

# Transliteration, Transcription and Pronunciation

## CONTENTS

<b>Transliteration Scheme and Pronunciation</b>	<b>2</b>
Transliteration table	2
Pronunciation guide	5
<b>Transliteration and Transcription</b>	<b>15</b>
<b>Transliteration of Tamil Script</b>	<b>19</b>
Transliteration of Tamil vowels	20
The Tamil ‘hermaphrodite letter’	20
Transliteration of Tamil consonants	21
Transliteration of Grantha consonants used in Tamil	28
<b>Transliteration of Devanagari Script</b>	<b>30</b>
Transliteration of Sanskrit vowels	31
Transliteration of Sanskrit consonantal diacritics	32
The classification of Sanskrit consonants	33

This PDF document is available for free download from:

[www.happinessofbeing.com/Transliteration.pdf](http://www.happinessofbeing.com/Transliteration.pdf)

An HTML copy of it is available at:

[www.happinessofbeing.com/transliteration.html](http://www.happinessofbeing.com/transliteration.html)

## Transliteration Scheme and Pronunciation

The transliteration scheme that I use is based upon several closely related schemes, namely the [International Alphabet of Sanskrit Transliteration \(IAST\)](#), the scheme used in the [Tamil Lexicon](#), the [National Library at Kolkata romanization scheme](#), the [American Library Association and the Library of Congress \(ALA-LC\) transliteration schemes](#) and the more recent international standard known as '[ISO 15919 Transliteration of Devanagari and related Indic scripts into Latin characters](#)' (a detailed description of which is available [here](#)).

### Transliteration table

The following table summarises this transliteration scheme. In the first column I list all the diacritic and non-diacritic [Latin](#) characters that I use to transliterate the Tamil and Sanskrit alphabets, in the second column I give the [Tamil](#) letter that each such character represents [followed in square brackets where applicable by the [Grantha](#) letter that is optionally used in Tamil to denote the represented sound more precisely], in the third column I give the [Devanagari](#) letter that it represents, and in the last column I give an indication of its pronunciation or articulation.

In the Tamil and Devanagari columns, a dash (–) indicates that there is no exact equivalent in that script for the concerned letter in the other script. In the Tamil column, round brackets enclosing a letter indicates that it is pronounced and transliterated as such only in words borrowed from Sanskrit or some other language. Likewise, in the Devanagari column, round brackets enclosing a letter indicates that it is not part of the alphabet of classical Sanskrit, though it does occur either in Vedic Sanskrit or in some other Indian languages written in Devanagari.

### Vowels:

<i>a</i>	அ	अ	Short 'a', pronounced like 'u' in cut
<i>ā</i>	ஆ	आ	Long 'a', pronounced like 'a' in father
<i>i</i>	இ	इ	Short 'i', pronounced like 'e' in English
<i>ī</i>	ஈ	ई	Long 'i', pronounced like 'ee' in see
<i>u</i>	உ	उ	Short 'u', pronounced like 'u' in put
<i>ū</i>	ஊ	ऊ	Long 'u', pronounced like 'oo' in food
<i>r</i>	–	ऋ	Short vocalic 'r', pronounced like 'ri' in merrily

$\bar{r}$	–	ठ	Long vocalic ‘r’
l	–	ल	Short vocalic ‘l’, pronounced like ‘lry’ in revelry (not to be confused with the Tamil consonant ள, which is also transliterated as l)
$\bar{l}$	–	ड	Long vocalic ‘l’
e	எ	( $\bar{e}$ )	Short ‘e’, pronounced like ‘e’ in else
$\bar{e}$	ஏ	ए	Long ‘e’, pronounced like ‘ai’ in aid
ai	ஐ	ऐ	Diphthong ‘ai’, pronounced like ‘ai’ in aisle
o	ஓ	( $\bar{o}$ )	Short ‘o’, pronounced like ‘o’ in cot
$\bar{o}$	ஔ	ओ	Long ‘o’, pronounced like ‘o’ in dote
au	ஔ	औ	Diphthong ‘au’, pronounced like ‘ou’ in sound

**Consonantal diacritics:**

$\underline{k}$	ஃ	–	Tamil <a href="#">āydam</a> , indicating gutturalization of the preceding vowel, pronounced like ‘ch’ in loch
m	–	◌̣	Sanskrit <a href="#">anusvāra</a> , indicating nasalization of the preceding vowel, pronounced like ‘m’ or (when followed by certain consonants) ‘ñ’, ‘ṇ̃’, ‘ṇ’ or ‘ṇ̣’
h	–	◌̣	Sanskrit <a href="#">visarga</a> , indicating frication (or lengthened aspiration) of the preceding vowel, pronounced like ‘h’ followed by a slight echo of the preceding vowel

**Consonants:**

k	க்	क्	Velar plosive, unvoiced and unaspirated
kh	(க்)	ख्	Velar plosive, unvoiced but aspirated
g	க்	ग्	Velar plosive, voiced but unaspirated
gh	(க்)	घ्	Velar plosive, voiced and aspirated
ṅ	ங்	ङ्	Velar nasal
c	ச்	च्	Palatal plosive, unvoiced and unaspirated (pronounced like ‘c’ in cello or ‘ch’ in chutney)
ch	(ச்)	छ्	Palatal plosive, unvoiced but aspirated
j	ச் [ஐ]	ज्	Palatal plosive, voiced but unaspirated
jh	(ச்)	झ्	Palatal plosive, voiced and aspirated
ñ	ஞ்	ञ्	Palatal nasal
t	ட்	ट्	Retroflex plosive, unvoiced and unaspirated
th	(ட்)	ठ्	Retroflex plosive, unvoiced but aspirated
d	ட்	ड्	Retroflex plosive, voiced but unaspirated

<i>dh</i>	(ட்)	ढ्	Retroflex plosive, voiced and aspirated
<i>n</i>	ண்	ण्	Retroflex nasal
<i>t</i>	த்	त्	Dental plosive, unvoiced and unaspirated
<i>th</i>	(த்)	थ्	Dental plosive, unvoiced but aspirated
<i>d</i>	த்	द	Dental plosive, voiced but unaspirated
<i>dh</i>	(த்)	ध्	Dental plosive, voiced and aspirated
<i>n</i>	ந்	न्	Dental nasal
<i>p</i>	ப்	प्	Labial plosive, unvoiced and unaspirated
<i>ph</i>	(ப்)	फ्	Labial plosive, unvoiced but aspirated
<i>b</i>	ப்	ब्	Labial plosive, voiced but unaspirated
<i>bh</i>	(ப்)	भ्	Labial plosive, voiced and aspirated
<i>m</i>	ம்	म्	Labial nasal
<i>y</i>	ய்	य्	Palatal semivowel
<i>r</i>	ர்	र्	Dental tap (in Tamil phonology) or retroflex trill (in Sanskrit phonology)
<i>l</i>	ல்	ल्	Dental lateral approximant
<i>v</i>	வ்	व्	Labial semivowel
<i>ɾ</i>	ழ்	–	Retroflex central approximant (transliterated as <i>l</i> in the <a href="#">Tamil Lexicon</a> , and commonly transcribed as <i>zh</i> )
<i>l</i>	ள்	(ळ)	Retroflex lateral approximant
<i>t̥</i>	ற்	–	Alveolar plosive, unvoiced (pronunciation of ற only when it is muted, that is, not followed by a vowel)
<i>dr</i>	ற்	–	Alveolar plosive, voiced (pronunciation of ற only when it follows ண)
<i>r</i>	ற்	–	Alveolar trill (pronunciation of ற when it follows and precedes a vowel)
<i>n̄</i>	ண்	–	Alveolar nasal
<i>ś</i>	(ச்) [ʃ]	श्	Palatal aspirated sibilant, pronounced somewhat like ‘s’ in sure (or ‘sh’ in she)
<i>ʂ</i>	(ச்) [ʃ̣]	ष्	Retroflex aspirated sibilant, pronounced somewhat like ‘s’ in sure (or ‘sh’ in she), but with the tongue curled further back
<i>s</i>	ச் [ஸ்]	स्	Dental aspirated sibilant, pronounced like ‘s’ in see
<i>h</i>	க் [ஹ்]	ह्	Voiced glottal fricative

## Pronunciation guide

In the following guide to the pronunciation of Tamil and Sanskrit words as represented by this transliteration scheme, I will not venture to go too deep into the science of phonetics, which is a subject of which my understanding is very limited, but will attempt to offer at least a simple guide.

For those who wish to learn more about the phonetics of Tamil and Sanskrit (and also their phonologies, scripts, transliteration, grammar and so on), there is abundant (but not always entirely reliable) information available online, particularly in [Wikipedia](#) articles that can be accessed through the [language portal](#) such as [Tamil language](#), [Tamil phonology](#), [Tamil script](#), [Tamil grammar](#), [Sanskrit](#), [Vedic Sanskrit](#), [Śikṣā](#) (the science of Sanskrit phonetics), [Grantha script](#), [Devanagari script](#), [Devanagari transliteration](#) and [International Alphabet of Sanskrit Transliteration](#) (IAST).

More detailed information about [phonetics](#) in general (including detailed explanations of many of the technical terms that I have used here) is also available in [Wikipedia](#) and can be accessed through the [index of phonetics articles](#). In many of these articles the symbols of the [International Phonetic Alphabet](#) (IPA) are used, and a detailed list and explanation of these symbols are given in the [Wikipedia:IPA](#) article, which will be helpful to anyone who wants to learn exactly how Tamil and Sanskrit letters should be pronounced.

