
Tamil Linux HOWTO

V. Venkataramanan <venkat@tamillinux.org>

Initial conversion from LaTeX to Docbook XML: D Sivaraj

Copyright © 2002, 2003 V. Venkataramanan

Jan 2003

	Revision History	
Revision 1.0	2003-02-14	venkat
	Initial release, reviewed by LDP	
Revision 0.9	2003-1-21	venkat
	Changes made to comply to TDLP specs.	
Revision 0.8	2002-10-24	venkat
	First draft	

Abstract

This document will help set up a working Tamil Linux environment. This describes setting up fonts, keyboard drivers, editing and printing Tamil/bilingual documents, and working with the X Window system. The information is kept as generic as possible. When it pertains to a specific distribution (say RedHat or Debian), it is explicitly noted.

Table of Contents

About this HOWTO	2
Purpose/Scope of this HOWTO	2
Feedback	2
Copyright and License	2
Acknowledgements	2
Introduction	2
Fonts	3
TSCII	3
TAB	3
Miscellaneous fonts and encodings	3
Console Tamil	4
X Window	4
Installing fonts	4
Bitmapped fonts	4
TrueType fonts	5
Other Font Servers	7
Keyboard Drivers	7
tamil_kmap	8
tamilvp	8
KDE	8
Getting Localization Files	8
Choosing a Tamil locale	9
Choosing Tamil fonts for GUI	9
KDE Miscellaneous	9
GNOME	9
Printing	10
L ^A T _E X	10
Postscript	10

PDF	11
Word Processors, Office Packages	11
Viewing Web pages	11
Pango	11
Miscellaneous	11
A. Appendix of Tamil Font Encodings	12

About this HOWTO

Purpose/Scope of this HOWTO

This document will help set up a working Tamil Linux environment. Step-by-step instructions are provided for setting up fonts, editors, etc. This document also describes the essential instructions need to use web browsers, edit documents and print them.

The base URL of this document is: [<http://www.tamilinux.org/doc/>]

Feedback

Comments and suggestions about this document may be directed to the author (<venkat@tamilinux.org>)

Copyright and License

© 2002, 2003 V.Venkataramanan.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no invariant Front-Cover text, no Back-Cover text and no invariant sections.

A verbatim copy of the license can be obtained from the Free Software Foundation Website at <http://www.gnu.org/licenses/fdl.html>

Acknowledgements

Several postings by the following people were useful in writing this document. The following people are thanked for all their help:

Thuraiappah Vaseeharan, D. Sivaraj, Sivakumar Shanmugasundaram, Dinesh Nadarajah, Anbumani Subramanian, Ganesan Rajagopal, M.K. Saravanan,...

Introduction

Tamil is a member of the Dravidian languages. Its origin is in southern India and the language is written with non-Roman alphabets. So there is a need for special fonts, encoding, keyboard layout and drivers, besides localization, including currency, date format, etc. This document will give a complete overview of setting up and working in the Tamil Linux environment. There are several pieces of information and tools available for Linux in Tamil; this how-to will serve as a meta-index to all the scattered resources.

A word before you enter - most of the fonts, tools, RPMs and documents are being gathered under one site. So try the resources at <http://tamil.homelinux.org> before you embark on treasure-hunting.

Fonts

It can seem like anarchy. There are an unknown number of fonts, each encoded with their own tables, driven by arbitrary keyboard layouts and outputs. In my opinion, Tamil can seriously compete with any other language for maximum number of font tables. Added to this commotion are the dynamic fonts for the web pages, that enable anyone to get away with a non-standard font as long as his pages are viewable.

Adding to all these is the official Indian Standard Code for Information Interchange (ISCII), the Government of India sponsored “unifying” scheme to bring all Indian fonts under the Devanagari umbrella. Anyone familiar with the way the characters are written in Tamil and in Devanagari script will understand the lack of any rationale in this approach.

Needless to say, this is serving to only add to the confusion. A good analysis of this and the unicode for Tamil is once again written by Sivaraj and can be found at [http://www.tamil.net/people/sivaraj/tamil_unicode.html]. For those not familiar with the Tamil script, a good introduction written by Sivaraj is at [<http://www.tamil.net/people/sivaraj/write-tamil.html>].

