

Gujarati Script



4.1 Gujarati Code Chart

	0A8	0A9	0AA	0AB	0AC	0AD	0AE	0AF
0		ૐ 0A90	ઠ 0AA0	૨ 0AB0	ી 0AC0	ૐ 0AD0	જ 0AE0	૦ 0AF0
1	ં 0A81	ૐ 0A91	ડ 0AA1		ૃ 0AC1	ૃ 0AD1	ૃ 0AE1	ૃ 0AF1
2	ં 0A82		ઢ 0AA3	લ 0AB2	ૃ 0AC2	ૃ 0AD2	ૃ 0AE2	
3	ઃ 0A83	ૐ 0A93	ણ 0AA5	ળ 0AB3	ૃ 0AC3	ૃ 0AD3	ૃ 0AE3	
4		ૐ 0A94	ત 0AA4		ૃ 0AC4	ૃ 0AD4	ૃ 0AE4	
5	અ 0A85	ક 0A95	થ 0AA5	વ 0AB5	ં 0AC5		ૃ 0AE5	
6	આ 0A86	ખ 0A96	દ 0AA6	શ 0AB6			ૃ 0AE6	ૃ 0AF6
7	ઇ 0A87	ગ 0A97	ધ 0AA7	પ 0AB7	ે 0AC7		ૃ 0AE7	ૃ 0AF7
8	ઈ 0A88	ઘ 0A98	ન 0AA8	સ 0AB8	ૈ 0AC8		ૃ 0AE8	ૃ 0AF8
9	ઉ 0A89	ઙ 0A99		હ 0AB9	ૌ 0AC9		ૃ 0AE9	
A	ઊ 0A8A	ચ 0A9A	પ 0AA9	ં 0ABA			ૃ 0AEA	
B	ઋ 0A8B	છ 0A9B	ફ 0AAB		ો 0ACB		ૃ 0AEB	
C	ૃ 0A8C	જ 0A9C	ખ 0AAC	્ 0ABC	ૌ 0ACC		ૃ 0AEC	
D	ૐ 0A8D	ઝ 0A9D	ભ 0AAD	ડ 0ABD	ૃ 0ACD		ૃ 0AED	
E		ઞ 0A9E	મ 0AAE	ો 0ABE			ૃ 0AEE	
F	ૐ 0A8F	ટ 0A9F	ય 0AAF	િ 0ABF			ૃ 0AEF	



4.1.1 Gujarati Code Chart Details

Code Character Description
Point

Various signs

0A81 ँ GUJARATI SIGN
CANDRABINDU
0A82 ः GUJARATI SIGN
ANUSVARA
0A83 ोः GUJARATI SIGN
VISARGA

Independent vowels

0A85 અ GUJARATI LETTER A
0A86 આ GUJARATI LETTER AA

0A87 ઇ GUJARATI LETTER I
0A88 ઈ GUJARATI LETTER II
0A89 ઉ GUJARATI LETTER U
0A8A ઊ GUJARATI LETTER UU
0A8B ઋ GUJARATI LETTER
VOCALIC R
0A8C ૠ GUJARATI LETTER
VOCALIC L
0A8D એ GUJARATI VOWEL
CANDRA E
0A8E <reserved>
0A8F એ GUJARATI LETTER E
0A90 ઐ GUJARATI LETTER AI
0A91 ઔ GUJARATI VOWEL
CANDRA O
0A92 <reserved>
0A93 ઓ GUJARATI LETTER O
0A94 ઔ GUJARATI LETTER AU

Consonants

0A95 ક GUJARATI LETTER KA
0A96 ખ GUJARATI LETTER KHA
0A97 ગ GUJARATI LETTER GA
0A98 ઘ GUJARATI LETTER GHA
0A99 ઙ GUJARATI LETTER NGA
0A9A ચ GUJARATI LETTER CA
0A9B છ GUJARATI LETTER CHA

0A9C જ GUJARATI LETTER JA
0A9D ઝ GUJARATI LETTER JHA
0A9E ઞ GUJARATI LETTER NYA
0A9F ઠ GUJARATI LETTER TTA
0AA0 ડ GUJARATI LETTER TTHA
0AA1 ઢ GUJARATI LETTER DDA
0AA2 ઢ GUJARATI LETTER DDHA
0AA3 ણ GUJARATI LETTER NNA
0AA4 ત GUJARATI LETTER TA
0AA5 થ GUJARATI LETTER THA
0AA6 દ GUJARATI LETTER DA
0AA7 ધ GUJARATI LETTER DHA
0AA8 ન GUJARATI LETTER NA
0AA9 <reserved>
0AAA પ GUJARATI LETTER PA
0AAB ફ GUJARATI LETTER PHA
0AAC બ GUJARATI LETTER BA
0AAD ભ GUJARATI LETTER BHA
0AAE મ GUJARATI LETTER MA
0AAF ય GUJARATI LETTER YA
0AB0 ર GUJARATI LETTER RA
0AB1 <reserved>
0AB2 લ GUJARATI LETTER LA
0AB3 ળ GUJARATI LETTER LLA
0AB4 <reserved>
0AB5 વ GUJARATI LETTER VA
0AB6 શ GUJARATI LETTER SHA
0AB7 ષ GUJARATI LETTER SSA
0AB8 સ GUJARATI LETTER SA
0AB9 હ GUJARATI LETTER HA
0ABA ળ GUJARATI INVISIBLE
LETTER

Various signs

0ABC ળ GUJARATI SIGN NUKTA
• for extending the alphabet
to new letters
0ABD ળ GUJARATI SIGN AVAGRAHA

Dependent vowel signs

0ABE ળ GUJARATI VOWEL SIGN AA
0ABF ળ GUJARATI VOWEL SIGN I
• stands to the left of the consonant



0AC0	ी	GUJARATI VOWEL SIGN II
0AC1	ु	GUJARATI VOWEL SIGN U
0AC2	ू	GUJARATI VOWEL SIGN UU
0AC3	ृ	GUJARATI VOWEL SIGN VOCALIC R
0AC4	ॄ	GUJARATI VOWEL SIGN VOCALIC RR
0AC5	ँ	GUJARATI VOWEL SIGN CANDRA E
0AC6		<reserved>
0AC7	े	GUJARATI VOWEL SIGN E
0AC8	ै	GUJARATI VOWEL SIGN AI
0AC9	ॉ	GUJARATI VOWEL SIGN CANDRA O
0ACA		<reserved>
0ACB	ो	GUJARATI VOWEL SIGN O
0ACC	ौ	GUJARATI VOWEL SIGN AU

Various signs

0ACD	्	GUJARATI SIGN HALANT
0ACE		<reserved>
0ACF		<reserved>
0AD0	ॐ	GUJARATI OM
0AD1	◌̣	GUJARATI STRESS SIGN UDATTA •Used with Sanskrit text
AD2	◌̤	GUJARATI STRESS SIGN ANUDATTA •Used with Sanskrit text
AD3	◌̥	GUJARATI GRAVE ACCENT •Used with Sanskrit text
AD4	◌̦	GUJARATI ACUTE ACCENT

Generic additions

0AE0	ॠ	GUJARATI LETTER VOCALIC RR
0AE1	ॡ	GUJARATI LETTER VOCALIC LL •Used with Sanskrit text