Each diacritic mark used in this transliteration scheme indicates a specific quality of pronunciation. A [macron](#) above any vowel ( $\bar{a}$ ,  $\bar{i}$ ,  $\bar{u}$ ,  $\bar{r}$ ,  $\bar{l}$ ,  $\bar{e}$  and  $\bar{o}$ ) indicates that it is [long](#) (and a [breve](#) above any vowel, such as  $\breve{a}$ ,  $\breve{i}$  or  $\breve{u}$ , indicates that it is particularly short, though it does not actually occur as a diacritic in this scheme). Except in the case of  $h$ , which indicates that the preceding vowel is aspirated, an [underdot](#) below any consonant ( $\underset{\cdot}{t}$ ,  $\underset{\cdot}{th}$ ,  $\underset{\cdot}{d}$ ,  $\underset{\cdot}{dh}$ ,  $\underset{\cdot}{n}$ ,  $\underset{\cdot}{l}$  and  $\underset{\cdot}{s}$ ) or any ‘vowel’ ( $\underset{\cdot}{r}$ ,  $\underset{\cdot}{\bar{r}}$ ,  $\underset{\cdot}{l}$  and  $\underset{\cdot}{\bar{l}}$ ) indicates that it is [retroflex](#) (as also does the [overdot](#) above the Tamil consonant  $\dot{r}$ ). A macron below any consonant other than  $k$  (namely  $\underset{\bar{t}}$ ,  $\underset{\bar{d}}$ ,  $\underset{\bar{r}}$  and  $\underset{\bar{n}}$ ) in a Tamil word indicates that it is [alveolar](#). An ‘h’ appended to any other consonant ( $kh$ ,  $gh$ ,  $ch$ ,  $jh$ ,  $\underset{\cdot}{th}$ ,  $\underset{\cdot}{dh}$ ,  $\underset{\cdot}{ph}$  and  $\underset{\cdot}{bh}$ ) in a Sanskrit word indicates that it is [aspirated](#).

However, no such general rule applies to any of the other diacritic marks, namely those on  $\underset{\bar{k}}$ ,  $\underset{\bar{m}}$ ,  $\underset{\bar{h}}$ ,  $\underset{\bar{n}}$ ,  $\underset{\bar{\tilde{n}}}$ ,  $\underset{\bar{r}}$  and  $\underset{\bar{s}}$ , so I will explain below what

each of them indicates while discussing each of these characters individually.

### Vowels:

A short vowel is pronounced for a single unit of sound duration called a *mātra*, and a long vowel or diphthong is pronounced for two such units. The first vowel, *a* (अ, आ), is pronounced like ‘u’ in ‘pun’ or ‘a’ in ‘above’, and *ā* (आ, आः) is the same sound pronounced twice as long, like ‘a’ in ‘after’ or ‘father’. The next vowel, *i* (इ, ई), is pronounced like ‘i’ in ‘in’ or ‘e’ in ‘English’, and *ī* (ई, ईः) is the same sound pronounced twice as long, like ‘ee’ in ‘see’. The next vowel, *u* (उ, ऊ), is pronounced like ‘u’ in ‘put’, and *ū* (ऊ, ऊः) is the same sound pronounced twice as long, like ‘oo’ in ‘food’. The next vowel in Tamil, *e* (ஏ), is pronounced like ‘e’ in ‘else’, and *ē* (ஏ, ஏ) is the same sound pronounced twice as long, like ‘ai’ in ‘aid’.

Unlike Tamil, in Sanskrit there is no short *e*, but prior to the long *ē* (ए) there are four other vowels (though the last of these is only classified for the sake of symmetry, since it is never actually used), namely *r* (ऋ), which is pronounced almost like ‘ri’ (as in ‘merrily’) or somewhere between ‘ri’ and ‘ru’ (with a short ‘u’ as in ‘put’), *ṛ* (ॠ), which is the same sound pronounced twice as long (somewhat like ‘ri’ in ‘marine’), *l* (ऌ), which is pronounced almost as ‘lri’ (like ‘lry’ in ‘revelry’), and *ḷ* (ॡ), which is theoretically the same sound pronounced twice as long.

In both Tamil and Sanskrit, *ē* (ए, ए) is followed by the diphthong *ai* (ऐ, ऐ), which is pronounced like ‘ai’ in ‘aisle’. In Tamil this is followed by *o* (ஔ), which is pronounced like ‘o’ in ‘cot’, and then in both languages comes *ō* (ஔ, औ), which is the same sound pronounced twice as long, like ‘o’ in ‘dote’. The final vowel is another diphthong, *au* (ஔ, औ), which is pronounced like ‘ou’ in ‘sound’.

### Consonantal diacritics:

In Tamil the next letter is the ‘hermaphrodite’ *āytam*, *k* (ஃ), which is pronounced somewhat like a guttural ‘k’, ‘g’ or ‘h’ (or ‘ch’ in the Scottish word ‘loch’) appended to the preceding vowel, and which only occurs between a short vowel and one of the ‘hard class’ consonants.

In Sanskrit the fourteen vowels are followed by two consonantal diacritics, the *anusvāra*, *m* (ं), which nasalises the vowel to which it is appended, and which may therefore be pronounced either like ‘m’ or

(when it is not at the end of a sentence) like any form of ‘n’, depending upon which consonant follows it, and the [visarga](#), ḥ (:), which aspirates the vowel to which it is appended, and which is therefore pronounced somewhat like ‘h’, but often followed by a slight echo of the preceding vowel. Thus *namaḥ* is pronounced ‘*namahā*’, and *śāntiḥ* pronounced ‘*śāntihī*’.

### Consonants:

In Sanskrit the consonants begin with five groups of five [stop consonants](#) (known in Sanskrit as *sparśa* or ‘touch’ consonants, since they are formed by complete contact of the organs of utterance), each group consisting of four [plosives](#) (oral stops) followed by one [nasal stop](#). Each group of four plosives consists of a pair of [voiceless](#) (unvoiced or *aghōṣa*) and a pair of [voiced](#) (*ghōṣa*) consonants, and each pair consists of one unaspirated (*alpaprāṇa* or ‘slight breath’) and one [aspirated](#) (*mahāprāṇa* or ‘great breath’) consonant. Thus the order of plosives within each group is unvoiced and unaspirated, unvoiced and aspirated, voiced and unaspirated, and voiced and aspirated.

In Tamil the consonants also begin with the same five groups of stop consonants, but each group of four plosives is represented by a single letter (all of which are collectively known as the [val-l-inam](#) or ‘hard class’ of consonants). As I will explain in more detail in the section on the transliteration of Tamil consonants, these five ‘hard class’ consonants (and the sixth one, *ṭ* [ṭ̣]), which is placed near the end of the [Tamil alphabet](#)) may be either unvoiced or voiced, because their exact pronunciation is determined by their position in a word and the letters that precede or follow them. In words of pure Tamil origin they are never aspirated, but in [loanwords](#) from Sanskrit and other languages they are aspirated where appropriate. In the Tamil alphabet each of these ‘hard class’ consonants is followed by its corresponding nasal (which are collectively known as the [mel-l-inam](#) or ‘soft class’ of consonants), so whereas Sanskrit has five nasal consonants (*ṅ, ṅ̄, ṇ, ṇ̄* and *m*), Tamil has six (*ṅ, ṅ̄, ṇ, ṇ̄, m* and *ṇ*).

With the exception of the sixth group of stop consonants in Tamil (namely *ṭ* [ṭ̣] and *ṇ* [ṇ̣]), which are [alveolar](#) and therefore belong phonetically between the third and fourth group, but which are placed separately at the end of the Tamil alphabet), in both Sanskrit and Tamil the

five groups of stop consonants are arranged phonetically according to their [place of articulation](#), beginning from the back of the mouth (velar) and ending with the lips (labial), so their sequence is [velar](#) (*kaṅṭhya* or ‘guttural’), [palatal](#) (*tālavya*), [retroflex](#) (*mūrdhanya* or ‘cerebral’), [dental](#) (*dantya*) and [labial](#) (*ōṣṭhya*).

Thus in both languages the first group of stop consonants are [velar](#), which means that they are pronounced with the back part of the tongue against the [velum](#) (the [soft palate](#) at the back of the mouth). The first velar consonant in both languages is *k* (क, क्), which is an unaspirated and unvoiced plosive, pronounced like ‘k’ in ‘skip’ (though in Tamil क् can also be pronounced as ‘g’ or ‘h’, according to the rules that I will explain later in the section on the transliteration of Tamil consonants). In Sanskrit this is followed by three more velar plosives, the aspirated unvoiced *kh* (ख), which is pronounced like ‘k’ but with a stronger exhalation (like the more strongly aspirated ‘k’ in ‘kip’), the unaspirated voiced *g* (ग, ग्), which is pronounced like ‘g’ in ‘game’, and the aspirated voiced *gh* (घ), which is pronounced like ‘g’ but with a stronger exhalation. The final consonant in this group is the [velar nasal](#), *ṅ* (ङ, ङ्), which is pronounced like ‘ng’ in ‘sing’.

The second group of stop consonants are [palatal](#), which means that they are pronounced with the body of the tongue raised against the [hard palate](#) (the middle part of the roof of the mouth). The first palatal consonant in both languages is *c* (च, च्), which is an unaspirated and unvoiced plosive, pronounced like ‘ch’ in ‘church’ (though in Tamil च् can also be pronounced as ‘j’ or ‘s’, according to the rules that I will explain in the section on the transliteration of Tamil consonants). In Sanskrit this is followed by three more palatal plosives, the aspirated unvoiced *ch* (छ), which is pronounced like ‘ch’ but with a stronger exhalation, the unaspirated voiced *j* (ज, ज्), which is pronounced like ‘j’ in ‘jug’, and the aspirated voiced *jh* (झ), which is pronounced like ‘j’ but with a stronger exhalation. The final consonant in this group is the [palatal nasal](#), *ñ* (ञ, ञ्), which is pronounced like ‘ni’ in ‘onion’ or ‘ny’ in ‘canyon’.

The first of these palatal consonants, *c* (च, च्), is often transcribed (both in Tamil and in Sanskrit words) as ‘ch’, since it is pronounced like ‘ch’ in many English words such as ‘chair’ (and also in ‘chutney’, which English



has borrowed from an Urdu and Hindi word, *caṭnī*), but its correct transliteration is only ‘c’, since in all precise schemes for transliterating Indic scripts the post-consonantal ‘h’ is reserved for indicating that the consonant to which it is appended is aspirated. Therefore the transliteration ‘ch’ represents the second of the Sanskrit palatal consonants (छ), which is aspirated. Thus, though some frequently used Sanskrit words such as *aruṇācala*, *cit* and *vicāra* are commonly transcribed as ‘Arunachala’, ‘chit’ and ‘vichara’ respectively, when transliterated precisely the ‘ch’ sound should be represented by ‘c’.

