

TamilNet Transcription: version 1.0

Content:

CONTENT:	1
1.0 INTRODUCTION:	2
2.0 GENERAL INSTRUCTIONS TO USERS:	3
2.1 TYPING TEXT IN TAMIL	3
2.2 AUTOMATED TRANSCRIPTION FOR THE TAMIL TEXT:	3
2.3 TYPING TAMIL TEXT IN ROMAN SCRIPT:	3
2.4 AUTOMATION OF THE TRANSCRIPTION TO TAMIL SCRIPT:	3
2.5 AUTOMATION TO MADRAS TAMIL LEXICON TRANSLITERATION	4
3.0 TRANSCRIPTION TABLES	5
3.1 TRANSCRIPTION AND AUTOMATION KEY: TAMIL TO ROMAN	5
3.1.1 Vowels	5
3.1.2 Aaytham	6
3.1.3 Consonants	6
3.1.3.1 Hard Consonants	6
3.1.3.2 Soft Consonants.....	8
3.1.3.3 Medium Consonants	9
3.1.4 Grantha characters	10
3.2 TRANSCRIPTION AND AUTOMATION KEY: ROMAN TO TAMIL	11
3.2.1 Vowels:.....	11
3.2.2 Aaytham	11
3.2.3 Consonants	12
3.2.3.1 Consonants: hard	12
3.2.3.2 Consonants: soft.....	13
3.2.3.3 Consonants: medium.....	13
3.2.4 Grantha characters	14
4.0 MADRAS TAMIL LEXICON TRANSLITERATION SYSTEM	15
4.1 VOWELS.....	15
4.2 AAYTHAM.....	15
4.3 CONSONANTS	15
4.3.1 Consonants: hard	15
4.3.2 Consonants: soft.....	16
4.3.3 Consonants: medium.....	16
4.4 GRANTHA CHARACTERS.....	16
5.0 DOCUMENT HISTORY	16

1.0 Introduction:

- A simple transcription system to present Tamil text in Roman script, especially for the users of Tamil/English languages.
- Two-way computer automation of transcription: from text in Tamil script to transcription in Roman script and vice versa.
- Linking the system with conventional Tamil-Roman transliteration.
- Facilitating storage and retrieval of information.
- Enhancement of the use of Tamil language in electronic communication.

2.0 General Instructions to Users:

2.1 Typing text in Tamil

- For the time being, use Bamini font for typing the text in Tamil.
- Those who don't have Bamini font, can download it from the Internet.
- Please note that Bamini font may not show certain unused or rarely used syllables such as **ங் னு னூ க்ஷ** (ngu, gnu, gnoo, ksh) etc.,
- The TamilNet Transcription will eventually be developed to adopt Unicode and other widely used character encodings for input and retrieval.
- When automated to Roman script, the system is designed to capitalize the first letter after every full stop followed by a space. If not needed, you can deactivate it.
- When there is a necessity to capitalize any letter in any part of the text, type a full stop immediately before that letter in the Tamil text. Such full stops not followed by a space will not be shown in the automation.

2.2 Automated transcription for the Tamil text:

- Any text typed in Tamil can be automated to produce a transcription in Roman script.
- The Transcription is designed primarily for English readers.
- Users are requested to consult the tables for familiarizing themselves with the phonology of the transcription.

2.3 Typing Tamil text in Roman script:

- Any normal computer keyboard in Roman script can be used.
- Users are requested to familiarize themselves with the transcription tables before typing.
- If you wish to present an English word or passage in the Tamil text, separate that English word or passage by typing @@ (two at signs) on both sides of the text. Example: Aanaiyi'ravu @(Elephant Pass)@ paravaik kadal. When automated to Tamil, the text in between the signs will be retained in English. The signs will not be shown in the automation. Example: ஆனையிறவு (Elephant Pass) பரவைக் கடல்.

2.4 Automation of the transcription to Tamil script:

- The English transcription in Roman script that is typed according to the TamilNet system can always be automated to Tamil script.