0AE2	ॢ	GUJARATI VOWEL SIGN VOCALIC L •Used with Sanskrit text
0AE3	ॣ	GUJARATI SIGN VOCALIC LL •Used with Sanskrit text
0AE4	।	GUJARATI SIGN PURNA VIRAM
0AE5	॥	GUJARATI SIGN DEERGH VIRAM

Digits

0AE6	૦	GUJARATI DIGIT ZERO
0AE7	૧	GUJARATI DIGIT ONE
0AE8	૨	GUJARATI DIGIT TWO
0AE9	૩	GUJARATI DIGIT THREE
0AEA	૪	GUJARATI DIGIT FOUR
0AEB	૫	GUJARATI DIGIT FIVE
0AEC	૬	GUJARATI DIGIT SIX
0AED	૭	GUJARATI DIGIT SEVEN
0AEE	૮	GUJARATI DIGIT EIGHT
0AEF	૯	GUJARATI DIGIT NINE
0AF0	૦	GUJARATI ABBREVIATION SIGN
0AF1	₹	GUJARATI RUPEE CURRENCY SIGN
AF6	ૠ	GUJARATI FRACTIONAL NUMERAL /PA/ ¼
AF7	ૡ	GUJARATI FRACTIONAL NUMERAL /Addho/ ½
AF8	ૢ	GUJARATI FRACTIONAL NUMERAL /Pono/ ¾

[Note : Code points 0A80 - 0ACF of Gujarati Code set have one-to-one phonetic correspondence for vowels, consonants and vowel modifiers with 0900 - 094F of Devanagari Code Set]



4.1.2 Gujarati Script Details

General Gujarati Background

Gujarati, and the other Indian languages such as Punjabi, Hindi, Oriya, Bengali, Assamese, Telugu, Kannada, Malayalam and Tamil, have originated from the ancient Brahmi script, which is phonetic in nature. The alphabet in each of these languages may vary somewhat but they all share a common phonetic structure. The differences between the scripts are primarily in their written forms where different combinations come into play.

Technical Characteristics

Gujarati Alphabet Characteristics

Gujarati alphabets utilize 94 symbols altogether, which can be categorized into the groupings shown below. Some of these characters are not as often used as others, and can only be found in some specific styles of writing such as manuscripts of pre-modern period. Thousands of such manuscripts which would need to be digitized in future, are available in 'Hastapatra Bhandar' of Gujarat.

Some of the Gujarati characters that are in daily use are not found in the Unicode Chart for Gujarati. Some of these characters can be constructed by using other independent elements of the script.

It is important to remember that the set of legitimate characters in Gujarati, exceed the number of characters available in the Unicode Chart on this date.

The graphic representation of these alphabets can be found in the reference section, together with the naming convention according to the Gujarati [Industry Standard Institute], and other standards registered with International Standards Organization(s).

Consonants

Gujarati character set provides 34 (+2 compound-ksha, gna) consonants.

There are 34 consonants (2 others are compound characters) which represent 34 consonant sounds. Further the sound of any consonant is dependent on whether it comes at the start or the end of a

syllable. In addition to the 34 consonants, there are 2 special symbols representing compound characters comprised of the combination of 2 consonants. These two are not treated as consonants lexically, although they always function like consonants.

Vowels

Gujarati character set provides 14 vowels.

Vowels are represented by a single symbol. Some vowels appear as the combination of one vowel and a dependent vowel sign. But that should not be mistaken as a complex / non-independent vowel (e.g. +U0A94 appears as U0A85 + U0ACC. But it is an independent vowel).

Though a vowel sign may occur after, before, above or below the consonant, the vowel sound always follows the consonant.

In addition to the vowels mentioned, there are the two special symbols which function as sound modifiers. They are specified separately in the "Numerals and Special Symbols" section.

In order to allow building all vowel combinations as described above with less complexity for Gujarati text processing applications, all vowel symbols and vowel-related special symbols are defined and grouped accordingly as shown in the following section of Character Set Considerations.

Tone Marks

Gujarati character set provides 1 tone mark. The 'avagraha' is used to indicate stress on the characters that precede / follow it.

The default tone of any written syllable is determined by the context.

Note : There are no tone marks in Gujarati. The category can be changed to elongation mark then the description above (replace stress with elongation) is valid.

Numerals and Special Symbols

Gujarati character set provides 10 numerals. There are 3 other symbols used for representing fractions. These are called 'pa' (.25), 'adadho' (.50) and 'poNo'



(.75). The symbols for these have been recommended for inclusion in the Unicode Chart for Gujarati.

Western numerals are now also widely used in Gujarati writing. There are no special symbols for punctuation, abbreviation, repetition, etc.

The Gujarati language has a separate set of notations for accounting.

Punctuation Marks

Unlike many other Indian scripts (Hindi, Marathi, Bengali), Gujarati does not use a line above characters to indicate word-units. Words are separated by spaces. Spaces in the text indicate the ends of words [not sentences] and are not used as a form of explicit punctuation. Thus individual words have to be recognized by scanning the text for space boundaries. Hence there is no complexity (as in Thai) in designing applications which are required to recognize words (for example, dictionaries, spell-checking, or search programs).

The following punctuation marks are used in Gujarati:

- | | |
|--|---|
| 1. Alpa Virama | represented by ‘,’ (comma) |
| 2. Ardha Viram | represented by ‘;’ (semi colon) |
| 3. Purna Viram | represented by ‘.’ (fullstop) |
| 4. Prashnarth Chinnha | represented by ‘?’ (question mark) |
| 5. Ashcharya Chinnha | represented by ‘!’ (exclamation mark) |
| 6. Avtaran Chinnha | represented by ‘ ’ ’ (single quotation marks) |
| 7. Avtaran Chinnha | represented by ‘ “ ” ’ (double quotation marks) |
| 8. Mahaviram Chinnha | represented by ‘:’ (colon) |
| 9. Marks at the end of sentence first single and second double line of pre-modern verse. | represented by ‘ ’ (single danda) and ‘ ’ (double danda) |

Case

Written Gujarati has no equivalent to capital letters as practised in English writing. The Shift / Caps Lock should be used to show / type the characters not available in the normal (un-capitalized) keyboard.

Writing Direction

The basic direction of writing Gujarati is from left to right and top to bottom, the same as English. However, vowels may be written before, below, above, or after their related consonants. Even when used like this, they are pronounced after the consonant.

Currency and Ancient Signs

Gujarati character set provides 1 currency sign and 4 ancient signs.

Gujarati currency unit is rupiyo, and its symbol resembles the Hindi digit ‘૩’ – it is not a part of the Unicode Chart.

In addition, there are also four ancient symbols.

Character Set Considerations

The previous background about Gujarati language will ease the task of understanding the Gujarati character set. Here we build upon the previous information to give a detailed description of the Gujarati Character set.

