKI LETTER THA
1129F ᴈ KHOJKI LETTER DA
112A0 ᴇ KHOJKI LETTER DDDA
112A1 ᴊ KHOJKI LETTER DHA
112A2 ᴦ KHOJKI LETTER NA
112A3 ᴦ KHOJKI LETTER PA
112A4 ᴋ KHOJKI LETTER PHA
112A5 ᴊ KHOJKI LETTER BA
112A6 ᴊ KHOJKI LETTER BBA
112A7 ᴧ KHOJKI LETTER BHA
112A8 ᴦ KHOJKI LETTER MA
112A9 ᴏᴏ KHOJKI LETTER YA
112AA ᴋ KHOJKI LETTER RA
112AB ᴦ KHOJKI LETTER LA
112AC ᴔ KHOJKI LETTER LLA
112AD ᴔ KHOJKI LETTER VA
112AE ᴔ KHOJKI LETTER SHA
112AF ▨ <reserved>
112B0 ᴔ KHOJKI LETTER SA
112B1 ᴣ KHOJKI LETTER HA

Dependent vowel signs

112B2 ᴏ KHOJKI VOWEL SIGN AA
112B3 ᴏᴦ KHOJKI VOWEL SIGN I
112B4 ᴏᴦ KHOJKI VOWEL SIGN II
112B5 ᴏ KHOJKI VOWEL SIGN U
112B6 ▨ <reserved>

112B7 ᴏ KHOJKI VOWEL SIGN E
112B8 ᴏ KHOJKI VOWEL SIGN AI
112B9 ᴏᴦ KHOJKI VOWEL SIGN O
112BA ᴏᴦ KHOJKI VOWEL SIGN AU

Various signs

112BB ᴏᴦ KHOJKI SIGN VIRAMA
112BC ᴏ KHOJKI SIGN NUKTA
112BD ᴏ KHOJKI SIGN SHADDA

Punctuation

112BE ᴦ KHOJKI DANDA
112BF ᴦ KHOJKI DOUBLE DANDA
112C0 ः KHOJKI WORD SEPARATOR
112C1 � KHOJKI SECTION MARK
112C2 ᴷ KHOJKI DOUBLE SECTION MARK
112C3 ⬚ KHOJKI VERSE MARKER

Table 2: Names list for Khojki

3

3   Characters Proposed

Number of Characters   The 66 letters proposed here comprise the core Khojki character set.

Character Names   The naming convention for Khojki characters follows that used for Devanagari. Characters for the Sindhi implosive consonants GGA, JJA, DDDA and BBA are derived from their Devanagari analogues. There is no standard transliteration for these characters. Alignment with the names of analogous Devanagari characters will faciliate the recognition of characters.

Allocation   Khojki is currently allocated in the Supplementary Multilingual Plane (SMP) at the range U+11280..U+112CF.

3.1   Character Set

Consonants   There are 37 consonant letters:

ᛞ	KHOJKI LETTER KA	�	KHOJKI LETTER TTHA	ᛗ	KHOJKI LETTER BBA
	KHOJKI LETTER KHA		KHOJKI LETTER DDA		KHOJKI LETTER BHA
	KHOJKI LETTER GA		KHOJKI LETTER DDHA		KHOJKI LETTER MA
	KHOJKI LETTER GGA		KHOJKI LETTER NNA		KHOJKI LETTER YA
	KHOJKI LETTER GHA		KHOJKI LETTER TA		KHOJKI LETTER RA
	KHOJKI LETTER NGA		KHOJKI LETTER THA		KHOJKI LETTER LA
	KHOJKI LETTER CA		KHOJKI LETTER DA		KHOJKI LETTER LLA
	KHOJKI LETTER CHA		KHOJKI LETTER DDDA		KHOJKI LETTER VA
	KHOJKI LETTER JA		KHOJKI LETTER DHA		KHOJKI LETTER SHA
	KHOJKI LETTER JJA		KHOJKI LETTER NA		KHOJKI LETTER SA
	KHOJKI LETTER JHA		KHOJKI LETTER PA		KHOJKI LETTER HA
	KHOJKI LETTER NYA		KHOJKI LETTER PHA		
	KHOJKI LETTER TTA		KHOJKI LETTER BA		

Vowels   There are 10 independent vowels:

	KHOJKI LETTER A		KHOJKI LETTER U		KHOJKI LETTER O
	KHOJKI LETTER AA		KHOJKI LETTER UU		KHOJKI LETTER AU
	KHOJKI LETTER I		KHOJKI LETTER E		
	KHOJKI LETTER II		KHOJKI LETTER AI		

Vowel Signs   There are 8 dependent vowel signs:

	KHOJKI VOWEL SIGN AA		KHOJKI VOWEL SIGN U		KHOJKI VOWEL SIGN O
	KHOJKI VOWEL SIGN I		KHOJKI VOWEL SIGN E		KHOJKI VOWEL SIGN AU
	KHOJKI VOWEL SIGN II		KHOJKI VOWEL SIGN AI		

Various Signs   There are 5 various signs:

◌̊ KHOJKI SIGN CANDRABINDU   ◌⁄ KHOJKI SIGN VIRAMA   ◌̈ KHOJKI SIGN SHADDA

◌̇ KHOJKI SIGN ANUSVARA   ◌̤ KHOJKI SIGN NUKTA

Punctuation   There are 6 punctuation signs:

ⅼ KHOJKI DANDA   ⚊ KHOJKI SECTION MARK

ⅼⅼ KHOJKI DOUBLE DANDA   ⚌ KHOJKI DOUBLE SECTION MARK

⁚ KHOJKI WORD SEPARATOR   ▢ KHOJKI VERSE MARKER

3.2   Basis for Character Set and Glyph Shapes

Basis for Character Set   Given the absence of a standard for Khojki, this proposal intends to provide an idealized character set that may be used to encode any Khojki document completely. The character set proposed here is based upon the inventory of Khojki characters presented by George Stack's *A Grammar of the Sindhi Language* (as represented in George Grierson's *Linguistic Survey of India*: figures 1–3); Ali S. Asani in *The Harvard Collection of Ismaili Literature in Indic Languages* (figures 4–9); and Christopher S. Shackle and Zawahir Moir in *Ismaili Hymns from South Asia* (figures 10–11). Included in this character set are characters that are not enumerated in official Khojki primers, but are nonetheless attested in manuscripts and printed books, as well as in character charts. The presence of characters with which readers of Khojki are not necessarily familiar should not be a distraction, but rather should be understood as a letter of the script that at some historical period and within a particular region was identified as being part of the Khojki repertoire.

These characters are:

1. **LETTER II**   Generally, Khojki does not distinguish between independent forms of I and II. The character ꠁ KHOJKI LETTER I is used to represent both. However, Figure 4 shows the glyph ꠄ as an alternate form of II. The character is included here in order to accommodate the attestation.

2. **LETTER UU**   Khojki typically does not distinguish between independent forms of U and UU. The character ꠊ KHOJKI LETTER U is used to represent both. Figure 4, however, shows glyphs for UU that are distinct from ꠊ. The character is included here in order to accommodate the attestation.

3. **LETTER JHA**   Character inventories of Khojki show the glyph ꠣ used to represent the letter JHA /ʤʰ/. However, they also show this glyph as representing KHOJKI LETTER DHA (/dʰ/). Although the form ꠣ is used to represent both *jha* and *dha*, it is more commonly used to write DHA and is, therefore, assigned here to KHOJKI LETTER DHA. The use of a single glyph to represent two characters requires a solution for representing both characters uniquely. Figure 6 shows the glyph ꠛ for JHA, but this character is not found in traditional character inventories. Should the need arise to encode the character *jha* from a Khojki manuscript, using KHOJKI LETTER DHA is unacceptable; the appearance may be captured, but the semantic value of the character would disappear. Thus, JHA is proposed for encoding in order to accommodate the attestation.

4. **LETTER SHA**   Although the character � SHA may be represented as ꠡ SA + ◌̈ NUKTA, it is proposed as an atomic character. Unlike other characters that are produced using NUKTA for the purpose of representing sounds not found in Sindhi, Khojki SHA represents /ʃ/, a sound native to Sindhi.

Basis for Glyph Shapes   The digitized Khojki characters proposed here are based on the forms of the metal fonts developed by Laljibhai Devraj in 1903. Devraj used the fonts to produce the first printed versions of Ismaili devotional literature (*ginān*) at his Khoja Sindhi Printing Press in Bombay. Although different publishers in Sindh, Gujarat, and Mumbai printed works in Khojki, the majority of publications were produced by Devraj. Devraj's font, therefore is universally recognized by readers of Khojki. This digitized Khojki font contains characters not originally found in metal fonts, but additional characters have been designed with the spirit of Devraj's style, and it is hoped that this will appeal to the Ismaili community and will mirror the 'standard' appearance of Khojki to which they are accustomed.

