

en.wikipedia.org

Sinhala script - Wikipedia

30-38 minutes

<div>Sinhala script (Sinhalese)</div> <div>සිංහල අක්ෂර මාලාව</div> <div>Siṁhala Akṣara Mālāva</div>	
<div>සිංහල</div>	
Type	Abugida
Languages	Sinhala , Pali , Sanskrit
Time period	C. 700–present
Parent systems	<div>Proto-Sinaitic alphabet</div> <ul style="list-style-type: none"> Phoenician alphabet Aramaic alphabet Brāhmī Kadamba Sinhala script (Sinhalese)
Sister systems	<div>Goykanadi</div> <div>Cham alphabet</div> <div>Tigalari alphabet</div> <div>Malayalam script</div> <div>Dhives akuru</div>
Direction	Left-to-right
ISO 15924	Sinh, 348
Unicode alias	Sinhala
Unicode range	<ul style="list-style-type: none"> U+0D80–U+0DFF Sinhala U+111E0–U+111FF Sinhala Archaic Numbers
<p>This article contains IPA phonetic symbols. Without proper rendering support, you may see question marks, boxes, or other symbols instead of Unicode characters. For an introductory guide on IPA symbols, see Help:IPA.</p>	

Brahmic scripts
The Brahmic script and its descendants
Northern Brahmic [show]
Southern Brahmic [show]
<ul style="list-style-type: none"> v t



Sinhala script (*Sinhala*: සිංහල අක්ෂර මාලාව; *Sinhala Akṣara Mālāva*), also known as **Sinhalese script**, is a [writing system](#) used by the [Sinhalese people](#) and most [Sri Lankans](#) in [Sri Lanka](#) and elsewhere to write the [Sinhala language](#), as well as the [liturgical languages](#), [Pali](#) and [Sanskrit](#).^[1] The Sinhalese Akṣara Mālāva, one of the [Brahmic scripts](#), is a descendant of the ancient [Ancient Indian Brahmi script](#) and closely related to the South Indian [Kadamba alphabet](#).^{[2][1]}

Sinhala script is an [abugida](#) written from left to right. Sinhala letters are ordered into two sets. The core set of letters forms the ***śuddha siṃhala*** alphabet (Pure Sinhala, ශුද්ධ සිංහල), which is a subset of the ***miśra siṃhala*** alphabet (Mixed Sinhala, මිශ්‍ර සිංහල).

History^[edit]

Sinhala script is a [Brahmi derivate](#), and was imported from Northern India, around the 3rd century BCE,^[3] but was influenced at various stages by South Indian scripts, manifestly influenced by the early [Grantha script](#).^[1]

Pottery from the 6th century BCE has been found in [Anuradhapura](#), with lithic inscriptions dating from 2nd century BCE written in [Prakrit](#).^[4]

By the 9th century CE, [literature](#) written in Sinhala script had emerged and the script began to be used in other contexts. For instance, the [Buddhist literature](#) of the [Theravada-Buddhists](#) of Sri Lanka, written in [Pali](#), used Sinhala script.

In 1736 the Dutch were the first to print with Sinhala [type](#) on the island. The resulting type followed the features of that of the native Sinhala script practiced on palm leaves. The Dutch created type was monolinear and geometric in fashion with no separation between words in early documents. During the second half of the 19th century, during the [Colonial period](#), a new style of Sinhala letterforms emerged in opposition to the monolinear and geometric form being high contrast in appearance and having varied thicknesses. This high contrast type gradually replaced the monolinear type as the preferred style which continues to be used in the present day. The high contrast style is still preferred for text typesetting in printed newspapers, books and magazines in Sri Lanka.^[5]

Today, the alphabet is used by over [16,000,000 people](#) to write Sinhala in very diverse contexts, such as [newspapers](#), [TV commercials](#), [government](#) announcements, [graffiti](#), and [schoolbooks](#).

[Sinhala](#) is the main language written in this script, but rare instances of [Sri Lanka Malay](#) are recorded.^[*citation needed*]

Structure^[edit]

ක = ka

ක + ි = කි = ki

ක + ් = ක් = k

The basic form of the letter k is ක "ka". For "ki", a small arch called *ispilla* is placed over the ක: කි. This replaces the inherent /a/ by /i/. It is also possible to have no vowel following a consonant. In order to produce such a pure consonant, a special marker, the *hal kirīma* has to be added: ක්. This marker suppresses the inherent vowel.

Sinhala script is an [abugida](#) written from left to right. It uses [consonants](#) as the basic unit for word construction as each consonant has an [inherent vowel](#) (/a/), which can be changed with a different vowel stroke. To represent different sounds it is necessary to add vowel strokes, or diacritics called පිළි Pili, that can be used before, after, above or below the base-consonant. Most of the Sinhala letters are [curlicues](#); straight lines are almost completely absent from the alphabet, and it does not have joining characters. This is because Sinhala used to be written on [dried palm leaves](#), which would split along the veins on writing straight lines. This was undesirable, and therefore, the round shapes were preferred. Upper and lower cases do not exist in Sinhala.^[5]

Sinhala letters are ordered into two sets. The core set of letters forms the ***śuddha siṃhala*** alphabet (Pure Sinhala, ශුද්ධ සිංහල), which is a subset of the ***miśra siṃhala*** alphabet (Mixed Sinhala, මිශ්‍ර සිංහල). This "pure" alphabet contains all the graphemes necessary to write Eḷu (classical Sinhala) as described in the classical grammar Sidatsaṅgarā (1300 AD).^[6] This is the reason why this set is also called ***Eḷu hōdiya*** ("Eḷu alphabet" එළු හොදිය). The definition of the two sets is thus a historic one. Out of pure coincidence, the phoneme inventory of present-day colloquial Sinhala is such that yet again the *śuddha* alphabet suffices as a good representation of the sounds.^[6] All native [phonemes](#) of the Sinhala spoken today can be represented in *śuddha*, while in order to render special Sanskrit and Pali sounds, one can fall back on *miśra siṃhala*. This is most notably necessary for the [graphemes](#) for the [Middle Indic](#) phonemes that the [Sinhala language](#) lost during its history,

such as [aspirates](#).^[6]