Gujarati Word Formation Description: Information and Statistics (Characteristics)

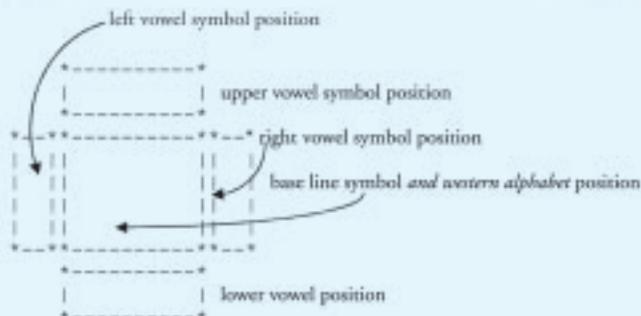
It is important to remember that the set of legitimate characters in Gujarati, exceed the number of characters available in the Unicode Chart on this date.

A line of Gujarati text can be considered as being logically divided into three horizontal parallel lines:

- 1.The base line, on which consonants and independent vowels are written.
- 2.The line below the base line, used for writing dependent (lower) vowels .



3. The line above the base line, used for writing dependent (upper) vowels.



Each individual character or symbol appears in one form only, independent of where it is placed. But when independent consonants are combined to form compound characters, they change shape. The Unicode 3.0 book claims that these rendering rules are same as those of Devanagari – but this is not so.

Combinations

The Unicode Standard 3.0 Chapter 9 subsection 9.4 Page 226 for Gujarati specifies that the rendering behaviour is the same as that of Devanagari. However some of these rules do not apply to the rendering of conjuncts and compound words. Some additional specifications and rules may be required.

See The Annexure - I

Collation background

General

- There is no official document giving the sort order for Gujarati.
- Different printed dictionaries do not follow the same order. They vary primarily with respect to the ordering of vowel signs, the nukta sign, and some consonant conjuncts.
- Sometimes, the same document may contain different sorting of some vowel signs. (It is very important to identify & define any variations in Gujarati sorting order.)

Observations

1. The independent vowels are, in order:

- U+0A85 (letter A) to U+0A8B (letter vocalic R)

- U+0AE0 (letter vocalic RR)
- U+0A8D (letter vowel candra E)
- U+0A8F (letter E) to U+0A91 (letter vowel candra O)
- U+0A93 (letter O)
- U+0A94 (letter AU)

2. The consonants, in order, are:

- U+0A95 (letter KA) to U+0AA8 (letter NA)
- U+0AAA (letter PA) to U+0AB0 (letter RA)
- U+0AB2 (letter LA)
- U+0AB5 (letter VA) to U+0AB9 (letter HA)
- U+0AB3 (letter LLA)

3. Combined with the observations by Gujarati readers, and from looking at Gujarati dictionaries, the un-modified character precedes the modified character where “modified” is the association of that character with any of a number of modifiers such as nukta, vowel signs etc. This means:

- a non-nukta character (i.e. one without a nukta) precedes its nukta modified form (i.e. one with a nukta).
- similarly for the chandra bindu, anuswara, visarga, and avagraha.

The various signs (generically called “modifiers” above) can be applied to the independent vowels, the consonants, and the dependent vowels. All of the signs are not applicable in all cases. As noted above, the un-modified character precedes the modified character in terms of sorting. Within the modifier set, the following order works:

- nukta
- anuswara
- chandra bindu
- visarga
- consonant + dependent vowel
- avagraha
- halant or virama

4. Dependent vowels like the modifiers above, are combining characters. They generally combine with the consonants.



As said earlier though some vowel appear as the combination of one vowel & and a dependent vowel sign, they should not by mistake, considered as a complex/non-independent vowels.

The following order works for the dependent vowels:

1. vowel sign aa
2. vowel sign i
3. vowel sign ii
4. vowel sign u
5. vowel sign uu
6. vowel sign vocalic r
7. vowel sign vocalic rr
8. vowel sign candra e
9. vowel sign e
10. vowel sign ai
11. vowel sign candra o
12. vowel sign o
13. vowel sign au

Recommended Groupings

These groups are in sort order:

- **Signs:** the Om symbol (given the highest priority since it is a revered religious symbol) followed by the danda, and the double danda.
- **Digits:** Gujarati digits.
- **Independent vowels:** independent vowel followed by modified independent vowel.
- **Consonants:** consonant followed by modified consonant followed by consonant combining with dependent vowels.

Dependent vowels and modifiers are not given a separate place in the collation order. Essentially they are combining marks and as such should not occur independently in any Gujarati text stream. If they do occur, by error, then they will be treated as any undefined characters - after the defined sequences and in code-point order.

Collation Order

Here is the Gujarati sort order:

Signs

- U+0AD0 - Om
- U+0964 - Danda
- U+0965 - Double Danda

Digits

- U+0AE6 - Zero
- U+0AE7 - One
- U+0AE8 - Two
- U+0AE8 - Three
- U+0AEA - Four
- U+0AEB - Five
- U+0AEC - Six
- U+0AED - Seven
- U+0AEE - Eight
- U+0AEF - Nine

Independent Vowels

- U+0A85 - Letter A
 - + U0A82 (anuswara)
 - + U0A81 (chandrabindu)
 - + U0A83 (visarga)
 - + U0ABD (avagraha)
- U+0A86 - Letter AA
 - + U0A82 (anuswara)
 - + U0A81 (chandrabindu)
 - + U0A83 (visarga)
 - + U0ABD (avagraha)
- U+0A87 - Letter I
 - + U0A82 (anuswara)
 - + U0A81 (chandrabindu)
 - + U0A83 (visarga)
 - + U0ABD (avagraha)
- U+0A88 - Letter II
 - + U0A82 (anuswara)
 - + U0A81 (chandrabindu)
 - + U0A83 (visarga)
 - + U0ABD (avagraha)
- U+0A89 - Letter U
 - + U0A82 (anuswara)



+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)
U+0A8A - Letter UU
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)
U+0A8B - Letter Vocalic R
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)
U+0AE0 - Letter Vocalic RR
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)
U+0A8D - Vowel Candra E
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)
U+0A8F - Letter E
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)
U+0A90 - Letter AI
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)
U+0A91 - Vowel Candra O
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)
U+0A93 - Letter O
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)

U+0A94 - Letter AU
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
+ U0ABD (avagraha)

Consonants

U+0A95 - Letter KA
+ U0ABC (nukta)
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
consonant + dependent vowel
+ U0ABD (avagraha)
+ U0ACD (halant)
+U0ABE
+U0ABF
+ U0AC0
+ U0AC1
+ U0AC2
+ U0AC3
+ U0AC4
+ U0AC5
+ U0AC7
+ U0AC8
+ U0AC9
+ U0ACB
+ U0ACC

U+0A96 - Letter KHA
+ U0ABC (nukta)
+ U0A82 (anuswara)
+ U0A81 (chandrabindu)
+ U0A83 (visarga)
consonant + dependent vowel
+ U0ABD (avagraha)
+ U0ACD (halant)
+ U0ABE
+ U0ABF
+ U0AC0
+ U0AC1
+ U0AC2
+ U0AC3
+ U0AC4
+ U0AC5



+ U0AC7
 + U0AC8
 + U0AC9
 + U0ACB
 + U0ACC
 U+0A97 - Letter GA
 + U0ABC (nukta)
 + U0A82 (anuswara)
 + U0A81 (chandrabindu)
 + U0A83 (visarga)
 consonant + dependent vowel
 + U0ABD (avagraha)
 + U0ACD (halant)
 + U0ABE
 + U0ABF
 + U0AC0
 + U0AC1
 + U0AC2
 + U0AC3
 + U0AC4
 + U0AC5
 + U0AC7
 + U0AC8
 + U0AC9
 + U0ACB
 + U0ACC