3.3 Characters Not Proposed

Although the character set proposed for Khojki is an idealized inventory intended to enable the complete encoding of Khojki documents, there are some characters that have been identified in Khojki documents, but are not proposed for encoding at present for one or more of the following reasons: (a) insufficient information regarding the characters and their properties; (b) the possibility of representing a character with another of similar or equal function.

1. **VOWEL SIGN UU**   Although ꧯ KHOJKI LETTER UU is proposed for encoding based upon attested use of the character, sources attesting a unique character for VOWEL SIGN UU have not been identified. The general practice is to use ◌ KHOJKI VOWEL SIGN U for dependent forms of both U and UU. Space has been reserved in the script block should a unique character for this sign be identified.

2. **Characters for *kṣa*, *jña*, *tra*, and *dra***   Khojki charts traditionally show the characters ꠇ *kṣa*, ꠇ *jña*, ꠇ *tra*, and ꠇ *dra*. These characters are consonant ligatures, and while their appearance differs from that of their constituent characters, they should be handled as such. The inclusion of these four characters in traditional character charts for Khojki is analogous to the inclusion of the conjuncts क्ष *kṣa*, त्र *tra*, and ज्ञ *jña* in charts of Devanagari for languages such Hindi.

3. **Pre-composed characters for consonant + NUKTA**   There are several Khojki letters that are composed of a base letter + NUKTA: ꠇ A + NUKTA; ꠇ GA + NUKTA (see section 4.8). It is not necessary to encode each of these uniquely. The only pre-composed character proposed for encoding is ꠇ KHOJKI LETTER SHA.

4. **Various signs derived from Devanagari**   Authors of Khojki instructional materials at times attempted to fix the deficiencies of the script by borrowing orthographic practices from scripts such as Devanagari. The specimen below illustrates such uses. One such script primer, ꠇꠇ ꠇꠇꠇꠇ ꠇꠇꠇꠇ (*Sindhī pahelī copaḍī*), introduces VOWEL SIGN VOCALIC R and various diacritics to represent conjunct conjuncts with RA. These characters are not proposed at present until a determination can be made as to the necessity of encoding them.

```
P Ism K 17                              Bombay 1966 s./ 1910

Siṅdhī pahelī copaḍī. 1st edition.

Publisher: Mûkhī Lāljībhāî Devrāj, Manager of the Bombay Shia
Imami Ismaili Khoja Jamat Dept of Education, for the Siṅdhī
Textbook Committee at the Khoja Siṅdhī Printing Press; 2000
copies; 92 pp. in various paginations; 17 x 13; covers missing;
binding in fair condition; paper in good condition; saṅkaḍīûṅ
(table of contents): 2.1-3.

According to publisher's note (pp. 1.3-4), since the Khojki
script lacks the facility to make "true" conjunct characters, the
publication uses Sanskrit ones. Teachers are requested to famil-
iarize themselves with these forms before proceeding with
instruction. Special uses: (1) "〜" under the first member of a
conjunct consonant suppresses the implicit short "a" vowel; (2)
if the letter "r" is the second member of a conjunct, the super-
script sign "∧" is used on the first member i.e. "ઘ૮" for "pr";
(3) Subscript "◟" with the consonant "r" indicates "r" as the
first member of the conjunct i.e. "પૂર્વ" for "pûrva"; (4) the
vowel "ṛu" is indicated by subscript "」" ; (5) the following
conjuncts introduced: ن = Gujarati શ્ય ; ન = Gujarati ખ્ય ;
   શ્ન = Gujarati ઘ્ન ; ઋ = Gujarati ત્ર ;
The script does not use ":" to separate individual words.
```

5. **Digits**   There are no unique Khojki digits. When digits appear in Khojki texts, they are generally Gujarati forms: ૦ ૧ ૨ ૩ ૪ ૫ ૬ ૭ ૮ ૯ (U+0AE6..U+0AEF).

6. **Number forms and unit marks**   Fraction signs and currency marks are found in Khojki documents. The forms and semantics of these characters are identical to the North Indic number forms proposed for inclusion in the "Common Indic Number Forms" block (U+A830..U+A83F). See section 4.11 for additional information.

4   The Writing System

4.1   General Features

Classification   The Khojki script is an abugida of the Brahmic type and is written from left to right.

Syllabic Formation   The formation of syllables in Khojki follows the pattern common to north Indic scripts. Consonant letters bear the inherent vowel *a* when unaccompanied by a vowel sign. The inherent vowel is changed by applying a vowel sign to the consonant or suppressed by applying VIRAMA to produce the bare consonant. All dependent vowel signs are written above, below, or to the right of the consonant (including ◌ſ KHOJKI VOWEL SIGN I).

4.2   Inadequacies of the Script

The Khojki script has several incongruencies between glyphs and characters:

1. *Redundancy*   One character represented by multiple glyphs

2. *Ambiguity*   Multiple characters represented by a single or by multiple glyphs

3. *Deficiency*   Ligature glyph forms composed of characters absent from the script

Redundancy   As Khojki was developed to represent the Sindhi language, it has unique characters for representing sounds found in Sindhi, but not in other neighboring language groups. These sounds are the four voiced implosive consonants: /ɠ/ (velar), /ʄ/ (palatal), /ɗ/ (dental), and /ɓ/ (labial). When the Khojki script was used outside of Sindh and by scribes unfamiliar with Sindhi, the phonetic value of the characters for the implosive consonants were unrealized. With the deletion of the implosive consonants, the glyphs for these characters were reassigned to represent characters for other consonants within the same articulation class. This deletion resulted in the use of multiple glyphs for a single character:

- ꠃ KHOJKI LETTER GGA /ɠ/ came to represent /g/, already represented by ꠉ KHOJKI LETTER GA. ꠃ was also used for writing the consonant conjunct (GA + VIRAMA + RA) /gr/.

- ꠚ KHOJKI LETTER JJA /ʄ/ came to represent /dʒ/, already represented as ꠎ KHOJKI LETTER JA. In some cases ꠚ was also used to write ꠎ̇ (JA + VIRAMA + NUKTA) /z/.

- ꠐ KHOJKI LETTER DDDA /ɗ/ came to represent /d/, already represented by ꠒ KHOJKI LETTER DA. The duplication of glyphs for DA led to the duplicate use of ꠒ for /dʰ/ (ꠖ KHOJKI LETTER DHA).

- ꠪ KHOJKI LETTER BBA /ɓ/ came to represent /b/, already represented with (ꠢ KHOJKI LETTER BA). In turn, ꠢ was used synonymously with ꠪ to write KHOJKI LETTER BHA /bʰ/.

Other issues with redundancy are:

- ꠖ was used to represent not only /dʰ/ (KHOJKI LETTER DHA), but also /dʒʰ/ (ꠏ KHOJKI LETTER JHA).

- ꠘ, the glyph for the ligature *tra* (TA + VIRAMA + RA) was at times used to represent /t/ (ꠕ KHOJKI LETTER TA).

**Ambiguity**   In some cases, one glyph is used to represent multiple characters.

- The glyph ꠂ is used for the independent form of both I and II. As II is attested as a unique character in some Khojki manuscripts, it is encoded as an independent character and represented with the glyph ꠣꠣ.

- The glyph ꠄ is used for the independent form of both U and UU. The sign ◌ represents both dependent forms of these vowel letters. As UU is attested as a unique character in some Khojki manuscripts, it is encoded as an independent character and represented with the glyph ꠅꠅ.

- The character ꠆ E is also represented as ꠄ̀.

- The character ꠄ̀ AI is also represented as ꠆ꠄ.

- The character ꠅ O is also represented as ꠅ̀.

- The character ꠅ̀ AU is also represented as ꠆ꠅ and ꠆ꠅ.

- The glyph ꠊ is used to represent both KA and a letter *QA (/q/, used to transcribe ق ARABIC LETTER QAF).

- In some manuscripts, the glyph ꠟ (assigned here to KHOJKI LETTER NNA) represents the characters DDA, *DDDHA, and NNA. Furthermore, given allophonic variation between DDA and *DDDHA, it is difficult to determine precisely which of these characters is being represented by ꠟ.

- In most sources, the character DDA is represented by the glyph ⟨ꙮ⟩; however in some this glyph represents *DDDHA. This ambiguity has been addressed by assigning unique glyphs to DDA and NNA; while the character *DDDHA has not been proposed for encoding (see section 3.3).

- **6** was used to represent both KHOJKI LETTER O and KHOJKI LETTER U.

**Deficiency**   The script lacks unique glyphs to represent sounds not found in Sindhi:

- In order to represent Arabic sounds, new characters are created by adding the diacritic NUKTA to an existing character (see section 4.8).