Most phonemes of Sinhala can be represented by a *śuddha* letter or by a *miśra* letter, but normally only one of them is considered correct. This one-to-many mapping of [phonemes](#) onto [graphemes](#) is a frequent source of [misspellings](#).^[7]

While a phoneme can be represented by more than one grapheme, each grapheme can be pronounced in only one way, with the exceptions of the inherent vowel sound, which can be either [a] (stressed) or [ə] (unstressed), and "ඉ" where the consonant is either [v] or [w] depending on the word. This means that the actual [pronunciation](#) of a word is almost always clear from its orthographic form. Stress is almost always predictable; only words with [v] or [w] (which are both allophones of "ඉ"), and a very few other words need to be learnt individually.

Some pronunciation exceptions in Sinhala:

- කරනවා - to do - [kəɾənə'waː] (not ['kəɾənə'waː])
- හතළිහ - forty - ['haʈʈəli'ha] (not ['haʈʈəli'ha])

Diacritics[\[edit\]](#)



The two shapes of the *hal kirīma* for p (left) and b (right).

In Sinhala the diacritics are called පිලි *pili* (vowel strokes). දිග *diga* means "long" because the vowel is sounded for longer and දෙක *deka* means "two" because the stroke is doubled when written.

			Using the consonant 'k' + 'vowel' as an example:			
පිලි <i>pilla</i>	Name	Transliteration	Formation	Compound form	ISO 15919	IPA
ට	හල් කිරීම	Hal kirīma	ක්	ක්	k	[k]
ඉ	Inherent /a/ (without any <i>pili</i>)		ක් + අ	ක	ka	[ka]
ආ	ඇලපිල්ල	Ælapilla	ක් + ආ	කා	kā	[kaː]
ඈ	ඈදය	Ædaya	ක් + ඇ	කෑ	kæ	[kæ]
ඉග	දිග ඈදය	Diga ædaya	ක් + ඇ	කෑ	kǣ	[kæː]
ඊ	ඉස්පිල්ල	Ispilla	ක් + ඉ	කි	ki	[ki]
ඊග	දිග ඉස්පිල්ල	Diga ispilla	ක් + ඊ	කි	kī	[kiː]
උ	පාපිල්ල	Pāpilla	ක් + උ	කු	ku	[ku], [kʷ]
උග	දිග පාපිල්ල	Diga pāpilla	ක් + උ	කු	kū	[kuː]
ඌ	ගැටය සහිත ඇලපිල්ල	Gæṭa sahita ælapilla	ක් + උ + උ	කෑ	kru	[kru]
ඌඳ	ගැටය සහිත ඇලපිල්ල දෙක	Gæṭa sahita ælapili deka	ක් + උ + උ	කෑඳ	krū	[kruː]
ඌඳු	ගයනුකිත්ත	Gayanukitta	Used in conjunction with kombuva for consonants.			
ඌඳු	දිග ගයනුකිත්ත	Diga gayanukitta	Not in contemporary use			
ඌඳු	කොම්බුව	Kombuva	ක් + ට	කෙ	ke	[ke]
ඌඳු	කොම්බුව සහ හල්කිරීම	Kombuva saha halkirīma	ක් + ට්	කේ	kē	[keː]
ඌඳු	කොම්බු දෙක	Kombu deka	ක් + ට්	කේ	kai	[kʌj]
ඌඳු	කොම්බුව සහ ඇලපිල්ල	Kombuva saha ælapilla	ක් + ඔ	කෝ	ko	[ko]
ඌඳු	කොම්බුව සහ හල්ඇලපිල්ල	Kombuva saha halælapilla	ක් + ඕ	කෝ	kō	[koː]
ඌඳු	කොම්බුව සහ ගයනුකිත්ත	Kombuva saha gayanukitta	ක් + ඔ	කෝ	kau	[kʌʋ]

Non-vocalic diacritics[\[edit\]](#)

The **anusvara** (often called *binduva* 'zero') is represented by one small circle ◌̣ (Unicode 0D82),^[8] and the **visarga** (technically part of the *miśra* alphabet) by two ◌̣◌̣ (Unicode 0D83). The inherent vowel can be removed by a special **virama** diacritic, the *hal kirīma* (◌̣̣̣̣), which has two shapes depending on which consonant it attaches to. Both are represented in the image on the right side. The first one is the most common one, while the second one is used for letters ending at the top left corner.

Letters^[edit]

Śuddha set^[edit]

The *śuddha* graphemes are the mainstay of Sinhala script and are used on an everyday-basis. Every sequence of sounds of Sinhala of today can be represented by these graphemes. Additionally, the *śuddha* set comprises graphemes for **retroflex** ◌̣̣̣̣ and ◌̣̣̣̣, which are no longer phonemic in modern Sinhala. These two letters were needed for the representation of E|u, but are now obsolete from a purely phonemic view. However, words which **historically** contain these two phonemes are still often written with the graphemes representing the retroflex sounds.

