

THE SINHALA WRITING SYSTEM: A GUIDE TO TRANSLITERATION

James W. Gair
W.S Karunatillake (Karunatilaka)

This is a reprinting in PDF form for the net and CD, with a new introduction, of the Transliteration Guide by the same authors that appeared in *Literary Sinhala Inflected Forms with a Transliteration Guide to Sinhala Script*, South Asia Program, Cornell University, 1976.



SINHAMEDIA
P.O. Box 1027
Trumansburg NY 14886
snhmedia @baka.com
© 2006

INTRODUCTION

About the Sinhala Script and Roman Transliteration:

The Sinhala writing system is used primarily for writing Sinhala, an official language of Sri Lanka and the language of a majority of its population. It is unique to Sri Lanka, and is also used in that country for writing Pali, the canonical language of Theravada Buddhism, and sometimes Sanskrit.

The Sinhala transliteration guide that is included here was originally appended to *Literary Sinhala Inflected Forms* by James Gair and W.S. Karunatillake that was published in 1976 by the South Asia program at Cornell University. It was intended strictly as a transliteration guide for the use of those wishing to represent materials in Sinhala script in Roman transliteration, especially those with a limited familiarity with that script. The Finder List (W.4) is an unusual feature created with that audience in mind as an additional aid to make locating forms easier, but may be puzzling to those to whom that script is familiar. It arranges simple and complex glyphs strictly according to shape, without consulting what they represent. It was created by asking a number of people with no acquaintance with Sinhala to sort cards bearing the glyphs according to perceived resemblance, and thus it was intended especially to help beginners in locating and identifying them. However, the guide was also intended to be useful to those who were familiar with the script but wanted a source for a consistent Roman transcription.

We should also mention here that Sinhala has the property that linguists call "diglossia". with very different written (Literary) and spoken varieties. Until quite recently, virtually all written materials were in the Literary variety, even where speech was represented as in novels and stories. Though this has now changed, the result was that there was lacking a single way of representing the spoken material in Sinhala script, so that one may encounter some variation in different texts, which could also affect Romanization.

There are actually two recognized forms of the Sinhala script, one a subset of the other. The full set is known as the *Miśra Hōḍiya*, or 'mixed alphabet', and includes the symbols necessary for writing loanwords or texts from Sanskrit and Pali, most notably the aspirated consonants, which do not occur in spoken Sinhala (see below). A subset of this inventory is called the *Eḷu Hōḍiya*, from the classical language Eḷu, as described in the classical grammar *Sidatsañgarā*, and that smaller alphabet is still regarded by many as the authoritative, truly Sinhala one, appearing in current school charts and grammars even today. The character for /c/ (◌) commonly omitted, since it is not listed in the *Sidatsañgarā*, and the *Eḷu Hōḍiya* lacks some characters necessary for writing Sanskrit, most notably the characters for the aspirated consonants (*kh*, *gh*, etc.) found in many other South Asian scripts. Classical Eḷu excludes these, reflecting the loss of aspiration that occurred before the earliest written texts (ca. 3rd. century BCE). Interestingly, the classical alphabet, with < c > added along with a few rules for some positionally determined vowel alternations and a few symbols for loanwords, suffices as a good representation of current Colloquial Sinhala. Table I at the beginning of the transliteration guide that follows this introduction gives the full inventory of basic symbols, with those included in the classical alphabet in italics.

General Principles:

Above, we have used 'alphabet' to refer to the script, and that is essentially correct, since each sound is represented, as compared to logo- ideographic scripts (like Chinese) or syllabaries (like Cree), in which single symbols are assigned to syllables. In fact, there are relatively few "pure" examples of these, and most scripts involve some degree of mixture, though one structural characteristic is dominant. The major South Asian scripts, including Sinhala, retain a structural pattern retained from the ancient Brahmi script from which they are descended: they represent the individual sounds, like an alphabet, but these are, loosely speaking, spatially "clumped" into complex units that generally represent syllable length sound patterns. The elements of these combinations are of two kinds:

1. Basic symbols (nuclei): consonants and vowels
2. Modifications (satellites): also consonants and vowels, but commonly include some "operator" element(s) as for eliminating a vowel that would otherwise occur, etc.

Modifications are written above, below, or to the right or left of the basic symbols, i.e., in a nucleus satellite(s) pattern. Consonants are almost always written with basic symbols, though there are a few modifications that represent consonants. Vowels, however, with virtually no exception, are written with basic symbols only when the begin a word, otherwise with modifications applied to the symbol representing the consonant that they follow. To illustrate, using Roman capitals as basic symbols and small letters as modifications, it is as if 'ego' were spelled EG_o and 'solo' as S_ol_o. One other feature of most South Asian scripts is that there is one "inherent" vowel, in this case 'a', that is implied by any basic consonant symbol unless it is removed by a special operator modification or displaced by a modification representing a different vowel. To continue our example using * as the vowel-removing modification, L would be 'la' but L* would be simply 'l', so that 'lo' would be Lo but 'al' would be AL*. Similarly, 'pat', 'pit', and 'pet' would be PT*, P_iT*, and P_eT* respectively, while 'sofa' and 'ida' would be S_oF and ID.