- The consonant ligature ⟨ꙮ⟩, which is composed of KA + VIRAMA + SSA; however, Khojki does not have a character SSA.

4.3   Supression of Inherent Vowel

As most Indic scripts, Khojki has the special sign *virāma*, which is written as a combining character with a consonant letter to supress the consonant's inherent vowel *a*. The ○/ KHOJKI SIGN VIRAMA is written to the right of the consonant, not beneath as is typical in most Indic scripts, eg. Khojki ꙮ/ *k* and Devanagari क् *k*.

4.4   Consonant-Vowel Ligatures

The addition of dependent signs to consonant letters results in consonant-vowel ligatures for all consonants.

VOWEL SIGN I   Unlike the practice in other Indic scripts, KHOJKI VOWEL SIGN I is written to the right of the consonant, not to the left. The KHOJKI VOWEL SIGN I has two forms: ○ꙮ and ○ꙮ.

- ○ꙮ is used with characters that have right descenders: ꙮ → ꙮ *ci*; ꙮ → ꙮ *vi*. Or that possess a final right stroke that terminates at the baseline: ꙮ → ꙮ *ki*; ꙮ → ꙮ *di*; ꙮ → ꙮ *hi*.

- ○ꙮ is used with characters that have rounded right edges: ꙮ → ꙮ *chi*; ꙮ → ꙮ *ṅi*; ꙮ → ꙮ *ṭi*; ꙮ → ꙮ *ḍi*; ꙮ → ꙮ *ri*;

VOWEL SIGN II   The VOWEL SIGN II has two forms: ○ꙮ and ○ꙮ:

- ○ꙮ is used with characters possessing right descenders: ꙮ → ꙮ *cī* and with characters with final strokes that terminate at the baseline: ꙮ → ꙮ *kī*.

- ○ꙮ is used with characters that have rounded right edges: ꙮ → ꙮ *chī*;

VOWEL SIGN U   With characters that possess a right descender, the ○ VOWEL SIGN U is generally written as part of the descender: ꙮ → ꙮ *cu*. Otherwise, the vowel sign is written as an extension of the basic final stroke of a character: ꙮ → ꙮ *ku*, ꙮ → ꙮ *chu*, etc.

4.5   Consonant Conjuncts

Consonant clusters are typically limited to two consonants and are represented in four ways:

1. *Ligatures*   There are four consonant ligatures in Khojki:

     ৬   *kṣa*   =   ঽ KA   +   ◌/ VIRAMA   +   *SSA

     ঝ   *jña*   =   ঽ JA   +   ◌/ VIRAMA   +   ২ঃ NYA

     ঽ   *tra*   =   ঽ TA   +   ◌/ VIRAMA   +   ৬ RA

     ঽ   *dra*   =   ঽ DA   +   ◌/ VIRAMA   +   ৬ RA

2. *Full-Form of Initial Consonant + Visible Virama*   The most common method of representing conso-
nant conjuncts is by writing the full forms of the initial consonant and marking it with VIRAMA; eg.
pra = ঽ/৬ PA + VIRAMA + ZWNJ + RA

3. *Full-Form of Initial Consonant + Sub-Form of Final Consonant* This type of consonant conjunct is
typically C + RA.

```
Ms Ism K 15                                              [190-?]


Local paper; 44 folios; 65; 17 x 14 (13 x 10); 11 lines; multiple
sections bound with two sets of sewing: side-sewn and stab-sewn
with cords pierced through and laced into laminated layers form-
ing thin paper covers; spine rounded and backed; black ink.

  ᴍ = ḍ, ṛ, ṇ; ᴗ = dh; special form of conjuncts: ᴋ  = br;
   ᴋ = kr; ᴗ = sr; ᴗ = mr; ᴗ⎮= ny; ᴋᴏ) = ky; ᴗᴏ = ly; ᴠᴠᴏ = ty;
   ᴊᴏ= jy; ᴗ = sv; ᴗᴏ⎮= gn.
```

4. *Half-Form of Final Consonant*   In some cases, consonant conjuncts are written with the full-form of
the initial consonant and a half-form of the second consonant.[1] For instance, the conjunct *kya* may be
represented as ঽ০) (০০) YA → ০) half-YA). See section 5.4 for discussion.

```
Ms Ism K 15                                              [190-?]


Local paper; 44 folios; 65; 17 x 14 (13 x 10); 11 lines; multiple
sections bound with two sets of sewing: side-sewn and stab-sewn
with cords pierced through and laced into laminated layers form-
ing thin paper covers; spine rounded and backed; black ink.

  ᴍ = ḍ, ṛ, ṇ; ᴗ = dh; special form of conjuncts: ᴋ  = br;
   ᴋ = kr; ᴗ = sr; ᴗ = mr; ᴗ⎮= ny; ᴋᴏ) = ky; ᴗᴏ = ly; ᴠᴠᴏ = ty;
   ᴊᴏ= jy; ᴗ = sv; ᴗᴏ⎮= gn.
```

4.6   Nasalization

Nasalization is indicated by ◌̊ KHOJKI SIGN CANDRABINDU and ◌̇ KHOJKI SIGN ANUSVARA.

4.7   Consonant Gemination

Geminated consonants are marked by ◌̋ KHOJKI SIGN SHADDA. The sign is written above the consonant that
is doubled, eg. LA + LA is written ঽ̋ (as below), not as a conjunct character. This practice is similar to that in

---
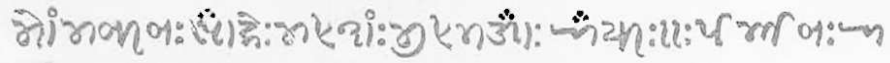
[1]   Asani 1992: 72, 73.

Gurmukhi, where the ◌̆ U+0A71 GURMUKHI SIGN ADDAK is used to denote doubled consonants, and also in the Arabic script, where the ◌̆ U+0651 ARABIC SHADDA is used.

4.8   Creation of New Characters

The Khojki script may be extended by producing characters to represent sounds not native to Sindhi. This is done by writing the sign *nukta* above a Khojki character whose phonetic value most closely matches that of the sound being accommodated. The *nukta* is written using the above-base combining character ◌̊ KHOJKI SIGN NUKTA. The NUKTA may be written with vowel letters, vowel signs, and consonant letters. In Khojki, *nukta* is used primarily to represent Arabic letters and sounds, eg. ◌́ A + NUKTA is used to represent ع U+0639 ARABIC LETTER AIN. The NUKTA has a variant, which is simply an inverted form: ◌̈.

The list of attested *nukta* letters in Khojki is:

- ◌̂ = ع U+0639 ARABIC LETTER AIN
- ◌́ = ع U+0639 ARABIC LETTER AIN
- ◌̂ = ع U+0639 ARABIC LETTER AIN
- ◌̂ = ق U+0642 ARABIC LETTER QAF
- ◌̈ = خ U+062E ARABIC LETTER KHAH
- ◌̂ = غ U+063A ARABIC LETTER GHAIN
- ◌̂ = ذ U+0630 ARABIC LETTER THAL, ز U+0632 ARABIC LETTER ZAIN, ض U+0636 ARABIC LETTER DAD, ظ U+0638 ARABIC LETTER ZAH
- ◌̂ = ط U+0637 ARABIC LETTER TAH
- ◌̈ = ف U+0641 ARABIC LETTER FEH
- ◌̂ = ح U+062D ARABIC LETTER HAH

4.9   Punctuation

Word Boundaries   Khojki separates words using ः KHOJKI WORD SEPARATOR. This character resembles characters such as ◌ः U+0903 DEVANGARI SIGN VISARGA. Khojki does not use VISARGA.

Sentence Boundaries   The ꠰ DANDA and ꠱ DOUBLE DANDA are used to mark the end of sentences.

11

End of line is typically marked with ፡ WORD SEPARATOR and ॥ DOUBLE DANDA:

DOUBLE DANDA is also used to mark verse sections. Typically, DOUBLE DANDA is written with WORD SEPA-RATOR to the left and right of verse numbers:

Section Marks   The section mark characters ⎯ KHOJKI SECTION MARK and ⁼ KHOJKI DOUBLE SECTION MARK appear frequently in Khojki manuscripts as punctuation that delimit the end of a section or another larger block of text. Both section marks are found in manuscripts. The KHOJKI SECTION MARK is generally used to mark the end of a sentence, while KHOJKI DOUBLE SECTION MARK is used to delimit larger blocks of text, such as paragraphs. Both generally extend to the margin of the text-block.
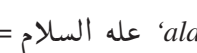
Verse Enumeration   The ⊡ KHOJKI VERSE MARKER is used to mark verse numbers in long Khojki texts. It is an above-base extending character that is written above numbers: ॥५२५॥ Its function may be considered similar to the Latin mark *vinculum*, which is written over numbers to indicate that they are to be considered a group.