Vowels	අ	ආ	ඈ	ඉ	ඊ	උ	ඌ	ඍ	ඎ	ඏ	ඐ	එ
Transliteration	a	ā	æ	ē	i	ī	u	ū	e	ē	o	ō
IPA	[a,ə]	[aː,a]	[æ]	[æː]	[i]	[iː]	[u]	[uː]	[e]	[eː]	[o]	[oː]

Consonants	ක	ග	ඟ	ච	ජ	ඤ	ඦ	ඞ	ට	ඨ	ඩ	ඪ	ණ	ඬ	ත	ඣ	ඵ	බ	භ	ඹ	භ්	ය	ර	ල	ළ	ව	ස	හ
Transliteration	k	g	ṅg	c	j	ṭ	ḍ	ṇ	ṇḍ	t	d	ṇḍ	p	b	m	m̐b	y	r	l	l̐	v	s	h					
IPA	[k]	[g]	[ʎḡ]	[tʃ~tʃ̌]	[dʒ~dǯ]	[t̪]	[d̪]	[n̪]	[ɳ̪d̪]	[t̪]	[d̪]	[ɳ̪d̪]	[p̪]	[b̪]	[m̪]	[m̐b̪]	[j̪]	[r̪]	[l̪]	[l̐]	[v̪]	[s̪]	[h̪]					

Vowels^[edit]

show																							
(Click on [show] on the right if you only see boxes below)																							
Vowels																							
short												long											
independent						diacritic						independent						diacritic					
අ	0D85	a	[a]			inherent		a	[a, ə]			ආ	0D86	ā	[aː]	ඈ	0DCF	ā	[aː]				
ඈ	0D87	æ/ā	[æ]	ඉ	0DD0	æ	[æ]					ඊ	0D88	ē	[æː]	උ	0DD1	ē	[æː]				
ඊ	0D89	i	[i]	උ	0DD2	i	[i]					ඌ	0D8A	ī	[iː]	ඍ	0DD3	ī	[iː]				
උ	0D8B	u	[u]	ඍ	0DD4	u	[u]					ඎ	0D8C	ū	[uː]	ඏ	0DD6	ū	[uː]				
ඍ	0D91	e	[e]	ඏ	0DD9	e	[e]					ඐ	0D92	ē	[eː]	එ	0DDA	ē	[eː]				
ඎ	0D94	o	[o]	ඐ	0DDC	o	[o]					එ	0D95	ō	[oː]	ඒ	0DDD	ō	[oː]				
Display this table as an image																							

Vowels come in two shapes: independent and **diacritic**. The independent shape is used when a vowel does not follow a consonant, e.g. at the beginning of a word. The diacritic shape is used when a vowel follows a consonant. Depending on the vowel, the diacritic can attach at several places. The diacritic for ◌̣̣̣̣ attaches above the consonant, the diacritic for ◌̣̣̣̣ attaches below, the diacritic for ◌̣̣̣̣ follows, while the diacritic for ◌̣̣̣̣ precedes. ◌̣̣̣̣ finally is marked by the combination of preceding ◌̣̣̣̣ and following ◌̣̣̣̣.

While <a,e,i,o> are regular, the diacritic for ◌̣̣̣̣ takes a different shape according to the consonant it attaches to. The most common one is represented on the image on the right for the consonant භ (p). The k-shape is used for some consonants ending at the lower right corner (ක (k), ග (g), ඟ (t), but not ඞ (n) or ඟ (h)). Combinations of ඟ (r) or ඟ (l) with ◌̣̣̣̣ have idiosyncratic shapes.^[9]

	ූ (u)	ූ̃ (u)
ප (p)	භූ	භූ̃
ක (k)	කූ	කූ̃
ර (r)	රූ	රූ̃
ළ (l)	ලූ	ලූ̃

The vocalic diacritics for *u* and *ū* vary according to the consonant to which they are attached.

Consonants[[edit](#)]

show									
(Click on [show] on the right if you see only boxes below)									
Plosives									
	voiceless					voiced			
		Unicode	translit.	IPA		Unicode	translit.	IPA	
velar	ක	0D9A	ka	[ka]	ග	0D9C	ga	[ga]	velar
retroflex	ඣ	0DA7	ṭa	[ṭa]	ඪ	0DA9	ḍa	[ḍa]	retroflex
dental	ත	0DAD	ta	[ta]	ද	0DAF	da	[da]	dental
labial	ප	0DB4	pa	[pa]	බ	0DB6	ba	[ba]	labial
Other letters									
		Unicode	translit.	IPA		Unicode	translit.	IPA	
fricatives	ස	0DC3	sa	[sa]	හ	0DC4	ha	[ha]	fricatives
affricates	ච (ච)	(0DA0)	(ca)	([t͡ʃa])	ජ	0DA2	ja	[d͡ʒa]	affricates
nasals	ම	0DB8	ma	[ma]	න	0DB1	na	[na]	nasals
liquid	ල	0DBD	la	[la]	ර	0DBB	ra	[ra]	liquid
glide	ව	0DC0	va	[va]	ය	0DBA	ya	[ja]	glide
retroflex	ණ	0DAB	ṇa	[ṇa]	ළ	0DC5	ḷa	[ḷa]	retroflex
Display this table as an image									

The *śuddha* alphabet comprises 8 [plosives](#), 2 [fricatives](#), 2 [affricates](#), 2 [nasals](#), 2 [liquids](#) and 2 [glides](#). Additionally, there are the two graphemes for the retroflex sounds /ɭ/ and /ɻ/, which are not phonemic in modern Sinhala, but which still form part of the set. These are shaded in the table.

The voiceless affricate (ච [t͡ʃa]) is not included in the *śuddha* set by purists since it does not occur in the main text of the Sidatsaṅgarā. The Sidatsaṅgarā does use it in examples though, so this sound did exist in Eḷu. In any case, it is needed for the representation of modern Sinhala.^[6]

The basic shapes of these consonants carry an inherent /a/ unless this is replaced by another vowel or removed by the *hal kirīma*.