Some Special Features of Sinhala and Some Variants in Transcription:

The transliterations given in the guide are generally consistent with the conventions commonly used for South Asian languages, but there are a few features of Sinhala that are often not included in fonts intended for use in representing them (such as South Asia Times).

The symbol æ_ɹ represents a vowel similar to that of 'hat' in many English dialects (technically a "low-front" vowel). It is uncommon in the languages of South Asia and thus the scripts for them do not include a symbol for it. Thus transliteration fonts also lack a dedicated character. It is commonly represented by the letter 'æ', which does occur in some international keyboards (<option > on the Macintosh). It may also be represented by unlauded 'a' (ä) (< Option u + u >), especially in German sources. Failing these, one can

resort to the sequence 'ae', since that is not an occurring sequence in Sinhala, and so will not cause ambiguity.

Vowel length may be represented by a macron above the letter ('ā', for example), or by a colon, thus 'a:'. It is sometimes also represented by doubling the letter : 'aa'. This is common in Spoken language texts and in some linguistic treatments.

The 'half-nasals' (technically "prenasalized voiced stops"), written **ඞ** **ඟ** **ච** **ඳ** and **ඞ**, contrast in Sinhala with clusters of nasal plus consonant: Thus *kandə* 'hill' versus *kañdə* 'tree trunk' (for the ə vowel, see the Note on Pronunciation below). They are commonly represented by a "breve" ("half moon" symbol ˇ) above the nasal. Thus 'ñd', 'ñg' (sometimes 'ñg'), etc. They are as far as we know unique to Sinhala and Dhivehi (Maldivian), and thus are also commonly missing in Roman fonts for South Asia. An alternate way of representing them uses superscripts: thus ^ˆg', ^ˆd', ^ˆb' etc.

Retroflex Consonants, **ඞ**, **ච** etc., are commonly represented by an underdot: 'ṛ', 'ṁ', etc. They are also found represented by capital letters, particularly in language texts and where the font lacks the underdot. Since Sinhala script does not have capitalization, this generally causes no ambiguity.

A Note on Pronunciation:

Since this guide was intended strictly as a brief reference aid for transliteration, it did not deal with pronunciation. For this, one must consult other sources, such as the relevant sections of Fairbanks, Gair and De Silva *Colloquial Sinhalese (Sinhala)* or the fuller description in the Introduction to the later text by W.S. Karunatillake *An Introduction to Spoken Sinhala* (Colombo, Gunasena 1992 and 1998). One or two points are worth mention here, however, particularly for those with some familiarity with other South Asian scripts:

1. Aspirated consonants are generally not distinguished from their unaspirated counterparts in pronunciation (as mentioned earlier, Sinhala lost aspiration before the earliest written documents ca. 3c BCE), though some people may attempt to distinguish them with varying consistency in formal situations with texts that include Sanskrit or Pali forms.
2. Retroflex (**ඞ**) and dental (**ඟ**) nasals are never distinguished in pronunciation, and have not been since before the 13th Century, though they are very often retained in texts in the literary language. The case is the same with dental and retroflex laterals (**ච** and **ච**).
3. **ඳ** is transliterated always as *a*, but the pronunciation varies between [a] and [ə] (the latter is similar to the final sound in "sofa" in many English dialects) in a way generally predictable by position in a rather complex manner. For transliteration, of course, as oppose to transcription, this does not matter. (Among other sources describing the

distribution of the variants are Fairbanks, Gair and De Silva 5,5.1 and 6,12, Karunatillake p. xxx and xxix, cited above, or Gair and Paolillo, *Sinhala*, Lincom Europa, München and Newcastle 1997, 1.4).

4 අඞ is transliterated ā (or a:) but may represent /aa/ (/a:/) or /a/ in the spoken language, again in a generally predictable but complex fashion according to position (Fairbanks, Gair and De Silva 7,12.4) and 6,12, Gair and Paolillo 1.4, or Karunatillake p. xxix). In reading literary Sinhala aloud, however, some may render අඞ consistently long even where it would be short in the Colloquial language. Thus මව්මව් 'maternal uncle (dative) is colloquial /maamaTə/ but may be read aloud as /maamaaTə/.

5. In reading Literary Sinhala aloud, in general the same inventory of sounds that are used in the spoken language are employed. Although as mentioned earlier, some people, especially those with a scholarly bent, may attempt to distinguish the aspirated consonants with varying success, they are usually pronounced as their unaspirated counterparts. Thus, for example, both ක and ක් are generally rendered as /k/.

A fuller account of South Asian scripts in general, and the Sinhala script in particular, can be found in part VI of Peter T. Daniels and William Bright, eds. *The World's Writing Systems*, Oxford University Press 1996 pp. 408 ff.

ABOUT THE TRANSLITERATION GUIDE

How the Guide is Organized:

The guide has the following sections:

- W.1 A Note on Form
- W.2 Basic Symbols (Nuclei)
 - W.21 Consonants
 - W.22 Vowels
 - W.23 Mixed or Combined symbols
- W.3 Modifications (Satellite Symbols)
- W.4 The Sinhala Alphabetical Order
- W.5 A Finder List, with basic and some combined symbols arranged by shape similarity alone, without reference to their sound values.
- W.6 A Chart of consonant-vowel combinations.