Other Punctuation   Printed Khojki employs Latin punctuation marks, such as comma, which are already encoded in the UCS:

4.10   Abbreviations

Khojki has two abbreviations for Arabic benedictions. These are not proposed for encoding because they can be represented using other characters.

1. ‏‎ = عله السلام *'alai-hi as-salām* 'peace be upon him'.

2. ‏ﱟ‎ = صلى الله عليه و سلّم *ṣallā allāhu 'alai-hi wa sallam* 'God bless him and grant him salvation'.

```
Ms Ism K 13                                              [190-?]


Foreign paper, watermark: made in England for Shaik Ahmad Shaik
Dawood of Bombay; 190 folios; 376; 21 x 17 (16 x 13); 12 - 13
lines; multiple sections, stab-sewn; binding in poor condition;
covers missing; black ink.

Same scribe as Ms Ism K 6, Ms Ism K 8, Ms Ism K 11 and Ms Ism K
12; ৬ = û, o; ३ = z; ২৭ = j; س = shaddah; ़ৢ = abbreviation
for Arabic benediction ᶜalai-hi as-salām "peace be upon him";
়ৡ = abbreviation for Arabic benediction ṣalla allāhu ᶜalai-hi
wa sallam "God bless him and grant him salvation."
```

4.11   Number Forms and Unit Marks

Number forms and unit marks are found in Khojki documents (see Figure 17). The most common characters are signs for writing fractions: eg. ⌐ U+A833 ɴᴏʀᴛʜ ɪɴᴅɪᴄ ꜰʀᴀᴄᴛɪᴏɴ ᴏɴᴇ ꜱɪxᴛᴇᴇɴᴛʜ, ‖ U+A831 ɴᴏʀᴛʜ ɪɴᴅɪᴄ ꜰʀᴀᴄᴛɪᴏɴ ᴏɴᴇ ʜᴀʟꜰ, etc.; and currency marks: eg.) U+A838 ɴᴏʀᴛʜ ɪɴᴅɪᴄ ʀᴜᴘᴇᴇ ᴍᴀʀᴋ. These characters may be represented using those in the Common Indic Number Forms block (U+A830..U+A83F). See Pandey (2007) for more information on these characters.

4.12   Variants

1. ᴋʜᴏᴊᴋɪ ʟᴇᴛᴛᴇʀ ᴅʜᴀ is also written as ৪.

2. ᴋʜᴏᴊᴋɪ ʟᴇᴛᴛᴇʀ ʙᴀ is also written as ৭.

5   Implementation

5.1   Encoding Model

The encoding model for Khojki may be based on the model implemented for Devanagari. There may be a requirement to accommodate an idiosyncratic method of rendering consonant conjuncts for the script (see section 5.4).

5.2　Collation

The collating order for Khojki is based on Sindhi. The script may be sorted according to both the Devanagari-based and Persian-based arrangement of letters.

Devanagari-Based Sort Order　Independent vowel letters are sorted before consonant letters. Vowels with special signs (CANDRABINDU and ANUSVARA) appear at the head of the vowel order.

The collating order for special signs and independent vowels in Khoji is:

aṁ	*aṃ*	*a*	*ā*	*i*	*ī*	*u*	*ū*	*e*	*ai*	*o*	*au*

Dependent vowel signs are sorted in the same position as their independent shape. Consonant-vowel combinations are sorted first by consonant letter and then by vowel sign. A consonant with *virāma* is sorted last.

kaṁ	*kaṃ*	*ka*	*kā*	*ki*	*kī*	*ku*	*ke*	*kai*	*ko*	*kau*	*k*

The pattern for consonants is as follows:

ka	*kha*	*ga*	*ga*	*gha*	*ṅa*	*ca*	*cha*	*ja*	*ja*	*jha*	*ña*	*ṭa*

ṭha	*ḍa*	*ḍha*	*ṇa*	*ta*	*tha*	*da*	*da*	*dha*	*na*	*pa*	*pha*	*ba*

ba	*bha*	*ma*	*ya*	*ra*	*la*	*ḷa*	*va*	*śa*	*sa*	*ha*

Persian-Based Sort Order　Since the Sindhi language is now commonly written in an extended form of the Arabic script, Khojki letters may be sorted according to the Arabic arrangement of characters as extended for Sindhi.

ba	*ba*	*bha*	*ta*	*tha*	*ṭa*	*ṭha*	*pa*	*pha*	*ja*	*ja*	*jha*	*ña*

ca	*cha*	*da*	*dha*	*da*	*ḍa*	*ḍha*	*ra*	*sa*	*śa*	*ka*	*kha*	*ga*

ga	*gha*	*ṅa*	*la*	*ḷa*	*ma*	*na*	*ṇa*	*va*	*ha*	*ya*

5.3 Character Properties

The properties for Khojki characters in the Unicode Character Database format are:

```
11280;KHOJKI SIGN CANDRABINDU;Mn;0;NSM;;;;;N;;;;;
11281;KHOJKI SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
11282;KHOJKI LETTER A;Lo;0;L;;;;;N;;;;;
11283;KHOJKI LETTER AA;Lo;0;L;;;;;N;;;;;
11284;KHOJKI LETTER I;Lo;0;L;;;;;N;;;;;
11285;KHOJKI LETTER II;Lo;0;L;;;;;N;;;;;
11286;KHOJKI LETTER U;Lo;0;L;;;;;N;;;;;
11287;KHOJKI LETTER UU;Lo;0;L;;;;;N;;;;;
11288;KHOJKI LETTER E;Lo;0;L;;;;;N;;;;;
11289;KHOJKI LETTER AI;Lo;0;L;;;;;N;;;;;
1128A;KHOJKI LETTER O;Lo;0;L;;;;;N;;;;;
1128B;KHOJKI LETTER AU;Lo;0;L;;;;;N;;;;;
1128C;KHOJKI LETTER KA;Lo;0;L;;;;;N;;;;;
1128D;KHOJKI LETTER KHA;Lo;0;L;;;;;N;;;;;
1128E;KHOJKI LETTER GA;Lo;0;L;;;;;N;;;;;
1128F;KHOJKI LETTER GGA;Lo;0;L;;;;;N;;;;;
11290;KHOJKI LETTER GHA;Lo;0;L;;;;;N;;;;;
11291;KHOJKI LETTER NGA;Lo;0;L;;;;;N;;;;;
11292;KHOJKI LETTER CA;Lo;0;L;;;;;N;;;;;
11293;KHOJKI LETTER CHA;Lo;0;L;;;;;N;;;;;
11294;KHOJKI LETTER JA;Lo;0;L;;;;;N;;;;;
11295;KHOJKI LETTER JJA;Lo;0;L;;;;;N;;;;;
11296;KHOJKI LETTER JHA;Lo;0;L;;;;;N;;;;;
11297;KHOJKI LETTER NYA;Lo;0;L;;;;;N;;;;;
11298;KHOJKI LETTER TTA;Lo;0;L;;;;;N;;;;;
11299;KHOJKI LETTER TTHA;Lo;0;L;;;;;N;;;;;
1129A;KHOJKI LETTER DDA;Lo;0;L;;;;;N;;;;;
1129B;KHOJKI LETTER DDHA;Lo;0;L;;;;;N;;;;;
1129C;KHOJKI LETTER NNA;Lo;0;L;;;;;N;;;;;
1129D;KHOJKI LETTER TA;Lo;0;L;;;;;N;;;;;
1129E;KHOJKI LETTER THA;Lo;0;L;;;;;N;;;;;
1129F;KHOJKI LETTER DA;Lo;0;L;;;;;N;;;;;
112A0;KHOJKI LETTER DDDA;Lo;0;L;;;;;N;;;;;
112A1;KHOJKI LETTER DHA;Lo;0;L;;;;;N;;;;;
112A2;KHOJKI LETTER NA;Lo;0;L;;;;;N;;;;;
112A3;KHOJKI LETTER PA;Lo;0;L;;;;;N;;;;;
112A4;KHOJKI LETTER PHA;Lo;0;L;;;;;N;;;;;
112A5;KHOJKI LETTER BA;Lo;0;L;;;;;N;;;;;
112A6;KHOJKI LETTER BBA;Lo;0;L;;;;;N;;;;;
112A7;KHOJKI LETTER BHA;Lo;0;L;;;;;N;;;;;
112A8;KHOJKI LETTER MA;Lo;0;L;;;;;N;;;;;
112A9;KHOJKI LETTER YA;Lo;0;L;;;;;N;;;;;
112AA;KHOJKI LETTER RA;Lo;0;L;;;;;N;;;;;
112AB;KHOJKI LETTER LA;Lo;0;L;;;;;N;;;;;
112AC;KHOJKI LETTER LLA;Lo;0;L;;;;;N;;;;;
112AD;KHOJKI LETTER VA;Lo;0;L;;;;;N;;;;;
112AE;KHOJKI LETTER SHA;Lo;0;L;;;;;N;;;;;
112AF;<reserved>
112B0;KHOJKI LETTER SA;Lo;0;L;;;;;N;;;;;
112B1;KHOJKI LETTER HA;Lo;0;L;;;;;N;;;;;
112B2;KHOJKI VOWEL SIGN AA;Mn;0;NSM;;;;;N;;;;;
112B3;KHOJKI VOWEL SIGN I;Mc;0;L;;;;;N;;;;;
112B4;KHOJKI VOWEL SIGN II;Mc;0;L;;;;;N;;;;;
112B5;KHOJKI VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
112B6;<reserved>
112B7;KHOJKI VOWEL SIGN E;Mn;0;NSM;;;;;N;;;;;
112B8;KHOJKI VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
```

```
112B9;KHOJKI VOWEL SIGN O;Mn;0;NSM;;;;;N;;;;;
112BA;KHOJKI VOWEL SIGN AU;Mn;0;NSM;;;;;N;;;;;
112BB;KHOJKI SIGN VIRAMA;Mc;9;NSM;;;;;N;;;;;
112BC;KHOJKI SIGN NUKTA;Mn;7;NSM;;;;;N;;;;;
112BD;KHOJKI SIGN SHADDA;Mn;0;NSM;;;;;N;;;;;
112BE;KHOJKI DANDA;Po;0;L;;;;;N;;;;;
112BF;KHOJKI DOUBLE DANDA;Po;0;L;;;;;N;;;;;
112C0;KHOJKI WORD SEPARATOR;Po;0;L;;;;;N;;;;;
112C1;KHOJKI SECTION MARK;Po;0;L;;;;;N;;;;;
112C2;KHOJKI DOUBLE SECTION MARK;Po;0;L;;;;;N;;;;;
112C3;KHOJKI VERSE MARKER;Po;0;L;;;;;N;;;;;
```