Prenasalized consonants[[edit](#)]

show									
(Click on [show] on the right if you see only boxes below)									

Prenasalized consonants							
	nasal	obstruent	prenasalized consonant	Unicode	translit.	IPA	
velar	ඛ	ග	ඟ	0D9F	<i>ṅga</i>	[ˈŋa]	velar
retroflex	ඣ	ඩ	ඬ	0DAC	<i>ṇḍa</i>	[ˈɖa]	retroflex
dental	න	ඳ	ඤ	0DB3	<i>ṇḍa</i>	[ˈɖa]	dental
labial	ම	ඞ	ඬ	0DB9	<i>ṁba</i>	[ˈmba]	labial
Display this table as an image							

The [prenasalized consonants](#) resemble their plain counterparts. <ṁb> is made up by the left half of <m> and the right half of , while the other three are just like the grapheme for the plosive with a little stroke attached to their left.^[10] Vowel diacritics attach in the same way as they would to the corresponding plain plosive.

Miśra set[edit]

The *miśra* alphabet is a [superset](#) of *śuddha*. It adds letters for [aspirates](#), [retroflexes](#) and [sibilants](#), which are not phonemic in today's Sinhala, but which are necessary to represent non-native words, like [loanwords](#) from [Sanskrit](#), Pali or [English](#). The use of the extra letters is mainly a question of prestige. From a purely phonemic point of view, there is no benefit in using them, and they can be replaced by a (sequence of) *śuddha* letters as follows: For the *miśra* aspirates, the replacement is the [plain śuddha](#) counterpart, for the *miśra* [retroflex liquids](#) the corresponding *śuddha* [coronal](#) liquid,^[11] for the [sibilants](#), <s>.^[12] ඤ (ṇ) and ඣ (ṇ) cannot be represented by *śuddha* graphemes but are found only in fewer than 10 words each. ො fa can be represented by ප pa with a Latin <f> inscribed in the cup.

Vowels	ඳෘ	ඳෘෘ	ඳෙ	ඳො	ඳු	ඳූ
Transliteration	ṛ	ṝ	ai	au	ī	ī̄
IPA	[ri,ru]	[riː,ruː]	[aɪ]	[aʊ]	[li]	[liː]

Consonants	ක	ඝ	ඞ	ඤ	ඤ	ඤ	ඤ	ඤ	ඤ	ඤ	ඤ	ඤ	ඤ	ඤ	ඤ	ඤ
Transliteration	kh	gh	ṇ	ch	jh	ñ	ṭh	ḍh	ṭh	ḍh	n	ph	bh	ś	ṣ	f
IPA	[k]	[g]	[ɳ]	[tʃ~tɕ]	[dʒ~dʒ]	[ɳ]	[ʈ]	[ɖ]	[ʈ]	[ɖ]	[n]	[p]	[b]	[ʃ]	[ʃ]	[f]

Vowels[edit]

show

(Click on [show] on the right if you see only boxes below)

Vocalic diacritics																		
	independent				diacritic					independent				diacritic				
diphthongs	ඳේ	0D93	ai	[ai]	ඳේ	0DDb	ai	[ai]		ඳො	0D96	au	[au]	ඳො	0DDE	au	[au]	diphthongs
syllabic r	ඳෘ	0D8D	ṛ	[ur]	ඳෘ	0DD8	ṛ	[ru,ur]		ඳෘෘ	0D8E	ṝ	[ruː,urː]	ඳෘෘ	0DF2	ṝ	[ruː,urː]	syllabic r
syllabic l	ඳු	0D8F	ī	[li]	ඳු	0DDF	ī	[li]		ඳූ	0D90	ī̄	[liː]	ඳූ	0DF3	ī̄	[liː]	syllabic l
Display this table as an image																		

There are six additional vocalic diacritics in the *miśra* alphabet. The two [diphthongs](#) are quite common, while the "syllabic" ṛ is much rarer, and the "syllabic" ī is all but obsolete. The latter are almost exclusively found in loanwords from Sanskrit.^[13]

The *miśra* <ṛ> can also be written with *śuddha* <r>+<u> or <u>+<r>, which corresponds to the actual [pronunciation](#). The *miśra*

syllabic ⟨l⟩ is obsolete, but can be rendered by *śuddha* ⟨l⟩+⟨i⟩.^[14] *Miśra* ⟨au⟩ is rendered as *śuddha* ⟨awu⟩, *miśra* ⟨ai⟩ as *śuddha* ⟨ayi⟩.

Note that the transliteration of both ළ and ෝ is ⟨l⟩. This is not very problematic as the second one is extremely scarce.

Consonants^[edit]

show									
(Click on [show] on the right if you see only boxes below)									
Extra miśra plosives									
	voiceless					voiced			
		Unicode	translit.	IPA		Unicode	translit.	IPA	
velar	ක	0D9B	kha	[ka]	ඝ	0D9D	gha	[ga]	velar
retroflex	ඬ	0DA8	ṭha	[ṭa]	ඪ	0DAA	ḍha	[ḍa]	retroflex
dental	ඵ	0DAE	tha	[t̪a]	ප	0DB0	dha	[d̪a]	dental
labial	ඵ	0DB5	pha	[pa]	භ	0DB7	bha	[ba]	labial
Other additional miśra graphemes									
		Unicode	translit.	IPA		Unicode	translit.	IPA	
sibilants	ශ	0DC1	śa	[sa]	ඪ	0DC2	ṣa	[sa]	sibilants
aspirate affricates	ඡ	0DA1	cha	[tʃa]	ඣ	0DA3	jha	[d͡ʒa]	aspirate affricates
nasals	ඤ	0DA4	ṇa	[ɲa]	ඥ	0DA5	jṇa	[d͡ʒɲa]	nasals
other	ඹ	0D9E	ṇa	[ɲa]	ඹ	0DC6	fa	[fa, φa, pa]	other
other	ජ	0DA6	ṇja ^[15]	[nd͡ʒa]	ජ	n/a	fa	[fa, φa, pa]	other
Display this table as an image									