An Important Note on Form:

The transliteration guide was produced on a computer in the "ancient time" before generally available desktops computers and varied computer fonts, so that only the background Roman text could be entered in that way. The Sinhala, which gave both typewritten and handwritten variants, had to be laboriously pasted in. Retyping it completely now so as to include computer fonts would have required more time and effort than we could devote to it and would also have required regenerating the handwritten forms if they were to be included. Nevertheless, it seemed that some users might still find the guide useful as it was if we created a digitally accessible form. Accordingly, we have simply scanned the original form so as to make it available, after making some corrections.

Symbols are given here in their careful handwritten forms, with the typewriter-generated forms given for comparison (except, of course, where the typewriter lacks the symbol). We were able to add Table I, however, which presents the basic symbols in a later, more book-like font, to which one may refer for comparison, while finding the handwritten and typed equivalents, including the modifications and their use in what follows. Note that the "half-nasals" (ආඤ්ඤා, ආඤ්ඤා, etc.) are given in Table I, but do not given in the original chart and are located in the guide under the appropriate modification or under W.23 Combined and Blended Symbols.

The original section numbering "W.1, W2....." has thus been retained ("W" for "Writing System" for cross-reference from the grammar to which it was linked). The Sinhala typing and handwritten examples were mostly done by Heloise Perera, and the original formatting for the computer by Gregory Pearson.

W.1: Table I The Sinhala Alphabet-Basic Symbols
(Letters included in the *Eḷu Hōḍiya* are in italics in the Roman)

Vowels						Consonants					
අ	ආ	ඇ	ඈ	ඉ	ඊ	ක	ඛ	ග	ඝ	ඞ	ඟ
<i>a</i>	<i>ā</i>	<i>a</i>	<i>ā</i>	<i>i</i>	<i>ī</i>	<i>ka</i>	<i>kha</i>	<i>ga</i>	<i>gha</i>	<i>ṅa, ṅa</i>	<i>ṅga</i>
උ	ඌ	ඍ	ඎ	ඏ	ඐ	ච	ඡ	ඣ	ඤ	ඦ	ට
<i>u</i>	<i>ū</i>	<i>u</i>	<i>ū</i>	<i>ī</i>	<i>ī</i>	<i>ca</i>	<i>cha</i>	<i>ja</i>	<i>jha</i>	<i>ṅa</i>	<i>ṅja</i>
එ	ඒ	ඓ	ඔ	ඕ	ඖ	ට	ඨ	ඩ	ඪ	ණ	ඬ
<i>e</i>	<i>ē</i>	<i>ai</i>	<i>o</i>	<i>ō</i>	<i>au</i>	<i>ṭa</i>	<i>ṭha</i>	<i>ḍa</i>	<i>ḍha</i>	<i>ṇa</i>	<i>ṇḍa</i>
						ත	ඵ	ද	ධ	න	ඳ
						<i>ta</i>	<i>tha</i>	<i>da</i>	<i>dha</i>	<i>na</i>	<i>ṅda</i>
						ප	ඵ	බ	භ	ම	ඹ
						<i>pa</i>	<i>pha</i>	<i>ba</i>	<i>bha</i>	<i>ma</i>	<i>m̃ba</i>
						ය	ර	ල	ව		
						<i>ya</i>	<i>ra</i>	<i>la</i>	<i>va</i>		
						ශ	ෂ	ස	හ	ඒ	
						<i>śa</i>	<i>ṣa</i>	<i>sa</i>	<i>ha</i>	<i>ḷa</i>	

W.2. The Basic Symbols in Written and Typed Form.

ක	ka	ඛ	kha	ග	ga	ඝ	gha	ඞ	ṅa, ṅa
ච	ca	ඡ	cha	ඣ	ja	ඤ	jha	ඦ	ṅa
ට	ṭa	ඨ	ṭha	ඩ	ḍa	ඪ	ḍha	ණ	ṇa
ත	ta	ඵ	tha	ද	da	ධ	dha	න	na
ප	pa	ඵ	pha	බ	ba	භ	bha	ම	ma

ය	ya	ර	ra	ල	la	ව	va		
ශ	śa	ෂ	ṣa, śa	ස	sa	හ	ha	ඒ	ḷa

The consonants in the second and fourth columns down to the dotted line are known as aspirated consonants. They are single consonant symbols in Sinhala, but transliterate as consonant plus h (followed as usual by a). The aspirate symbols are rarer than the non-aspirates since they are not used at all in writing some kinds of Sinhala.

Note that Sinhala has no capital/lower case distinction.