5.4 Outstanding Issues

5.4.1 Encoding Model

1. **Half-form of final letter of a consonant conjunct**   Consonant conjuncts in Khojki are generally written with a visible VIRAMA. For instance, the conjunct *kya* is typically represented as 𑈈/𑈀:

    𑈈 KA + ◌𑈾 VIRAMA + ZWNJ + 𑈀 YA

    Alternately, this conjunct may also be written in a rather unconventional manner: the full-form of the initial consonant letter and the half-form of the second consonant letter:[2] 𑈈𑈀. Despite the visual disjunction, 𑈈𑈀 represents *kya*, not *kay*, and is semantically identical to 𑈈/𑈀. Given this, the expected character sequence would be

    𑈈 KA + ◌𑈾 VIRAMA + ZWJ + 𑈀 YA   (*kya*)

    not

    𑈈 KA + 𑈀 YA + ◌𑈾 VIRAMA + ZWJ   (*kay*)

    Here, the appearance of the conjunct is the reverse of the semantic value of the characters. The letter KA remains visually unchanged and unmarked by VIRAMA, but the elision of its inherent vowel is indicated by the half form of YA. Thus, the rendering behavior of VIRAMA in this case affects not the character it is written after (KA), but the one that follows it (YA): 𑈀 → 𑈀 (semantically YA, but visually half-YA).

    This behavior cannot be produced by the typical method of producing conjuncts by applying VIRAMA to the consonant letter in question.

5.4.2 Characters

1. **Consonant letter DDDHA**   Some Khojki manuscripts use 𑈍 KHOJKI LETTER DDA (/ɖ/) to represent *ṛa* (/ɽ/), which is analogous to Devanagari ड़ U+095C DEVANAGARI LETTER DDDHA. The sound is an allophonic variant of DDA, which appears intervocalically. Other sources represent this sound using the glyph 𑈇. Although 𑈇 is assigned here to KHOJKI LETTER LLA, it may be possible to propose it as KHOJKI LETTER DDDHA, despite the fact that the glyph 𑈇 generally represents LLA in other Indic scripts. The reassignment of 𑈇 from LLA to DDDHA would be further supported by the fact that Sindhi does not have the sound /ḷ/ that is generally represented by the letter LLA in several north Indic scripts.

---

[2] Asani 1992: 72, 73.

The following specimen shows ◌ used for writing *ṛa*:

```
P Ism K 19                                    Bombay 1985 s./1929

Dash Avatār nāno pûro. 3rd edition.

Publisher: Rahemtulā Vīrjī, honorary secretary, The Recreation
Club Institute, Press Dept. at the Khoja Sindhi Printing Press;
500 copies; 20 pp.; 22 x 17; paper covers; in good condition;
same as P Ism K 31.

Does not use ":" to separate individual words;  ◌ = ṛ;  ◌ = ḍ
◌  = ṇ.
```

The following specimen shows ◌ as being analogous to ળ U+0AB3 GUJARATI LETTER LLA (*ḷa*):

```
vowel "ṛu" is indicated by subscript " ◌ " ; (5) the following
conjuncts introduced: ◌ = Gujarati ચ◌ ;  ◌   = Gujarati ળ  ;
  ◌ = Gujarati ધ ;   ◌   = Gujarati ન
```

2. **Consonant letter ssA** Khojki does not have a character ssA to represent *ṣa* (/ṣ/), analogous to ष U+0937 DEVANAGARI CHARACTER SSA. It does, however, have the glyph ◌ KSSA (Devanagari क्ष), which is a consonant conjunct produced with the sequence KA + VIRAMA + SSA in most Indic scripts. The absence of ssA in Khojki, therefore, requires a solution to the composition of ◌. One approach may be to define a script-specific rule that produces the ligature KSSA using ◌ KHOJKI LETTER SHA: KA + VIRAMA + SHA (Devanagari क्श *kśa*); but, this approach introduces a host of other problems regarding character semantics. Another solution might be to encode ◌ as an atomic character.

3. **Consonant letter QA** Khojki *ka* is generally represented by the glyph ◌, but in some manuscripts and printed books it takes the form ◌. These forms at times may appear together within a single document. However, in some documents the form ◌ is used to represent a transcription of ق ARABIC LETTER QAF (/q/). In this case, the glyph represents a unique Khojki character QA, analogous to क़ U+0958 DEVANAGARI LETTER QA. Yet, in other sources, the sound /q/ is represented using KA + NUKTA as ◌. The following specimen shows ◌ used to uniquely represent *QA:

```
P Ism K 14                                    Bombay 1962 s./1905

Elam sār saṅgrah rāg mālā

Publisher: Lāljībhāî Devrāj, Khojā Siñdhī Chāpakhānû; 147 pp. in
three paginations; 23 x 16; hardboard covers, binding in poor condi-
tion; pages in fair condition; frontispiece: full portrait of Aga
Khan III in official regalia; preface on the necessity of publishing
this book (pp. 1.1-3); tavsīlo (index): pp. 2.135-138;
list of publications of the Khoja Siñdhī Printing Press pp. 3.1-6.

  ◌   = gāf;  ◌ = g, gr;  ◌ = ʿain;  ◌ = 1?;  ◌ = vowel-less
consonant.
```

and the same glyph used to represent KA:

```
P Ism K 11                                          Bombay 1959 s./1902

Ārtī:dhûā:vakhat:trejī:sāñjījā:ćogaḍīā:5

Publisher: Khojā Lāljībhāî Devrāj, Khojā Siñdhī Ćhāpākhānû for
the Imāmī Ismāilī school (nisāl). 64 pp.; 16 x 12; covers miss-
ing; binding in poor condition; tafsīlo (index): pp. 63-64;
reverse side of title page includes list of sale publications of
the Khojā Siñdhī Printing Press and an announcement of the pub-
lication of other works.

    ऊ = z;  ꣑ = j;  ꣒ = k;  ᷓ = ᶜain;  ꣓ = g;  ꣔ = gain;
shaddah for doubled consonant or emphasis.
```

A specimen from a printed book shows KA represented not only with 𑈈, but also with 𑈉. Both glyphs are used in the word *kalāma* (𑈉𑈌𑈓 / 𑈈𑈌𑈓) and are semantically identical.
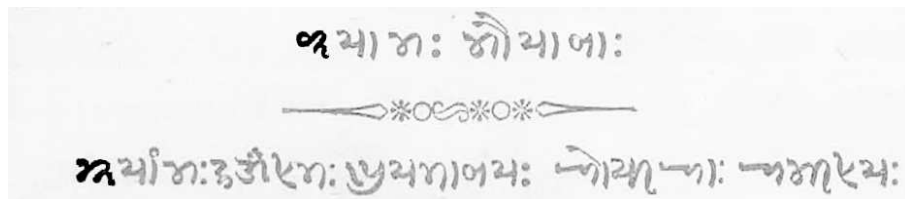
4. **VERSE MARKER**   The mark ⬚ is written over verse numbers in long Khojki texts. As described in section 4.9, the mark is an extending character that spans all of the digits in a number. It is possible that this specialized rendering behavior would make implementation of this function difficult. It may be possible to use ◌̅ U+0305 COMBINING OVERLINE to represent the Khojki verse marker. The drawback to this solution is that COMBINING OVERLINE would visually represent the verse marker, but it would not capture the semantic properties of the Khojki character.