Consonant conjuncts^[edit]



Certain combinations of graphemes trigger special [ligatures](#). Special signs exist for an ජ (r) following a consonant (inverted arch underneath), a ජ (r) preceding a consonant (loop above) and a ය (y) following a consonant (half a ය on the right). ^[11] ^[16] ^[17] Furthermore, very frequent combinations are often written in one stroke, like *ddh*, *kv* or *kś*. If this is the case, the first consonant is not marked with a *hal kirīma*. ^[11] ^[13] ^[17] The image on the left shows the [glyph](#) for *śrī*, which is composed of the letter *ś* with a ligature indicating the *r* below and the vowel *ī* marked above. Most other conjunct consonants are made with an explicit virama, called *al-lakuna* or *hal kirīma*, and the [zero-width joiner](#) as shown in the following table, some of which may not display correctly due to limitations of your system. Some of the more common are displayed in the following table. Note that although modern Sinhala sounds are not aspirated, aspiration is marked in the sound where it was historically present to highlight the differences in modern spelling. Also note that all of the combinations are encoded with the *al-lakuna* (Unicode U+0DCA) first, followed by the zero-width joiner (Unicode U+200D) except for touching letters which have the zero-width joiner (Unicode U+200D) first followed by the *al-lakuna* (Unicode U+0DCA). Touching letters were used

in ancient scriptures but are not used in modern Sinhala. Vowels may be attached to any of the ligatures formed, attaching to the rightmost part of the glyph except for vowels that use the *kombuva*, where the *kombuva* is written before the ligature or cluster and the remainder of the vowel, if any, is attached to the rightmost part. In the table below, appending "o" (*kombuva saha ælepilla* – *kombuva* with *ælepilla*) to the cluster "ky" /kja/ only adds a single code point, but adds two vowel strokes, one each to the left and right of the consonant cluster.

IPA	Letters	Unicode	Combined	Unicode	Type
/kja/	කය	U+0D9A U+0DCA U+0DBA	කය්	U+0D9A U+0DCA U+200D U+0DBA	<i>yansaya</i>
/kjo/	කයො	U+0D9A U+0DCA U+0DBA U+0DCC	කයො	U+0D9A U+0DCA U+200D U+0DBA U+0DCC	<i>yansaya</i>
/gja/	ගය	U+0D9C U+0DCA U+0DBA	ගය්	U+0D9C U+0DCA U+200D U+0DBA	<i>yansaya</i>
/kra/	කර	U+0D9A U+0DCA U+0DBB	කර්	U+0D9A U+0DCA U+200D U+0DBB	<i>rakāransaya</i>
/gra/	ගර	U+0D9C U+0DCA U+0DBB	ගර්	U+0D9C U+0DCA U+200D U+0DBB	<i>rakāransaya</i>
/rka/	රක	U+0DBB U+0DCA U+0D9A	රක්	U+0DBB U+0DCA U+200D U+0D9A	<i>rēpayā</i>
/rga/	රග	U+0DBB U+0DCA U+0D9C	රග්	U+0DBB U+0DCA U+200D U+0D9C	<i>rēpayā</i>
/kja/	කය්	U+0D9A U+0DCA U+0DBA U+0DCA U+0DBB	කය්	U+0D9A U+0DCA U+200D U+0DBA U+0DCA U+200D U+0DBB	<i>yansaya + rakāransaya</i>
/gja/	ගය්	U+0D9C U+0DCA U+0DBA U+0DCA U+0DBB	ගය්	U+0D9C U+0DCA U+200D U+0DBA U+0DCA U+200D U+0DBB	<i>yansaya + rakāransaya</i>
/rkja/	රකය	U+0DBB U+0DCA U+0D9A U+0DCA U+0DBA	රකය්	U+0DBB U+0DCA U+200D U+0D9A U+0DCA U+200D U+0DBA	<i>rēpayā + yansaya</i>
/rgja/	රගය	U+0DBB U+0DCA U+0D9C U+0DCA U+0DBA	රගය්	U+0DBB U+0DCA U+200D U+0D9C U+0DCA U+200D U+0DBA	<i>rēpayā + yansaya</i>
/kva/	කව	U+0D9A U+0DCA U+0DC0	කව්	U+0D9A U+0DCA U+200D U+0DC0	conjunct
/kfa/	කෑ	U+0D9A U+0DCA U+0DC2	කෑ	U+0D9A U+0DCA U+200D U+0DC2	conjunct
/gd ^h a/	ගධ	U+0D9C U+0DCA U+0DB0	ගධ්	U+0D9C U+0DCA U+200D U+0DB0	conjunct
/t ^h a/	ඨ	U+0DA7 U+0DCA U+0DA8	ඨ්	U+0DA7 U+0DCA U+200D U+0DA8	conjunct
/tt ^h a/	ඨඨ	U+0DAD U+0DCA U+0DAE	ඨඨ්	U+0DAD U+0DCA U+200D U+0DAE	conjunct
/tva/	ඨව	U+0DAD U+0DCA U+0DC0	ඨව්	U+0DAD U+0DCA U+200D U+0DC0	conjunct
/dd ^h a/	ද්ධ	U+0DAF U+0DCA U+0DB0	ද්ධ්	U+0DAF U+0DCA U+200D U+0DB0	conjunct
/dva/	ද්ව	U+0DAF U+0DCA U+0DC0	ද්ව්	U+0DAF U+0DCA U+200D U+0DC0	conjunct
/nda/	න්ද	U+0DB1 U+0DCA U+0DAF	න්ද්	U+0DB1 U+0DCA U+200D U+0DAF	conjunct
/nd ^h a/	න්ධ	U+0DB1 U+0DCA U+0DB0	න්ධ්	U+0DB1 U+0DCA U+200D U+0DB0	conjunct
/mma/	මම	U+0DB8 U+0DCA U+0DB8	මම	U+0DB8 U+200D U+0DCA U+0DB8	touching