Examples:

කල කල
kala

සත සත
sata

කහ කහ
khaga

බර බර
bara

මන මන
mana

ඝට ඝට
ghata

W.22 Basic Vowel Symbols

අ අ a ආ ආ ā ඇ ඇ ä (ඈ) ඈ ඈ ä (ඈ) ඉ ඉ i ඊ ඊ ī

උ උ u ඌ ඌ ū ළ ළ ɛ ළ ළ ɛ̄ ළ ළ ɛ̄ ළ ළ ī

ඒ ඒ e ඒ ඒ ē ඔ ඔ ai ඔ ඔ o ඔ ඔ ō ඔ ඔ au

Note that there are alternate symbols for ä and ā in transliteration.

Basic vowel symbols are used almost only when the vowel begins a word. Hence:

අත අත
ata

ඉර ඉර
ira

ඔබ ඔබ
oba

උල උල
ula

ඊය ඊය
īya

ආම ආම
āma

අදා ආද
āda

උරා උරා
ūra

එකා එකා
aika

අගල අගල
agala

ඉඩම ඉඩම
idama

උපල උපල
upala

එතන එතන
etana

ඔරසා ඔරසා
aurasa

Otherwise, vowels following consonants are written with modifications added to the consonant symbols.

W.23 Combined and Blended Consonant Symbols

Consonant combinations are sometimes represented by two basic symbols written together without space.

Examples:

රකතා rakta compare with සකතා sakata

උපතා upta compare with උපතා upata

නසනා nasna compare with නසනා nasana

Sometimes the component symbols may be changed in shape and may not be easily recognizable in the combinations. The one usually encountered are:

ඟ ŋga

ඳ ŋda

ඳා ŋda

the ŋ and ṅ are known as half nasals and only occur in those combinations. They should be kept distinct from n and m in transliteration. Note that the first three are made with the symbol c plus o ga da da.

𑀓𑀭𑀢𑀺 m̄ba

𑀓𑀭𑀢𑀺 kva

𑀓𑀭𑀢𑀺 ttha

𑀓𑀭𑀢𑀺 nda

𑀓𑀭𑀢𑀺 dva

𑀓𑀭𑀢𑀺 j̄ña

𑀓𑀭𑀢𑀺 kṣa

𑀓𑀭𑀢𑀺 ṅca

𑀓𑀭𑀢𑀺 ṅja

𑀓𑀭𑀢𑀺 ṭtha

𑀓𑀭𑀢𑀺 ddha

𑀓𑀭𑀢𑀺 lu

NOTE: Another way of writing consonant combinations and double consonants is to use the modification ' or ' on the first consonant symbol. (See the first modification in Section W.3.) This is very common nowadays, particularly in printing and typing.

W.3 Modifications

Modifications appear above, below or to the right or left of a basic symbol.

• } when added to consonant symbols remove the vowel a. These two signs are equivalent, but used with different types of consonant symbols:

˘ is used with letters of a generally circular shape like ට t, ම m;
 ˙ is used elsewhere, e.g. ක k, ග g.

Examples:

අක	අක	අච	අච
	at		at
හළ	හළ	ගම	ගම
	hal		gam
වක	වක		
	vak		
පක	පක		
	pan		

When a consonant symbol marked with one of these modifications ˘ or ˙ precedes another consonant symbol, the result will be a consonant combination or a double consonant:

පට්ට	පට්ට
	paṭṭa
කල්ක	කල්ක
	kalka
වත්ත	වත්ත
	vatta

Note that if the symbol ෙ precedes the consonant symbol with ˘ or ˙ they have a different significance (see below under ෙ).

•) indicates ā (unless ෙ precedes the consonant symbol, for which see under ෙ below). e.g. කා ka, තා ta, රා ra, බා ba.

Examples:

හාන	හාන hāna
මාමා	මාමා māmā
තාත්තා	තාත්තා tāttā
බාප්පා	බාප්පා bāppā

ඳ් ද් is ඳ plus ට hence ḍā.

ඳ් indicates ä with any consonant but ඳ or one with the modification ට (see below) ට: pä, ට: kä, ට: t̄ä

Examples:

ආහ	ආහ āhā
බාත	බාත bāta

With ඳ or a consonant with the modification ට (see below), it almost always indicates u.

රුව	රුව ruva
ධ්‍රුව	ධ්‍රුව dhruwa

With ට lu it changes u to u: ට: lū.

However, in some words, ට may indicate æ where used together with ට just as it does with other consonants. This ambiguity does not cause much difficulty since such words are English borrowings.

Examples:

බ්‍රැන්ඩි	බ්‍රැන්ඩි brændi 'brandy'
ට්‍රැක්ටර්	ට්‍රැක්ටර් trækt̄ar 'tractors'

ඳ් indicates ā with any consonant but ඳ or one with the modification ට (see below) ට: pā, ට: t̄ā

Examples:

ପାଳା ପାଳା
pāla

କାମା କାମା
kāma

With ଠ or a consonant with the modification ୠ (see below) it almost always indicates ū:

ରୂପା ରୂପା
rūpa

ଗ୍ରୁ ଗ୍ରୁ
grū

However, in some words ୠ may indicate æ when used together with ୠ just as it does with other consonants. This ambiguity does not cause much difficulty since such words are English borrowings.

