

Transcription Rules and Sikhri Font

Gurmukhī Script

Written documentation is perhaps the most effective and commonly used method of guarding and transmitting sacred knowledge. Therefore, the examination of linguistics within a scripture is imperative within the realm of theological studies. Script, a systematic process of alphabets utilized as representations for audible sounds, is the primary medium for this documentation. An exploration of the Gurū Granth Sāhib¹, the revelatory canonized scripture as authenticated by the Gurū, would be inadequate without an understanding of its distinctive communicative script, Gurmukhī.

Origin

Literally meaning ‘of the Gurū-oriented²,’ Gurmukhī is an exceptionally phonetic script that traces its roots to the Pañjāb region of South Asia. Though many of the letters have been confirmed as predating the Gurū-Prophet period (1469-1708), it was Gurū Nānak Sāhib (Nānak I) who first utilized these characters in a systematic manner. Gurū Aṅgad Sāhib (Nānak II) proceeded to institutionalize these characters into a single, coherent script.

While Gurmukhī was particularly developed as a means to record and convey the Divine Word, it also innovatively empowered the oppressed society throughout South Asia. Before the advent of the script, a majority of the population was prohibited from learning the language of religious texts³. The Sikh doctrine rejected ideas of social inequalities and embraced ideas of personal and direct interaction with the Gurū Granth Sāhib⁴, a revelatory text imparting the experience and process of experience with Divinity. The content of the Gurū Granth Sāhib shares timeless and universal wisdom and consists of various languages, contributors, and musical prescriptions. The script, however, is singular in its expression.

Linguistics

Gurmukhī is an extraordinarily phonetic script that allows it the flexibility to accurately detail all spoken languages – every letter, syllable, and word is pronounced as inscribed. Several vernaculars throughout Pañjāb, as well as Braj (precursor to Hindī), Fārsī (later Urdū), and Pañjābī, adopted the Gurmukhī script upon its introduction.

In Gurmukhī, each audible sound prevalent in regions that the Gurū Granth Sāhib’s contributors resided in or traveled to, was assigned a single character-letter and incorporated into the script. Three vowel bases were created and then further partitioned into ten distinct vowel characters. Tonal variations indicated changes in pitch, and subscripts were utilized to signify compound consonants. Symbols indicative of nasalization were developed, while characters also serve as representatives for an assortment of other sounds.

The Gurmukhī script has also adopted several characters subsequent to the Gurū-Prophet period (1469-1708), primarily from Perso-Arabic. However, because the Perso-Arabic influence followed the time of its compilation, these letters have not been utilized in the primary source of the Sikhs, the Gurū Granth Sāhib.

Because of shortcomings in the English language in general, and the Roman script in particular, it has been difficult to develop a technique with the capacity to consistently and accurately transfer text from Gurmukhī characters to Roman characters. This shortcoming has posed a problem Sikhs and non-Sikhs alike as not all individuals aspiring to study Sikh texts possess an operational knowledge of Gurmukhī. Although numerous attempts have been made in the last several years to develop such a system for transferring the two scripts, modern linguistics standards and thorough research have finally made Gurmukhī-Roman transcription a reality.

Transcription is the formulaic method of replicating a script through the substitution of alphabet characters while preserving all audible sounds in the original script. The majority of the characters used to transcribe Gurmukhī into English have been derived from a universally recognized system of linguistics. For those Gurmukhī letters lacking an equivalent representation in the Roman alphabet, a symbol comparable to a character with an associable sound has been generated to represent those alphabets. In order to read Gurmukhī in English characters, one must first read and understand the guidelines for transcription.

As each Gurmukhī letter represents a unique audible sound, each Gurmukhī letter has been assigned its own symbol for transcription into the Roman script.

Endnotes

¹ All examples and conclusions included in this entire document are based on internal evidence from the Gurū Granth Sāhib, the revelatory scriptural canon compiled by the Gurū-Prophet.

² The term Gurmukhī is often incorrectly assumed to mean ‘from the mouth of the Gurū.’

³ The caste system, an institutionalized apartheid, has been rampant in South Asia for several centuries. In this society, only the highest-caste Brahman is permitted to read and interpret the scriptural texts written in the Sanskrit script. To combat this notion of elitism, Gurmukhī was established as a script for the people, embracing the spoken language in an adequate yet simple alphabet.

⁴ Before Gurū Gobind Singh Sāhib’s (Nānak X) demise in 1708, the spiritual authority of the Sikh nation was conferred upon the Gurū Granth Sāhib. Nine years earlier, the Tenth Nānak had given physical authority of the Sikhs to the Khālsā Panth, the order of committed Sikhs. As the Khālsā Panth operates in accordance with the Gurū Granth Sāhib, these two institutions jointly serve as Gurū Granth-Panth, the present Gurū of the Sikhs.