Let us ignore the anarchy for a moment and get a picture of the frequently used font encodings. There are two main contenders and luckily they will converge soon. The first and most popular one is the Tamil Standard Code for Information Interchange (TSCII), developed by volunteers throughout the world, and the other, TAmil Monolingual (TAM), and TAmil Bilingual (TAB) encodings, were proposed by the Tamil Nadu Government. Once again, TAM is of limited use in an OS environment and we can safely ignore that. Almost all Linux efforts are in TSCII (Console, KDE, GNOME localizations).

TSCII

TSCII is a glyph-based, 8-bit bilingual encoding. It uses a unique set of glyphs; the usual lower ASCII set. Roman letters with standard punctuation marks occupy the first 128 slots and the Tamil glyphs occupy the *upper ASCII* segment (slots 128-256). A good overview of the early font encoding schemes and a the rationale behind the TSCII approach can be found at <http://www.geocities.com/Athens/5180/tscii.html>.

The home URL for TSCII volunteers is <http://www.tamil.net/tscii>. This site discusses the TSCII encoding and provides tools including fonts, keyboard drivers, editors and inter-conversion tools for various platforms. The font encoding table according to TSCII-1.6 can be found at <http://www.tamil.net/tscii/charset16.gif>.

The current version of TSCII is 1.6, and a revision is expected anytime now that will fix some anomalies in using various slots for encoding. This version 1.7 will be fully backward compatible with 1.6 and is expected to gain popularity. The TSCII discussion group [<http://groups.yahoo.com/group/tscii>] currently brainstorms on modifications to TSCII-1.6. You may be able to participate in the discussions by becoming a member. You may also be able to download various beta tools from there. The font encoding table according to TSCII-1.7 (draft) can be found at [http://groups.yahoo.com/group/tscii/files/charset17_a.gif].

TAB

TAB is a character based bilingual standard proposed by the government of Tamil Nadu. The TAB bilingual encoding table can be found at <http://www.tamilnet99.org/annex4.htm>. Tools for TAB encoding (mostly restricted to the Windows platform) can also be downloaded in the vicinity of this page.

Miscellaneous fonts and encodings

There are too many types, and unfortunately they are not documented well. It is beyond the scope of this document to discuss them.

Console Tamil

This so far has been a one man effort - once again by Sivaraj. He has written a set of console tools for Tamil that include a monospace font, keyboard driver and locale setup. In his words:

You can use it with Lynx to read any TSCII-based web sites or Pico to email in TSCII. Some characters may be disoriented, since I try to fit all the characters in an 8x16 cell. But it is still readable.

—Sivaraj

The tools can be downloaded here [<http://www.tamil.net/people/sivaraj/>]. Follow the instructions in the REAME file to install and use.

X Window

Welcome! This is where you will find the most useful tools for Tamil. Even for basic users, it is now possible to have close to a total Tamil-localized office suite. Tamil GUI is achieved in KDE or GNOME environment with localization settings (more about this later in this document), and Tamil character input is achieved using keymanager programs. But first you need to get some fonts to do all this.

Installing fonts

Linux, by default, uses “pcf” fonts and one can also use “bsd” fonts; these are bitmapped fonts that display under X and can be printed. But, as is common with all bitmapped fonts, these are not always WYSIWYG in print. For high-quality printing you need “Type-I” fonts (Adobe), with Ghostscript you need PS fonts and for “afm” fonts (American Font metrics) are used. But most of the Tamil fonts that are freely available are TrueType (ttf). We will see next how to get all these fonts working.

Bitmapped fonts

A bitmapped font is a matrix of dots; because of this, these fonts are device-independent. A 75 dpi font, which is good enough for displaying, is still a 75 dpi font in your 1200 dpi printer. So usually bitmapped fonts are created for a specific purpose, such as for displaying on a monitor or for printing. Linux usually uses `bd`f or `pcf` font for console or X display. Fonts like those created by `dvips` or `dvi` are printer-related bitmapped fonts. These fonts occupy large sizes, but programs circumvent this by dynamically creating them as and when they are needed, and at a specific resolution.

You can get bitmapped Tamil fonts for various applications from: [<http://www.tamil.net/tscii/tscii-tools/tsciifonts.tar.Z>]

When an application makes a font request to the X Server, XFree86 looks for fonts in specific directories. This means that when you add fonts to your system and you want them to be recognized by X Server, you need to tell X about the location of these fonts. Simply add a directory to your font path with the commands:

```
mkfontdir
xset fp+ <directory>
```

where the family directory is the name of the directory where you have fonts. Once you have done this you have to ask the server to get this registered for the session, with the command

xset fp rehash

Since you will want these commands to run automatically, you should put them in your `.xinitrc` file (or possibly your `.Xclients` or `.xsession` file -- this depends on how you start X. Another way to have the commands set automatically is edit `XF86Config`. For example, to add `/usr/share/fonts/myfonts` to the font path when X is started, edit `XF86Config` like this:

```
...
Section "Files"
...
FontPath /usr/share/fonts/myfonts
...
EndSection
...
```

The advantage of editing `XF86Config` is that the resulting changes are system wide.