Examples:

ଗ୍ରାମ ଗ୍ରାମ
græm 'grams'

ଡ୍ରାମ ଡ୍ରାମ
dræm 'drams'

ୠ indicates ā and is used only with ଠ: ଠ ରା

Examples:

ରାଜା ରାଜା
rāḷa

ବାରାଣସୀ ବାରାଣସୀ
varāḷi

ୡ indicates ā̄ and is used only with ଠ: ଠ ରା̄

Examples:

ରାଣା ରାଣା
rāḅa

ୢ indicates ū and is used with ଠ ta, ଠ ka, ଠ ga, ଠ ṅga, ଠ bha, ଠ śa: ଠ tu ଠ ku ଠ gu ଠ ṅgu ଠ bhu
ଠ śu.

Examples:

තුගු	තු	තුරු	තුරු	භුජ	භුජ
	tuṅgu		kuru		bhuja
භුභ	භු	ගුරු	ගුරු	වාගු	වාගු
	subha		guru		vaṅgu

ඊ indicates \bar{u} and is used with ත ta, ක ka, ග ga, ඟ ṅga, භ bha, ශ śa: තු tū, කු kū, ගු gū, ඟු ṅgū, භු bhū, ශු śū.

Examples:

තූල	තූ	ගුභ	ගුභ	සූල	සූ
	tūla		gūha		sūla
කූප	කූ	භූප	භූප	හූඟු	හූ
	kūpa		bhūpa		hūṅgū

උ indicates \bar{u} . It is used with all consonants that do not take ඌ, or ණ for \bar{u} . e.g. පු pu ඳු du ටු tu

Examples:

පුර	පුර	උභුක	උභුක
	pura		ubuka
දුටු	දුටු	උකුට	උකුට
	duṭu		ukuta

ආ indicates \bar{a} . It is used with the same consonants as උ e.g. ඹු බු ඹු දු ඹු හු

Examples:

ඹුරු	ඹුරු	ඹුව	ඹුව
	būru		hūva

ඊ indicates \bar{i} , e.g. කි ki ගි gi පි pi

Examples:

කිරි	කිරි	විලි	විලි
	kiri		vili
පිටි	පිටි	නිදි	නිදි
	piṭi		nidi

~ ୧ indicates ī, e.g. କି di ବି vi ଓ hi

Examples:

କିରି	କିରି	ବିବା	ବିବା
	kiri		bina
ହିୟା	ହିୟା	ନିତି	ନିତି
	hiya		niti

୧ ୧ preceding a consonant symbol indicates e following the consonant. (Note that this is true only if another modification does not follow the relevant consonant symbol, for which see the subsequent listings.) e.g. କେ ke ପେ pe ଘେ he

Examples:

କେତା	କେତା	ପେରା	ପେରା	କେହେଲ	କେହେଲ
	keta		pera		kehel
ବେଦା	ବେଦା	ହେଦି	ହେଦି	ବେଲେକ	ବେଲେକ
	veda		hedī		belek

୧...୧ indicates ē, e.g. ଝେ ତେ ଝେ ମେ

୧...୧ } Examples:

୧...୧	୧...୧	୧...୧	୧...୧
୧...୧	୧...୧	୧...୧	୧...୧
	୧...୧		୧...୧
	pēli		veda

୧...୧ indicates o, e.g. କୋ କୋ କୋ ତୋ ଲୋ ଲୋ

୧...୧ } Examples:

କୋଳା	କୋଳା	ଲୋକା	ଲୋକା
	kola		loku
ପୋଟା	ପୋଟା	କୋରା	କୋରା
	poṭa		tora

୧...୧ indicates ō, e.g. କୋ କୋ କୋ ଲୋ ଲୋ

୧...୧ } Examples:

ଲୋକା	ଲୋକା	କୋପା	କୋପା
	lōka		kōpa

... indicates au, e.g. ගෞ gau ලො lau

... Examples:

ගෞරව ගෞරව gaurava ලොකීක ලොකීක laukika

... indicates ai, e.g. වෛ vai තෛ tai

Examples:

දෛව දෛව daiva සෛල සෛල śaila

... by itself following a consonant symbol indicates l. (If ෆ precedes see above.) e.g. ක් ක්

Examples:

ක්ෂ ක්ෂ ක්ෂ kṣta

... indicates ḷ, e.g. ක්ෂ kḷ

... indicates ṛ, e.g. ක්ෂ kṛ ක්ෂ tṛ

Examples:

ක්ෂ ක්ෂ ක්ෂ kṛmi ක්ෂ ක්ෂ kṛna

... indicates ṝ, e.g. ක්ෂ kṝ ක්ෂ pṝ

Examples:

ක්ෂ ක්ෂ ක්ෂ kṝra

... indicates ṛ preceding a consonant.

Examples:

අකි අකි arka සරි සරි sarpa

indicates r following a consonant; the r precedes any vowel implied by or attached to the consonant symbol.

Examples:

අක්‍ර අඞ ව්‍ර ව්‍ර ශ්‍රී ශ්‍රී
akra vapra śrī

When r is used, u and ū are written with : and : :
kru ඞ: krū

ඞ ඞ indicates y following a consonant; the y precedes any vowel implied by or attached to the consonant symbol.