Transcription Rules

I. Consonants

		C1	C2	C3	C4	C5
R1	Gurmukhī Transcription	ੳ -	ਅ a	ੲ -	ਸ s	ਹ h
R2	Gurmukhī Transcription	ਕ k	ਖ kh	ਗ g	ਘ gh	ਙ ṅ
R3	Gurmukhī Transcription	ਚ c	ਛ ch	ਜ j	ਝ jh	ਞ ṇ
R4	Gurmukhī Transcription	ਟ ṭ	ਠ ṭh	ਡ ḍ	ਢ ḍh	ਣ ṇ
R5	Gurmukhī Transcription	ਤ t	ਥ th	ਦ d	ਧ dh	ਨ n
R6	Gurmukhī Transcription	ਪ p	ਫ ph	ਬ b	ਭ bh	ਮ m
R7	Gurmukhī Transcription	ਯ y	ਰ r	ਲ l	ਵ v	ੜ ṛ

Consonant Chart

- A. The 32 consonants begin from ‘s’ (ਸ) and continue through ‘r’ (ੜ).
- B. In pronunciation, if a consonant stands independent of a vowel character, that consonant will be followed by an implicit ‘a.’
- The above statement is only true for consonants devoid of vowel characters. However, if the consonant is the final letter of the word, then the implicit ‘a’ is omitted in transcription.
Example 1: ਨਦਰ (na + da + ra = nadar)
Example 2: ਸਤ (sa + ta = sat)
 - The implicit ‘a’ is applicable to the end of single character words.
Example 1: ਨ (na)
Example 2: ਤ (ta)
- C. There are instances where the implicit ‘a’ is overlooked. In these cases, the ‘a’ can be omitted as it aids in pronunciation. The inclusion of the implicit ‘a’ is not inaccurate, but instead, more functional.

Example 1: ਦਰਸਨ (da + ra + sa + na = darsan)

Example 2: ਕਰਤਾ (ka + ra + tā = kartā)

- D. When a consonant is followed by a vowel character supported by a vowel base, an implicit 'a' will be included and the visible vowel symbol will be indicated by a corresponding dieresis (ī or ū) vowel.

Example 1: ਸਉ (sa + u = saū)

Example 2: ਗਇ (ga + i = gai)

II. Vowels

Vowel Character	Transcription	Inscription Position
ਅ	a	Before letter
ਾ	ā	After letter
ਃ	i	Before letter
਴	ī	After letter
ੁ	u	Below letter
ੂ	ū	Below letter
ੇ	e	Above letter
ੈ	ai	Above letter
ੋ	o	Above letter
ੌ	au	Above letter

Vowel Chart

- A. There are three vowel bases (ੳ, ਅ, ਏ).
- B. In addition to the implicit 'a' that follows consonants independent of vowels, there are ten vowel characters (see Vowel Chart).
- C. Of the three vowel bases, only 'ਅ' can stand independent of another vowel character.
- Example 1: ਅਗਨ (a + ga + na = agan)
- Example 2: ਅਕਾਲ (a + kā + la = akāl)

- D. If a vowel character is directly followed by another vowel base and character, then the implicit ‘a’ will be omitted.

Example 1: ਜਾਇ (jā + i = jāī)

Example 2: ਕੀਓ (kī + o = kīo)

- E. If a vowel character combines with a vowel base to produce the first letter in the word, there will be no ‘a’ preceding the vowel.

Example 1: ਉਪਰਹੁ (ū + pa + ra + hu = ūparahu)

Example 2: ਆਸਣੁ (ā + sa + ṇu = āsaṇu)

- F. If a vowel character is present with a consonant rather than a vowel base, there will be no ‘a’ between the consonant and vowel.

Example 1: ਸੂਰਾ (sū + rā = sūrā)

Example 2: ਮੋਹਿ (mo + hi = mohi)

III. ṭipī (°)

- A. When a ṭipī is representing one of the five sounds (ṇ, ṇ̇, ṇ̈, n, m), it will not be followed by an implicit ‘a.’

Example 1: ਖੰਡ (kha + ṇḍa = khaṇḍ)

Example 2: ਚੰਗਾ (ca + ṇgā = caṅgā)

- B. If a ṭipī is representing one of the five sounds (ṇ, ṇ̇, ṇ̈, n, m), and the letter following the ṭipī falls within rows 2-6 (Consonant Chart), then the ṭipī will make the sound of the last letter in the corresponding row.