TrueType fonts

You may get TrueType fonts for TSCII, TAB and TSCII.7 encoding from the download section of <http://tamil.homelinux.org/>. Alternate sources for these fonts are

TSCII - <http://www.tamil.net/tscii/>

TAB - <http://www.tamilnet99.org/> and <http://www.thinnai.com>

TSCII-1.7 (experimental) - <http://groups.yahoo.com/group/tscii/files/>

Installing these fonts are either too easy or too difficult. Too easy if you have one of the latest distributions, like RedHat7.x or Mandrake7.x. This is because RedHat (and Mandrake, maybe SuSE) come with xfs pre-packaged. It is also easy to find xfs for Debian, but as far as I know, Debian does not come with xfs packaged.

Debian users are now redirected to this mini-howto on TrueType fonts in Debian - <http://www.linuxdoc.org/HOWTO/mini/TT-Debian-3.html>

There is also another utility, `xfstt`, which is easier to install and use, but xfs is becoming popular as it can handle Adobe Type1 in addition to TrueType fonts.

If you do not have either of these, consider getting either xfs (not to be confused with Silicon Graphics (SGI) sponsored XFS journaling file system) from <http://www.xfree86.org>.

or `xfstt` from <http://www.dcs.ed.ac.uk/home/jec/programs/xfstt/>. You may also get `xfstt` binaries from <http://independence.seul.org/>, or reading an article about `xfstt` in the Linux Gazette at [<http://www.linuxgazette.com/issue28/ayers1.html>]

Installing TrueType Fonts

You need to run these commands as root. If you are currently logged in as a normal user, you can use `su` to do this now.

You should now have xfs availability, otherwise use the steps in the previous section to obtain it.

In some distributions like Mandrake, installing TrueType fonts is a cakewalk. Just go to `DrakConf` and use the font install utility - follow a few easy steps there and you'll have them all.

Put your TrueType fonts in whatever directory you want. For example, `/usr/share/tamiltt`.

From within the directory containing your new fonts, type:

ttmkfdir -m 50 -o fonts.scale

This makes a file that will contain the necessary information about the fonts for the xfs server. The option `-m 50` specifies the magnification for the fonts; I have seen some Tamil fonts working well only with `-m 100`.

Then type:

mkfontdir

Now you can add the new directory to your xfs search path. Red Hat (and Red Hat-like) distributions come with a neat utility to do this called **chkfontpath**. Run `chkfontpath` like this:

chkfontpath --add /usr/share/tamiltt

This will add the new font directory to your font path.

(Other users, who have an xfs font server, without `ttf` support, can do this by editing their xfs configuration file.

If xfs is already installed on your system, you should see which port it is running on. You can do this with the following command:

ps ax grep xfs

Then check your XFree86 font path with this command:

xset -q

If your font path includes something like “`unix:/port number`,” where port number is the port on which the server is running, then you already have xfs set up properly. Otherwise, you should add it to your XFree86 font path with these commands:

xset fp+ <unix:/port number>

xset fp rehash

Note

The port number is a numerical value, something like 7100.

You can add the fontpath permanently by editing your `.xinitrc`. To add it system-wide, edit your `XF86Config` file (either under `/etc/X11/XF86Config`, `/etc/X11/XF86Config-4`, `/etc/XF86Config`, or `/usr/X11R6/lib/X11/XF86Config`), by adding the following line to the `Files` section:

FontPath "unix:/port number"

Here is an example of how it should look:

```
...
Section Files
...

FontPath "unix/:-1"
...
```

EndSection

...

If xfs is already properly installed, then you can restart it like this as root:

service xfs restart

After restarting xfs, it is a good idea to restart your X session.

As most of the users in Tamil will be doing this, let me summarize the essential steps.