6   References

Asani, Ali S. 1992. *The Harvard Collection of Ismaili Literature in Indic Languages: A Descriptive Catalog and Finding Aid.* Boston: G. K. Hall & Co.

Grierson, George A. 1905. *The Linguistic Survey of India.* Vol. VIII. Indo-Aryan family. North-western group: pt. I. Specimens of Sindhī and Lahndā. Calcutta: Office of the Superintendent of Government Printing, India.

Khakee, Gulshan. 1981. "The Dasa Avatara of Pir Shams as Linguistic and Literary Evidence of the Early Development of Ismailism in Sind." In *Sind Through the Centuries.* Proceedings of an International Seminar held in Karachi in Spring 1975 by the Department of Culture, Government of Sind. Pages 143–155. Edited by Hamida Khuhro. Karachi: Oxford University Press.

Pandey, Anshuman. 2007. "Proposal to Encode North Indic Number Forms in ISO/IEC 10646." ISO/IEC JTC1/SC2/WG2 N3367 L2/07-354. October 7, 2007. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3367.pdf`

Shackle, Christopher and Zawahir Moir. 1992. *Ismaili Hymns from South Asia: An Introduction to the Ginans.* London: School of Oriental and African Studies, University of London.

Roman characters.	Dēvanāgarī.	Khudāwādī.*	Shikārpurī.	Sakkar.	THATTAI.		Lārrī.	Waṅgāī.	Rajjāi.	Khwājāa.	MAIMONA.		Sewhāṇī Bhābhiriā.	Southern Lahndī.
					Luhānāi.	Bhāṭiās.					Thattā.	Haidarābād.		
a	अ	m	m	m	⁄⁄∧	m	m	ૐ	ɲ	⌐	⌐	⌐	᠈m	⅏
ā	आ	m	m	m	⁄⁄∿	m	m	⸚	ɲ	⌐⎸	⌐	⌐	᠈m	⅏
i	इ	†6	†6	†6	°.₁	⍭	⋑	⩴	⫶	⋎	⸭	⋑	⊂	6
ī	ई.	†6	†6	†6	°.₀₁	⍭	⋥	⩴	⫶	ɟ̈	⸭	⋑	⊂	6
u	उ	m	m	m	6	6	6	6	6	᠊ᢖ	⊂	6	⊙	⅏
ū	ऊ	m	m	m	6	6	6	6	6	ᢖ	⊂	6	⊙	⅏
ē	ए	m	m	m	⋑	⍭	⋎	⊏	⫶⎸	⸦⸦	⊱	⋎	᠈m	6
ai	ऐ	m	m	m	⁄⁄∿	m	m⋥	⸚⊱⸽	⍬⋮	⌐⋎	⌐⋮	⋑⋑	᠈m	⅏⌐
ō	ओ	m	m	m	6	6	⸦	⋎	⊂	⸦ᢖ	⊂	⋎	᠈m	⅏
au	औ	m	m	m	6	6	6	6	6	⌐ᢖ	ᙗ⊂	⌐6	᠈m	⅏⅏

* The Khudāwādī is the character used at Haidarābād and by most of the respectable merchants throughout the country. That of the Shikārpur and Sakkar merchants differs but little from it.
† These are, however, never written in commencing a word. In lieu of them ㎜ is used.

Figure 1: [Inventory of Khojki vowel letters from Grierson (1905: 15).

Roman characters.	Dévanāgari.	Khudāwādi.	Shikārpuri.	Sakkar.	Thaṭṭai. Lohānāe.	Thaṭṭai. Bhāṭiāa.	Lāṛai.	Wāṅgai.	Rājai.	Khwājae.	Maimons. Thaṭṭa.	Maimons. Haidarābād.	Sewhāṇī Bhābhirāe.	Southern Lahndā.
ka	क													
kha	ख													
ga	ग													
gga	ॻ													
gha	घ													
ṅa	ङ													
cha	च													
chha	छ													
ja	ज													
jja	ॼ													
jha	झ													
ña	ञ													
ṭa	ट													
ṭra	ॾ													
ṭha	ठ													
ṛa	ड़													
ḍḍa	ॺ													
ḍra	ॻ													
ḍha	ढ													
ṇa	ण													

Figure 2: [Inventory of Khojki consonant letters from Grierson (1905: 16).

20

Roman characters	Dēvanāgarī	Khudāwādī	Shikārpurī	Sakkar	THAṬṬĀI. Luhānās	Bhātiās	Lārāi	Waṅgāi	Rājāi	Khwājās	MAIMONS. Thaṭṭa	Haidarābād	Sewhāṇī Bhābhirās	Southērn Lahndā
ta	त													
tha	थ													
da	द													
dha	ध													
na	न													
pa	प													
pha	फ													
ba	ब													
bba	ॿ													
bha	भ													
ma	म													
ya	य													
ra	र													
la	ल													
va	व													
śa	श													
sha	ष													
sa	स													
ha	ह													
Numerals				1	2		3	4	5	6	7	8	9	10

Figure 3: [Inventory of Khojki consonant letters from Grierson (1905: 17).

Table showing correspondence of
intial vowels in the Roman, Urdu,
Sindhi, Gujarati and Khojki scripts

Roman	Urdu	Sindhi	Gujarati	Khojki#
a	ا	ا	અ	⌐, ⌐
ā	آ	آ	આ	⌐ा, ⌐ा
i*	اِ	اِ	ઈ	ॐ [३०] (ᔉ र?)
ī*	ايِ	ايِ	ઈ	ॐ [२०] (३९?)
u*	اُ	اُ	ઉ	ऽ [८]
ū*	اوُ	اوُ	ઊ	ऽ [८] (ᘊ?)
o	او	او	ઓ	ᖑ [→ो, ᘊो] (ᔉ?)
e	اے	اى	એ	ᖷ [→ो] (ᗉ?)
ai	اے	اى	ઐ	[→ᖷ, →ो]
au	او	او	ઔ	[→ऽ, →ᖷ, →ो]

*On account of ambiguities in the Khojkī script, initial and
independent short or long "i" vowels in Khojkī texts are trans-
literated as "î". Similarly, short or long "u" vowels are trans-
literated as "û".

#[   ] documented variant; (?) variant listed by G. Allana in
Sindhī Ṣūratkẖatī (Hyderabad, Pakistan: Sindhī Zabān Publica-
tions, 1964) or G. Stack in **A Grammar of the Sindhi Language**
(Bombay, 1849); (   ) alternate sound employed.

Figure 4: Inventory of Khojki independent vowel letters (from Asani 1992: 678).

**Table showing correspondence of
non-initial vowels in the Roman, Urdu,
Sindhi, Gujarati and Khojki scripts**

Roman	Urdu	Sindhi	Gujarati	Khojki#
bad	بَد	بَد	બદ	ઝદ [ઝદ]
bād	باد	باد	બાદ	ઝાદ
din	دن،	دن،	દિન	દન
dīn	دين	دين	દીન	ધન
but*	بُتّ	بُتّ	બુત	ભુન
būt*	بوت	بوت	બૂત	ભુન
so	سو	سو	સો	ભો [ભદ]
be	بے	بی	બે	ભે [ભ]
hai	بَے	بَی	ઐ	[ઝ", ઝ–ઝ]
nau	نوَ	نوَ	નૌ	[ભૌ, ભ–ભ]

*On account of ambiguities in the Khojki script, short and long
"u" vowels in Khojki texts are transliterated as "û".

#[   ] documented variant; (?) variant listed by Allana or Stack;
(   ) alternate sound employed.

Figure 5: Inventory of Khojki dependent vowel signs (from Asani 1992: 679).

**Table showing correspondence of
consonants in the Roman, Urdu,
Sindhi, Gujarati and Khojki scripts**

Roman	Urdu	Sindhi	Gujarati	Khojki#
b	ب	ب	બ	ચ [m, મ]
ḅ		ب		મ
bh	ک	پ	બ	ત, મ [ભ]
t	ت	ت	ત	ч
th	ته	ٹ	થ	પ [म]
ṭ	ط	ٹ	ટ	ટ, ટ, (મ?)
ṭh	ط	ٹ	δ	ઠ
ṣ	ث	ٹ	(મ)	(પ)
p	پ	پ	પ	ч (પ?)
ph	ک	ق	ફ	ટ
j	ج	ج	૪	ા, મ, ૩ (૩?)
dy		ج		૩ (૩?)

