SCRIPT GRAMMAR FOR PUNJABI LANGUAGE
Prepared by

Technology Development for Indian Languages (TDIL) Programme of DIT, GoI in co-ordination with

C-DAC, GIST Pune.

Following guidelines have been followed while preparing this script grammar.

1. This document contains information especially as to how shapes are formed which is filled out by hand.
2. This document is written by Punjabi Experts
3. Insofar as Section 8 is concerned, following things should be noted:
    1. LIGATURES:
        o   Dead ligatures i.e. ligatures which are dysfunctional in the language have not been used.
        o   CHC cases have been tested and checked.
        o   More complex clusters such as CHCHC etc have been generated out from the corpus and presented for checking

    2. VARIANTS
    Variants haven been handled in the script grammar and where two variants exist concurrently for the same shape and are deemed as viable, one of them has been entered in the document and the other has been provided separately as a variant. Uniformity has been maintained i.e. all stacked variants haven been bunched together whereas the non-stacked variant have been grouped together.
5. Items such as the History of the Language and the evolution of the script has been supplied in the form of an Appendix.

SCRIPT GRAMMAR TEMPLATE
LANGUAGE: PUNJABI
SCRIPT: GURMUKHI

Instructions

Please read through these instructions before you fill in the template:
1. This template will contain information especially as to shapes which will need to be filled out by hand.
2. Please print out the template and fill it in completely.
3. Once it is complete, please have it validated by an expert.
4. Subsequent to the validation, please get the document checked and validated by the statutory certifying body in your State.

0. 1. Name of Expert *AMANDREET SINGH*

0.2. Name of Evaluator

1. Name of the language and its representation in the 3 letter mnemonic
Name of the Language: PUNJABI
Alpha-3 code: pan

2. Name of the statutory board governing the language
The name and address /tel number/email of the statutory body

A scanned/hard copy of the statutes laid down. (Please append)

3. Identification of the writing system(s) used to inscribe the given language
The name(s) of the script system(s) used. GURMUKHI

4. Short Historical Picture of the Language and the Script used.

PLEASE PROVIDE THE DATA IN APPENDIX.

5. Modifications brought to the writing system by a given language in terms of addition of characters and deprecation of other characters.
You need to enumerate here the character set of the language preferably as per the sorting order. **Please write the set as you want it to be displayed on screen**

CONSONANTS

VOWELS

MATRAS

DIACRITICS

1. ADDAK ਁ

Addka is used with geminated consonants.
Addak is placed on the preceding syllable,
Addak cannot be at the beginning of a word.

2. VIRAMA (HALANTA)
Please specify where the halanta is used.

੍

3. NASALS

Bindi ਂ

Tippi ਁ

Bindi and Tippi are allographic variants of the nasal.
The rules for placement of Bindi and Tippi are as under:
      The initial forms of *u, uu* i.e. vowels take Bindi;
      u, uu after a consonant i.e. matras take Tippi;
      all other short vowels/matras take Tippi;
      all other long vowels/matras take Bindi.

Addak Bindi ਁ

Is addak bindi possible in modern Gurmukhi ?   N O
Please give examples

4. Visarga: ਃ
NOT USED Frequently

NUMERALS
Please write the numerals in Gurmukhi. as you would like them represented in Gurmukhi

OTHERS

YAKASH

EK ONKAR ੴ

UDATTA

6. The structure of the writing system of the language

Tick whichever is appropriate: √Abugida. **X** Abjad

7. Rule ordering of the characters within the syllable (only for abugidas)
Case of Addak and Bindi Tippi and Addak Bindi needs to be discussed.

8. Script Pertinent Description of the syllabic clusters
8.1. BASIC SET OF CHARACTERS
The basic set of characters has been provided in this inventory.
These are arranged as per their class: CONSONANT / VOWEL / MATRAS /
DIACRITICS
The allographs are presented at the end.