1. Become root.
2. Download and copy some `ttf` fonts into a directory (say `/usr/share/fonts/tamiltt`).
3. Go to that directory and do a **`ttmkfdir -m 50 -o fonts.scale`** (use the `-m 100` option if your fonts do not budge).
4. Do a **`mkfontdir .`** (Notice that you need to specify the directory either absolutely or with a dot).
5. Do a **`chkfontpath --add /usr/share/fonts/tamiltt .`** (Remember this command is available only in Red Hat-like distributions. If you can run this successfully, skip the remaining steps and restart the X server).
6. Do **`ps ax | grep xfs`** and get the xfs port known.
7. Check your font path: **`xset -q`**

If your font path includes something like “unix:/port number”, (something like “unix: 7100”), add this to your xfont path:

`xset fp+ unix: port number`

`xset fp rehash`

8. It is a good idea to restart the X Server.
9. If everything works fine, update your `.xinitrc` file, wherever it is.
10. Have fun!

Other Font Servers

There is another project, X-TrueType Server, worth looking into, at <http://www.io.com/~kazushi/xtt/>.

Another interesting project with broader scope is FreeType; check <http://www.freetype.org>.

I personally feel xfs is a great utility; it can handle Type1 fonts (very useful if you use programs like GIMP). Besides, a stand alone xfs server is not attached to X server. This means that you can deliver these fonts for remote X displays. I use this feature extensively with VNC Server running in my host and VNC Viewer running locally in Windows. It's something of a luxury having a Tamil Linux desktop while working for my employer.

Keyboard Drivers

Once again, lack of standards shows up here. There are quite a few Tamil keyboard layouts, the traditional typewriter keyboard; then with the surge of internet arrived the romanized transliteration keyboards; later

the TAmil-Nadu government played its part by prescribing a tamilnet99 keyboard. These are only a few to talk about; we have a few others which do not fall into any of these “standards.”

There are two Tamil keyboard drivers for the X Window System, both of them set to tamilnet99 standards (see tamilnet99 website [<http://www.tamilnet99.org>] for the details on the keymap). You will be able to download both the keydrivers from the Yahoo! tamilnix group files section [<http://groups.yahoo.com/group/tamilnix/files>].

tamil_kmap

The first driver is *tamil_kmap*, created by Vasee. It is based on the original version of Siva. It is operable under both TSCII 1.6 and TAB encodings. The detailed installation instructions are given in the README file in the package. It is very simple to install. First, untar the package into a temporary directory. Then type:

```
cp ta /usr/X11R6/lib/X11/xkb/symbols/
```

```
then: cp Compose /usr/X11R6/lib/X11/locale/iso8859-1
```

and put the shell script `setkb` into a directory on your system `PATH`. You may need to become **root** to copy these files into these directories.

To use the Tamil keyboard, type **setkb tscii** or **setkb tab**. From inside the keyboard driver you will be able to switch between the two standards, and also between Roman and Tamil fonts.

tamilvp

The other keyboard driver, *tamilvp* (vp for Visaip Palakai) is written and maintained by Dinesh. As indicated above, you may download that from the Yahoo! *tamilnix group* file section. It is available as rpm (I have not tried it out yet). Just install the rpm and files will be in appropriate locations. To run the program type **tamilvp** and you will get the GUI cell to choose between Tamil (TSCII 1.6 or TAB) and English.

KDE

Historically, the K Desktop Environment (KDE) was the first full Tamil user interface. Though far from complete, KDE was there for Tamil, and Tamil among the Indic languages, for the first time. Under KDE, with your localization properly set to Tamil, you may be able to do almost everything (from editing files, to browsing the web and e-mail, to administrative tasks such as user management and task scheduling) with a Tamil user interface.

Getting Localization Files

For the newbie, it is very easy to search the web for Tamil KDE localizations RPMs. They are usually labelled something like `kde-i18n-Tamil-2.0-1mdk.i586.rpm`. `i18n` is just that: i(nternationalizatio)n, 18(18letters). `Tamil` is the localization setting corresponding to the Tamil language. `mdk` signifies the package for Mandrake distribution. Then comes the most important part; `2.0-1`, the KDE version number. Your base KDE version and this should be the same, so when downloading, make sure that you get the proper localized menus for the proper KDE version. `i586` signifies the pre-compiled binaries for the intel 586 platforms. Make sure that you get the proper binary (there are usually source rpms and rpms for other platforms such as alpha). If you are a newbie you are better off using GUI based rpm installer such as GNORPM or KPackage. First do a test install and check if your system has all the needed packages. If not go to the same source from where you downloaded the Tamil localization and get them. After making sure that you installed all dependencies, install the `kde-i18n-tamil` package as well.