In Sanskrit a commonly occurring consonant cluster is *jñ* (for which the Devanagari character is ज्ञ, which is a [ligature](#) of ज् [j] and ञ् [ñ]), but it is not pronounced exactly as it is spelt. In north India *jña* (ज्ञ) tends to be pronounced like ‘gya’, whereas in south India when it occurs in initial position (as for example in ज्ञान, *jñāna*) the *j* (ज्) is hardly pronounced (and hence in Tamil *jñāna* is spelt as it is pronounced, namely ஞானம் [*jñānam*]), whereas in the middle of a word (as for example in *ajñāna* or *prajñāna*) the *j* (ज्) is pronounced somewhat like *g* (which in Tamil is indicated by [gemination](#) of ஞ் [ñ], as for example in *ajñāna*, which is spelt அஞ்ஞானம் [*aññānam*]). In this respect *jñ* is similar to the cognate cluster ‘gn’ in English, because in initial position (as for example in ‘gnaw’ or ‘gnosis’) the ‘g’ is silent, whereas in the middle of a word (as for example in ‘agnostic’ or ‘diagnosis’) the ‘g’ is pronounced.

The third group of stop consonants are [retroflex](#) (as indicated in transliteration by the diacritic [underdot](#)), which means that they are pronounced by curling the tip of the tongue back to point up towards (or in the case of these stop consonants, to actually touch) the roof of the mouth, just behind the [alveolar ridge](#) (for which reason they are called in Sanskrit *mūrdhanya*, ‘head’ or ‘cerebral’ consonants). When articulating any of the five retroflex *sparśa* or ‘touch’ consonants (*t*, *th*, *ḍ*, *ḍh* and *ṇ*) or the Tamil retroflex lateral approximant, *ḷ* (ள), the tip of the tongue actually touches the roof of the mouth, but when articulating Tamil retroflex central approximant, *ṛ* (ṛ), or the Sanskrit retroflex sibilant, *ṣ* (ष्, ष), no contact is made.

The first retroflex consonant is *t* (ṭ, ṭ), which is an unaspirated and unvoiced retroflex plosive, pronounced like an English ‘t’ but with the

tongue curled up (though in Tamil ್ can also be pronounced as ‘d’ with the tongue curled up, according to the rules that I will explain in the section on the transliteration of Tamil consonants). In Sanskrit this is followed by three more retroflex plosives, the aspirated unvoiced *th* (ꣳ), which is pronounced like an English ‘t’ but with the tongue curled up and with a stronger exhalation, the unaspirated voiced *d* (ꣳ, ꣳ), which is pronounced like an English ‘d’ but with the tongue curled up, and the aspirated voiced *dh* (ꣳ), which is pronounced like an English ‘d’ but with the tongue curled up and with a stronger exhalation. The final consonant in this group is the [retroflex nasal](#), *n* (ꣳ, ꣳ), which is pronounced like an English ‘n’ but with the tongue curled up.

The fourth group of stop consonants are [dental](#), which means that they are pronounced with the tongue touching or close to the upper teeth (unlike the corresponding consonants in English, which are pronounced with the tongue touching or close to the alveolar ridge, just above the upper teeth). The first dental consonant is *t* (ꣳ, ꣳ), which is an unaspirated and unvoiced dental plosive, pronounced like an English ‘t’ but with the tongue touching the teeth (though in Tamil ꣳ can also be pronounced as ‘d’ with the tongue touching the teeth, according to the rules that I will explain in the section on the transliteration of Tamil consonants). In Sanskrit this is followed by three more dental plosives, the aspirated unvoiced *th* (ꣳ), which is pronounced like an English ‘t’ but with the tongue touching the teeth and with a stronger exhalation, the unaspirated voiced *d* (ꣳ, ꣳ), which is pronounced like an English ‘d’ but with the tongue touching the teeth, and the aspirated voiced *dh* (ꣳ), which is pronounced like an English ‘d’ but with the tongue touching the teeth and with a stronger exhalation. The final consonant in this group is the [dental nasal](#), *n* (ꣳ, ꣳ), which is pronounced like an English ‘n’ but with the tongue touching the teeth.

The fifth group of stop consonants are [labial](#) (or more precisely, [bilabial](#)), which means that they are pronounced with the lips. The first labial consonant is *p* (ꣳ, ꣳ), which is an unaspirated and unvoiced labial plosive, pronounced like ‘p’ in ‘spun’ (though in Tamil ꣳ can also be pronounced as ‘b’, according to the rules that I will explain in the section on the transliteration of Tamil consonants). In Sanskrit this is followed by three more labial plosives, the aspirated unvoiced *ph* (ꣳ), which is

pronounced like ‘p’ but with a stronger exhalation (like the more strongly aspirated ‘p’ in ‘pun’), the unaspirated voiced *b* (ब, ब्र), which is pronounced like ‘b’ in ‘bat’, and the aspirated voiced *bh* (भ), which is pronounced like ‘b’ but with a stronger exhalation. The final consonant in this group is the [bilabial nasal](#), *m* (म, म्र), which is pronounced like an English ‘m’.

These five groups of stop consonants are followed by a group of oral [sonorants](#) that in Sanskrit are called *antastha* (*antaḥ-stha*), which means ‘standing between’ (since they are sounds that considered to be standing between vowels and true consonants), and in Tamil are called the [itai-y-inam](#) or ‘medial class’ of consonants. Since the Sanskrit term *antastha* is sometimes translated in English as ‘semivowel’, this group of oral sonorants are often loosely described as semivowels, but they are more accurately described as [approximants](#), because only two of them (*y* and *v*) are truly [semivowels](#) and the rest (*r*, *l*, *ṛ* and *l̥*) are [liquids](#).

In Tamil there are six of these ‘medial class’ consonants (*y*, *r*, *l*, *v*, *ṛ* and *l̥*), but in Sanskrit there are only the first four members of this group. The first of these is *y* (य, य्र), which is the palatal semivowel (also called the [palatal approximant](#) or palatal central approximant). The second is *r* (र, र्र), which according to [Tamil phonology](#) is a dental [tap](#) (though phonetically it is described as the [alveolar tap](#)), but according to [Sanskrit phonology](#) is a retroflex [trill](#) (though phonetically it is described as the [alveolar trill](#)). The third is *l* (ल, ल्र), which is traditionally described as the dental ‘l’ or dental [lateral approximant](#) (though phonetically it is described as the [alveolar lateral approximant](#)). The fourth is *v* (व, व्र), which is the labiodental semivowel (also called the [labiodental approximant](#)).

Each of these four oral sonorants is pronounced more or less like its counterpart in English, except that the voiced labiodental approximant, *v* (व, व्र), is often pronounced slightly more like an English ‘w’ (the [voiced labiovelar approximant](#)) [or in Sanskrit loanwords in Tamil like ‘v’ with a slight ‘u’ sound before it], particularly when it follows a mute consonant in Sanskrit (so for example *īśvara* is pronounced ‘īśwara’ [or ‘[īśūvara](#)’], and *svāmi* is pronounced ‘swāmi’ [or ‘[sūvāmi](#)’]), or colloquially in certain Tamil words, such as the respectful greeting [வணக்கம்](#) (*vaṇakkam*), which is often pronounced *waṇakkam*.

In Tamil these four are followed by two more liquid sonorants, both of which are retroflex. The first is the [retroflex central approximant](#), *ṛ* (ṛ), which is pronounced by simultaneously curling the tongue back so that its tip points up towards but does not touch the roof of the mouth, and spreading it sideways so that its sides touch the sides of the upper [alveolar ridge](#), thereby causing the air to flow [centrally](#) (that is, over the centre of the tongue). The resulting sound (which is the final ‘l’ in ‘Tamil’ and which is often transcribed as ‘zh’) is somewhere between an ‘r’ and an ‘l’, but phonetically it is classified as an ‘r’ rather than an ‘l’, since it is not a [lateral](#) sound but a [rhotic](#) one (and hence, though the *Tamil Lexicon* transliterates it as *l*, some scholars [such as the authors of *A Dravidian Etymological Dictionary*] transliterate it as an ‘r’ with either one or two dots below [*r* or *ṛ*, and I transliterate it as *ṛ*, for a reason that I will explain in the section on the transliteration of the Tamil script). The second is the [retroflex lateral approximant](#), *ḷ* (ḷ, ḷ), which is pronounced by curling the tongue back so that its tip touches the roof of the mouth, thereby causing the air to flow [laterally](#) (that is, past both sides of the tongue).

Thus in Tamil there are two pairs of [liquids](#), each of which consists of one [rhotic](#) (or ‘r’-like) sound and one [lateral](#) (or ‘l’-like) sound, the distinction between them being that the first pair, *r* (ṛ) and *l* (ḷ), are dental (that is, pronounced with the tongue close to or touching the upper teeth), whereas the second pair, *ṛ* (ṛ) and *ḷ* (ḷ), are retroflex (that is, pronounced by curling the tip of the tongue back to point up towards or touch the roof of the mouth).

In Tamil these six oral sonorants or ‘medial class’ (*iṭaiyiṇa*) consonants (*y*, *r*, *l*, *v*, *ṛ* and *ḷ*) are followed by the last two letters of the [Tamil alphabet](#), namely the sixth ‘hard class’ (*valliṇa*) consonant, *ṛ* (ṛ), which is the [alveolar trill](#), and the sixth ‘soft class’ (*melliṇa*) consonant, *ṛ* (ṛ), which is the [alveolar nasal](#). Thus, though the pronunciation of this ‘hard class’ *ṛ* (ṛ) is similar to that of the ‘medial class’ *r* (ṛ), it is somewhat harder, and hence it is phonetically classified as a [trill](#), as opposed to the ‘medial class’ *r*, which is classified as a [tap](#).

As I will explain in more detail in the section on the transliteration of Tamil consonants, when this hard *ṛa* (ṛ) is muted (that is, when its inherent vowel sound, *a*, is suppressed, as indicated by the addition of a [pulli](#) or

diacritic dot above it, ण̣), it is pronounced  $ṭ$  (which is a [voiceless alveolar plosive](#)), when it is [geminated](#) (that is, doubled as ण̣ण̣), it is pronounced *tra*, and when it follows the mute form of the final nasal,  $n$  (ண்), it is pronounced *dra* (which is a combination of a [voiced alveolar plosive](#) and an [alveolar trill](#)).