#[   ] documented variant; (?) variant listed by Allana or Stack;
(   ) alternate sound employed.

Figure 6: Inventory of Khojki consonants: *b* to *dy* (from Asani 1992: 680).

24

Roman	Urdu	Sindhi	Gujarati	Khojki‡
jh				
ny			म	२: (२५?)
ć			२|	ज, (ए?)
ćh			७	९१, (ल?)
ḥ			(ड)	(२)
kh			(ज्ञ)	(स)
d			६	६, २
dh			६	२१, २, (३?)
ḍ				६
ḍ			S	म, म (३,३,२१?)
ḍh			६	२ (२?)
ż			(ज्ञ)	(ज, २५)
r			२	८
ṛ			७	म, ७ [म]
ṛh				
z			३५	ज [२५]
s			म	१५

Figure 7: Inventory of Khojki consonants: *jh* to *s* (from Asani 1992: 681).

Roman	Urdu	Sindhi	Gujarati	Khojki♯
sh	شؔ	شؔ	શ	...
ṣ	ص	ص	(સ)	(ય)
ẕ	ض	ض	(ઙ)	(ઝ઼, ઝ)
ṭ	ط	ط	(ત, ટ)	(ય, ટ)
ẓ	ظ	ظ	(ઙ)	(ઝ઼, ઝ)
ʿ	ع	ع	(અ)	... (ન)
ġ	غ	غ	(ગ)	... (ગ)
f	ف	ف	ફ	...
q	ق	ق	(ક)	... (ય)
k	ک	ک	(ક)	... [ન]
kh	کھ	ک	ખ	... (ખ?)
g	گ	گ	ગ	ગ [ગ, ગ઼]
ṅg		گ		...
gh	گھ	گھ	ઘ	ઘ (ઝ?)
ng		گ	ઙ	... (ર?)
l	ل	ل	લ	... [ન]

Figure 8: Inventory of Khojki consonants: *sh* to *l* (from Asani 1992: 682).

Roman	Urdu	Sindhi	Gujarati	Khojki#
m

 |

 | મ |

 |
| n |

 |

 | ન | ન,બ |
| ṇ | |

 | ણા | ᨶ (ᨶ:?) |
| w/v |

 |

 | વ | વ [વ] |
| h |

 |

 | ૭ | ર [ζ] |
| y |

 |

ય	૦ટે,૦૦૬, (ﻭ?)
ksh			ક્ષ	ક્ષૂ , (ક્ષ ?)
gn			જ્ઞ	૩ન
tr			ત્ર	ત્ર (ત્ર ,ट?)
dhr			ધ્ર	૩

Nasalization is indicated by ñ or m̐

Figure 9: Inventory of Khojki consonants: *m* to *dhr* (from Asani 1992: 683).

27

Table 1

Khojkī Script: Vowel System (early 20th century)

(a) Initial and Independent

a	ā	i	ī	u	ū	o	e	au	ay

(b) Medial and Final
(with the Consonant र)

a	ā	i	ī	u	ū	o	e	au	ay

Table 2

Khojkī Script: Consonants (early 20th century)

k	kh	g	gh	ṅg	ṅg
c	ch	j/jh	j/jh	ny	dy
ṭ	ṭh	ḍ	ḍh	ṇ	
t	th	d	dh	n	ḍ
p	ph	b	bh	m	ḍ
y	r	l	v	s	sh
dr	tr	h	r/l		

28

	1		2		3		4		5
⊃	*a*	⊃))	*lā*	ñ)	*to*	᾽ᾳ)	*hā*	⋎	*r*
⊃⟍	*ā*	ᵒᾷ	*i*	ᴦ	*u*	ᴜᴄ	*e*	⊃ᴜ	*g²*
ᴣᴢ	*k*	⤳	*kh*	ᴣᴧ	*g¹*	ᴱᴜ	*gh*	ᴣ8	*ṅ*
ᴜᴦ	*c*	ᴢᴇ	*ch*	ᴣᴦ	*j¹*	ᴣᴦ	*j²*	ᴢ8	*ñ*
ᴢ	*ṭ*	ᴣ	*ṭh*	ᴦᴍ	*ḍ*	ᴜᴏ	*ḍh*	ᴦᴍ	*ṇ*
ᴨ	*t*	ᴜᴸ	*th*	ᴱ	*d¹*	ᴏᴢ	*dh*	ᴏᴢ	*n*
ᴧ	*p*	ᴧ	*ph*	ᴥ	*b¹*	ᴍᴍ	*bh*	ᴣᴦ	*m*
ᴏᴏ)	*y*	ᴧ	*r*	ᴣᴦ	*l*	ᴄᴧ	*v*	ᴦᴦ	*d²/dh*
ᴨ	*s*	ᴣ	*dr*	ᴧ	*tr*	ᴣᴦ	*h*	ᴣᴦ	*b²/bh*

Figure 10: Traditional arrangement of Khojki characters, beginning with the traditional Sindhi invocation ᴨᴣᴧ) ñ)ᴣᴧ*ᴋ* *allā tohār* (from Shackle and Moir 1992: 35).

⊃	*a*	⊃)	*ā*	ᵒᾷ	*i*	ᴦ	*u*	ᴜᴄ	*e*
ᴼ	*o*	ᴣᴢ	*k*	⤳	*kh*	ᴣᴦ/ᴱᴜ	*g*	ᴣᴦ/ᴱᴜ	*gh*
ᴣ8	*ṅ*	ᴜᴦ	*c*	ᴢᴇ	*ch*	ᴣᴦ	*j/jh*	ᴣᴦ	*j/jh*
ᴢ8	*ñ*	ᴢ	*ṭ*	ᴣ	*ṭh*	ᴦᴍ/ᴜᴏ	*ḍ*	ᴦᴍ/ᴜᴏ	*ḍh*
ᴦᴍ	*ṇ*	ᴨ	*t*	ᴧ	*tr*	ᴜᴸ	*th*	ᴱ/ᴏᴢ	*d*
ᴣ	*dr*	ᴱ/ᴏᴢ	*dh*	ᴏ	*n*	ᴧ	*p*	ᴧ	*ph*
ᴧ/ᴍ	*b*	ᴧ/ᴍ	*bh*	ᴣᴦ	*m*	ᴏᴏ)	*y*	⋎	*r*
ᴣᴦ	*l*	ᴄᴧ	*v*	ᴨ	*s*	ᴣᴦ	*h*		

Figure 11: Arrangement of Khojki characters based on the traditional Indic phonological pattern (from Shackle and Moir 1992: 36).

Roman	Devana-gari	Gujarati	Sindhi	D MS 1815	Kx MS 1737	Khojaki Primer 1932
a	अ					
ā	आ					
i	इ					
ī	ई					
u	उ					
ū	ऊ					
e	ए					
o	ओ					
ai	ऐ					
au ou	औ					

Figure 12: Comparison of Khojki with other scripts: *a* to *au* (from Khakee 1981).

Roman	Devana-gari	Gujarati	Sindhi	D MS 1815	Kx MS 1737	Khojaki Primer 1932
k	क					
kh	ख					
g	ग					
gh	घ					
c	च					
ch	छ					
j	ज					
jh	झ					
ṭ	ट					
ṭh	ठ					
ḍ	ड					

Figure 13: Comparison of Khojki with other scripts: *ka* to *ḍa* (from Khakee 1981).

Roman	Devana-gari	Gujarati	Sindhi	D MS 1815	Kx MS 1737	Khojaki Primer 1932
ḍh	ट	૭	ૐ	—	૭	૭
ṇ	ण	૭	૭	—	૭ [8]	૭
t	त	૮	૭	त	૮	૮
th	थ	૫	૭	૫	૫	૫
d	द	૬	>	૭	૬	૬
dh	ध	૬ં	૭	૭	૭ [9]	૭૭
n	न	૭	૭	न	न	૭
p	प	૭	૭	૫	૫ [10]	૫
f	फ	૬	૭	૭	— [11]	૭
ph	—	—	૭	—	—	—
b	ब	૭	૭	ब	૫	૫

Figure 14: Comparison of Khojki with other scripts: *ḍha* to *ba* (from Khakee 1981).