Letter names[\[edit\]](#)

The Sinhala *śuddha* graphemes are named in a uniform way adding *-yanna* to the sound produced by the letter, including vocalic diacritics.^{[9][19]} The name for the letter අ is thus *ayanna*, for the letter ආ *āyanna*, for the letter ක *kayanna*, for the letter ඞ *kāyanna*, for the letter ඞෙ *keyanna* and so forth. For letters with *hal kirīma*, an *epenthetic* *a* is added for easier pronunciation: the name for the letter ඞ is *akyanna*. Another naming convention is to use *al-* before a letter with suppressed vowel, thus *alkayanna*.

Since the extra *miśra* letters are phonetically not distinguishable from the *śuddha* letters, proceeding in the same way would lead to confusion. Names of *miśra* letters are normally made up of the names of two *śuddha* letters pronounced as one word. The first one indicates the sound, the second one the shape. For example, the aspirated ඞ (kh) is called *bayanu*

kayanna. *kayanna* indicates the sound, while *bayanu* indicates the shape: ක (kh) is similar in shape to බ (b) (*bayunu* = *like bayanna*). Another method is to qualify the *miśra* aspirates by *mahāprāṇa* (ආ: *mahāprāṇa kayanna*) and the *miśra* retroflexes by *mūrdhaja* (ඃ: *mūrdhaja layanna*).

Stroke order[edit]



This section **needs expansion**. You can help by [adding to it](#). (February 2019)

Each Sinhala letter has a specific stroke order and method of writing.

Numerals[edit]

Sinhala had special symbols to represent numerals, which were in use until the beginning of the 19th century. This system is now superseded by [Hindu–Arabic numeral system](#).^{[19][20]}

Sinhala Illakkam ([Sinhala Archaic Numbers](#))

Sinhala Illakkam were used for writing numbers prior to the fall of [Kandyan Kingdom](#) in 1815. These digits did not have a zero instead the numbers had signs for 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 1000. These digits and numbers can be seen primarily in Royal documents and artefacts.

1	2	3	4	5	6	7	8	9	10	20	30	40	50	60	70	80	90	100	1000
□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□

Sinhala Lith Illakkam ([Sinhala Astrological Numbers](#))

Prior to the fall of Kandyan Kingdom all calculations were carried out using Lith digits. After the fall of the [Kandyan Kingdom](#), Sinhala Lith Illakkam were primarily used for writing horoscopes. However, there is evidence that they were used for other purposes such as writing page numbers etc. The tradition of writing degrees and minutes of zodiac signs in horoscopes continued into the 20th century using different versions of Lith Digits. Unlike the Sinhala Illakkam, Sinhala Lith Illakkam included a 0.

0	1	2	3	4	5	6	7	8	9
□	□	□	□	□	□	□	□	□	□

Neither the [Sinhala numerals](#) nor U+0DF4 ෙ Sinhala punctuation [kuddaliya](#) is in general use today. The kuddaliya was formerly used as a full stop.^[21]

Transliteration[edit]

Sinhala [transliteration](#) (Sinhala: රෝම අකුරින් ලිවීම *rōma akurin livīma*, literally "Roman letter writing") can be done in analogy to [Devanāgarī transliteration](#).

Layman's transliterations in Sri Lanka normally follow neither of these. Vowels are transliterated according to English spelling equivalences, which can yield a variety of spellings for a number of phonemes. /i:/ for instance can be ⟨ee⟩, ⟨e⟩, ⟨ea⟩, ⟨i⟩, etc. A transliteration pattern peculiar to Sinhala, and facilitated by the absence of phonemic aspirates, is the use of ⟨th⟩ for the [voiceless dental plosive](#), and the use of ⟨t⟩ for the [voiceless retroflex plosive](#). This is presumably because the retroflex plosive /ʈ/ is perceived the same as the English [alveolar plosive](#) /t/, and the Sinhala dental plosive /t̪/ is equated with the English [voiceless dental fricative](#) /θ/.^[22] Dental and retroflex voiced plosives are always rendered as ⟨d⟩, though, presumably because ⟨dh⟩ is not found as a representation of /ɖ/ in English orthography.

Use for the Pali language[edit]

Many of the oldest manuscripts in the Pali language are written in the Sinhala script. *Miśra* consonants are used to represent Pali phonemes that have no Sinhala counterpart. The following table lays out the Sinhala representations of Pali consonants with their standard academic Romanizations:

class	unaspirated unvoiced	aspirated	voiced	aspirated voiced	nasal
velar	ක (<i>ka</i>)	ඛ (<i>kha</i>)	ග (<i>ga</i>)	ඝ (<i>gha</i>)	ඞ (<i>ṇa</i>)

palatal	ච (<i>ca</i>)	ඡ (<i>cha</i>)	ජ (<i>ja</i>)	ඣ (<i>jha</i>)	ඤ (<i>ña</i>)
retroflex	ඪ (<i>ṭa</i>)	ඬ (<i>ṭha</i>)	ඬ (<i>ḍa</i>)	ඬ (<i>ḍha</i>)	ඹ (<i>ṇa</i>)
dental	ත (<i>ta</i>)	ථ (<i>tha</i>)	ද (<i>da</i>)	ධ (<i>dha</i>)	න (<i>na</i>)
labial	ප (<i>pa</i>)	ඵ (<i>pha</i>)	බ (<i>ba</i>)	භ (<i>bha</i>)	ම (<i>ma</i>)

unordered	ය (<i>ya</i>)	ර (<i>ra</i>)	ල (<i>la</i>)	ව (<i>va</i>)	ස (<i>sa</i>)	හ (<i>ha</i>)	ළ (<i>ḷa</i>)
-----------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------