The pronunciation of the final Tamil nasal,  $n$  (ண்), is virtually the same as that of the [dental nasal](#),  $n$  (ந்), though it is phonetically classified as an [alveolar nasal](#), which means that it is pronounced by touching the tip of the tongue against the upper [alveolar ridge](#) rather than the upper teeth. However, in practice this distinction hardly exists (and hence in the [International Phonetic Alphabet](#) the dental and alveolar nasals are represented by the same symbol), and both these Tamil nasals are used to transcribe the Sanskrit dental nasal,  $n$  (न) – the one that is used in each case depending upon which vowels or consonants precede or follow it – so when transliterating Tamil words of Sanskrit origin, I often do not distinguish  $n$  (ண்) but transliterate it according to the Sanskrit spelling as *n*.

In the Tamil alphabet, after this final Tamil consonant,  $na$  (ன), six [Grantha](#) characters (five consonants and one consonantal ligature) are appended for optional use when writing loanwords from Sanskrit or other languages. These six Grantha characters are ஜ (*ja*), ஸ (*śa*), ஶ (*ṣa*), ஸ (*sa*), ஹ (*ha*) and க்ஷ (*kṣa*), which are pronounced exactly like their [Devanagari](#) counterparts, ज (*ja*), श (*śa*), ष (*ṣa*), स (*sa*), ह (*ha*) and क्ष (*kṣa*).

In Sanskrit the four oral sonorants (*y*, *r*, *l* and *v*) are followed by four [fricatives](#), of which the first three are [sibilants](#), namely the [voiceless palatal fricative](#),  $ś$  (श्, ष, श्), the [voiceless retroflex fricative](#),  $ṣ$  (ष्, ष, ष), and the voiceless dental fricative,  $s$  (श्, ष, स). The palatal  $ś$  (श्) is pronounced somewhat like ‘s’ in ‘sure’ or ‘sh’ in ‘she’; the retroflex  $ṣ$  (ष्) is pronounced like  $ś$  but with the tongue curled back to point up at the roof of the mouth; and the dental  $s$  (स) is pronounced like ‘s’ in ‘see’.

The retroflex  $ṣ$  (ष्) is often transcribed as ‘sh’, and the palatal  $ś$  (श्) is sometimes transcribed thus, but since in all precise schemes for transliterating Indic scripts the post-consonantal ‘h’ is used only to distinguish [aspirated](#) (*mahāprāna* or ‘great breath’) consonants from their unaspirated (*alpaprāna* or ‘small breath’) counterparts, and since all the

three Sanskrit sibilants are aspirated and have no unaspirated counterparts, none of them should be transliterated as 'sh'.

The final letter of the Sanskrit alphabet is the [voiced glottal fricative](#), *h* (ह, ह्र, ह्), which is pronounced somewhat like 'h' in 'happy', but with more resonance of the vocal cords.

## Transliteration and Transcription

There are two basic methods that can be employed when writing in [Latin script](#) words from a language whose original script is not Latin based, such as Tamil or Sanskrit, namely precise [transliteration](#) or simple [transcription](#).

Though the terms ‘transcription’ and ‘transliteration’ are often used interchangeably, in a technical sense transcription means the writing of the sounds of one language in the script of another language (and though strictly phonetic transcription employs the use of a technical code such as the [International Phonetic Alphabet](#), simple transcription employs no code other than the basic alphabet of the language in which it is written and is therefore less precise), whereas transliteration means the writing of the script of one language in the script of another language using diacritic marks (or some other device) where necessary to indicate precisely how each word is spelt in the original script.

Thus when a word from a language such as Tamil or Sanskrit is transcribed in Latin script for English-speaking readers, no diacritical characters are used to indicate precisely how it is spelt in its original script (or exactly how it should be pronounced), so it is written using only the twenty-six Latin characters of the English alphabet to indicate approximately how it is pronounced. But when a word from such a language is transliterated in Latin script, a specific (and usually internationally recognised) code employing diacritical characters is used to indicate precisely how the word is spelt in its original script (and also ideally how it should be pronounced).

The principle that I generally follow is just to transcribe the names of people or places and their associated titles whenever I write them in plain type, and to transliterate them and all other Tamil or Sanskrit words, including the titles of poems, books and other texts, whenever I write them in italic type. For example, I transcribe names such as [ரமண](#) (*ramaṇa*), [அருணாசல](#) (*aruṇācala*) and [திருவண்ணாமலை](#) (*tiruvannāmalai*) and titles such as [ஸ்ரீ](#) (*śrī*) and [பகவான்](#) (*bhagavān*) in plain type as ‘Ramana’, ‘Arunachala’, ‘Tiruvannamalai’, ‘Sri’ and ‘Bhagavan’ respectively, except when they are part of the title of a text, but I transliterate all other Tamil or Sanskrit words according to the following principles:

Firstly, I try to strike a balance between precision and user-friendliness, because a precise system of transliteration does not always indicate how a non-Latin script should be pronounced, since many scripts include letters whose pronunciation varies according to context. This is particularly true of the six [‘hard class’ consonants](#) in Tamil (as I will explain in more detail below), and also of the [anusvāra](#) (*m̐*) in Sanskrit. Therefore I transliterate these consonants according to their actual pronunciation, rather than according to a strict one-to-one system of transliteration.

Moreover Tamil has its own system of transcribing Sanskrit words, which in many cases involves certain euphonic changes that alter the original pronunciation, so if Sanskrit words used in Tamil are transliterated exactly as they are spelt in Tamil, they may become unrecognisable. Therefore I generally transliterate such words as they are spelt in Sanskrit, unless there is any particular reason for transliterating them as they are spelt in Tamil (as for example in the case of [ahandai](#), the Tamil form of *aham̐tā*, meaning ‘ego’, which I often transliterate as it is spelt in Tamil, since this is also how it is usually pronounced and since it is word that Sri Ramana uses so frequently in his Tamil writings).

Secondly, since the scripts of Indian languages such as Tamil or Sanskrit do not have any capital letters, I avoid capitalising the initial character of any Tamil or Sanskrit word, except when such a word occurs in the title of a text or at the beginning of an English sentence.

(Incidentally, with the exception a few words such as ‘God’, which are conventionally spelt with an initial capital, I also avoid capitalising the initial character of any English word that is used to denote the one absolute reality, such as ‘self’, ‘being’, ‘consciousness’ or ‘reality’, because there is something intrinsically dualistic about attempting to distinguish between a capitalised ‘Self’ and a lesser ‘self’, when in reality there are no two selves. When used by Sri Ramana, ‘self’ is a metaphysical term that often defies definition, and is intended to defy it. For example, when we translate the term *ātma-vicāra* as ‘self-investigation’, ‘self-scrutiny’ or ‘self-enquiry’, it is unnecessary and undesirable to define whether the word ‘self’ denotes our real self or our false self, because though we may initially imagine that the ‘I’ we are scrutinising is our finite mind or ego, which appears to be our ‘self’, we will end up discovering that what exists and shines as ‘I’ is in



fact only the one infinite being-consciousness, which is our real ‘self’, just as a person may initially imagine that he is scrutinising a snake but will end up discovering that what he is actually looking at is only a rope.)

There are also no hyphens in the scripts of Indian languages, but when transliterating compound words, which occur frequently in such languages, I often use hyphens to separate the individual words within each compound. Thus for example in a compound term such as *ātmajñāna* (which means self-knowledge), I usually insert a hyphen between its two component words, *ātma* and *jñāna*, to indicate that it is a compound formed of these components.

I also follow the [Tamil Lexicon](#) (the comprehensive Tamil-Tamil-English dictionary published in seven volumes by the University of Madras between 1924 and 1939) in using hyphens to separate consonants that are inserted in Tamil compounds according to the rules of word-conjunction (known in Tamil as [punarcci](#) or [sandhi](#), which is a Tamil form of [a Sanskrit word that is spelt \*sam̐dhi\* but pronounced \*sandhi\*](#)). For example, in the compound name [திருவண்ணாமலை](#) (*tiru-v-aṇṇāmalai*), the ‘v’ is inserted to link euphonicly the two consecutive vowels, ‘u’ and ‘a’. This insertion of consonants for the sake of euphonic conjunction often occurs not only in compound words but also between consecutive words within a sentence, when the first of two words ends with a vowel, so where appropriate I separate such inserted consonants by hyphens.

However, when transliterating Tamil, I sometimes omit such conjunctive consonants, particular in the titles of poems or books, such as *Upadēśa-v-Undiyār* or *Guru-Vācaka-k-Kōvai*, which I transliterate simply as *Upadēśa Undiyār* and *Guru Vācaka Kōvai* respectively. Moreover, though it may be useful in some titles to separate certain words, it may not be necessary to separate other words, so for example I transliterate the title *Upadēśa-t-Taṇi-p-Pākkaḷ* (which means ‘Solitary Verses of Instruction’) as *Upadēśa Taṇippākkaḷ*, because it seems unnecessary and inelegant to split the compound word *taṇippākkaḷ* (which means ‘solitary verses’).

Moreover, as I mentioned above, I usually transliterate Tamil words of Sanskrit origin according to the spelling of the Sanskrit original, because this makes it easier for people who do not know Tamil to recognise them, and because in most cases this is more or less how they are pronounced in

Tamil. Thus, for example, I often transliterate the title of the Tamil poem *Ēkāṇma Pañcakam* as *Ēkātma Pañcakam*, but I generally transliterate the title of the Tamil poem *Āṇma-Viddai* as it is spelt and pronounced in Tamil, rather than as *Ātma-Vidyā*, as it would be spelt and pronounced in Sanskrit.

## Transliteration of Tamil Script

When transliterating Tamil words, I use a modified form of the transliteration scheme used in the [Tamil Lexicon](#), which was later incorporated in the [National Library at Kolkata romanization scheme for all Indic scripts](#) and more recently in the international standard known as ‘[ISO 15919 Transliteration of Devanagari and related Indic scripts into Latin characters](#)’ (a detailed description of which is available [here](#)).

The reason why I use a modified form of the [Tamil Lexicon](#) transliteration scheme rather than its original form is that it uses a single Latin character (with or without an appropriate diacritic mark) or pair of Latin characters (in the case of the two diphthongs, ஐ [ai] and ஔ [au]) to represent each individual letter of the [Tamil alphabet](#), and hence it is not a strictly phonetic scheme, because as I will explain in more detail below, the pronunciation of the six ‘[hard class](#)’ [Tamil consonants](#) varies according to the context in which each of them is used. Therefore, I transliterate each such consonant according to its actual pronunciation in each particular context.