INSTRUCTIONS
In case you do not see any issues just tick the VALID box. In case you see issues
tick invalid and provide the necessary correction for the combination in question.
In case a particular character is not used in your script, please cross it out
In case you feel a particular character from your script has been left out, please
specify the same.

8.1.1. CONSONANT SET: VALID / INVALID
Basic Consonants arranged as per their vargas

ਕ	ਖ	ਗ	ਘ	ਙ
ਚ	ਛ	ਜ	ਝ	ਞ
ਟ	ਠ	ਡ	ਢ	ਣ
ਤ	ਥ	ਦ	ਧ	ਨ
ਪ	ਫ	ਬ	ਭ	ਮ
ਯ	ਰ	ਲ	ਲ਼	ਵ
ਸ	ਸ਼	ਹ		

Nukta Consonants VALID
For Perso-Arabic Loans

ਖ਼ ਗ਼ ਜ਼ ਫ਼ ਸ਼ ਲ਼

For flapped forms

ੜ

8.1.2.VOWEL SET:

ਅ ਆ ਇ ਈ ਉ ਊ ਏ ਐ ਓ ਔ

8.1.3. MATRA SET

ਾ ਿ ੀ ੁ ੂ ੇ ੈ ੋ ੌ

Active Catenator(s) i.e. Displaced Matra(s):

CATENATOR	POSITION	EXAMPLE
ਿ	consonant	ਕਿ

8.1.4. DIACRITICS

ਂ: Bindi

ੰ: Tippi

see above for rules of bindi and tippi

Addak Bindi ੑ

Please give istance where addak bindi is possible

NO

8.1.5.1. ALLOGRAPHS OF ਰ

NOTE: Rakars will be automatically generated out in the CHC list. The present inventory is just for validating the different forms that exist in your script.

RAFARS:
NONE IN GURMUKHI

RAKARS
PLEASE PROVIDE LIST: *NO*

8.1.5.2. Any other Allographs. *NO*
Please mention below with substantiating evidence.

8.1.6. PUNCTUATION MARKERS
Please specify the punctuation markers specific to the character set omitting the markers taken from the Latin set such as . , ; : " ' ( ) [ ] etc.

NONE

Please remember that if you use Purna and Deergha Virama (full-stop/danda), as per Unicode norms, you will have to use at present the characters provided in Devanagari codechart: 0964, 0965 ।, ॥ till as such time this regulation is removed.

8.1.7. NUMERALS/DIGITS
Please specify the numbers for your script. Is the following VALID/ INVALID

੦੧੨੩੪੫੬੭੮੯

If not valid please give the correct form/forms
Please specify if the English (Latino-Arabic set: 0,1,2,3,4,5,6,7,8,9) is used in
official communications yes

8.1.8.
OTHER SYMBOLS (religious, currency markers etc. included in Unicode)
(Pl. see Unicode chart appended at the end)

EK ONKAR ੴ

YAKASH ̥

UDATT ̣

Khanda

Rupee Sign
DANDA'   0964
DOUBLE DANDA 0965

8.2. CONSONANT+MATRA COMBINATIONS

This set is divided into three parts:

CM:The combination of Consonant and Matra

CMD(Bindi/Tippi) i.e. Consonant+Matra+Nasal

- In case you do not see any issues just tick the VALID box. In case you see issues tick invalid and provide the necessary correction for the combination in question.
- Please do not forget that some combinations are dead clusters but are still needed by the font designer to generate out the grammar.
- In case you feel a particular Consonant+Matra combination has been left out, please specify the same.
- In case a particular character combination is not used in your script, please cross it out

C8.2.1.CM:/VALID/INVALID

	ਕ	ਖ	ਗ	ਘ	ਙ	ਚ	ਛ	ਜ	ਝ	ਞ
ਾ										
ਿ										
ੀ										
ੁ										
ੂ										
ੋ										
ੌ										
ੇ										
ੈ										