- At present the automated Tamil text is shown in Bamini as well as in Unicode fonts. It will eventually be extended to TSCII (Tamil Standard Code for Information Interchange).
- Both Bamini and Unicode fonts can be downloaded from the Internet.
- Most of the Operating Systems such as Windows, MacOS and Linux support Unicode by default. However, you may need to customise the settings for Unicode character display and ensure that a Unicode enabled web-browser is used and a Unicode font with Tamil glyphs is installed in your system.

2.5 Automation to Madras Tamil Lexicon Transliteration

- Both the text in Tamil script and the transcription in Roman script, can be automated to produce the Madras Tamil Lexicon transliteration.
- The Transliteration thus produced follows the version of the text in Tamil script.

3.0 Transcription Tables

3.1 Transcription and automation key: Tamil to Roman

3.1.1 Vowels

உயிர்/uyir

Tamil script	Roman Alphabet	Pronunciation guide	Tamil word	Transcription
அ	a	as <i>u</i> in <i>nun</i> . near to <i>a</i> in <i>apple</i>	அறிவு அரிசி	a'rivu arisi
ஆ	aa	as <i>a</i> in <i>arm</i>	ஆடு	aadu
இ	i	as <i>i</i> in <i>ink</i> as <i>i</i> in <i>imp</i>	இலை இழு	ilai izhu
ஈ	ee	as <i>ee</i> in <i>see</i>	ஈரம்	eeram
உ	u	as <i>u</i> in <i>put</i> . Half <i>u</i> is possible at certain instances of the end of words	உறவு ஆறு	u'ravu aa'ru
ஊ	oo	as <i>oo</i> in <i>loop</i>	ஊர்	oor
எ	e	as <i>e</i> in <i>enter</i> . as <i>a</i> in <i>among</i>	எச்சம் எறும்பு	echcham e'rumpu
ஏ	ea	as <i>ea</i> in <i>early</i> . In certain instances, elongated sound of <i>e</i> or <i>a</i> as in <i>end</i> , <i>angel</i> etc., is possible. There is no exact equivalent in English pronunciation.	ஏணி ஏன்	ea'ni ean
ஐ	ai	as <i>i</i> in <i>ivory</i>	ஐவர்	aivar
ஓ	o	as <i>o</i> in <i>on</i>	ஓட்டகம்	oddakam
ஔ	oa	as <i>oa</i> in <i>loan</i>	ஓட்டம்	oaddam
ஔ	au	as <i>ow</i> in <i>cow</i>	ஔவை	auvai

3.1.2 Aaytham

ஆய்தம்

Tamil script	Roman Alphabet	Pronunciation guide	Tamil word	Transcription
ஃ	:h	near to consonant <i>h</i>	எஃகு	e:hku

3.1.3 Consonants

மெய்/mey

3.1.3.1 Hard Consonants

வல்லினம் / vallinam

Tamil script	Roman Alphabet	Pronunciation guide	Tamil word	Transcription
க்	k	as <i>c</i> in <i>cat</i> . Pronunciation may shift towards <i>g</i> or <i>h</i> in certain instances. This is not differentiated in transcription due to ambiguity and dialectal differences (Ex: <i>akam: aham/agam; Thangkam/ Thanggam</i>).	கல் பக்கம் அகம் தங்கம்	kal pakkam akam thangkam
ச்	ch/s	as <i>ch</i> in <i>chin</i> . Pronunciation may shift to <i>s</i> in certain instances which will be shown in the transcription. Shift to <i>s</i> in the beginning of words and shift towards <i>j</i> when <i>ch</i> follows <i>gn</i> are not transcribed due to ambiguity and dialectal differences. Ex.: <i>cheppu /seeppu</i> and <i>vagnchi / vagnji</i> .	சரி அச்சம் அரசன் வம்சம் சீப்பு வஞ்சி	chari achcham arasan vamsam cheppu vagnchi
ட்	d	as <i>d</i> in <i>did</i> or near to <i>tt</i> in <i>putt</i>	படம் பட்டம்	padam paddam
த்	th	near to <i>th</i> in <i>thin</i> . Pronunciation may shift towards <i>dh</i> (<i>ð</i>)	தயிர் சத்தம்	thayir chaththam