I also differ from the [Tamil Lexicon](#) scheme in the transliteration of the retroflex consonant ழ, the mute form of which is the final ‘l’ in the word ‘Tamil’ ([தமிழ்](#), *tamiṛ*), but which in many other words is traditionally transcribed as *zha*, as in Tiruchuzhi ([திருச்சுழி](#), *tiru-c-cuṛi*, the birthplace of Sri Ramana), *vazhi* ([வழி](#), *vari*, which means path, way or means) and *ezhuttu* ([எழுத்து](#), *eṛuttu*, which means a letter or alphabetical character). Whereas the [Tamil Lexicon](#) transliterates this consonant as *la*, I transliterate it as *ra*, because it is not a [lateral approximant](#) (a typical ‘l’-like sound) but a [central approximant](#), which means that it is actually closer to an ‘r’-like sound than an ‘l’-like sound (though when muted it is often pronounced with a somewhat more ‘l’-like quality [almost as a fusion of a retroflex ‘r’ and ‘l’], and in some spoken dialects it is wrongly pronounced as ள [*la*], which creates confusion between certain pairs of words such as [அழி](#) [*ari*, meaning ‘destroy’] and [அளி](#) [*ali*, meaning ‘give’]).

This consonant ழ (*ra*) is technically described as the [retroflex central approximant](#), in contrast to the subsequent consonant, ள (*la*), which is the [retroflex lateral approximant](#). Therefore, since the standard diacritic used to distinguish the [retroflex consonants](#) is the [underdot](#) (as in *t, th, d, dh, n, l*

and *ṣ*), *ṛ* should logically be transliterated as *ra* (as some scholars transliterate it), but since *r* is used to transliterate the seventh vowel in Sanskrit (the retroflex *ṛ*), to avoid confusion I chose instead to transliterate *ṛ* as *ṛa*.

### Transliteration of Tamil vowels

Since the pronunciation of each of the twelve Tamil vowels is fixed, the [Tamil Lexicon](#) transliteration scheme is phonetic with respect to them, so I transliterate each vowel accordingly as follows: அ *a*, ஆ (ா) *ā*, இ (ி) *i*, ஈ (ீ) *ī*, உ (ு) *u*, ஊ (ூ) *ū*, எ (ெ) *e*, ஏ (ே) *ē*, ஐ (ை) *ai*, ஒ (ொ) *o*, ஔ (ோ) *ō*, ஔ (ௌ) *au*.

The five vowels with a [macron](#) above are each a long form of the same vowel without a macron, so for example அ (*a*) is a short sound while ஆ (*ā*) is a long sound. The symbol that I have given in brackets after each vowel except அ (*a*) is the typical diacritic form that that vowel takes when it is combined with a preceding consonant (though the actual diacritic form of some vowels varies according to the consonant with which it is combined), the dotted circle in each case representing the position of the consonant in the resulting ligature or compound set of characters. The first vowel, அ (*a*), has no such diacritic form, because it is inherent in the default form of each consonant, and whenever a consonant is not followed by any vowel sound, the absence of a vowel is indicated by a [pulli](#) (◌̣), a diacritic dot placed above the consonant. Thus, for example, the thirteen possible forms of the first consonant க (*ka*) are as follows: க் (*k*), க (*ka*), கா (*kā*), கி (*ki*), கீ (*kī*), கு (*ku*), கூ (*kū*), கெ (*ke*), கே (*kē*), கை (*kai*), கொ (*ko*), கோ (*kō*) and கௌ (*kau*).

### The Tamil ‘hermaphrodite letter’

There is one Tamil letter called the [āyṭam](#), which is considered to be neither a pure vowel nor a pure consonant, and is therefore described as the [ali-y-eruttu](#) or ‘hermaphrodite letter’. It is hence the thirteenth letter in the Tamil alphabet, being placed after the twelve vowels and before the eighteen consonants, and it is pronounced somewhat like a guttural ‘k’, ‘g’ or ‘h’ (or ‘ch’ in the Scottish word ‘loch’) appended to the preceding vowel. It is written as ெ and it is transliterated as *ḳ*.

In words of Tamil origin it is rarely used, occurring only in about fifty

such words, some of which are alternative spelling of other words, such as அஃகுடி (*akdu*), which is an alternative spelling of அகுடி (*adu*), meaning ‘that’, and it always occurs only after a short open initial syllable (that is, an initial syllable consisting of only a short vowel or a short consonant-vowel combination, without any closing mute consonant) and before a ‘hard class’ consonant.

### Transliteration of Tamil consonants

There are eighteen pure Tamil consonants, plus another six consonantal characters borrowed from the [Grantha](#) script that are optionally used in Tamil to write words of Sanskrit origin. The original eighteen Tamil consonants are divided into three groups of six, namely the [val-l-inam](#) or ‘hard class’ of consonants, which consists of six [plosives](#) (oral stops), க (*ka*), ச (*ca*), ட (*ta*), த (*ta*), ப (*pa*) and ற (*ra*, a trill, whose muted form, *t̤*, is a plosive), the [mel-l-inam](#) or ‘soft class’ of consonants, which consists of six corresponding [nasals](#), ங (*ṅa*), ஞ (*ṅa*), ண (*ṇa*), ந (*na*), ம (*ma*) and ன (*na*), and the [itai-y-inam](#) or ‘medial class’ of consonants, which consists of six oral [sonorants](#) (two [semivowels](#) and four [liquids](#)), ய (*ya*), ர (*ra*), ல (*la*), வ (*va*), ழ (*ṛa*) and ள (*la*).

The pronunciation of each of the six soft consonants (ங, ஞ, ண, ந, ம and ன) and six medial consonants (ய, ர, ல, வ, ழ and ள) is more or less fixed, as is the pronunciation of each of the six Grantha characters, ஜ (*ja*), ஶ (*śa*), ஷ (*ṣa*), ஸ (*sa*), ஹ (*ha*) and க்ஷ (*kṣa*), but the pronunciation of each of the six hard consonants (க, ச, ட, த, ப and ற) varies, and hence I transliterate them as far as possible according to their actual pronunciation in each context.

In the case of words of Tamil origin, the actual pronunciation of each of these six hard consonants depends largely upon its position in a word and whether or not it is conjoined to another consonant, because each of them has an [unvoiced](#) and a [voiced](#) form of pronunciation, and some of them also have a [fricative](#) or [trilled](#) form of pronunciation. In simpler terms, the unvoiced pronunciation of each of these hard consonants (*k*, *c*, *t*, *t*, *p* and *t̤* respectively) is hardest, their voiced pronunciation (*g*, *j*, *d*, *d*, *b* and *d̤* respectively) is softer, and the fricative pronunciation of the first two (*h* and *s* respectively) and the trilled pronunciation of the final one (*r*) are still softer.

More precisely, ‘unvoiced’ (or ‘[voiceless](#)’) means a sound that is pronounced without resonance of the vocal cords (as in the normal English pronunciation of ‘k’, ‘ch’, ‘t’ or ‘p’), whereas ‘[voiced](#)’ means one that is pronounced with resonance of the vocal cords (as in the normal English pronunciation of ‘g’, ‘j’, ‘d’ or ‘b’), and ‘[fricative](#)’ means one that is pronounced by forcing the breath through a narrow opening between two organs of articulation (as in the normal English pronunciation of ‘h’ or ‘s’), whereas ‘trilled’ means one that is pronounced with vibration of (in this case) the tongue.

The hardest (unvoiced) pronunciation of the [velar](#) க is *ka*, its softer (voiced) pronunciation is *ga*, and its softest (fricative) pronunciation is *ha*. The hardest (unvoiced) pronunciation of the [palatal](#) ச is *ca* (pronounced somewhat like the English ‘chu’ in ‘chum’ or ‘chutney’), its softer (voiced) pronunciation is *ja*, and its softest (fricative) pronunciation is *sa*. The hardest (unvoiced) pronunciation of the [retroflex](#) ட is *ṭa* and its softest (voiced) pronunciation is *ḍa*. The hardest (unvoiced) pronunciation of the [dental](#) த is *ta* and its softest (voiced) pronunciation is *da*. The hardest (unvoiced) pronunciation of the [labial](#) ப is *pa* and its softest (voiced) pronunciation is *ba*. And the hardest (unvoiced) pronunciation of the [alveolar](#) ற is *ṭa* (though this is actually used only when it is muted), its softer (voiced) pronunciation is *ḍra*, and its softest (trilled) pronunciation is *ra*.

The rules that determine how each of these ‘hard class’ consonants should be pronounced in any given context are defined in [Tolkāppiyam](#), the most ancient extant Tamil grammar, and are based upon natural principles of [euphony](#), which are as follows: The pronunciation of each of these consonants is hardest (unvoiced) when it is the initial letter of a word, when it is muted (that is, when its inherent vowel sound is suppressed), when it is [geminated](#) (that is, when its basic consonantal sound is lengthened by duplication) or when it follows any other muted ‘hard class’ consonant; it is softer (voiced) when it occurs ‘post-nasally’ (that is, after any muted ‘soft class’ consonant, ன், ஞ், ண், ற், ழ or ள்) or when it follows any muted ‘medial class’ consonant (such as ற் or ற்); and it is softest (either fricative or trilled, or if it has no such form, voiced) when it occurs [intervocally](#) (that is, between two vowels).

Thus க is pronounced *ka* when it is the initial letter of a word, *k* when it is muted (க்), *kka* when it is geminated (க்க), *ka* when it follows any other muted hard consonant (such as ட or த்), *ga* when it follows a muted soft consonant (as in the frequently occurring cluster க்ங், which is pronounced *nga*) or a muted medial consonant (such as ங or ற்), and *ha* when it follows a verb. Likewise ச is pronounced *ca* (or arbitrarily *sa*, as in fact it is customarily pronounced in many if not most cases, though strictly speaking this contravenes the ancient rule described here) when it is the initial letter of a word, *c* when it is muted (ச்), *cca* when it is geminated (ச்ச), *ca* when it follows any other muted hard consonant, *ja* when it follows a muted soft consonant (as in the frequently occurring cluster ச்ஞ், which is pronounced *ñja*), and *sa* when it follows a verb. ட is not the initial letter of any word of Tamil origin, but it is pronounced *t* when it is muted (ட்), *tta* when it is geminated (ட்ட), and *da* when it follows either a muted soft consonant (as in the frequently occurring cluster ட்ளட, which is pronounced *nda*) or a verb. த is pronounced *ta* when it is the initial letter of a word, *t* when it is muted (த்), *tta* when it is geminated (த்த), and *da* when it follows either a muted soft consonant (as in the frequently occurring cluster த்ஞ், which is pronounced *nda*), a muted medial consonant or a verb. ப is pronounced *pa* when it is the initial letter of a word, *p* when it is muted (ப்), *ppa* when it is geminated (ப்ப), *pa* when it follows any other muted hard consonant, and *ba* when it follows either a muted soft consonant (as in the frequently occurring cluster ப்ம, which is pronounced *mba*, or in the clusters ப்ளப and ப்லப, which are pronounced respectively *nba* and *nba*) or a verb.