	ਟ	ਠ	ਡ	ਢ	ਣ	ਤ	ਥ	ਦ	ਧ	ਨ
◌ਾ										
ਿ◌										
◌ੀ										
◌ੁ										
◌ੂ										
◌ੇ										
◌ੈ										
◌ੋ										
◌ੌ										

	ਪ	ਫ	ਬ	ਭ	ਮ	ਯ	ਰ	ਲ	ਲ਼	ਵ	ਸ਼	ਸ	ਹ
◌ਾ													
ਿ◌													
◌ੀ													
◌ੁ													
◌ੂ													
◌ੇ													
◌ੈ													
◌ੋ													
◌ੌ													

	ਢ	ਖ਼	ਗ	ਜ	ਤ
ੋ					
ਿ					
ੀ					
ੁ					
ੂ					
ੇ					
ੈ					
ੋ					
ੌ					

VALID

	ਕ	ਖ	ਗ	ਘ	ਙ	ਚ	ਛ	ਜ	ਝ	ਞ
ਂ/ਁ										
ਾ ਂ/ਁ										
ਿ ਂ/ਁ										
ੀ ਂ/ਁ										
ੁ ਂ/ਁ										
ੂ ਂ/ਁ										
ੇ ਂ/ਁ										
ੈ ਂ/ਁ										
ੋ ਂ/ਁ										
ੌ ਂ/ਁ										

INVALID

Table 1 — consonant column headers: ਟ | ਠ | ਡ | ਢ | ਣ | ਤ | ਥ | ਦ | ਧ | ਨ

	ਟ	ਠ	ਡ	ਢ	ਣ	ਤ	ਥ	ਦ	ਧ	ਨ
◌ਂ/◌ੰ										
◌ਾਂ ◌ਂ/◌ੰ										
ਿ◌ ◌ਂ/◌ੰ										
◌ੀ ◌ਂ/◌ੰ										
◌ੁ ◌ਂ/◌ੰ										
◌ੂ ◌ਂ/◌ੰ										
◌ੇ ◌ਂ/◌ੰ										
◌ੈ ◌ਂ/◌ੰ										
◌ੋ ◌ਂ/◌ੰ										
◌ੌ ◌ਂ/◌ੰ										

Table 2 — consonant column headers: ਪ | ਫ | ਬ | ਭ | ਮ | ਯ | ਰ | ਲ | ਲ਼ | ਵ | ਸ਼ | ਸ | ਹ

	ਪ	ਫ	ਬ	ਭ	ਮ	ਯ	ਰ	ਲ	ਲ਼	ਵ	ਸ਼	ਸ	ਹ
◌ਂ/◌ੰ													
◌ਾਂ ◌ਂ/◌ੰ													
ਿ◌ ◌ਂ/◌ੰ													
◌ੀ ◌ਂ/◌ੰ													
◌ੁ ◌ਂ/◌ੰ													
◌ੂ ◌ਂ/◌ੰ													
◌ੇ ◌ਂ/◌ੰ													
◌ੈ ◌ਂ/◌ੰ													
◌ੋ ◌ਂ/◌ੰ													
◌ੌ ◌ਂ/◌ੰ													

	ਖ	ਗ	ਜ	ਢ	ਝ
ਂ/ਁ					
ਾ ਂ/ਁ					
ਿ ਂ/ਁ					
ੀ ਂ/ਁ					
ੁ ਂ/ਁ					
ੂ ਂ/ਁ					
ੇ ਂ/ਁ					
ੈ ਂ/ਁ					
ੋ ਂ/ਁ					
ੌ ਂ/ਁ					

8.3. CONSONANT+CONSONANT CLUSTERS
8.3.1.CHC

This is by far the most important inventory and comprises the basic 2 consonant conjuncts of the script. At present all the conjunct shapes you see are provided by the existing font for your script.