Roman	Devana-gari	Gujarati	Sindhi	D MS 1815	Kx MS 1737	Khojaki Primer 1932
bh	भ	૭	૭	૫	૭ [12]	૫
m	म	૫	૭	૫	૫	૫
y	य	૨	૭	૫	— [13]	૭
r	र	૨	૭	૨	૭	૭
ru	—	—	—	૭	—	—
l	ल	૭	૭	૭	૫	૫
v	व	૭	૭	૭	व	व
v / s	श	૨૦	૭	श	— [14]	૭
ṣ	ष	૭	૭	—	— [14]	—
s	स	૨૦	૭	૨૦	૫	૫
h	ह	૭	૭	૭	૭	૭

Figure 15: Comparison of Khojki with other scripts: *bha* to *ha* (from Khakee 1981).

Roman	Devana-gari	Gujarati	Sindhi	D MS 1815	Kx MS 1737	Khojaki Primer 1932
r	ड़	—	—	—	म [8]	—
ks	क्ष	—	—	—	—	—
gn	ज्ञ	—	—	—	—	—
tr	त्र	—	—	—	—	—
dv	—	—	—	—	—	—
dy	—	—	—	—	[15]	—
ny	—	—	—	—	[16]	—
z	—	—	—	—	— [17]	—
dhr	—	—	—	—	—	—
sr	—	—	—	—	—	—

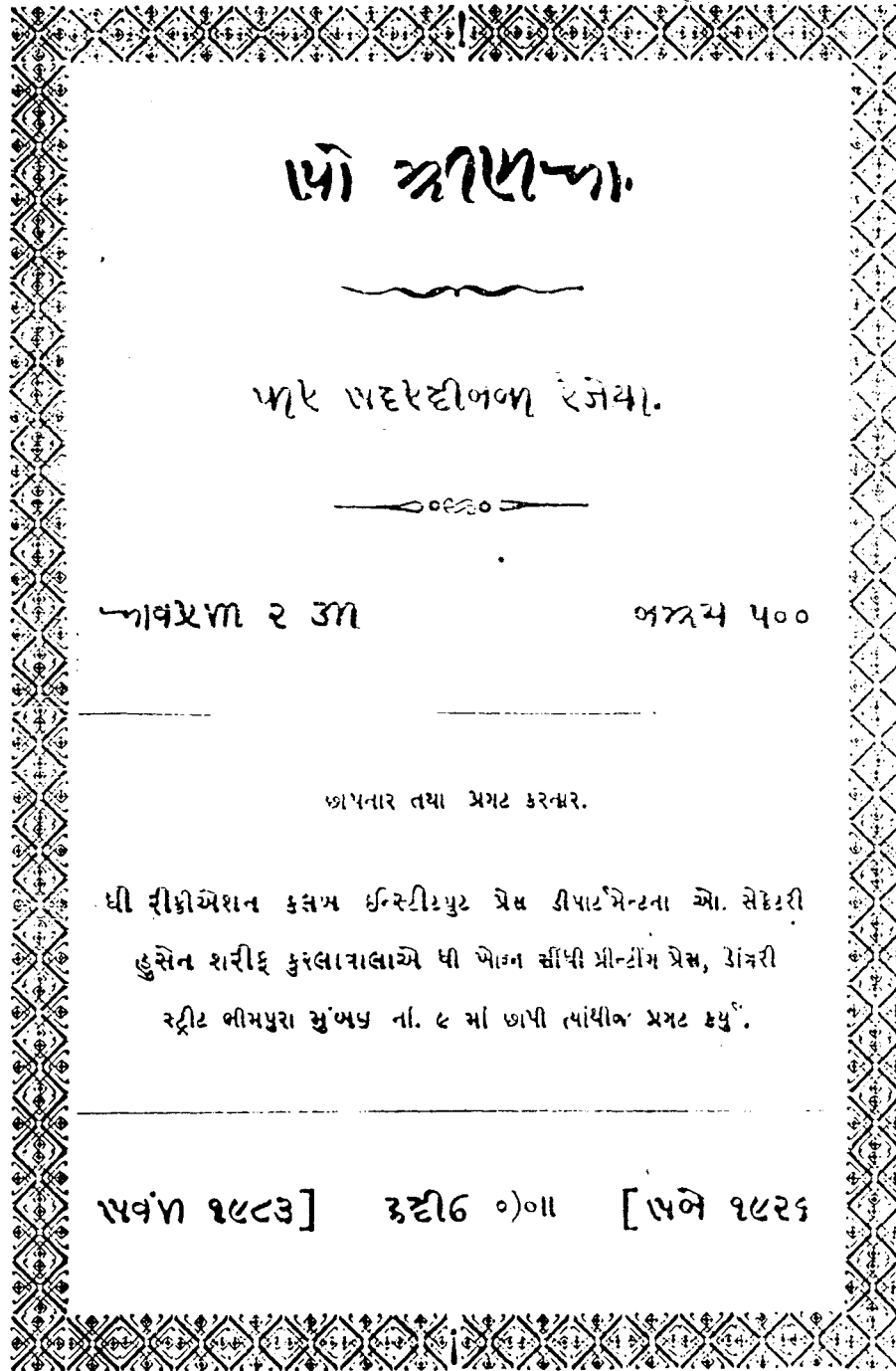Figure 16: Comparison of Khojki with other scripts: (*ra* to *sra* (from Khakee 1981).

Figure 17: Title page of the 𑈶𑈴 𑈀𑈬𑈭𑈲𑈬 (*So Kīrīā*) printed in Khojki and Gujarati (from Shackle and Moir 1992: 33). North Indic number forms and the U+A838 NORTH INDIC RUPEE MARK are used to indicate the price of the item.

Figure 18: Pages from a Khojki book printed with the metal fonts cut by Laljibhai Devraj.

Ms Ism K 1   [Bhadneshwar] [192-?]

Foreign wove paper; watermark: E.R. & Co., Made in Austria, chariot; 96 folios; 184; 20 x 16 (17 x 13); 12 - 16 lines; stab/sewn binding laced on hard board with half-cloth covering with decorative paper; black ink.

Scribe: Khojā bhāî Ladhū Devjī (65; 179-80); does not use short "i" vowel; 𑈈 = ḍ ṛ ṇ; 𑈇 = j; 𑈇 = h; 𑈈 = m; 𑈈 = o; 𑈈 = pr.

Pp. 41-43 torn, no other significant damage; **tafsilo**: p. 182 [f. 91v]; collected from Bhadneshwar Jamat, 12th July 1970; sequence of works from pp. 70-102 corresponds to that in P Ism K 27, pp. 19-30.

[**Garbīs**.   <u>Ref</u>: D 61-64 I 177 N 9,44,81,100 P 300]

1-2
(a.t.)   ejī:ād:gūr:samas:mûnivar:jā:ṇ
                        𑈈𑈈𑈈 (Khojki script)

10 verses, G, Pīr Shams

2-4
(rāg ćāl)   āj:ānañd:pāmeā:manke:sat:gûr
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

4-6
(same rāg as above)   gûr:āveā:bījī:rātke:avī
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

6-8
(same rāg as above)   gûr:āvatā:sarve:ratake:ganān
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

8-11
(same rāg as above)   ke:tame:amīras:pījo:dhanane:rātake
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

11-13
(same rāg)   maññhe:varasete:abharat:nûrake
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

13-15
(same rāg)   tame:ramajo:dhanane:rātake
                        𑈈𑈈𑈈 (Khojki script)

19 verses, G, Pīr Shams

15-18
(same rāg)   tame:japajo:dhanane:rātake
                        𑈈𑈈𑈈 (Khojki script)

22 verses, G, Pīr Shams

18-20
(rāg ćaltī)   bhûlo:bhûlo:te:bhamaraṇo
                        𑈈𑈈𑈈 (Khojki script)

22 verses, G, Pīr Shams

20-22
(rāg ṭhāl/ḍhāl?)   sat:mārg:samas:pīrake
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

22-25
(same rāg)   sate:pāvaṛ:pīo:dhanane:rātke
                        𑈈𑈈𑈈 (Khojki script)

21 verses, G, Pīr Shams

25-26
(rāg ćaltī)   ćeto:ćeto:te:ćañćal:ćetīarelol
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

26-29
(rāg ṭhāl/ḍhāl?)   gûrû:āveā:sātmī:rātake
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

29-31
(rāg dohrā)   satgûr:samas:kaere:gāfalo
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

31-32
(rāg ćaltī)   gūr:kahere:gāfalo:sābhaṛo
                        𑈈𑈈𑈈 (Khojki script)

18 verses, G, Pīr Shams

33-35
(rāg ṭhāl/ḍhāl?)   tathī:āṭhamī:āveā:gāmanā:lokake
                        𑈈𑈈𑈈 (Khojki script)

19 verses, G, Pīr Shams

Figure 19: Khojki text written by Ali S. Asani for his *The Harvard Collection of Ismaili Literature in Indic Languages* (1992: 56–57).