Like these other hard consonants, the final hard consonant, ற, also has several [allophones](#) or variant forms of pronunciation. Like ட (*ta*), it is never the initial letter of a word. Its default pronunciation is considered to be *ra* (in which *r* is a trilled ‘r’, described technically as an [alveolar trill](#)), but its mute form (ற்) is pronounced *t* (or sometimes slightly more like *r*, depending upon which consonant it precedes, and when it is used in the transliteration of a word of Sanskrit origin, it can also be pronounced *d* or *l*). Its geminated form (ற்ற) is pronounced *tra*, and the cluster ற்ற is pronounced *ndra*, the extra *d* sound being a natural euphonic increment.

In earlier forms of Tamil this consonant ற was probably treated as

being essentially the alveolar plosive,  $t$  (ṭ), phonetically belonging between the retroflex plosive,  $t$  (ḷ), and the dental plosive,  $t$  (ṭ), but since it never occurs at the beginning of a word and is therefore pronounced  $t$  only when it is muted, it has come to be considered to be basically an alveolar  $r$ , as it is pronounced when followed by any vowel, and hence it is commonly called the *valliṇa* ‘*ra*’ (hard class ‘*ra*’), or more colloquially as the *periya* ‘*ra*’ (big ‘*ra*’). However, if we consider it to be essentially the alveolar  $t$ , it is clearly a plosive like the other five *valliṇa* or ‘hard class’ consonants ( $k$ ,  $c$ ,  $t$ ,  $t$  and  $p$ ), and it follows the same basic rule concerning its pronunciation, namely that it is unvoiced ( $t$ ) when it is muted, voiced ( $dr$ ) when it occurs post-nasally, and trilled ( $r$ ) when it occurs intervocalically. According to this rule, it should also be unvoiced when it is geminated ( $tt$ ), as it is in certain dialects such as Sri Lankan Tamil (and also in Malayalam), but in most Tamil dialects it is actually pronounced  $tr$ .

However, these rules for the pronunciation of the six hard consonants are usually followed only in the case of words of Tamil origin, and are seldom followed in the case of words of Sanskrit origin, which are usually pronounced more or less as they would be in Sanskrit. For example, if புத்தி (which means intellect) were a word of Tamil origin, it would be pronounced *putti*, but since it is of Sanskrit origin, it is pronounced *buddhi*.

The reason why the normal rules for the pronunciation of these six hard consonants are thus generally not followed in the case of words of Sanskrit origin is that each of the first five Tamil plosive consonants represents a series of four Sanskrit consonants, and some of them are also used to transcribe other Sanskrit consonants, as also is the muted alveolar plosive, ṭ ( $t$ ). That is, க ( $ka$ ) represents the series of four Sanskrit [velar](#) plosives, क ( $ka$ ), ख ( $kha$ ), ग ( $ga$ ) and घ ( $gha$ ), and is also used to transcribe the Sanskrit voiced glottal fricative, ह ( $ha$ ); ச ( $ca$ ) represents the series of four Sanskrit [palatal](#) (or postalveolar) plosives, च ( $ca$ ), छ ( $cha$ ), ज ( $ja$ ) and झ ( $jha$ ), and is also used to transcribe the three Sanskrit unvoiced sibilant fricatives, श ( $śa$ ), ष ( $ṣa$ ) and स ( $sa$ ); ட ( $ta$ ) represents the series of four Sanskrit [retroflex](#) plosives, ट ( $ṭa$ ), ठ ( $ṭha$ ), ड ( $ḍa$ ) and ढ ( $ḍha$ ), and its mute form (ḷ) is often used to transcribe the muted Sanskrit retroflex sibilant, ढ ( $ṣ$ ); த ( $ta$ ) represents the series of four Sanskrit [dental](#) plosives, त ( $ta$ ), थ ( $tha$ ), द ( $da$ ) and ध ( $dha$ ); and ப ( $pa$ ) represents the series of four Sanskrit [bilabial](#)



plosives, प (*pa*), फ (*pha*), ब (*ba*) and भ (*bha*).

The Tamil [alveolar](#) plosive, ற (*ra*), does not have any equivalent in Sanskrit, but its mute form, ற் (*t*), is used to produce a euphonic transcription of certain mute consonants in Sanskrit. For instance, the Sanskrit words *sadguru* (a compound of *sat* and *guru*, which means ‘real teacher’ or ‘teacher of reality’), *sadbhāva* (a compound of *sat* and *bhāva*, which means ‘real being’ or ‘state of being’), *alpa* (which means small, little or insignificant) and *kalpanā* (which means fabrication, imagination, mental creation or illusion) are transcribed in Tamil as சற்குரு (*caṭkuru*), சற்பாவம் (*caṭpāvam*), அற்பம் (*aṭpam*) and கற்பனை (*kaṭpanai*) respectively, but are pronounced *sadguru*, *sadbhāvam*, *alpa* and *kalpanai*.

Certain words that are spelt the same in Tamil are actually derived from different words in Sanskrit, and hence they are pronounced in each case according to the Sanskrit word from which they are derived. For example, பாவம் is pronounced *bhāvam* when it is [the Tamil form of \*bhāva\*](#) (a Sanskrit word with many meanings such as becoming, being, state, nature, feeling, thought, idea, imagination or meditation), but *pāvam* when it is [the Tamil form of \*pāpa\*](#) (a Sanskrit word that means a sinful action or the result of such an action).

One other Tamil consonant is commonly used to transcribe more than one sound in Sanskrit, namely the [palatal nasal](#), ஞ (*ñā*), which is used to transcribe both its Sanskrit equivalent, ञ (*ñā*), and the consonant cluster ज्ञ (*jñā*). Though ज्ञ (*jñā*) is a ligature of ज् (*j*) and ञ (*ñā*), when it occurs in initial position (as for example in *jñāna*) the ज् (*j*) is hardly pronounced, whereas in the middle of a word (as for example in *ajñāna*) the ज् (*j*) is pronounced somewhat like *g* (as I explained earlier in the section on pronunciation). This is reflected in Tamil by the fact in initial position ज्ञ (*jñā*) is transcribed as ஞ (*ñā*), whereas in the middle of a word it is transcribed as ஞ்ஞ (*ññā*), so for example ज्ञान (*jñāna*) is transcribed as ஞானம் (*ñānam*) whereas अज्ञान (*ajñāna*) is transcribed as அஞ்ஞானம் (*aññāna*). However, to avoid confusion, when either ஞ (*ñā*) or ஞ்ஞ (*ññā*) are used in Tamil to represent ज्ञ (*jñā*) in a word of Sanskrit origin, I transliterate them as *jñā* in accordance with the actual spelling of the original word.

In Tamil, loanwords from Sanskrit can take one of two forms, [தற்சமம்](#) (*tatsamam*) or [தற்புவம்](#) (*tadbhavam*), the former meaning words that are borrowed without any alteration in pronunciation (or in spelling, other than any adaptation necessitated by the availability of a different range of vowels and consonants in Tamil script, and by any appropriate change to the form of the nominative case-ending) and the latter meaning words that are borrowed with certain permitted alterations in pronunciation and spelling. Some examples of the permitted alterations that can occur in *tadbhava* loanwords are given in [Indo-Aryan loanwords in Tamil](#), but many of the ‘changes’ listed there (such as *ahaṃ* to *akam*, *agati* to *akati*, *adhikāra* to *atikāra*, *adhika* to *atika*, *aśuddha* to *acutta*, *ahaṃkāram* to *akankāram*, *ahambhāva* to *akampāvam*, *ārambha* to *ārampam*, *kārttika* to *kārttikai*, *śani* to *cani* and *budha* to *putan*) are not actual *tadbhava* modifications but are merely adaptations necessitated by the restrictions imposed by Tamil script (and in a few cases by Tamil nominative case-endings) and therefore appear to be *tadbhava* changes only when each Tamil plosive (‘hard class’ consonant) is transliterated by a single unvoiced Latin consonant rather than by whichever one of the range of consonant sounds that they each truly represent.

Therefore, though the one-to-one transliteration scheme used in the [Tamil Lexicon](#) and recommended in the international standard ‘[ISO 15919](#)’ is useful in that it indicates exactly how a word is spelt in Tamil, its value is limited because it does not serve to indicate how a word is actually pronounced (except to people who already know both the principles that determine how each of the six hard consonants should be pronounced in any given context and the principles that determine how Sanskrit words should be transcribed in Tamil). Therefore, though I transliterate the twelve Tamil vowels, the six soft consonants (except [ஃ](#)) and the six medial consonants (except [ழ](#)) strictly according to the [Tamil Lexicon](#) scheme, I transliterate the six hard consonants and [ஃ](#) according to their actual pronunciation in each particular context, and [ழ](#) according to a more accurate phonetic representation.