INSTRUCTIONS:

Basically germination is alone permissible in Gurmukhi which is shown by the addak In addition Halanta is used with certain combinations to generate out ligatures.
The combination of Consonant+Consonant is shown by the Halanta as a theoretical construct.

Thus ਕੁਕ =ੱਕ which implies that the addak will fall on the preceding syllable.

Please mark with X the slots where combinations are not permissible as per Gurmukhi writing.

Please do NOT forget that the conjunct shapes should be in conformity with norms laid down by the statutory bodies of your state.

Apart from the Geminated consonants, PUNJABI only the following ligatures. These are regularly formed with three letters only:ਹ(h), ਰ(r), and ਵ(v).

ਰ(r) and ਵ(v) are used to make consonant clusters and behave similarly

ਹ(h) is a tone marker and raises the tone.
Examples:

ਪਰ *pər* ("but") vs. ਪ੍ਰਬੰਧ (*prəbənd*, "management, government")

ਸਵੇਰ (*səvēr*, "morning") vs. ਸੂਰਗ (*svərəg*, "heaven")

Subjoined ਹ(h): behaves the same way as the regular ਹ(h) in non-word-initial positions. The regular ਹ(h) is pronounced at the beginning of words but not in other positions, where it instead raises the tone. The difference in usage is that the regular ਹ is used after vowels and the subscript version when there is no vowel, and is attached to consonants.

All other ligatural forms are theoretically possible but do NOT exist in Punjabi

	ਰ	ਵ	ਹ
ਕ	ਕ੍ਰ	ਕ੍ਵ	ਕ੍ਹ
ਖ	ਖ੍ਰ	ਖ੍ਵ	ਖ੍ਹ
ਗ	ਗ੍ਰ	ਗ੍ਵ	ਗ੍ਹ
ਘ	ਘ੍ਰ	ਘ੍ਵ	ਘ੍ਹ
ਙ	ਙ੍ਰ	ਙ੍ਵ	ਙ੍ਹ
ਚ	ਚ੍ਰ	ਚ੍ਵ	ਚ੍ਹ
ਛ	ਛ੍ਰ	ਛ੍ਵ	ਛ੍ਹ
ਜ	ਜ੍ਰ	ਜ੍ਵ	ਜ੍ਹ
ਝ	ਝ੍ਰ	ਝ੍ਵ	ਝ੍ਹ

8.3.2 CONSONANT+CONSONANT+CONSONANT CLUSTERS CHCHC

NONE EXIST IN GURMUKHI

8.3.3. CONSONANT+CONSONANT+CONSONANT+CONSONANT CLUSTERS: CHCHCHC

NONE EXIST IN GURMUKHI

9. COLLATION ORDER OF THE CHARACTERS: LEXICAL / DICTIONARY SORTING ORDER

List all the basic characters of the language in the expected sort-order. A sample sort order is provided below. Please provide an exhaustive collation order for your language. If there is any change in the sort order, please specify:

ਂ ੰ ਅ ਆ ਇ ਈ ਉ ਊ ਏ ਐ ਓ ਔ ਕ ਖ ਗ ਗ਼ ਘ ਙ ਚ ਛ ਜ ਜ਼ ਝ ਞ ਟ ਠ

ਡ ੜ ਢ ਣ ਤ ਥ ਦ ਧ ਨ ਪ ਫ ਫ਼ ਬ ਭ ਮ ਯ ਰ ਲ ਲ਼ ਵ ਸ਼ ਸ ਹ ਾ ਿ ੀ ੁ ੂ

ੇ ੈ ੋ ੌ ੍

ੴ ੧ੳ ਓ ਉ ਊ ੳ ਅ ਆ ਐ ਔ ੲ ਇ ਈ ਏ ਸ ਹ ਕ

ਕ ਖ ਗ ਘ ਙ ਚ ਛ ਜ ਝ ਞ ਟ ਠ ਡ ਢ ਣ ਤ ਥ ਦ ਧ ਨ

ਪ ਫ ਬ ਭ ਮ ਯ ੍ ਰ ਲ ਵ ੜ ਾ ਿ ੀ ੁ ੂ ੇ ੈ ੋ ੌ ਂ

ੱ ਃ ੰ ੑ

10. HOMOGRAPHIC IDENTITIES WITHIN THE CHARACTER SET.

Please provide a list of look alikes. Each set of homographs will be proposed as a pair. In extreme cases even three homographs are permissible Add more columns if so required.