As we can see from the above list that the collation sequence for all the consonants remains same, instead of repeating the same sequence for all the consonants, we are listing below the order of the consonants only. (Here note that the letter LLA(+U0AB3) has code position next to LA in Unicode chart but in collation sequence it comes after letter HA(+0AB9))

U+0A98 - Letter GHA
 U+0A99 - Letter NGA
 U+0A9A - Letter CA
 U+0A9B - Letter CHA
 U+0A9C - Letter JA
 U+0A9D - Letter JHA
 U+0A9E - Letter NYA
 U+0A9F - Letter TTA
 U+0AA0 - Letter TTHA
 U+0AA1 - Letter DDA

U+0AA2 - Letter DDHA
 U+0AA3 - Letter NNA
 U+0AA4 - Letter TA
 U+0AA5 - Letter THA
 U+0AA6 - Letter DA
 U+0AA7 - Letter DHA
 U+0AA8 - Letter NA
 U+0AAA - Letter PA
 U+0AAB - Letter PHA
 U+0AAC - Letter BA
 U+0AAD - Letter BHA
 U+0AAE - Letter MA
 U+0AAF - Letter YA
 U+0AB0 - Letter RA
 U+0AB2 - Letter LA
 U+0AB5 - Letter VA
 U+0AB6 - Letter SHA
 U+0AB7 - Letter SSGA
 U+0AB8 - Letter SGA
 U+0AB9 - Letter HA
 U+0AB3 - Letter LLA



Annexure - I

Gujarati Script

Encoding Principles. The writing system followed by Gujarati is the same as that of Devanagari, i.e. they constitute a cross between syllabic writing systems and phonemic writing systems (alphabets). The effective unit of Gujarati writing system is the orthographic syllable, consisting of a consonant and vowel (CV) core and, optionally, one or more preceding consonants, with a canonical structure of ((C)C)CV. The orthographic syllable need not correspond exactly with a phonological syllable, especially when a consonant cluster is involved, but the writing system is built on phonological principles and tends to correspond quite closely to pronunciation.

The orthographic syllable is built up of alphabetic pieces, the actual letters of Gujarati script. These pieces consist of three distinct character types : consonant letters, independent vowels and dependent vowel signs. In a text sequence, these characters are stored in logical (phonetic) order.

Principles of the Script

Rendering Gujarati Characters : Gujarati characters, like characters from many other scripts, can combine or change shape depending on their context. A character's appearance is affected by its ordering with respect to other characters, the font used to render the character, and the application or system environment. These variables can cause the appearance of Gujarati characters to differ from their nominal glyphs (used in the code charts).

Additionally, a few Gujarati characters cause a change in the order of the displayed characters. This rendering is not commonly seen in non-Indic scripts and occurs independently of any bidirectional character reordering that might be required.

Consonant Letters. Each consonant letter represents single consonantal sound but also has the peculiarity of having an inherent vowel, generally the short vowel /a/ in Gujarati. Thus U+0A95 Gujarati Letter KA represents not just /k/ but also /ka/. In the presence

of a dependent vowel, however, the inherent vowel associated with a consonant letter is overridden by the dependent vowel.

Consonant letters may also be rendered as half-forms, which are presentation forms used to depict the initial consonant in consonant clusters. These half-forms do not have an inherent vowel. Their rendered forms in Gujarati often resemble the full consonant but are missing the vertical stem, which marks a syllabic core. (The stem glyph is graphically and historically related to the sign denoting the inherent /a/ vowel.)

Some Gujarati consonant letters have alternative presentation forms whose choice depends upon neighboring consonants. This variability is especially notable for U+0AB0 Gujarati Letter RA, which has numerous different forms, both as the initial element and as the final element of a consonant cluster. Only the nominal forms, rather than the contextual alternatives, are depicted in the code chart.

The traditional Gujarati alphabetic encoding order for consonants follows articulatory phonetic principles, starting with velar consonants and moving forward to bilabial consonants, followed by liquids and then fricatives. ISCII and the Unicode Standard both observe this traditional order.

Independent Vowel Letters : The independent vowels in Gujarati are letters that stand on their own. The writing system treats independent vowels as orthographic CV syllables in which the consonant is null. The independent vowel letters are used to write syllables that start with a vowel.

Dependent Vowel Signs (Matras) : The dependent vowels serve as the common manner of writing noninherent vowels and are generally referred to as vowel signs, or as matras in Sanskrit. The dependent vowels do not stand alone; rather, they are visibly depicted in combination with a base letterform. A single consonant, or a consonant cluster, may have a dependent vowel applied to it to indicate the vowel quality of the syllable, when it is different from the inherent vowel. Explicit appearance of a dependent



vowel in a syllable overrides the inherent vowel of a single consonant letter.

The greatest variation among different Indic scripts is found in the way that the dependent vowels are applied to base letterforms. Gujarati has a collection of nonspacing dependent vowel signs that may appear above or below a consonant letter, as well as spacing dependent vowel signs that may occur to the right or to the left of a consonant letter or consonant cluster. Other Indic scripts generally have one or more of these forms, but what is a nonspacing mark in one script may be a spacing mark in another. Also, Gujarati like some other Indic scripts has single dependent vowels that are indicated by two or more glyph components-and those glyph components may surround a consonant letter both to the left and right or may occur both above and below it.

The Gujarati script has only one character denoting a left-side dependent vowel sign: U+0ABF Gujarati Vowel Sign I.

A one-to-one correspondence exists between the independent vowels and the dependent vowel signs. Independent vowels are sometimes represented by a sequence consisting of the independent form of the vowel /a/ followed by a dependent vowel sign. For example, Figure-1 illustrates this relationship (see the notation formally described in the “Rules for Rendering” later in this section).

The combination of the independent form of the default vowel /a/ (in the Gujarati script, U+0A85 Gujarati Letter A) with a dependent vowel sign may be viewed as an alternative spelling of the phonetic information normally represented by an isolated independent vowel form.

Figure 1. Dependent Versus Independent Vowels

/a/ +Dependent Vowel	Independent Vowel
$A_n + I_{vs} \rightarrow I_{vs} + A_n$	I_n
અ + િ → િ	ઈ
$A_n + U_{vs} \rightarrow A_n + U_{vs}$	U_n
અ + ં → ં	ઉ

Higher-level text processes may choose to consider these alternative spellings equivalent in terms of information content, but such an equivalence is not stipulated by this standard.

Halant : Gujarati employs a sign known as halant, or vowel omission sign. It is also informally as referred to as *khodo* in Gujarati. A halant sign U+0ACD Gujarati Sign Halant) nominally serves to cancel (or kill) the inherent vowel of the consonant to which it is applied. The halant functions as a combining character, with its shape varying from script to script. When a consonant has lost its inherent vowel by the application of halant, it is known as dead consonant; in contrast, a live consonant is one that retains its inherent vowel or is written with an explicit dependent vowel sign. In the Unicode Standard, a dead consonant is defined as a sequence consisting of a consonant letter followed by a halant. The default rendering for a dead consonant is to position the halant as a combining mark bound to the consonant letterform.