Thus I transliterate the eighteen Tamil consonants as follows:

- [ஃ](#) (*valliṇa* velar plosive or fricative) as *ka*, (*kha*), *ga*, (*gha*) or *ha*
- [ழ](#) (*melliṇa* velar nasal) as *ṇa*

- ூ (*valliṇa* palatal plosive or fricative) as *ca*, (*cha*), *ja*, (*jha*), (*śa*), (*ṣa*) or *sa*
- ௃ (*melliṇa* palatal nasal) as *ñā* or (*jñā*)
- ௄ (*valliṇa* retroflex plosive) as *ṭa*, (*ṭha*), *ḍa* or (*ḍha*)
- ௅ (*melliṇa* retroflex nasal) as *ṇa*
- ெ (*valliṇa* dental plosive) as *ta*, (*tha*), *da* or (*dha*)
- ே (*melliṇa* dental nasal) as *na*
- ை (*valliṇa* bilabial plosive) as *pa*, (*pha*), *ba* or (*bha*)
- ௉ (*melliṇa* bilabial nasal) as *ma*
- ொ (*iṭaiyiṇa* palatal central approximant) as *ya*
- ோ (*iṭaiyiṇa* dental tap) as *ra*
- ௌ (*iṭaiyiṇa* dental lateral approximant) as *la*
- ் (*iṭaiyiṇa* labiodental central approximant) as *va*
- ௎ (*iṭaiyiṇa* retroflex central approximant) as *ṛa* [though it is transliterated in the [Tamil Lexicon](#) as *ṛa* and is commonly transcribed as *zha*]
- ௏ (*iṭaiyiṇa* retroflex lateral approximant) as *ḷa*
- ௐ (*valliṇa* alveolar plosive or trill) as *ra* [ṛ as *ṛ*, *r*, (*d*) or (*l*); ṛṛ as *ṛra*; ௐ as *ndra*]
- ௑ (*melliṇa* alveolar nasal) as *ṇa*

In this list, the first transliteration for each consonant (which in the case of each of the first five plosives is its unvoiced form) other than ௎ (*ṛa*) is the transliteration used in the [Tamil Lexicon](#) and recommended by ‘ISO 15919’, and each transliteration enclosed in brackets is one that I use only when transliterating a Tamil word of Sanskrit origin.

The main consonant sounds in Sanskrit and other Indian languages that are missing in Tamil are the ten aspirated plosives (*kh*, *gh*, *ch*, *jh*, *ṭh*, *ḍh*, *th*, *dh*, *ph* and *bh*) and the distinction made between the three varieties of sibilant, the palatal *ś*, the retroflex *ṣ* and the dental *s*. Though there are no separate Tamil letters for the unaspirated voiced plosives (*g*, *j*, *ḍ*, *d* and *b*), the voiced glottal fricative (*h*) or any basic sibilant (*s*), these sounds do exist in Tamil as [allophones](#) (variant forms) of the first five hard consonants, ூ (*k*), ூ (*c*), ௄ (*t*), ெ (*t*) and ை (*p*).

Two consonant sounds, the retroflex ௎ (*ṛ*) and alveolar ௐ (*r*), occur only in Tamil and Malayalam (and perhaps in some other minor or older

Dravidian languages) but not in Sanskrit or most other Indian languages, and one consonant, the alveolar nasal ண (*ṇ*), is unique to Tamil (and older forms of Malayalam), though in pronunciation it is virtually the same as ण (*ṇ*).

Unlike Grantha, Devanagari and most other Indic scripts, Tamil script does not form consonant clusters into [ligatures](#) (single characters that each represent a cluster of two or more letters), so in this respect it is much easier to learn than most other Indic scripts.

### Transliteration of Grantha consonants used in Tamil

As I mentioned above, there are also six [Grantha](#) characters that are used in Tamil, particularly in words of Sanskrit origin. Grantha is an ancient script that was used in south India to write Sanskrit, and it predates [Devanagari](#), which is nowadays the standard script used for writing Sanskrit. The use of Grantha in south India was replaced by Devanagari only in the late nineteenth and early twentieth century, when it became the custom to print Sanskrit texts in Devanagari.

Since the Tamil script does not have separate characters to represent all the Sanskrit letters, it borrowed the following six characters from Grantha: ஜ (*ja*), ஸ (*śa*), ஷ (*ṣa*), ண (*ṇa*), ஹ (*ha*) and க்ஷ (*kṣa*). However the use of these Grantha letters is optional, because each of them can be transcribed using a Tamil letter (or pair of Tamil letters), and ஸ (*śa*) in particular is used very rarely nowadays. ஜ (ज or *ja*), ஸ (श or *śa*), ஷ (ष or *ṣa*) and ண (ण or *ṇa*) are each usually transcribed using the Tamil letter ச (*ca*); the muted retroflex sibilant, ஷ் (ष् or *ṣ*), is often transcribed using the muted Tamil retroflex plosive, ட் (*t*); the glottal fricative, ஹ (*h* or *ha*), is usually transcribed using the Tamil velar fricative, க (*ka*); and க்ஷ (क्ष or *kṣa*), which is a ligature of க் (क् or *k*) and ஷ் (ष् or *ṣ*), is often transcribed using the Tamil cluster ட்ச (*tca*).

However, even when such Tamil characters are used to transcribe Sanskrit words containing these letters, I usually transliterate them according to the Sanskrit letters that they represent, in order to indicate clearly in each case what the original Sanskrit word is. For example, in Tamil the Sanskrit word मोक्ष (mōkṣa), which means ‘liberation’, can be written either as மோக்ஷம் (*mōkṣam*) or as [மோட்்சம்](#) (*mōtcam*), but whichever way it happens to be written, I would usually transliterate it as

*mōkṣam*. Likewise, in Tamil the Sanskrit compound word आत्मनिष्ठा (*ātma-niṣṭhā*), which means ‘self-abidance’, can be written either as ஆத்மநிஷ்டை (*ātma-niṣṭhai*) or as ஆன்மநிட்டை (*ānma-niṭṭāi*), but whichever way it happens to be written, I would usually transliterate it as *ātma-niṣṭhā* (or *ātma-niṣṭhai*).

Besides கூ (kṣa) and its various verb-bound forms, there is one other consonantal ligature that Tamil has borrowed from Grantha, namely ஸ்ரீ (*śrī*), for which there is no exact equivalent in pure Tamil script (but which is sometimes transliterated as ஸ்ரீ [*śrī*] in certain names such as Srinivasan).

### Transliteration of Devanagari Script

As I explained above, since the late nineteenth century [Devanagari](#) has become the standard script for writing Sanskrit, partly due to standardisations that took place under the influence of the British education policy in India, and partly due to the subsequent fact that Devanagari was adopted as the principal script in which Sanskrit texts were printed. Prior to that, such texts were written not only in Devanagari (in many northern regions of India) and [Grantha](#) (in many southern regions of India), but also in several other regional scripts such as [Bengali](#), [Gujarati](#), [Kannada](#), [Malayalam](#), [Oriya](#) and [Telugu](#), and in older scripts such as [Brahmi](#), [Gupta](#) and [Sarada](#), because most Indic scripts other than Tamil have characters that denote every sound in the Sanskrit alphabet.

Though the Sanskrit alphabet has more vowels than and nearly twice as many consonants as the Tamil alphabet, it is actually simpler to transliterate phonetically, because each vowel or consonant has a more or less fixed pronunciation, unlike the ‘hard class’ consonants in Tamil. However, the Tamil script is simpler to read and write than Devanagari, because it uses [ligatures](#) only for some consonant-vowel compounds, whereas Devanagari uses ligatures not only for consonant-vowel compounds but also for all consonant clusters, which are numerous and often quite complex, since a single ligature can represent a cluster of several consonants. However, though the numerous ligatures in Devanagari make it somewhat difficult to read or write, this is not a problem when it is transliterated into Latin script using diacritics, because each member of a ligature is then represented by separate Latin letter or pair of letters.

When transliterating Sanskrit words in Latin script, I use a slightly modified form of the [International Alphabet of Sanskrit Transliteration](#) (IAST), upon which both the [National Library at Kolkata romanization scheme](#) and ‘[ISO 15919](#)’ are based. I deviate from IAST only in regard to two long Sanskrit vowels, ए and ओ, which I transliterate with a macron above as *ē* and *ō* respectively, and the [anusvāra](#),<sup>7</sup> (a diacritic dot placed above a vowel or consonant-vowel compound indicating that it is nasalised), which I transliterate either as *m̄* or as some other nasal (*ñ*, *ṅ*, *ṇ*, *n* or *m*), depending upon which consonant follows it (as I will explain in more detail below).

In IAST these two long vowels, ए and औ, are transliterated without a macron as *e* and *o* respectively, because in Sanskrit there are no short equivalents of them, but the National Library at Kolkata romanization scheme and ‘ISO 15919’ both recommend that they should each be transliterated with a macron above, as I do, both for the sake of uniformity and to avoid confusion with the short *e* and *o* in [Dravidian languages](#) such as Tamil, which have both a short *e* and *o* and a long  $\bar{e}$  and  $\bar{o}$ .

The [anusvāra](#) (ँ) is transliterated in IAST as *ṃ* (with an underdot), but ‘ISO 15919’ recommends that it should be transliterated as *ṁ* (with an overdot), and hence I use *ṁ* as its default transliteration, though in many cases I transliterate it as whichever nasal it is actually pronounced in that context.

### Transliteration of Sanskrit vowels

Thus I transliterate the fourteen Sanskrit vowels as follows: अ *a*, आ (ऀ)  $\bar{a}$ , इ (ः) *i*, ई (ः)  $\bar{i}$ , उ (ः) *u*, ऊ (ः)  $\bar{u}$ , ऋ (ः) *r*, ॠ (ः)  $\bar{r}$ , ऌ (ः) *l*, ॡ (ः)  $\bar{l}$ , ए (ः)  $\bar{e}$ , ऐ (ः) *ai*, ओ (ः)  $\bar{o}$ , औ (ः) *au*.

The seven vowels with a macron above are each pronounced long, as are the two diphthongs, ऐ (*ai*) and औ (*au*), so for example अ (*a*) is a short sound while आ ( $\bar{a}$ ) is a long sound. The symbol that I have given in brackets after each vowel except अ (*a*) is the diacritic form that that vowel takes when it is combined with a preceding consonant, the dotted circle in each case representing the position of the consonant in the resulting ligature. The first vowel अ (*a*) has no such diacritic form, because it is inherent in the default form of each consonant, and whenever a consonant is not followed by any vowel sound, the absence of a vowel is indicated by a [virāma](#) (◌), an oblique diacritic stroke placed below the consonant (though this is used much less frequently than the equivalent diacritic [pulli](#) or overdot [◌̇] in Tamil, because except at the end of a sentence a mute consonant in Devanagari should always be joined to the following consonant to form a ligature). Thus, for example, the fifteen possible forms of the first consonant क (*ka*) are as follows: क (*k*), क (*ka*), का ( $k\bar{a}$ ), कः (*ki*), की ( $k\bar{i}$ ), कु (*ku*), कू ( $k\bar{u}$ ), कृ (*kr*), कृ ( $k\bar{r}$ ), क्ल (*kl*), क्ल ( $k\bar{l}$ ), के ( $k\bar{e}$ ), कै (*kai*), को ( $k\bar{o}$ ) and कौ (*kau*).