Example 1: ਅੰਤ (a + nta = ant)

Example 2: ਸੰਗਤਿ (sa + ṇga + ti = saṅgati)

1. When a ṭipī represents one of the five sounds (ṇ, ṇ̇, ṇ̈, n, m), and the letter emphasized by the ṭipī represents the same sound (ṇ, ṇ̇, ṇ̈, n, m), then it will create a double sound (compound) when spoken, and should be transcribed accordingly.

Example 1: ਅੰਮ੍ਰਿਤ (a + mma + ri + ta = ammr̥it)

Example 2: ਮੰਨੀਐ (ma + nnī + ai = mannīai)

- C. If a ṭipī is at the end of the word, and therefore unsupported by another consonant, then the default sound is ‘aṇ.’

Example 1: ਸੈਭੰ (sai + bhaṇ = saibhaṇ)

Example 2: ਕਰਮੰ (ka + ra + maṇ = karmaṇ)

1. If a ṭipī concludes a term in a compound word, then the ṭipī will create the ‘ṇ̈’ sound.

Example 1: ਅਹੰਕਾਰ (a + haṇ + kā + ra = ahaṇkāṇ)

Example 2: ਅਹੰਬੁਧਿ (a + hañ + bu + dhi = ahañbudhi)

D. All vowel nasalizations are denoted by ‘m.’

Example: ਤੂੰ (tūm)

E. The ṭipī has two possible connotations – it commonly symbolizes an additional sound incorporating one of five characters (ṇ, ṇ̄, ṇ̄̄, ṇ̄̄̄, m), and sometimes denotes a nasalization (ṇ̄). Identification of correct usage in various instances can only be determined with knowledge of the language.

IV. Bindī (º)

A. When creating a ‘n’ sound, there will be no implicit ‘a’ sound following the ‘n’ sound.

Example: ਸਾਂਤ (sā + nt = sānt)

B. All vowel nasalizations are denoted by ‘m.’

Example 1: ਤੂੰਹੈਂ (tū + haim̄ = tūhaim̄)

Example 2: ਹਾਂ (hām̄)

C. The bindī can play two different roles. Predominantly, it calls for vowel nasalization, and occasionally, it forms an ‘n’ sound. One can only know the bindī’s role with an understanding of the language.

V. Visarg (º)

A. The visarg is represented by an ‘h.’ It is pronounced as a softer ‘h,’ and an implicit ‘a’ precedes the ‘h.’

Example 1: ਦ੍ਰਿੜੰਤਣਃ (dri + ṛa + nta + ṇa + ḥa = driṛantaṇaḥ)

Example 2: ਲਿਖੁਣਃ (li + kha + ha + ṇa + ḥa = likhhaṇaḥ)

B. The visarg emerged as a Sanskrit influence and only appears at the end of words.

VI. Subscripts

Subscript	Origin	Example	Transcription
ੴ	ੴ	ਵੁਰ੍ਹਿਅਾ	vurhiā
ਅਸ੍ਰ	ਛ	ਅਸ੍ਰੁਰਜ	ascaraj
ਬਿਸ੍ਰ	ਠ	ਬਿਸ੍ਰਾ	bisṭā
ਅਸ੍ਤ	ਥ	ਅਸ੍ਤੁ	ast
ਕ੍ਰਿਸ਼	ੜ	ਕ੍ਰਿਸ਼ੁ	krisn
ਦੁਖ੍ਯ	ਯ	ਦੁਖ੍ਯੰ	dukhyañ

ˆ	ˆ	ਪ੍ਰਾਣੀ	prāṇī
ੴ	ੴ	ਜਗਦੀਸ਼ਵਰਹ	jagdīsvarah

Subscript Chart

- A. There are eight subscripts used within the Gurū Granth Sāhib.
- B. The subscript appears below the consonant.
- C. The subscript is pronounced directly after the consonant above it; there is no implicit ‘a’ between the two.
- D. In pronunciation, if a vowel character supports a consonant and subscript, the subscript will follow the consonant but precede the vowel.

VII. Halant (ˆ)

- A. The halant calls for a tonal variation in specific cases; explicitly, an increase in pitch.
Example 1: ਜਿਨੀ (ji + nī = jinī)
Example 2: ਧਿਆਵਨਿ (dhi + ā + va + ni = dhiāvaṇi)
- B. When an equivalent word is present, but the halant is not inscribed, one should not modify tone.