For example, if C_n denotes the nominal form of consonant C, and C_d denotes the dead consonant form, then a dead consonant is encoded as shown in Figure 2.

Figure 2 Dead Consonants

$TA_n + HALANT_n$	\rightarrow	TA_d
ઠ + ં	\rightarrow	ઠ

Consonant Conjuncts. The Indic scripts are noted for a large number of consonant conjunct forms that serve as orthographic abbreviations (ligatures) of two or more adjacent letterforms. This abbreviation takes place only in the context of a consonant cluster. An orthographic consonant cluster is defined as a sequence of characters that represents one or more dead consonants (denoted C_d) followed by a normal, live consonant letter (denoted C) or an independent vowel letter.

Under normal circumstances, a consonant cluster is depicted with a conjunct glyph if such a glyph is available in the current font(s). In the absence of a



conjunct glyph, the one or more dead consonants that form part of the cluster are depicted using half-form glyphs. In the absence of half-form glyphs, the dead consonants are depicted using the nominal consonant forms combined with visible virama signs (see Figure 3)

Figure 3 Conjunct Formations

- (1) $GA_d + DHA_l \rightarrow GA_h + DHA_n$
 $ग़ + ध \rightarrow ग़ध$
- (2) $KA_d + KA_l \rightarrow K.KA_n$
 $क़ + क \rightarrow क़क$
- (3) $KA_d + SSHA_l \rightarrow K.SSHA_n$
 $क़ + ष \rightarrow क़ष$
- (4) $RA_d + RI_n \rightarrow RI_n + RA_{sup}$
 $ऱ + र \rightarrow ऱर$

A number of types of conjunct formations appear in these examples: (1) a half-form of GA in its combination with the full form of DHA; (2) a vertical conjunct K.KA; (3) a fully ligated conjunct K.SSHA, in which the components are no longer distinct; and (4) a rare conjunct formed with an independent vowel letter, in this case the vowel letter RI (also known as vocalic r). Note that in example (4) in Figure 3, the dead consonant RA_d is depicted with the nonspacing combining mark RA_{sup} (repha).

A well-designed Indic script font may contain hundreds of conjunct glyphs, but they are not encoded as Unicode characters because they are the result of ligation of distinct letters. Indic script rendering software must be able to map appropriate combinations of characters in context to the appropriate conjunct glyphs in fonts.

When an independent vowel appears as the terminal element of a consonant cluster, as in example (4) in Figure 3, the independent vowel should not be depicted as a dependent vowel sign, but as an independent vowel letterform.

Explicit Halant : Normally a halant character serves to create dead consonants that are, in turn, combined

with subsequent consonants to form conjuncts. This behavior usually results in a halant sign not being depicted visually. Occasionally, however, this default behavior is not desired when a dead consonant should be excluded from conjunct formation, in which case the virama sign is visibly rendered. To accomplish this goal, the Unicode Standard adopts the convention of placing the character U+200C Zero Width Non-joiner immediately after the encoded dead consonant that is to be excluded from conjunct formation. In this case, the halant sign is always depicted as appropriate for the consonant to which it is attached.

For example, in Figure 4, the use of Zero Width Non-joiner prevents the default formation of the conjunct form क़ष (K.SSHA_n).

Figure 4. Preventing Conjunct Forms

- $KA_d + ZWNJ + SSHA \rightarrow KA_d + SSHA_n$
 $क़ + ZWNJ + ष \rightarrow क़ष$

Explicit Half-Consonants : When a dead consonant participates in forming a conjunct, the dead consonant form is often absorbed into the conjunct form such that it is no longer distinctly visible. In other contexts, however, the dead consonant may remain visible as a half-consonant form. In general, a half-consonant form is distinguished from the nominal consonant form by the loss of its inherent vowel stem, a vertical stem appearing to the right side of the consonant form. In other cases, the vertical stem remains but some part of its right-side geometry is missing.

In certain cases, it is desirable to prevent a dead consonant from assuming full conjunct formation yet still not appear with an explicit halant. In these cases, the half-form of the consonant is used. To explicitly encode a half-consonant form, the Unicode Standard adopts the convention of placing the character U+200D Zero Width Joiner immediately after the encoded dead consonant. The Zero Width Joiner denotes a nonvisible letter that presents linking or cursive joining behavior on either side (that is, to



the previous or following letter). Therefore, in the present context, the Zero Width Joiner may be considered to present a context to which a preceding dead consonant may join so as to create the half-form of the consonant.

For example, if C_h denotes the half-form glyph of consonant C, then a half-consonant form is encoded as shown in Figure 5.

Figure 5. Half-Consonants

$$KA_d + ZWJ + SSHA_n \rightarrow KA_d + SSHA_n$$

$$ઠ + ZWJ + ષ \rightarrow ઠષ$$

- In the absence of the Zero Width Joiner, this sequence would normally produce the full conjunct form ઠઠ (K.SSHA_n).

This encoding of half-consonant forms also applies in the absence of a base letterform. That is, this technique may also be used to encode independent half-forms, as shown in Figure 6.

Consonant Forms. In summary, each consonant may be encoded such that it denotes a live consonant, a dead consonant that may be absorbed into a conjunct, or the half-form of a dead consonant (see Figure 7)

Figure 6 Independent Half-Forms

$$GA_d + ZWJ \rightarrow GA_h$$

$$ઠ + ZWJ \rightarrow ઠ$$

Figure 7 Consonant Forms

$$\text{ઠ} \rightarrow \text{ઠ} \quad KA_l$$

$$\text{ઠ} + \text{ઠ} \rightarrow \text{ઠ} \quad KA_d$$

$$\text{ઠ} + \text{ઠ} + \text{ZWJ} \rightarrow \text{ઠ} \quad KA_h$$

Rendering

Rules for Rendering. The following provides more formal and detailed rules for minimal rendering of Gujarati as part of a plain text sequence. It describes the mapping between Unicode characters and the glyphs in a Gujarati font. It also describes the combining and ordering of those glyphs.

These rules provide minimal requirements for legibly rendering interchanged Gujarati text. As with any script, a more complex procedure can add rendering characteristics, depending on the font and application.

It is important to emphasize that in a font that is capable of rendering Gujarati, the set of glyphs is greater than the number of Gujarati Unicode characters.

Notation. In the next set of rules, the following notation applies:

- C_n Nominal glyph form of consonant C as it appears in the code charts.
- C_l A live consonant, depicted identically to C_n .
- C_d Glyph depicting the dead consonant form of consonant C.
- C_h Glyph depicting the half-consonant form of consonant C.
- L_n Nominal glyph form of a conjunct ligature consisting of two or more component consonants. A conjunct ligature composed of two consonants X and Y is also denoted $X.Y_n$.
- RA_{sup} A nonspacing combining mark glyph form of the U+0AB0 Gujarati Letter RA positioned above or attached to the upper part of a base glyph form. This form is also known as repha.
- RA_{sub} A nonspacing combining mark glyph form of the U+0AB0 Gujarati Letter RA positioned below of attached to the lower part of a base glyph form.
- V_{vs} Glyph depicting the dependent vowel sign form of a vowel V.
- $HALANT_n$ The nominal glyph form nonspacing combining mark depicting U+0ACD Gujarati Sign Halant.