Four of these Sanskrit vowels, ऋ (*r*), ॠ ( $\bar{r}$ ), ऌ (*l*) and ॡ ( $\bar{l}$ ), are transliterated using a diacritic form of a Latin consonant, because they are

vocalised retroflex forms of these consonant sounds. The two vocalic retroflex ‘r’s, ठ (ṛ) and ठ̄ (ṝ), are pronounced approximately as ‘ri’ (or sometimes more like ‘ru’) and ‘rī’ respectively, while the two vocalic retroflex ‘l’s, ढ (ḷ) and ढ̄ (ḹ), are pronounced approximately as ‘lri’ (somewhat like ‘lry’ in ‘revelry’) and ‘lrī’ respectively.

Of these four vowels, ठ (ṛ) occurs most frequently. Common examples of words in which it occurs include ठषि (ṛṣi), which is commonly transcribed as ‘rishi’, कृष्ण (kṛṣṇa), which is commonly transcribed as ‘Krishna’, and various words derived from the verb दृश् (dṛś), which means to see, look at, understand, intuit or know. ठ̄ (ṝ) occurs much less frequently (usually as the terminating vowel in a verbal root), and ढ (ḷ) still less frequently (mostly in certain forms of the verb कृप्, kṛp, which means to be well ordered, be adapted to, be suitable, correspond, happen, arrange, produce, cause or create), but ढ̄ (ḹ) does not occur in any word except ढ̄ (ḹ), which means Siva, mother or a divine female.

‘ISO 15919’ recommends that these four vowels should be transliterated with a small circle below instead of a dot below (as in IAST) in order to avoid confusion with two retroflex consonants, namely ṛa (for which the Devanagari character is ठ), which occurs in some languages such as Hindi and Panjabi, and ḷa (for which the Devanagari character is ढ), which occurs in Tamil (as ञ) and in most Indian languages other than classical Sanskrit (though it does occur in the older Vedic Sanskrit). However I transliterate them according to IAST with a dot below, because the characters ‘r’ and ‘l’ with a circle below are not yet available in any suitable font, and fortunately the Sanskrit retroflex vowel ढ (ḷ) occurs so rarely that it is unlikely to be confused in any transliteration with the muted Tamil retroflex consonant ञ (ḷ).

### Transliteration of Sanskrit consonantal diacritics

In the Sanskrit alphabet, the fourteen vowels are followed by two consonantal diacritics, the [anusvāra](#), ṁ (ṁ), and the [visarga](#), ḥ (ḥ). As I explained above, the *anusvāra* indicates that the vowel or consonant-vowel syllable above which it is placed is nasalised, so at the end of a sentence it is pronounced more or less like ‘m’, whereas before a plosive consonant it is pronounced as the corresponding nasal. The *visarga* indicates an aspiration of the preceding vowel and is therefore pronounced somewhat



like ‘h’ in English, but often followed by a slight echo of the preceding vowel (so after an ‘a’, for example, it may be pronounced ‘hă’, or after an ‘i’, ‘hī’). It seldom occurs except at the end of a sentence, because before most consonants it merges or mutates into some other letter according to the rules of [sandhi](#) or letter-conjunction.

Though the default transliteration of the *anusvāra* is *m̐* (or *m̐* in IAST), before many consonants I transliterate it as *m̐* only when there is a particular need to indicate the precise spelling of a word. That is, when it occurs at the end of a word, before a labial consonant (*p*, *ph*, *b*, *bh* or *m*) or before any non-*sparśa* consonant (*y*, *r*, *l*, *v*, *ś*, *ṣ*, *s* or *h*), I usually transliterate it as *m̐* (or sometimes simply as *m*), but when it occurs before a non-labial *sparśa* consonant (whether plosive or nasal) I usually transliterate it as the corresponding nasal, because that is how it is actually pronounced (and also how it is transcribed in Tamil script). For example, *ahamkāra* and *ahantā* (which both mean ‘ego’) are pronounced *ahamkāra* and *ahantā* respectively, *saṁcita* (which means ‘piled’ or ‘accumulated’) is pronounced *sañcita*, [samdhi](#) (which means ‘conjunction’) is pronounced [sandhi](#), and *saṁnyāsa* (which means ‘renunciation’) is pronounced *sannyāsa*. Thus before a velar consonant (*k*, *kh*, *g*, *gh* or *ṅ*) I usually transliterate the *anusvāra* as the velar nasal (*ṅ*), before a palatal consonant (*c*, *ch*, *j*, *jh* or *ñ*) as the palatal nasal (*ñ*), before a retroflex consonant (*t*, *th*, *ḍ*, *ḍh* or *ṇ*) as the retroflex nasal (*ṇ*), and before a dental consonant (*t*, *th*, *d*, *dh* or *n*) as the dental nasal (*n*).

A variant of the *anusvāra* is the [candrabindu](#) or ‘moon-dot’, ° (*m̐*), which is used in Vedic Sanskrit in place of the *anusvāra*, and which like the *anusvāra* indicates nasalization of the vowel or consonant-vowel syllable above which it is placed. The best known example of its use is in the sacred monosyllable [ॐ](#) (*ōm̐*). In IAST it is transliterated like the *anusvāra* as *m̐*, but ‘ISO 15919’ recommends that it should be transliterated as *m̐* (‘m’ with a moon-dot centred above).

### The classification of Sanskrit consonants

Like the Sanskrit vowels, the Sanskrit consonants are arranged phonetically, being divided into five groups of five [stop consonants](#), each consisting of four [plosives](#) and one [nasal](#), followed by two groups of four other consonants, the first consisting of four oral [sonorants](#) (which can be

roughly described as [approximants](#)), two of which are [semivowels](#) (*y* and *v*) and the other two of which are [liquids](#) (*r* and *l*), and the second consisting of four [fricatives](#), the first three of which are [sibilants](#) and the last of which is [glottal](#).

The five groups of five stop consonants are called *sparsā* or ‘touch’ consonants, because they are formed by the organs of utterance (either the tongue and part of the mouth, or both the lips) actually touching each other. In each of these five groups, the first four are plosives (oral stops), of which the first two are unvoiced (or ‘voiceless’) while the second two are [voiced](#), and in each of these two pairs the first is unaspirated while the second is [aspirated](#) (that is, pronounced with *mahāprāṇa* or a full breath, like ‘h’ in the English word ‘house’, and indicated in transliteration by an ‘h’ following the lead consonant). The fifth member of each of these five groups is a voiced and unaspirated nasal.

Thus for example in the first group, which are the [velar](#) or ‘throat’ (*kaṅṭhya*) stops, क (*ka*) is unvoiced and unaspirated, ख (*kha*) is unvoiced but aspirated, ग (*ga*) is voiced but unaspirated, घ (*gha*) is voiced and aspirated, and ङ (*ṅa*) is the velar nasal.

Each of the other four groups of *sparsā* consonants follows this same pattern. The second group are the [palatal](#) (*tālavya*) stops, च (*ca*), छ (*cha*), ज (*ja*), झ (*jha*) and ञ (*ṅa*). The third group are the [retroflex](#) or ‘cerebral’ (*mūrdhanya*) stops, ट (*ṭa*), ठ (*ṭha*), ड (*ḍa*), ढ (*ḍha*) and ण (*ṇa*). The fourth group are the [dental](#) (*dantya*) stops, त (*ta*), थ (*tha*), द (*da*), ध (*dha*) and न (*na*). The fifth group are the [labial](#) (*ōṣṭhya*) stops, प (*pa*), फ (*pha*), ब (*ba*), भ (*bha*) and म (*ma*).

In the Sanskrit alphabet, these five groups of *sparsā* consonants are followed by two other groups, each consisting of four consonants, namely the four oral sonorants or [approximants](#) (called the *antastha* or ‘standing between’ consonants), य (*ya*), र (*ra*), ल (*la*) and व (*va*), and the four [fricatives](#) (called the *ūṣma* or ‘hot’ consonants), which are the three sibilants, श (*śa*), ष (*ṣa*) and स (*sa*), and the glottal fricative, ह (*ha*).

Each of these final eight consonants are further classified along with one of the five groups of *sparsā* consonants according to their [place of articulation](#), as follows:

- Since ह (*ha*) is articulated with the [glottis](#) in the throat, it is classified

along with the first group of *sparśa* consonants, the five velar stops, क (*ka*), ख (*kha*), ग (*ga*), घ (*gha*) and ङ (*ṅa*), which are articulated with the tongue touching the soft palate at the back of the mouth, and together they are called the *kaṅṭhya* or ‘throat’ consonants.

- Since य (*ya*) and श (*śa*) are articulated with the body of the tongue close to the hard palate, they are classified along with the second group of *sparśa* consonants, the five palatal stops, च (*ca*), छ (*cha*), ज (*ja*), झ (*jha*) and ञ (*ña*), which are articulated with the body of the tongue touching the hard palate, and together they are called the *tālavya* or ‘palatal’ consonants.
- Since र (*ra*) and ष (*ṣa*) are articulated with the tip of the tongue curled back close to the roof of the mouth, they are classified along with the third group of *sparśa* consonants, the five retroflex stops, ट (*ṭa*), ठ (*ṭha*), ड (*ḍa*), ढ (*ḍha*) and ण (*ṇa*), which are articulated with the tip of the tongue curled back so that it touches the roof of the mouth, and together they are called the *mūrdhanya* or ‘cerebral’ consonants.
- Since ल (*la*) and स (*sa*) are articulated with the tongue close to the upper teeth, they are classified along with the fourth group of *sparśa* consonants, the five dental stops, त (*ta*), थ (*tha*), द (*da*), ध (*dha*) and न (*na*), which are articulated with the tongue touching the upper teeth, and together they are called the *dantya* or ‘dental’ consonants.
- Since व (*va*) is articulated with the lower lip close to the upper teeth, it is classified along with the fifth group of *sparśa* consonants, the five bilabial stops, प (*pa*), फ (*pha*), ब (*ba*), भ (*bha*) and म (*ma*), which are articulated with both lips touching each other, and together they are called the *ōṣṭhya* or labial consonants.