Examples:

තප්‍ය තප්‍ය ශක්‍ය ශක්‍ය
tapyā śakya
අඞ්‍යුදාය අඞ්‍යුදාය
abyudaya

ඞ ශ්‍රී is written after o to form the cluster ośrī. The entire cluster represents ry and is thus an equivalent for oś.

Examples:

ආර්‍ය්‍රී ආර්‍ය්‍රී කාර්‍ය්‍රී කාර්‍ය්‍රී
āryā kārāyā

ඞ ඞ indicates ṃ. If the preceding symbol is a vowel, the ṃ will follow that vowel: e.g. oṃ aṃ. If the preceding symbol is a consonant, the ṃ will follow any vowel implied by or attached to it.

Examples:

ගං ගං සංඝ සංඝ තන්ත්‍රීං තන්ත්‍රීං
gaṃ saṃgha tantrīṃ

ඞ ඞ indicates ḥ. If the preceding symbol is a vowel, the ḥ will follow that vowel: e.g. oḥ. If the preceding symbol is a consonant, the ḥ will follow any vowel implied by or attached to it: e.g. oḥ taḥ.

W.4 Alphabetical Order

Vowels precede consonants. The vowels occur in the order given in W.22. The consonants occur in the order of W.21 (reading row after row taking the top row first and proceeding to the last). ṇ follows the vowels; half-nasals follow ordinary nasal plus consonant*, so that ṇḍ precedes ṇḍ; ṇb precedes ṇb, etc. Hence:

අ, ආ, ඇ, ඈ, ඉ, ඊ, උ, ඌ, ඍ, ඎ, ඏ, ඐ, ඒ, ඔ, ඓ, ඔ, ඍ, ඎ, ඏ, ඐ, ඒ, ඔ, ඓ, ඔ, ඍ, ඎ, ඏ, ඐ, ඒ, ඔ, ඓ, ඔ, ඍ, ඎ, ඏ, ඐ, ඒ, ඔ, ඓ, ඔ, ඍ, ඎ, ඏ.

W.5 Finder List

The finder list which follows on pages 80 and 81 is ordered only by similarity of shape, without regard to what the symbols stand for. It includes basic symbols and others for which the meaning of the whole is not predictable from its parts.

Thus to find an unfamiliar symbol, scan the list for a group of symbols that resemble it generally in shape, and then look for the desired symbol among them. If it is not there, it is probably a basic symbol or a regular combination of a basic symbol and a modification, so consult sections W.2 and W.3.

W.6 Chart of Consonant-vowel Combinations

The way in which individual consonant symbols combine with vowel modifications is shown in the chart on pages 82-83.

*In the order used here, which is a common one, but there is some variation. In some dictionaries and glossaries, when o is an alternant for ඔ (usually before ක, ග, ඛ, or ඝ), or for ඉ (usually before ට, ධ, ඩ, or ඒ), it will be alphabetized in the place of those nasal symbols, but when it is not such an alternate, it will follow the vowels and precede all consonants, as in the order used here. Similarly, the order of the "half-nasals" (ṇ etc.) may vary in relation to the other nasals in different dictionaries.

Note also that o may sometimes be transliterated as ṇ when it represents the velar nasal (ඔ). Thus සිංහල may be found transliterated as sindhala as well as sinhala.

ආ a	උ u	ක ta	ච ඡa	ධ dha
ආ ā	උ ū	ක na	ච ma	ධ o
ආ ä, æ	උ la	ක ttha	ච va	ධ ̄o
ආ ā, ̄æ	උ ̄la	ක nda	ච ca	ධ au
උ da		ක na	ච nca	ධ ̄a
උ ñda		ක ka	ච dva	ධ ña
උ dra		ක jha	ච ba	ධ ddha
උ dā		ක nja	ච kha	ධ ̄, na
උ dya		ක ña		ධ m̄ba
		ක kva		
		ක kṣa		