- A halant character is not always depicted; when it is depicted, it adopts this nonspacing mark form.

Dead Consonant Rule. The following rule logically precedes the application of any other rule to form a dead consonant. Once formed, a dead consonant may be subject to other rules described next.

R1 When consonant C_n precedes a $HALANT_n$, it is considered to be dead consonant C_d . A consonant C_n that does not precede $HALANT_n$ is considered to be a live consonant C_l .

$$TA_n + HALANT_n \rightarrow TA_d \quad \text{ㄊ + ॠ \rightarrow ㄊ}$$

Consonant RA Rules. The character U+0AB0 Gujarati Letter RA takes one of a number of visual forms depending on its context in a consonant cluster. By default, this letter is depicted with its nominal glyph form (as shown in the code charts). In two contexts, it is depicted using a nonspacing glyph form that combines with a base letterform.

R2 If the dead consonant RA_d precedes either a consonant or an independent vowel, then it is replaced by the superscript nonspacing mark RA_{sup} , which is positioned so that it applies to the logically subsequent element in the memory representation.

$$\begin{aligned} RA_d + KA_l &\rightarrow KA_l + RA_{sup} && \text{Displayed Output} \\ \text{ર} + \text{ક} &\rightarrow \text{ક} + \text{ઠ} && \rightarrow \text{કઠ} \\ RA_d^1 + RA_d^2 &\rightarrow RA_d^2 + RA_{sup}^1 \\ \text{ર} + \text{ર} &\rightarrow \text{ર} + \text{ઠ} && \rightarrow \text{રઠ} \end{aligned}$$

R3 If the superscript mark RA_{sup} is to be applied to a dead consonant and that dead consonant is combined with another consonant to form a conjunct ligature, then the mark is positioned so that it applies to the conjunct ligature form as a whole.

$$\begin{aligned} RA_d + JA_d + NA_l &\rightarrow G.NA_n + RA_{sup} && \text{Displayed Output} \\ \text{ર} + \text{જ} + \text{ન} &\rightarrow \text{જન} + \text{ઠ} && \rightarrow \text{જનઠ} \end{aligned}$$

R4 If the superscript mark RA_{sup} is to be applied to a dead consonant that is subsequently replaced by its half-consonant form, then the mark is positioned so that it applies to the form that serves as the base of the consonant cluster.

$$\begin{aligned} RA_d + GA_d + GHA_l &\rightarrow GA_h + GHA_l + RA_{sup} && \text{Displayed Output} \\ \text{ર} + \text{ગ} + \text{ઘ} &\rightarrow \text{ગ} + \text{ઘ} + \text{ઠ} && \rightarrow \text{ગઘઠ} \end{aligned}$$

R5 Except for the dead consonant RA_d , when a dead consonant C_d precedes the live consonant RA_l , then C_d is replaced with its nominal form C_n , and RA is replaced by the subscript nonspacing mark RA_{sub} , which is positioned so that it applies to C_n .

$$\begin{aligned} THA_d + RA_d &\rightarrow THA_n + RA_{sub} && \text{Displayed Output} \\ \text{ઠ} + \text{ર} &\rightarrow \text{ઠ} + \text{ઠ} && \rightarrow \text{ઠઠ} \end{aligned}$$

R6 For certain consonants, the mark RA_{sub} may graphically combine with the consonant to form a conjunct ligature form. These combinations, such as the one shown here, are further addressed by the ligature rules.

$$\begin{aligned} PHA_d + RA_l &\rightarrow PHA_n + RA_{sub} && \text{Displayed Output} \\ \text{ફ} + \text{ર} &\rightarrow \text{ફ} + \text{ઠ} && \rightarrow \text{ફઠ} \end{aligned}$$

R7 If a dead consonant (other than RA_d) precedes RA_d , then the substitution of RA for RA_{sub} is performed as described above; however, the HALANT that formed RA_d remains so as to form a dead consonant conjunct form.

$$\begin{aligned} TA_d + RA_d &\rightarrow THA_n + RA_{sub} + HALANT_n && \text{Displayed Output} \\ \text{ઠ} + \text{ર} &\rightarrow \text{ઠ} + \text{ઠ} + \text{ઠ} && \rightarrow \text{ઠઠઠ} \end{aligned}$$

A dead consonant conjunct form that contains an absorbed RA_d may subsequently combine to form a multipart conjunct form.

$$\begin{aligned} T.RA_d + YA_l &\rightarrow T.R.YA_n \\ \text{ઠ} + \text{ય} &\rightarrow \text{ઠય} \end{aligned}$$



Modifier Mark Rules. In addition to vowel signs, three other types of combining marks may be applied to a component of an orthographic syllable or to the syllable as a whole: bindus, and svaras.

R8 The other modifying marks, bindus and svaras, apply to the orthographic syllable as a whole and should follow (in the memory representation) all other characters that constitute the syllable. In particular, the bindus should follow any vowel signs, and the svaras should come last. The relative placement of these marks is horizontal rather than vertical; the horizontal rendering order may vary according to typographic concerns.

$$KA_n + AA_{vs} + BINDU_n \\ \text{क} + \text{अ} + \text{◌} \rightarrow \text{कः}$$

Ligature Rules. Subsequent to the application of the rules just described, a set of rules governing ligature formation apply. The precise application of these rules depends on the availability of glyphs in the current font(s) being used to display the text.

R9 If a dead consonant immediately precedes another dead consonant or a live consonant, then the first dead consonant may join the subsequent element to form a two part conjunct ligature form.

$$HA_d + YA_l \rightarrow HYA_n \\ \text{ह} + \text{य} \rightarrow \text{ह्य}$$

R10 A conjunct ligature form can itself behave as a dead consonant and enter into further, more complex ligatures.

$$SA_d + TA_d + RA_n \rightarrow SA_d + T.R.A_n \rightarrow S.T.RA_n \\ \text{स} + \text{त} + \text{र} \rightarrow \text{स} + \text{त्र} \rightarrow \text{सत्र}$$

A conjunct ligature form can also produce a half-form.

$$K.SSHA_d + YA_l \rightarrow K.SSH_h + YA_n \\ \text{क्ष} + \text{य} \rightarrow \text{क्षय}$$

R11 If a nominal consonant or conjunct ligature form precedes RA_{sub} then the consonant or

ligature form may join with RA_{sub} to form a multipart conjunct ligature.

$$KA_n + RA_{sub} \rightarrow K.RA_{sub} \\ \text{क} + \text{◌} \rightarrow \text{क◌}$$

$$PHA_n + RA_{sub} \rightarrow PH.RA_n \\ \text{फ} + \text{◌} \rightarrow \text{फ◌}$$

R12 In some cases, other combining marks will also combine with a base consonant, either attaching at a nonstandard location or changing shape. In minimal rendering there are only two cases, in Gujarati RA_l , with U_{vs} or UU_{vs} .

$$RA_l + U_{vs} \rightarrow RU_n \\ \text{ર} + \text{◌} \rightarrow \text{રુ}$$

$$RA_l + U_{vs} \rightarrow RU_n \\ \text{ર} + \text{◌} \rightarrow \text{રૂ}$$

R12a In addition to these two cases there is one more case for the consonant JA_l in Gujarati, where combining mark attaches to the base consonant at the non standard location. Here when JA_l combines with II_{vs} or AA_{vs} , the shape of the resultant ligature changes to a nonstandard form.

$$JA_l + II_{vs} \rightarrow JII_l \\ \text{જ} + \text{◌} \rightarrow \text{જી}$$

$$JA_l + AA_{vs} \rightarrow JAA_l \\ \text{જ} + \text{◌} \rightarrow \text{જા}$$

Memory Representation and Rendering Order. The order for storage of plain text in Gujarati follows phonetic order; that is, a CV syllable with a dependent vowel is always encoded as a consonant letter C followed by a vowel sign V in the memory representation. This order is employed by the ISCII standard and corresponds with both the phonetic and keying order of textual data. (Fig. 8).