If you are not a newbie, you know it. Get KDE Tamil i18n files, and if you have time, get the sources and compile them!

KDE localization uses TSCII 1.6 encoding. This means that you will need at least one TSCII font. Read the section on fonts as to how to get it.

Choosing a Tamil locale

This section assumes that you have installed at least one TSCII font (preferably several, to jazz up your GUI) and the KDE Tamil localization package.

From Start, go to configuration > KDE > Personalization and choose default (c) location.

Note

Tamil/India is yet to be made available under countries/languages. Choose language >other >Tamil. Accept this. All changes will be activated, and will work on all windows opened subsequently.

Your user interface is now set in Tamil. If you see some garbage on the window header etc., pat yourself on the back. You are ready to see Tamil; move on!

Choosing Tamil fonts for GUI

Again, from Start go to configuration >KDE >LooknFeel. You will see a set of fonts for most (these are the ones used in display). Choose a Tamil font instead for all these. Accept.

Well done, you now see Tamil everywhere on your desktop. You are ready, with a fully operational Tamil system.

KDE Miscellaneous

As with every other project, KDE-Tamil also needs a lot of volunteers. Contact either Sivakumar or Vaseeharan (both of them can be reached through the egroup

Visit [<http://groups.yahoo.com/groups/tamilinux>] before you try KDE Tamil. If you want to convince yourself (and be bowled over), view the screenshots from tamillinux.org site.

KDE's i18n process is unicode-based. As a work around, Trolltech's QTsciiCodec class provides conversion to and from the Tamil TSCII encoding. This codec uses the mapping table found at [<http://www.geocities.com/Athens/5180/tscii.html>]. Unfortunately Tamil uses composed Unicode. As such, Unicode fonts cannot be used under KDE-TSCII; you need to have TSCII fonts. The TSCII codec was contributed to Qt by Hans Petter Bieker <bieker@kde.org>.

GNOME

GNOME Tamil localization works have just begun. There are few applications for which Tamil menus are translated, and are available. But it is yet to become the official member of GNOME i18n distribution.

In order to use them, download the currently available files from:

<http://www.tamillinux.org/gnome/gnome.html>

and put them into the directory `/usr/share/locale/ta/LCMESSAGES/`.

Under GNOME Control Panel you have set the fonts (both in Themes and the Window Manger applet) to a TSCII font.

You need to create binary messages from the `po` files. This is done as follows:

```
msgfmt xxx.po -o /usr/share/locale/ta/LCMESSAGES/xxx.mo.
```

Note that the binary messages files contain an extension `.mo` as opposed to `.po` for the text file.

In order for you see Tamil, you have to set the locale to Tamil.

If you are using bash as your shell, then enter the following line in your home directory.

```
export LANG=ta
export LANGUAGE=ta
export LC_ALL=ta
```

Restart the X server. You should see Tamil menus and dialogs in many of the GNOME enabled applications. Once again, please consider contributing to the Tamil GNOME Project; we need a lot of volunteers. Contact Dinesh <(n_dinesh@yahoo.com)> or through tamilnix yahoogroups.

Printing

This section is all about getting high-quality Tamil output in printing. While it is one issue to load a binary font and start using Tamil in Linux, if your work is to destroy the forests, you need high-quality printing too!

L^AT_EX

L^AT_EX is perhaps the mother of all typographic systems. It frees the author from the trivia of typesetting and concentrate on the content. It does not use the WYSWYG input, but the end result is great. Recent developments are centered toward internationalization. Unfortunately lack of unicode standard does not permit Tamil to be tried under the more ambitious Omega Project. Once again, workaround is the only way. A first step in Tamil has been attempted by Thuraiappah Vaseeharan. You may get the the package from the tamillinux.org site. The tar ball contains a great `readme` file that describes the installation and usage. The `tamiltex` package does a short work by keeping all related stuff under one directory (which means that you need to keep your work under the same directory to compile your source files). But the great thing about this package is that it is compatible with both TSCII and TAB encodings and the results are just what you would expect from a L^AT_EX package - great!

Postscript

Many Linux applications use Ghostscript to print, which means that you must have Ghostscript configured if you want to use Tamil in printed documents. If L^AT_EX is there, can PostScript be far away? Not thanks to Vasee. Set the environment variable `GS_FONTPATH` to point to your TrueType font directory. For example, I have:

```
GS_FONTPATH=/usr/local/share/fonts/tamiltt
export GS_FONTPATH
```

You should be able to view Tamil PostScript files.