VIII. Prolongation (ਅ)

- A. When at the end of a word and directly following a vowel character, then the ‘ਅ’ represents a prolongation of the sound preceding it.
Example 1: ਜੀਅ (jī + a = jīa)
Example 2: ਲੋਅ (lo + a = loa)

IX. Numerals

Roman	0	1	2	3	4	5	6	7	8	9
Gurmukhī	੦	੧	੨	੩	੪	੫	੬	੭	੮	੯

Cardinal Number Chart

X. Adoptions

- A. Fārsī (Persian) Influence

1. If an equivalent sound for a Fārsī-inspired letter exists in the Roman script, then that Roman character will be used in transcription.

Example 1: ਫ (f)

Example 2: ਜ (z)

2. If there is no corresponding sound already inscribed in the Roman alphabet, then the sound will be indicated with a line underneath the corresponding Roman letters.

Example 1: ਗ (gh)

Example 2: ਖ (kh)

3. The contemporary version of the Gurmukhī script has adopted five letters to represent Fārsī sounds. These sounds are indicated in Gurmukhī script with a dot somewhere near the bottom of the corresponding Gurmukhī letter: ਗ, ਜ, ਫ, ਖ, ਸ਼ (gh, z, f, kh, sh). These sounds were introduced to the Pañjābī language in the post-Gurū Granth Sāhib compilation period and, therefore, are not included in the text. However, secondary writings of Sikhs utilize these letters frequently.

B. Addak (º)

1. In transcription, the Gurmukhī characters emphasized by the addak will be written twice in the Roman script.

Example: ਅੱਗੀ (a + ggī = aggī)

2. If the Roman representation of the Gurmukhī letter utilizes more than one Roman alphabet and is emphasized by the addak, then only the first letter will be doubled.

Example: ਸਿੱਖੀ (si + kkhī = sikkhī)

3. If an addak is not present, no characters will be doubled.
4. The addak conveys which syllable should be stressed. Though used within secondary writings of the Sikhs, the addak came into use after the Gurū Granth Sāhib was compiled. Therefore, it is nonexistent within the text of the Gurū Granth Sāhib.

Sikhri Font

Sikhri is a True-Type/Open-Type Font. Currently, this font is only compatible with Microsoft Windows based systems and will not work on the Macintosh. This upgrade may be made in the future depending on need or demand.

Description

The font is organized such that all of the lowercase characters required for Gurbāñī

transcription are located directly on the keyboard. The keyboard keys for 'a' and 'A' will both produce lowercase letters representing distinct sounds found in Gurbāṇī. For example:

<u>Gurmukhī</u>	<u>Transcription</u>	<u>Keyboard</u>
ਅ	a	a
ਆ	ā	A

Where necessary, uppercase characters corresponding to each of the lowercase characters mapped to the keyboard can be found in one of the Extended Latin Character Sets. To access these characters, and/or the Fārsī and Sanskrit transcription characters, one can use the Insert Symbol command found in MS Word. When using other programs, the Microsoft Character Map provided with MS Windows can be used to select the desired character and paste it in where necessary.

Though certain symbols on the screen appear to have imperfections, printouts of these symbols emerge flawless. This is due to a Microsoft Windows rendering issue.

Keymap

G	T	K
ਉ	u	u
ਊ	ū	U
ਓ	o	o
ਅ	a	a
ਆ	ā	A
ਐ	ai	ai
ਔ	au	au
ਇ	i	i
ਈ	ī	I
ਏ	e	e
ਅਉ	aü	aK
ਅਇ	aĩ	aJ
ਸ	s	s
ਹ	h	h
ਕ	k	k
ਖ	kh	kh
ਗ	g	g
ਘ	gh	gh
ਙ	ṅ	O
ਚ	c	c
ਛ	ch	ch
ਜ	j	j
ਝ	jh	jh
ਞ	ñ	P
ਟ	ṭ	T
ਠ	ṭh	Th
ਡ	ḍ	D
ਢ	ḍh	Dh
ਣ	ṇ	N
ਤ	t	t
ਥ	th	th
ਦ	d	d
ਧ	dh	dh

G	T	K
ਨ	n	n
ਪ	p	p
ਫ	ph	ph
ਬ	b	b
ਭ	bh	bh
ਮ	m	m
ਯ	y	y
ਰ	r	r
ਲ	l	l
ਵ	v	v
ੜ	ṛ	R
ਯੋ	ਯੋ	F
ਗ੍ਰੰ	ਗ੍ਰੰ	E
ਸ਼	sh	sh
ਖ਼	kh	Q
ਗ਼	gh	G
ਜ਼	z	z
ਫ਼	f	f
ਲ਼	l	L
ਃ	ḥ	H
ੰ	m̐	M
ੰ	m̐	M
	ṙ	Z
	ṛ	C
	ṛ	B

G = Gurmukhī

T = Transcription

K = Keyboard

For other characters, use
“Insert Symbol” command in
MS Word.