थ tha ढ pa ढ ya ङ ga ञ i

थु lu थि śa, ṣa ढि r̄ ङ ha ञ ra

थि pha थि ja ढि r̄ ङ bha ञ ī

थे e थि ṅja ढि r̄ ya ङ śa

थे ē थि cha ढि sa ङ ṅga

थे ai थि ḷ थि gha ञ ja

थि tha थि ī ञ su

थि t̄tha ञ sū

थि dha ञ bhū

थि bhū

VOWELS		CONSONANTS															
Vowel	Vowel Symbol After Consonant	k	kh	g	gh	ṅg	c	ch	j	jh	ñ	t	tḥ	d	dḥ	n	ṅḍ
a	අ	ක	ඛ	ග	ඝ	ඟ	ච	ඡ	ජ	ඣ	ඤ	ට	ඨ	ඬ	භ	ඞ	ඣ
ā	ආ	කා	ඛා	ගා	ඝා	ඟා	චා	ඡා	ජා	ඣා	ඤා	ටා	ඨා	ඬා	භා	ඞා	ඣා
ae	ඈ	කැ	ඛැ	ගැ	ඝැ	ඟැ	චැ	ඡැ	ජැ	ඣැ	ඤැ	ටැ	ඨැ	ඬැ	භැ	ඞැ	ඣැ
āe	ඉ	කෑ	ඛෑ	ගෑ	ඝෑ	ඟෑ	චෑ	ඡෑ	ජෑ	ඣෑ	ඤෑ	ටෑ	ඨෑ	ඬෑ	භෑ	ඞෑ	ඣෑ
i	ඊ	කි	ඛි	ගි	ඝි	ඟි	චි	ඡි	ජි	ඣි	ඤි	ටි	ඨි	ඬි	භි	ඞි	ඣි
ī	උ	කී	ඛී	ගී	ඝී	ඟී	චී	ඡී	ජී	ඣී	ඤී	ටී	ඨී	ඬී	භී	ඞී	ඣී
u	ඌ	කු	ඛු	ගු	ඝු	ඟු	චු	ඡු	ඣු	ඤු	ටු	ඨු	ඬු	භු	ඞු	ඣු	ඤු
ū	ඍ	කූ	ඛූ	ගූ	ඝූ	ඟූ	චූ	ඡූ	ඣූ	ඤූ	ටූ	ඨූ	ඬූ	භූ	ඞූ	ඣූ	ඤූ
ṛ	ඎ	කෘ	ඛෘ	ගෘ	ඝෘ	ඟෘ	චෘ	ඡෘ	ජෘ	ඣෘ	ඤෘ	ටෘ	ඨෘ	ඬෘ	භෘ	ඞෘ	ඣෘ
ṝ	ඏ	කෘෘ	ඛෘෘ	ගෘෘ	ඝෘෘ	ඟෘෘ	චෘෘ	ඡෘෘ	ජෘෘ	ඣෘෘ	ඤෘෘ	ටෘෘ	ඨෘෘ	ඬෘෘ	භෘෘ	ඞෘෘ	ඣෘෘ
ḷ	ඐ	කා	ඛා	ගා	ඝා	ඟා	චා	ඡා	ජා	ඣා	ඤා	ටා	ඨා	ඬා	භා	ඞා	ඣා
ḹ	එ	කා	ඛා	ගා	ඝා	ඟා	චා	ඡා	ජා	ඣා	ඤා	ටා	ඨා	ඬා	භා	ඞා	ඣා
e	ඒ	කෙ	ඛෙ	ගෙ	ඝෙ	ඟෙ	චෙ	ඡෙ	ඣෙ	ඤෙ	ටෙ	ඨෙ	ඬෙ	භෙ	ඞෙ	ඣෙ	ඤෙ
ē	ඓ	කේ	ඛේ	ගේ	ඝේ	ඟේ	චේ	ඡේ	ඣේ	ඤේ	ටේ	ඨේ	ඬේ	භේ	ඞේ	ඣේ	ඤේ
ai	ඔ	කො	ඛො	ගො	ඝො	ඟො	චො	ඡො	ඣො	ඤො	ටො	ඨො	ඬො	භො	ඞො	ඣො	ඤො
o	ඕ	කෝ	ඛෝ	ගෝ	ඝෝ	ඟෝ	චෝ	ඡෝ	ඣෝ	ඤෝ	ටෝ	ඨෝ	ඬෝ	භෝ	ඞෝ	ඣෝ	ඤෝ
ō	ඖ	කෝ	ඛෝ	ගෝ	ඝෝ	ඟෝ	චෝ	ඡෝ	ඣෝ	ඤෝ	ටෝ	ඨෝ	ඬෝ	භෝ	ඞෝ	ඣෝ	ඤෝ
au	඗	කො	ඛො	ගො	ඝො	ඟො	චො	ඡො	ඣො	ඤො	ටො	ඨො	ඬො	භො	ඞො	ඣො	ඤො