Character Order		Glyph Order
$KA_n + I_{vs}$	\rightarrow	$I_{vs} + KA_n$
$\text{ક} + \text{◌}$	\rightarrow	ક◌

Gujarati also has some dependent vowels that must be depicted to the left side of their consonant letter,



the software that renders the Indic scripts must be able to reorder elements in mapping from the logical(character) store to the presentational(glyph) rendering. For example, if C_n denotes the nominal form of consonant C , and V_{vs} denotes a left-side dependent vowel sign form of vowel V , then a reordering of glyphs with respect to encoded characters occurs as just shown.

R13 When the dependent vowel l_{vs} is used to override the inherent vowel of a syllable, it is always written to the extreme left of the orthographic syllable. If the orthographic syllable contains a consonant cluster, then this vowel is always depicted to the left of that cluster. For example:

$TA_d + RA_l + l_{vs} \rightarrow T.RA_n + l_{vs} \rightarrow l_{vs} + T.RA_a$
 $त् + र + ङ \rightarrow त्र + ङ \rightarrow त्रि$

As Nukta is not used frequently in Gujarati we have not mentioned rules involving nukta. So whenever it occurs in Gujarati rendering rules remain same as those in Devanagari.

Sample Half-Forms. Table 1 shows examples of half-consonant forms that are commonly used with the Gujarati script. These forms are glyphs, not characters. They may be encoded explicitly using ZERO WIDTH JOINER as shown; in normal conjunct formation, they may be used spontaneously to depict a dead consonant in combination with subsequent consonant forms.

Table 1. Sample Half-Forms

ક	◌	ZWJ	ક	ન	◌	ZWJ	ન
ખ	◌	ZWJ	ખ	પ	◌	ZWJ	પ
ગ	◌	ZWJ	ગ	ભ	◌	ZWJ	ભ
ઘ	◌	ZWJ	ઘ	ઝ	◌	ZWJ	ઝ
ચ	◌	ZWJ	ચ	ષ	◌	ZWJ	ષ
જ	◌	ZWJ	જ	ટ	◌	ZWJ	ટ
ઝ	◌	ZWJ	ઝ	ડ	◌	ZWJ	ડ
ઞ	◌	ZWJ	ઞ	ણ	◌	ZWJ	ણ
ત	◌	ZWJ	ત	થ	◌	ZWJ	થ

થ ◌ ZWJ થ સ ◌ ZWJ સ
 ઘ ◌ ZWJ ઘ

Sample Ligatures. Table 2 shows examples of conjunct ligature forms that are commonly used with the Gujarati script. These forms are glyphs, not characters. Furthermore, individual fonts may provide fewer or more ligature forms than are depicted here.

Table 2 Sample Ligature

ક	◌	ક	ક	હ	◌	મ	હ
ક	◌	ત	ક	હ	◌	ય	હ
ક	◌	ર	ક	હ	◌	ૃ	હ
ક	◌	ષ	હ	શ્ર	◌	ૃ	શ્ર
જ	◌	ઞ	જ	ઠ	◌	ય	ઠ
ઠ	◌	ૃ	ઠ	ઠ	◌	મ	ઠ
ઠ	◌	ધ	ઠ	ઠ	◌	ર	ઠ
ઠ	◌	ઠ	ઠ				
ટ	◌	ટ	ટ				
ત	◌	ત	ત				
ત	◌	ર	ત				
ન	◌	ન	ન				
ફ	◌	ર	ફ				
શ	◌	ર	શ				

Sample Half-Ligature Forms. In addition to half-form glyphs of individual consonants, half-forms are also used to depict conjunct ligature forms. A sample of such forms is shown in Table 3. These forms are glyphs, not characters. They may be encoded explicitly using zero with joiner as shown; in normal conjunct formation, they may be used spontaneously to depict a conjunct ligature in combination with subsequent consonant forms.

Table 3 Sample Half-Ligature Forms

ક	◌	ષ	◌	ZWJ	ક
ત	◌	ત	◌	ZWJ	ત
ત્ર	◌	ર	◌	ZWJ	ત્ર
શ	◌	ર	◌	ZWJ	શ

Combining Marks. Gujarati, has a number of combining marks the could be considered diacritic. One class of these marks, shown as bindus, is



represented by U+0A81 Gujarati Sign Chandrabindu and U+0A82 Gujarati Sign Anusvara. These marks indicate nasalization or nasal closure of a syllable.

Digits. Gujarati scripts has a distinct set of digits and they are used with the ordinary text in practice.

Punctuation and Symbols. Gujarati script, in general uses the symbols for punctuation marks, as those of Latin script.

Encoding structure. The Unicode Standard organizes the nine principal Indic Scripts in block of 128 encoding points each. The first six columns are identical with ISCII-1988 encoding, except that the last 11 positions, which are unassigned or undefined in ISCII-1988 are used in the Unicode encoding.

The seventh column in each of these scripts, along with the last 11 position in the sixth column, represent additional character assignments in the Unicode Standard that are matched across all nine scripts. For example, positions U+xx66..U+xx6F and U+xxE6..U+xxEF code the Indic script digits for each script.

The eighth column in Unicode standard for each script is reserved for script-specific additions that do no correspond from one Indic script to the next. There are certain characters which are used in Gujarati only. The first one is currency symbol રૂ. and others are ૧, ૨, ૩ which are used to denote the value 0.25, 0.5 and 0.75 respectively when they follow any digit.

*(Courtesy : Prof. Sitanshu Y. Mehta
M.S. University of Baroda,
Baroda-390002
Tel. 0265-792959
E-mail : sityash@satyam.net.in
&
Sh. Jignesh Dholakia
Director IT Services, RC MSU
E-mail : jignesh_dholakia@yahoo.com)*

4.1.3 Typical Colloquial Sentences in Gujarati

GREETING

- ▶ Hello
નમસ્તે / નમસ્કાર / જયજય
નમસ્તે / નમસ્કાર / જયજય
Namste / Namaskar / Jay Jay
 - ▶ Good Morning
સુપ્રભાત
સુપ્રભાત
Suprabhat
 - ▶ Good Night
શુભરાત્રિ
શુભરાત્રિ
Shubharatri
 - ▶ Good Bye
આવજો
આવજો
Aavajo
 - ▶ Thanks
આભાર
આભાર
Aabhar
 - ▶ How are you
કેમ છો ?
કેમ છો ?
Kem Chho ?
 - ▶ I am fine thank you
હું મજામાં છું, આભાર. / મજામાં
હું મજામાં છું, આભાર. / મજામાં
Hu majaman chhu, Aabhar / Majaman
 - ▶ Sorry
દિલગીર છું.
દિલગીર છું.
Dilgir chhu.
- #### WEATHER
- ▶ It is cold
ખૂબ જ ઠંડી પડે છે. / ટાઢ છે.
ખૂબ જ ઠંડી પડે છે. / ટાઢ છે.
Khuba j thandi pade chhe. / Tadh chhe.