The vowels are a subset of those for writing Sinhala:

Independent	Romanization	Dependent (on ක <i>ka</i>)	Romanization
අ	<i>a</i>	ක	<i>ka</i>
ආ	<i>ā</i>	කා	<i>kā</i>
ඉ	<i>i</i>	කි	<i>ki</i>
ඊ	<i>ī</i>	කී	<i>kī</i>
උ	<i>u</i>	කු	<i>ku</i>
ඌ	<i>ū</i>	කූ	<i>kū</i>
එ	<i>e</i>	කෙ	<i>ke</i>
ඔ	<i>o</i>	කො	<i>ko</i>

The *niggaḥīta* is represented with the sign ො. Consonant sequences may be combined in ligatures in a manner identical to that described above for Sinhala.

As an example, below is the first verse from the [Dhammapada](#) in Pali in Sinhala script, followed by Romanization:

මනොපුබ්බභිගමා ධම්මා, මනොසෙට්ඨා මනොමයා;
 මනසා චෙ පදුට්ඨෙන, භාසති වා කරොති වා;
 තනො තං දුක්ඛමත්ථෙති, චක්ඛංච චහනො පදං.
Manopubbaṅgamā dhammā, manoseṭṭhā manomayā;
manasā ce paduṭṭhena bhāsati vā karoti vā;
tato naṃ dukkhamanveti cakkarīva vahato padarī.
 — *Yamaka-vaggo 1*

Relation to other scripts[\[edit\]](#)

Similarities

Sinhala is one of the [Brahmic scripts](#), and thus shares many similarities with other members of the family, such as the [Kannada](#), [Malayalam](#), [Telugu](#), [Tamil script](#) and [Devanāgarī](#). As a general example, /a/ is the inherent vowel in all these scripts.^[1] Other similarities include the diacritic for ⟨ai⟩, which resembles a doubled ⟨e⟩ in all scripts and the diacritic for ⟨au⟩ which is composed of preceding ⟨e⟩ and following ⟨ḷ⟩.

Script	⟨e⟩	⟨ai⟩	⟨au⟩
Sinhala	ෙ	ෙෙ	ෙූ
Malayalam	െ	ൈ	ൌ
Tamil	ெ	ை	ௌ
Bengali	ে	ৈ	ৌ
Devanāgarī	े	ै	ौ

Likewise, the combination of the diacritics for ⟨e⟩ and ⟨ā⟩ yields ⟨o⟩ in all these scripts.

Script	⟨e⟩	⟨ā⟩	⟨o⟩
--------	-----	-----	-----

Sinhala	ට	ඟ	ඞ
Malayalam	ഏ	ഐ	ഓ
Tamil	ஐ	ஓ	ஔ
Bengali	ঐ	ও	ঔ
Dēvanāgarī	ॐ	॑	॒

Differences

Sinhala alphabet differs from other Indo-Aryan alphabets in that it contains a pair of vowel sounds (U+0DD0 and U+0DD1 in the proposed Unicode Standard) that are unique to it. These are the two vowel sounds that are similar to the two vowel sounds that occur at the beginning of the English words *at* (අ) and *ant* (ඇ).^[23]

Another feature that distinguishes Sinhala from its sister Indo-Aryan languages is the presence of a set of five nasal sounds known as half-nasal or prenasalized stops.

භ	ඡ	ඪ	ඣ	ඹ
ṅga	ṅja	ṅḍa	ṅḍa	ṅba

Computer encoding^[edit]



Sinhala language support in [Linux](#). [Firefox](#) is shown in the background, with [mlterm](#) in the foreground with text having been entered into it by [ibus-m17n](#).

Generally speaking, Sinhala support is less developed than support for Devanāgarī, for instance. A recurring problem is the rendering of diacritics which precede the consonant and diacritic signs which come in different shapes, like the one for ⟨u⟩.

Sinhala does not come built in with [Windows XP](#), unlike [Tamil](#) and [Hindi](#). However, all versions of [Windows Vista](#) and [Windows 10](#) come with Sinhala support by default, and do not require external [fonts](#) to be installed to read Sinhala script. *Nirmala UI* is the default Sinhala font in windows 10. The newest version of Windows 10 has added support for [Sinhala Archaic Numbers](#) that were not supported by default in the previous version.

For [Mac OS X](#), Sinhala font and keyboard support can be found at web.nickshanks.com/typography/ and at www.xenotypetech.com/osxSinhala.html.

For [Linux](#), the [IBus](#), and [SCIM](#) input methods allow the use Sinhala script in applications with support for a number of key maps and techniques such as traditional, phonetic and assisted techniques.^[24] In addition, newer versions of Android mobile operating system also support both rendering and input of Sinhala script.

Unicode^[edit]

Sinhala script was added to the [Unicode](#) Standard in September 1999 with the release of version 3.0. This character allocation has been adopted in Sri Lanka as the [Standard](#) SLS1134.

The main Unicode block for Sinhala is U+0D80–U+0DFF. Another block, [Sinhala Archaic Numbers](#), was added to Unicode in version 7.0.0 in June 2014. Its range is U+111E0–U+111FF.