		<p>as <i>th</i> in <i>then</i> in certain instances.</p> <p>This is not differentiated in transcription due to ambiguity and dialectal differences.</p> <p>Ex.: <i>aanthai</i> / <i>aandhai</i>; <i>tharmam</i> / <i>dharmam</i>; <i>patham</i>/<i>padham</i>)</p>	<p>ஆந்தை தர்மம் பதம்</p>	<p>aanthai tharmam patham</p>
ப்	p	<p>as <i>p</i> in <i>put</i>.</p> <p>The pronunciation shift towards <i>b</i> in certain instances is not transcribed due to ambiguity and dialectal differences.</p> <p>Ex: <i>ampalam</i>/<i>ambalam</i> and <i>payam</i>/<i>bayam</i>.</p>	<p>பணம் அப்பா அம்பலம் பயம்</p>	<p>pa'nam appaa ampalam payam</p>
ற்	'r / t	<p>near to <i>r</i> in <i>rat</i>.</p> <p>This is a retroflex <i>r</i>, produced by the tip of the tongue touching the deeper part of the palate in a raised position.</p> <p>When doubled and when occurs as a single consonant the pronunciation becomes similar to <i>t</i>, which will be shown in the transcription.</p>	<p>பறவை குற்றம் இயற்கை</p>	<p>pa'ravai kuttam iyatkai</p>

3.1.3.2 Soft Consonants

மெல்லினம் / mellinam

Tamil script	Roman Alphabet	Pronunciation guide	Tamil word	Transcription
ங்	ng	as <i>ng</i> in <i>ring</i>	பங்கு	pangku
ஞ்	gn	as <i>ñ</i> in spanish <i>señor</i>	ஞாயிறு பஞ்சு	gnaayi'ru pagnchu
ண்	'n	Retroflex sound of <i>n</i> . Near to <i>ne</i> in <i>one</i>	பணம் வண்ணம்	pa'nam va'n'nam
ந்	n (Tamil -> Roman) :n (Roman -> Tamil)	as <i>n</i> in <i>no</i> . As a single consonant, when followed by <i>th</i> , it is pronounced similar to <i>n</i> in month. This is not differentiated in transcription.	நிலம் பந்து	nilam panthu
ம்	m	as <i>m</i> in <i>man</i>	மதி அம்மி	mathi ammi
ன்	n	pronunciation is similar to <i>n</i> in <i>sand</i> . The subtle difference in pronunciation between ன் and ந், outlined in the grammatical texts is not shown in the transcription.	மனம் வன்னி	manam vanni

3.1.3.3 Medium Consonants

இடையினம் / idaiyinam

Tamil script	Roman Alphabet	Pronunciation guide	Tamil word	Transcription
ய்	y	as <i>y</i> in <i>yard</i>	வயல் பாய்	vayal paay
ர்	r/t	The palatal sound of <i>r</i> , produced by the tip of the tongue stroking the front part of the palate. No exact equivalent in English. In certain instances, when preceded by <i>n</i> , it may be pronounced as <i>t</i> .	நீர் மரம் என்றை	neer maram entai
ல்	l	as <i>l</i> in <i>love</i>	இல்லம்	illam
வ்	v	as <i>v</i> in <i>voice</i>	வள்ளம் வெளவால்	va'llam vauvaal
ழ்	zh	Retroflex sound of <i>l</i> , produced by the tip of the tongue touching the deeper part of the palate in a raised position. No equivalent in English or in Indian languages other than Tamil and Malayalam.	பழம் வாழ்க்கை தமிழ்	pazham vaazhkkai Thamizh
ள்	l'	Palatal sound of <i>l</i> , produced by the tip of the tongue, gently touching the palate. Near to <i>l</i> in <i>temple</i>	குளம் பள்ளம்	ku'lam pa'llam