PDF

As of now, the only source to create PDF files is the PDF package. If you are able to successfully compile your source with the tamiltex package, use

pdflatex source.tex

to generate the PDF file. You should be able to view it, using xdvi or Adobe's Acroread for Linux.

Word Processors, Office Packages

Once TrueType fonts are installed properly, there is no problem using them in Abiword, Gnumeric or KOffice. However, StarOffice needs Type 1 fonts. (I hear the latest StarOffice supports TrueType fonts?). You can expect Type 1 Tamil fonts to be available shortly:-).

For receiving and sending email, KMail works well with TrueType fonts. You should also be able to use PINE with Sivaraj's console fonts and utils.

Viewing Web pages

Konquerer supports Tamil fonts neatly, once made at the proper scale under your font directory and served to X. Widely used Netscape, however, is a problem. Netscape uses only 75 dpi fonts for display. You might have noticed this even while viewing Roman fonts, and got annoyed seeing small fonts. That being the case with Roman, Tamil is impossible to comprehend under 75 dpi. This can, however, be fixed by specifying the appropriate resources in your `.xdefaults` file:

```
Netscape*documentFonts.sizeIncrement: 20
Netscape*documentFonts.xResolution*iso-8859-1: 150
Netscape*documentFonts.yResolution.iso-8859-1: 150
```

Remember that TSCII fonts are used as ISO-8859-1 fonts. The parameter `150` is arbitrary; I have seen some fonts scaling neatly under `100` itself (TSCparanar, for one) which is good enough for viewing. If you are still not satisfied with what you see, try using anti-aliasing under X.

Pango

Pango [<http://www.pango.org>] provides an open-source framework for the layout and rendering of internationalized text and uses Unicode for all of its encoding. It aims to eventually support output in all the major languages. When GNOME 2.0 comes out, the text rendering is expected to be by Pango. Pango is expected to be the panacea for complex font schemes like kanji, arabic/hebrew (bidirectional), so Tamil is no problem. Tamil is one of the early languages in Pango - right there in the first public version. Sivaraj provided TSCII support, which was later extended to TAB by Vikram.

Miscellaneous

For the latest news, views and tools in Tamil Linux:

<http://tamil.homelinux.org/>

Issues related to Tamil localization are mostly discussed at:

<http://groups.yahoo.com/group/tamilinux/>

Under the files section there you may get some tools, few HOWTOs (most of those issues are unified in this document already) and some tutorials.

If you want to read about Open Source (Free Software) history in Tamil, see:

<http://www.tamillinux.org/venkat/cover.html>

Ganesan Rajagopal is checking in CVS for Tamil locales under the Sourceforge project on Tamil Linux, you may be get them from:

<http://cvs.sourceforge.net/cgi-bin/viewcvs.cgi/tamillinux/locale-ta/>

There is a simple guide to setting up a working Tamil Linux environment, addressed to newbies, available at:

http://www.tamillinux.org/venkat/tamil_inst.html

A. Appendix of Tamil Font Encodings

There are several non-standard font encoding schemes for Tamil. Then there are a whole lot of fonts (used mostly by publishing houses in Tamil Nadu, such as Vikatan, Kumutham, thinamaNi, etc.) which do not comply with any of these. The three major font encoding schemes are;

TSCII (Tamil Standard Code of Information Interchange - currently running in beta version 1.7); the first efforts by volunteers throughout world.

TAB (Tamil Bilingual); proposed and approved by the Tamil Nadu government

TSCII 1.6 Encoding Table

vowels: a, aa/A, i, ii/I, u, uu/U, e, ee/E, ai, o, oo/O, au, aq

consonants: k, ng, c, ny, t, N, th, n^, p, m, y, r, l, v, zh, L, R, n

Position	character name	TSCII glyph
Characters 0-127 are as in the standard lower ASCII set		
128 80	c128	tamil numeral 0
129 81	c129	tamil numeral 1
130 82	baseline single quote	tamil numeral 2
131 83	florin	tamil numeral 3
132 84	baseline double quote	tamil numeral 4
133 85	ellipsis	tamil numeral 5
134 86	dagger (single)	tamil numeral 6
135 87	dagger (double)	tamil numeral 7
136 88	circumflex	tamil numeral 8