a. Unique characters

HOMOGRAPH1	HOMOGRAPH2	HOMOGRAPH3
ਹ	ਰ	
ਮ	ਸ	
ਘ	ਧ	ਯ
ੲ	ਦ	
ਣ	ੜ	
ਠ	ਤ	
ਥ	ਖ	

b. Conjunct characters.

COMPOSING CONSONANTS	RESULTING HOMOGRAPH	COMPOSING CONSONANTS	RESULTING HOMOGRAPH

11. Compliance with Unicode.
  i.   Is the character set compliant with Unicode: YES

UNICODE CODE CHART FOR GURMUKHI

0A00   Gurmukhi   0A7F

Various signs
0A01	ੰ	GURMUKHI SIGN ADAK BINDI
0A02	ਂ	GURMUKHI SIGN BINDI
0A03	ਃ	GURMUKHI SIGN VISARGA

Independent vowels
0A05	ਅ	GURMUKHI LETTER A
		= aira
0A06	ਆ	GURMUKHI LETTER AA
0A07	ਇ	GURMUKHI LETTER I
0A08	ਈ	GURMUKHI LETTER II
0A09	ਉ	GURMUKHI LETTER U
0A0A	ਊ	GURMUKHI LETTER UU
0A0B	▩	<reserved>
0A0C	▩	<reserved>
0A0D	▩	<reserved>
0A0E	▩	<reserved>
0A0F	ਏ	GURMUKHI LETTER EE
0A10	ਐ	GURMUKHI LETTER AI
0A11	▩	<reserved>
0A12	▩	<reserved>
0A13	ਓ	GURMUKHI LETTER OO
0A14	ਔ	GURMUKHI LETTER AU

Consonants
0A15	ਕ	GURMUKHI LETTER KA
0A16	ਖ	GURMUKHI LETTER KHA
0A17	ਗ	GURMUKHI LETTER GA
0A18	ਘ	GURMUKHI LETTER GHA
0A19	ਙ	GURMUKHI LETTER NGA
0A1A	ਚ	GURMUKHI LETTER CA
0A1B	ਛ	GURMUKHI LETTER CHA
0A1C	ਜ	GURMUKHI LETTER JA
0A1D	ਝ	GURMUKHI LETTER JHA
0A1E	ਞ	GURMUKHI LETTER NYA
0A1F	ਟ	GURMUKHI LETTER TTA
0A20	ਠ	GURMUKHI LETTER TTHA
0A21	ਡ	GURMUKHI LETTER DDA
0A22	ਢ	GURMUKHI LETTER DDHA
0A23	ਣ	GURMUKHI LETTER NNA
0A24	ਤ	GURMUKHI LETTER TA
0A25	ਥ	GURMUKHI LETTER THA
0A26	ਦ	GURMUKHI LETTER DA
0A27	ਧ	GURMUKHI LETTER DHA
0A28	ਨ	GURMUKHI LETTER NA
0A29	▩	<reserved>
0A2A	ਪ	GURMUKHI LETTER PA
0A2B	ਫ	GURMUKHI LETTER PHA
0A2C	ਬ	GURMUKHI LETTER BA
0A2D	ਭ	GURMUKHI LETTER BHA
0A2E	ਮ	GURMUKHI LETTER MA
0A2F	ਯ	GURMUKHI LETTER YA
0A30	ਰ	GURMUKHI LETTER RA
0A31	▩	<reserved>
0A32	ਲ	GURMUKHI LETTER LA
0A33	ਲ਼	GURMUKHI LETTER LLA
		≡ 0A32 ਲ 0A3C ਼
0A34	▩	<reserved>
0A35	ਵ	GURMUKHI LETTER VA
0A36	ਸ਼	GURMUKHI LETTER SHA
		≡ 0A38 ਸ 0A3C ਼
0A37	▩	<reserved>