Sinhala^{[1][2]} Official Unicode Consortium code chart (PDF)																
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+0D8x		ට	ඟ	ඞ		අ	ආ	ඇ	ඈ	ඉ	ඊ	උ	ඌ	ඍ	ඎ	ඏ
U+0D9x	ඐ	එ	ඒ	ඓ	ඔ	ඕ	ඖ				ක	ඛ	ග	ඝ	ඞ	ඟ
U+0DAx	ච	ඡ	ජ	ඣ	ඤ	ඦ	ට	ඨ	ඩ	ඪ	ණ	ඬ	න	ඵ	ඹ	ඳ

U+0DBx	ඨ	ඪ		ඳ	ඵ	ඬ	භ	ඹ	ඹ	ඹ	ඳ	ඳ		ඳ		
U+0DCx	ට	ට	ඨ	ඨ	ඨ	ට					ඳ					ට
U+0DDx	ට	ට	ට	ට	ට				ට	ට	ට	ට	ට	ට	ට	ට
U+0DEx																
U+0DFx			ට	ට	ට											

Notes

1. [^](#) As of Unicode version 13.0
2. [^](#) Grey areas indicate non-assigned code points

[Sinhala Archaic Numbers](#)^{[1][2]}

[Official Unicode Consortium code chart](#) (PDF)

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+111Ex																
U+111Fx																

Notes

1. [△] As of Unicode version 13.0
2. [△] Grey areas indicate non-assigned code points

See also[\[edit\]](#)

- [Sinhala Braille](#)
- [History of Sinhala software](#)
- Loanwords
- [Dutch loanwords in Sinhala](#)
- [English loanwords in Sinhala](#)
- [Portuguese loanwords in Sinhala](#)
- [Tamil loanwords in Sinhala](#)

References[[edit](#)]

1. ^ [Jump up to: ^a ^b ^c ^d](#) Daniels (1996), p. 408.
2. ^ [Jump up to: ^a ^b ^c ^d](#) Jayarajan, Paul M. (1976-01-01). [History of the Evolution of the Sinhala Alphabet](#). Colombo Apothecaries' Company, Limited.
3. ^ Daniels (1996), p. 379.
4. ^ [Jump up to: ^a ^b ^c ^d](#) Ray, Himanshu Prabha (2003-08-14). [The Archaeology of Seafaring in Ancient South Asia](#). Cambridge University Press. ISBN 9780521011099.
5. ^ [Jump up to: ^a ^b ^c ^d](#) "The Sinhala Script". Dalton Maag. Retrieved 26 August 2018.
6. ^ [Jump up to: ^a ^b ^c ^d](#) Gair and Paolillo 1997.
7. ^ Matzel (1983) p. 15, 17, 18
8. ^ [Jump up to: ^a ^b ^c ^d](#) Karunatilake (2004), p. xxxii
9. ^ Jayawardena-Moser (2004) p. 11
10. ^ Fairbanks et al. (1968), p. 126
11. ^ [Jump up to: ^a ^b ^c ^d](#) Karunatilake (2004), p. xxxi

12. [^] Daniels (1996), p. 410.
13. [^] [Jump up to: ^a ^b](#) Matzel (1983), p. 8
14. [^] Matzel (1983), p. 14
15. [^] This letter is not used anywhere, neither in modern nor ancient Sinhala. Its usefulness is unclear, but it forms part of the standard alphabet <<http://unicode.org/reports/tr2.html>>.
16. [^] Fairbanks et al. (1968), p. 109
17. [^] [Jump up to: ^a ^b](#) Jayawardena-Moser (2004), p. 12
18. [^] Fairbanks et al. (1968), p. 366
19. [^] ["Online edition of Sunday Observer – Business". *Sunday Observer*. Archived from \[the original\]\(#\) on 7 February 2009. Retrieved 21 September 2008.](#)
20. [^] ["Unicode Mail List Archive: Re: Sinhala numerals". *Unicode Consortium*. Retrieved 21 September 2008.](#)
21. [^] Roland Russwurm. ["Old Sinhala Numbers and Digits". *Sinhala Online*. Retrieved 23 September 2008.](#)
22. [^] Matzel (1983), p. 16
23. [^] https://www.isoc.org/inet97/proceedings/E1/E1_3.HTM
24. [^] [A screenshot showing some of the options](#)

Further reading^[edit]

- [Daniels, Peter T.](#) (1996). "Sinhala alphabet". *The World's Writing Systems*. Oxford, UK: Oxford University Press. [ISBN 0-19-507993-0](#).
- Fairbanks, G. W.; J. W. Gair; M. W. S. D. Silva (1968). *Colloquial Sinhalese (Sinhala)*. Ithaca, NY: South Asia Programm, Cornell University.
- Gair, J. W.; John C. Paolillo (1997). *Sinhala*. München, Newcastle: South Asia Programm, Cornell University.
- Geiger, Wilhelm (1995). *A Grammar of the Sinhalese Language*. New Delhi: AES Reprint.
- Jayawardena-Moser, Premalatha (2004). *Grundwortschatz Singhalesisch – Deutsch* (3 ed.). Wiesbaden: Harrassowitz.
- Karunatilake, W. S. (1992). *An Introduction to Spoken Sinhala* ([several new editions] ed.). Colombo.
- Matzel, Klaus (1983). *Einführung in die singhalesische Sprache*. Wiesbaden: Harrassowitz.

External links^[edit]

- [Scripts \(ISO 15924\) "Sinhala"](#)
- [Sinhala Unicode Characters](#)
- [Sinhala Unicode Characters](#)
- [Sinhala Unicode Character Code Chart](#)
- [Sinhala Archaic Numbers Unicode Character Code Chart](#)
- [Complete table of consonant-diacritic-combinations](#)

Online resources

- [Sinhala guide of the Sinhala Wikipedia \(in English\)](#)
- [Online Sinhala Unicode Writer](#)
- [Sinhala English Dictionary and Sinhala To Hindi Language Translator](#)
- [Sinhala Unicode Support Group](#)
- [Online Unicode Converter](#)