VOWELS

Vowel	Vowel Symbol After Consonant	t	th	d	dh	n	ñd	p	ph	b	bh	m	mb	y	r	l	v	ś	ṣ	s	h	ḷ
a	අ	ත	ඵ	ද	ඨ	න	ඤ	ප	ඵ	බ	භ	ම	ඹ	ය	ර	ල	ව	ශ	ඡ	ස	හ	ඬ
ā	ආ	තා	ඵා	දා	ඨා	නා	ඤා	පා	ඵා	බා	භා	මා	ඹා	යා	රා	ලා	වා	ශා	ඡා	සා	හා	ඬා
ae	ඈ	තෑ	ඵෑ	දෑ	ඨෑ	නෑ	ඤෑ	පෑ	ඵෑ	බෑ	භෑ	මෑ	ඹෑ	යෑ	රෑ	ලෑ	වෑ	ශෑ	ඡෑ	සෑ	හෑ	ඬෑ
āe	ආඈ	තාඈ	ඵාඈ	දාඈ	ඨාඈ	නාඈ	ඤාඈ	පාඈ	ඵාඈ	බාඈ	භාඈ	මාඈ	ඹාඈ	යාඈ	රාඈ	ලාඈ	වාඈ	ශාඈ	ඡාඈ	සාඈ	හාඈ	ඬාඈ
i	ඉ	ති	ඵි	දි	ඨි	නි	ඤි	පි	ඵි	බි	භි	මි	ඹි	යි	රි	ලි	වි	ශි	ඡි	සි	හි	ඬි
ī	ඊ	තී	ඵී	දී	ඨී	නී	ඤී	පී	ඵී	බී	භී	මී	ඹී	යී	රී	ලී	වී	ශී	ඡී	සී	හී	ඬී
u	උ	තු	ඵු	දු	ඨු	නු	ඤු	පු	ඵු	බු	භු	මු	ඹු	යු	රු	ලු	වු	ශු	ඡු	සු	හු	ඬු
ū	ඌ	තූ	ඵූ	දූ	ඨූ	නූ	ඤූ	පූ	ඵූ	බූ	භූ	මූ	ඹූ	යූ	රූ	ලූ	වූ	ශූ	ඡූ	සූ	හූ	ඬූ
ṛ	ඹ	තෑ	ඵෑ	දෑ	ඨෑ	නෑ	ඤෑ	පෑ	ඵෑ	බෑ	භෑ	මෑ	ඹෑ	යෑ	රෑ	ලෑ	වෑ	ශෑ	ඡෑ	සෑ	හෑ	ඬෑ
ṝ	ඹඹ	තෑඹඹ	ඵෑඹඹ	දෑඹඹ	ඨෑඹඹ	නෑඹඹ	ඤෑඹඹ	පෑඹඹ	ඵෑඹඹ	බෑඹඹ	භෑඹඹ	මෑඹඹ	ඹෑඹඹ	යෑඹඹ	රෑඹඹ	ලෑඹඹ	වෑඹඹ	ශෑඹඹ	ඡෑඹඹ	සෑඹඹ	හෑඹඹ	ඬෑඹඹ
ḷ	ඹ	තෑ	ඵෑ	දෑ	ඨෑ	නෑ	ඤෑ	පෑ	ඵෑ	බෑ	භෑ	මෑ	ඹෑ	යෑ	රෑ	ලෑ	වෑ	ශෑ	ඡෑ	සෑ	හෑ	ඬෑ
ḹ	ඹඹ	තෑඹඹ	ඵෑඹඹ	දෑඹඹ	ඨෑඹඹ	නෑඹඹ	ඤෑඹඹ	පෑඹඹ	ඵෑඹඹ	බෑඹඹ	භෑඹඹ	මෑඹඹ	ඹෑඹඹ	යෑඹඹ	රෑඹඹ	ලෑඹඹ	වෑඹඹ	ශෑඹඹ	ඡෑඹඹ	සෑඹඹ	හෑඹඹ	ඬෑඹඹ
e	ඊ	තෙ	ඵෙ	දෙ	ඨෙ	නෙ	ඤෙ	පෙ	ඵෙ	බෙ	භෙ	මෙ	ඹෙ	යෙ	රෙ	ලෙ	වෙ	ශෙ	ඡෙ	සෙ	හෙ	ඬෙ
ē	ඊඊ	තේඊඊ	ඵේඊඊ	දේඊඊ	ඨේඊඊ	නේඊඊ	ඤේඊඊ	පේඊඊ	ඵේඊඊ	බේඊඊ	භේඊඊ	මේඊඊ	ඹේඊඊ	යේඊඊ	රේඊඊ	ලේඊඊ	වේඊඊ	ශේඊඊ	ඡේඊඊ	සේඊඊ	හේඊඊ	ඬේඊඊ
ai	ඊඊ	තෙඊඊ	ඵෙඊඊ	දෙඊඊ	ඨෙඊඊ	නෙඊඊ	ඤෙඊඊ	පෙඊඊ	ඵෙඊඊ	බෙඊඊ	භෙඊඊ	මෙඊඊ	ඹෙඊඊ	යෙඊඊ	රෙඊඊ	ලෙඊඊ	වෙඊඊ	ශෙඊඊ	ඡෙඊඊ	සෙඊඊ	හෙඊඊ	ඬෙඊඊ
o	ඊ - ා	තොඊ	ඵොඊ	දොඊ	ඨොඊ	නොඊ	ඤොඊ	පොඊ	ඵොඊ	බොඊ	භොඊ	මොඊ	ඹොඊ	යොඊ	රොඊ	ලොඊ	වොඊ	ශොඊ	ඡොඊ	සොඊ	හොඊ	ඬොඊ
ō	ඊ - ා	තෝඊ	ඵෝඊ	දෝඊ	ඨෝඊ	නෝඊ	ඤෝඊ	පෝඊ	ඵෝඊ	බෝඊ	භෝඊ	මෝඊ	ඹෝඊ	යෝඊ	රෝඊ	ලෝඊ	වෝඊ	ශෝඊ	ඡෝඊ	සෝඊ	හෝඊ	ඬෝඊ
au	ඊ - ා	තොඊ	ඵොඊ	දොඊ	ඨොඊ	නොඊ	ඤොඊ	පොඊ	ඵොඊ	බොඊ	භොඊ	මොඊ	ඹොඊ	යොඊ	රොඊ	ලොඊ	වොඊ	ශොඊ	ඡොඊ	සොඊ	හොඊ	ඬොඊ