| 0A38 | ਸ | GURMUKHI LETTER SA |
| 0A39 | ਹ | GURMUKHI LETTER HA |

Various signs
0A3C	਼	GURMUKHI SIGN NUKTA
		= pairin bindi
		• for extending the alphabet to new letters

Dependent vowel signs
0A3E	ਾ	GURMUKHI VOWEL SIGN AA
		= kanna
0A3F	ਿ	GURMUKHI VOWEL SIGN I
		= sihari
		• stands to the left of the consonant
0A40	ੀ	GURMUKHI VOWEL SIGN II
		= bihari
0A41	ੁ	GURMUKHI VOWEL SIGN U
		= aunkar
0A42	ੂ	GURMUKHI VOWEL SIGN UU
		= dulainkar
0A43	▩	<reserved>
0A44	▩	<reserved>
0A45	▩	<reserved>
0A46	▩	<reserved>
0A47	ੇ	GURMUKHI VOWEL SIGN EE
		= lavान
0A48	ੈ	GURMUKHI VOWEL SIGN AI
		= dulavan
0A49	▩	<reserved>
0A4A	▩	<reserved>
0A4B	ੋ	GURMUKHI VOWEL SIGN OO
		= hora
0A4C	ੌ	GURMUKHI VOWEL SIGN AU
		= kanaura

Virama
| 0A4D | ੍ | GURMUKHI SIGN VIRAMA |

Sign
| 0A51 | ੑ | GURMUKHI SIGN UDAAT |

Additional consonants
0A59	ਖ਼	GURMUKHI LETTER KHHA
		≡ 0A16 ਖ 0A3C ਼
0A5A	ਗ਼	GURMUKHI LETTER GHHA
		≡ 0A17 ਗ 0A3C ਼
0A5B	ਜ਼	GURMUKHI LETTER ZA
		≡ 0A1C ਜ 0A3C ਼
0A5C	ੜ	GURMUKHI LETTER RRA
0A5D	▩	<reserved>
0A5E	ਫ਼	GURMUKHI LETTER FA
		≡ 0A2B ਫ 0A3C ਼

Reserved
For viram punctuation, use the generic Indic 0964 and 0965.
0A64	▩	<reserved>
		→ 0964 । devanagari danda
0A65	▩	<reserved>
		→ 0965 ॥ devanagari double danda

Digits
0A66	੦	GURMUKHI DIGIT ZERO
0A67	੧	GURMUKHI DIGIT ONE
0A68	੨	GURMUKHI DIGIT TWO
0A69	੩	GURMUKHI DIGIT THREE
0A6A	੪	GURMUKHI DIGIT FOUR

0A6B	੫	GURMUKHI DIGIT FIVE
0A6C	੬	GURMUKHI DIGIT SIX
0A6D	੭	GURMUKHI DIGIT SEVEN
0A6E	੮	GURMUKHI DIGIT EIGHT
0A6F	੯	GURMUKHI DIGIT NINE

Gurmukhi-specific additions
0A70	ੰ	GURMUKHI TIPPI
		• nasalization
0A71	ੱ	GURMUKHI ADDAK
		• doubles following consonant
0A72	ੲ	GURMUKHI IRI
		• base for vowels
0A73	ੳ	GURMUKHI URA
		• base for vowels
0A74	ੴ	GURMUKHI EK ONKAR
		• God is One
0A75	ੵ	GURMUKHI SIGN YAKASH