3.1.4 Granta characters

கிரந்த எழுத்துக்கள்

Tamil script	Roman Alphabet	Pronunciation guide	Tamil word	Transcription
ஜ்	j	as <i>j</i> in <i>June</i>	ஜெயம்	jeyam
ஸ்	s (Tamil -> Roman) 's (Roman -> Tamil)	as <i>s</i> in <i>sing</i>	வஸ்து	vasthu
ஷ்	sh	as <i>sh</i> in <i>crash</i>	புஷ்பம்	pushpam
ஹ்	h (Tamil -> Roman) 'h (Roman -> Tamil)	as <i>h</i> in <i>hat</i>	ஹரி	hari
க்ஷ்	ksh	combination of <i>k</i> and <i>sh</i>	லக்ஷ்மி	lakshmi
ஸ்ரீ	s'ree (Tamil -> Roman) s'ree/sree (Roman -> Tamil)	combination of palatal 's', r and ee	ஸ்ரீல ஸ்ரீ	s'reela s'ree

3.2 Transcription and automation key: Roman to Tamil

3.2.1 Vowels:

Roman	Tamil
a	அ
aa	ஆ
i	இ
ee	ஈ
u	உ
oo	ஊ
e	எ
ea	ஏ
ai	ஐ
o	ஓ
oa	ஔ
au	ஔ

3.2.2 Aaytham

Roman	Tamil	Notes
:h	ஃ	Used rarely in classical Tamil.

There is a feature called *a'lapedai*, for elongation of vowel sounds in classical Tamil. *Aadoou* (ஆஊஉ) has to be read with an elongation of sound for *oo*. The elongation is equivalent to *u*, means the total length of sound is equivalent to three *u*'s. Note this feature when classical texts are involved in transcription.

3.2.3 Consonants

3.2.3.1 Consonants: hard

Roman	Tamil	Notes
k kh g gh h	க	<p>Any of the listed Roman characters can be typed to produce க (<i>k</i>) in Tamil.</p> <p>In the beginning of words and when occurs as a single consonant, h will be automated to ஹ (<i>h</i>). In other instances, it will become க (<i>k</i>).</p> <p>However, when typed in Tamil, க (<i>k</i>) will always be automated to <i>k</i> in Roman.</p>
ch s	ச	<p><i>ch</i> will always be automated to ச (<i>ch</i>).</p> <p><i>s</i> will be automated to ஸ (<i>s</i>) in the beginning of words and when occurs as a single consonant. In other instances, it will be automated to ச (<i>ch</i>).</p> <p>When typing in Tamil, ச (<i>ch</i>) will be automated to <i>s</i> in Roman when it occurs between two vowels and when it follows soft or medium consonants. In other instances, it will be automated to <i>ch</i> in Roman.</p>
d	ட	
th dh	த	<p>Both can be automated to the Tamil character த (<i>th</i>).</p> <p>However, when typing in Tamil, த (<i>th</i>) will always be automated to <i>th</i> in Roman.</p>
p ph b bh	ப	<p>Any of the Roman characters listed can be typed for the automation to ப (<i>p</i>).</p> <p>While typing in Tamil, ப (<i>p</i>) will always be automated to <i>p</i> in Roman.</p>
'r t	ர	<p>With combinations of vowel, 'r will be automated to ர (<i>ra</i>) etc.</p> <p>When occurs as a single consonant and when doubled, <i>t</i> can be typed to produce ர.</p> <p>The vowel consonant ர (<i>ra</i>) etc., following an ர in doubled instances can be produced by typing <i>tt</i>.</p> <p>Even an 'r or 'r'r will be automated to ர or ரர etc., by default.</p> <p>When typing in Tamil, ர (<i>ra</i>) etc., will be automated to 'r in Roman. But, ர ('r) and ரர etc., will be automated to <i>t</i> and <i>tt</i> in Roman.</p>