- It is cool outside
બહાર ખૂબ જ ઠંડી છે. / બહાર બહુ ઠાઢ છે.
બહાર ખૂબ જ ઠંડી છે. / બહાર બહુ ઠાઢ છે.
Bhahar khuba j thandi chhe. / bahar bahu tadha chhe.
- It is hot
ખૂબ જ ઘામ છે. / ગરમી બહુ છે.
ખૂબ જ ઘામ છે. / ગરમી બહુ છે.
khuba j gham chhe. / garami bahu chhe.
- It is raining
વરસાદ પડે છે. / પાણી વરસે છે.
વરસાદ પડે છે. / પાણી વરસે છે.
Varasad pade chhe. / paani varase chhe.

GENERAL

- What is your name?
તમારું નામ શું છે ? / નામ શું તમારું ?
તમારું નામ શું છે ? / નામ શું તમારું ?
Tamaaru naam shu chhe ? / Naam shu tamaaru?
- My name is Ranjan
મારું નામ રંજન છે. / મારું નામ રંજન.
મારું નામ રંજન છે. / મારું નામ રંજન.
Maaru naam ranjan chhe. / maaru naam ranjan.
- Where do you live?
તમે ક્યાં રહો છો ? / ઘર ક્યાં તમારું ?
તમે ક્યાં રહો છો ? / ઘર ક્યાં તમારું ?
Tame kyan raho chho ? ghar kyan tamaru ?
- I live near Ghantaghar
હું ઘંટાઘર નજીક રહું છું.
હું ઘંટાઘર નજીક રહું છું.
Hug ghantaghar najeeek rahu chhu.
- How old are you?
તમારી ઉંમર શું છે ?
તમારી ઉંમર શું છે ?
Tamari ummar shu chhe ?
- That building is tall
પેલું મકાન ઉચું છે.
પેલું મકાન ઉચું છે.
Pelu mkaan uchu che.

- She is beautiful
તે સુંદર છે. / એ દેખાવડી છે.
તે સુંદર છે. / એ દેખાવડી છે.
Te sundar chhe. / E dekhavadi chhe.
- I like Bengali sweets
મને બંગાળી મીઠાઈ ભાવે છે.
મને બંગાળી મીઠાઈ ભાવે છે.
Mane bangali mithai bhav chhe.
- I love birds
મને પંખીઓ ગમે છે. / હું પક્ષીઓને ચાહું છું.
મને પંખીઓ ગમે છે. / હું પક્ષીઓને ચાહું છું.
Mane pankhio game chhe. / hu pakshione chahu chhu .
- Where is Railway station?
રેલ્વે સ્ટેશન ક્યાં છે ?
રેલ્વે સ્ટેશન ક્યાં છે ?
ralve stashan kyan chhe ?
- How far is the Bus Terminal from here?
બસ અડ્ડો અહીંથી કેટલો દૂર છે ?
બસ અડ્ડો અહીંથી કેટલો દૂર છે ?
bas addo ahithi ketalo dur chhe ?
- How long will it take to reach the Airport?
એરપોર્ટ પહોચતાં કેટલી વાર લાગશે ?
એરપોર્ટ પહોચતાં કેટલી વાર લાગશે ?
erport pahochata ketali var lagashe ?
- Is Mr. Raghunath there?
શું શ્રી રઘુનાથ છે ? / રઘુનાથભાઈ છે કે ?
શું શ્રી રઘુનાથ છે ? / રઘુનાથ ભાઈ છે કે ?
shu shri raghunath chhe ? / raghunathhai chhe ke ?
- Please tell him to call back as soon as he is free
જેવા એ નવરા થાય કે તરત જ ફોન કરવા કહેજો.
જેવા એ નવરા થાય કે તરત જ ફોન કરવા કહેજો.
Jeva e navara thay ke tarat j karava kahejo.
- How much will it cost?
તેની કિંમત શું છે ?
તેની કિંમત શું છે ?
teni kimat shu chhe ?



► Excuse me

માફ કરજો.

માફ કરજો

maaf karajo

► From which Platform can I get the train for Chandigarh?

મને ચંડીગઢ જવા માટે ગાડી કયા પ્લેટફોર્મ પરથી મળશે ?

મને ચંડીગઢ જવા માટે ગાડી કયા પ્લેટફોર્મ પરથી મળશે ?

Mane chandigadha java maate gaasi kaya platform parthi malashe ?

► Does this train stop at Aligarh?

શું આ ગાડી અલીગઢ ઉભી રહેશે ?

શું આ ગાડી અલીગઢ ઉભી રહેશે ?

shu aa gaadi aligadh ubhi raheshe ?

► How many kids do you have?

તમને કેટલાં બાલકો છે ?

તમને કેટલાં બાલકો છે ?

Tamane ketala balako chhe ?

► This gift is wonderful

આ ભેટ અદ્ભૂત છે.

આ ભેટ અદ્ભૂત છે.

Aa bhet adbhut chhe.

► It is really pretty

તે ખરેખર સુંદર છે.

તે ખરેખર સુંદર છે.

Te kharekhar sundar chhe.

► Food is delicious

ખાણું સ્વાદિષ્ટ છે.

ખાણું સ્વાદિષ્ટ છે.

khanu swadisht chhe.

► Congratulations

અભિનંદન

અભિનંદન

Abhinandan

► You look lovely

તમે ખૂબ જ સુંદર લાગો છો.

તમે ખૂબ જ સુંદર લાગો છો.

Tame khuba j sundar lago chho.

► Wish you happy new year

સાલમુબારક / નૂતન વર્ષાભિનંદન

સાલમુબારક / નૂતન વર્ષાભિનંદન

Salmubarak / nutan varshabhinandan.

► I wish you all the happiness

તું સહુ સુખ પામજે / સદા સુખી રહે

તું સહુ સુખ પામજે / સદા સુખી રહો

Tu sahu sukh pamje / sada sukhi raho.

► Congratulations on your marriage

લગ્ન-પ્રસંગે અભિનંદન.

લગ્ન-પ્રસંગે અભિનંદન.

lagna-prasange abhinandan.

► Keep your eyes wide open before marriage and half- shut afterwards

ખુલ્લી આંખે લગ્ન કરો, બાદમાં અડધી બીડી રાખો.

ખુલ્લી આંખે લગ્ન કરો, બાદમાં અડધી બીડી રાખો.

khulli aankhe lagna karo, badma addhi beedee rakho.

(Courtesy : Shri Sitanshu Y. Mehta

M.S. University of Baroda,

Baroda-390002

Tel. 0265-792959

E-mail : sityash@satyam.net.in)