3.2.3.2 Consonants: soft

ng	ங்	
gn	ஞ்	
'n	ண்	
n :n	ந்	<p>When preceds த் (<i>th</i>) and in the beginning of words, <i>n</i> will be automated to ந், ந etc.</p> <p>In other instances it will become ன்.</p> <p>In combinations of words such as வெந்தீர் and in certain instances of sanskrit words and personal names such as அந்நியர். விபுலாநந்தர் etc., ந் may occur in the middle of words. In such instances, users are requested to type :n to produce the Tamil script ந்.</p> <p>While typing in Tamil, both ந் and ன் are automated to <i>n</i> in Roman. They are not differentiated in the transcription.</p>
m	ம்	
n	ன்	<p>Except in the beginning of words and in combinations followed by th, <i>n</i> will be automated to ன். The said circumstances don't occur involving ன் in Tamil language. Therefore, ன் is safely automated from Roman <i>n</i> to Tamil ன் (see ந்).</p> <p>When typing in Tamil, ன் and ந் are automated to <i>n</i> in Roman.</p>

3.2.3.3 Consonants: medium

y	ய்	
r t	ர்	<p><i>r</i> will always be automated to ிர்.</p> <p>There are instances where ிர் (<i>r</i>) after ன் (<i>n</i>), is pronounced <i>t</i>. In words, such as அவன்ரை, என்ரை (avantai, entai) etc., <i>t</i> can be typed to automate ிர்.</p> <p>When typing in Tamil, ிர் can be automated to <i>r</i> in Roman.</p> <p>In instances where vowel consonants of ிர் follow ன், the automation will be <i>t</i> in Roman.</p>
l	ல்	
v w	வ்	<p>Both can be automated to வ in Tamil.</p> <p>However, வ will always be automated to <i>v</i> in Roman.</p>
zh	ழ்	
'l	ள்	

3.2.4 Grantha characters

j	ஜ்	
's s	ஸ்	<p>In the beginning of words and when occurs as a single consonant <i>s</i> will be automated to ஸ் in Tamil.</p> <p>In other instances, <i>s</i> will produce ஃ (<i>ch</i>) in automation.</p> <p>'s can be typed to produce ஸ் on any occasion.</p> <p>When typing in Tamil, ஸ் will always be automated to <i>s</i> in Roman.</p>
sh	ஷ்	
'h h	ஹ்	<p>In the beginning of words and when occurs as a single consonant, <i>h</i> will be automated to ஹ் in Tamil.</p> <p>In other instances <i>h</i> will be automated to க் (<i>k</i>).</p> <p>Typing '<i>h</i> will produce ஹ் on all occasions.</p> <p>When typing in Tamil ஹ் will always be automated to <i>h</i>.</p>
ksh	கஷ்	
'ree sree	ரீ	<p>As a combination of letters, both <i>s'ree</i> as well as <i>sree</i>, can be automated to ரீ.</p> <p>However, when typed in Tamil, ரீ can only be automated to <i>s'ree</i>.</p>

4.0 Madras Tamil Lexicon Transliteration System

4.1 Vowels

Tamil alphabet	Transliteration in Roman
அ	a
ஆ	ā
இ	i
ஈ	ī
உ	u
ஊ	ū
எ	e
ஏ	ē
ஐ	ai
ஓ	o
ஔ	ō
ஔள	au

4.2 Aaytham

Tamil alphabet	Transliteration in Roman
ஃ	ḷ

4.3 Consonants

4.3.1 Consonants: hard

Tamil alphabet	Transliteration in Roman
க	k
ச	c
ட	ṭ
த	t
ப	p
ற	ṛ

4.3.2 Consonants: soft

Tamil alphabet	Transliteration in Roman
ங்	ṅ
ஞ்	ñ
ண்	ṇ
ந்	n
ம்	m
ன்	ṅ

4.3.3 Consonants: medium

Tamil alphabet	Transliteration in Roman
ய்	y
ர்	r
ல்	l
வ்	v
ழ்	ḷ
ள்	ḷ

4.4 Grantha characters

ஜ்	j
ஸ்	s
ஷ்	ṣ
ஹ்	h
கஷ்	kṣ
ஸ்ரீ	sri

5.0 Document History

Date	Version	Comment
April 28, 2007	Beta 1	Trial
June 01, 2007	1.0	Improved and modified after feedbacks
June 02, 2007	1.0	Typo in A'risi. Changed to Arisi in table 3.1.1