137	89	per mil (thousand)	tamil numeral 9
138	8A	S caron	modifier for aa/A
139	8B	left single guillemet	modifier for I
140	8C	OE ligature	modifier for Ii/I
141	8D	c141	modifier for u
142	8E	c142	modifier for uu/U
143	8F	c143	modifier for e
144	90	c144	modifier for ee/E
145	91	open single quote	(left single guillemet)
146	92	close single quote	(right single guillemet)
147	93	open double quote	(left double guillemet)
148	94	close double quote	(right double guillemet)
149	95	bullet (large)	tamil numeral 10
150	96	en dash	tamil numeral 100
151	97	em dash	tamil numeral 1000
152	98	tilde	modifier for ai
153	99	unregistered trademark	tamil vowel a
154	9A	s caron	tamil vowel aa/A
155	9B	right single guillemet	tamil vowel i
156	9C	oe ligature	tamil vowel ii/Ai
157	9D	c157	tamil vowel u
158	9E	c158	tamil vowel uu/U
159	9F	Y diaeresis	tamil vowel e
160	A0	non-breaking space	(vacant)
161	A1	Spanish inverted !	tamil vowel ee/E
162	A2	cents	tamil vowel ai
163	A3	pounds	tamil vowel o
164	A4	intl. monetary symbol	tamil vowel oo/O
165	A5	yen	tamil vowel au
166	A6	broken bar	tamil vowel aq
167	A7	section symbol	tamil uyirmei ka
168	A8	diaeresis	tamil uyirmei nga
169	A9	copyright	copyright
170	AA	feminine ordinal	tamil uyirmei ca
171	AB	left double guillemet	tamil uyirmei nya
172	AC	logicalnot	tamil uyirmei ta
173	AD	soft hyphen (minus)	tamil uyirmei Na
174	AE	registered trademark	registered trademark
175	AF	macron	tamil uyirmei tha
176	B0	ring (also degrees)	tamil uyirmei n^a
177	B1	plus/minus	tamil uyirmei pa
178	B2	superscript 2	tamil uyirmei ma
179	B3	superscript 3	tamil uyirmei ya
180	B4	acute	tamil uyirmei ra
181	B5	micro symbol (or mu)	tamil uyirmei la
182	B6	pilcrow (paragraph)	tamil uyirmei va
183	B7	bullet (small)	bullet (small)
184	B8	cedilla	tamil uyirmei zha
185	B9	superscript 1	tamil uyirmei La
186	BA	masculine ordinal	tamil uyirmei Ra
187	BB	right double guillemet	tamil uyirmei na
188	BC	one-fourth	grantha letter ja
189	BD	one-half	grantha letter sha
190	BE	three-fourths	grantha letter sa

191	BF	Spanish inverted ?	grantha letter ha
192	C0	A grave	grantha letter ksha
193	C1	A acute	grantha letter sri
194	C2	A circumflex	tamil uyirmei ti/di
195	C3	A tilde	tamil uyirmei tii/dii
196	C4	A diaeresis	tamil uyirmei ku
197	C5	A ring	tamil uyirmei ngu
198	C6	AE ligature	tamil uyirmei cu
199	C7	C cedilla	tamil uyirmei nyu
200	C8	E grave	tamil uyirmei tu
201	C9	E acute	tamil uyirmei Nu
202	CA	E circumflex	tamil uyirmei thu
203	CB	E diaeresis	tamil uyirmei n^u
204	CC	I grave	tamil uyirmei pu
205	CD	I acute	tamil uyirmei mu
206	CE	I circumflex	tamil uyirmei yu
207	CF	I diaeresis	tamil uyirmei ru
208	D0	Icelandic Eth	tamil uyirmei lu
209	D1	N tilde	tamil uyirmei vu
210	D2	O grave	tamil uyirmei zhu
211	D3	O acute	tamil uyirmei Lu
212	D4	O circumflex	tamil uyirmei Ru
213	D5	O tilde	tamil uyirmei nu
214	D6	O diaeresis	tamil uyirmei kU
215	D7	multiply symbol	tamil uyirmei ngU
216	D8	O with oblique stroke	tamil uyirmei cU
217	D9	U grave	tamil uyirmei nyU
218	DA	U acute	tamil uyirmei tU
219	DB	U circumflex	tamil uyirmei NU
220	DC	U diaeresis	tamil uyirmei thU
221	DD	Y acute	tamil uyirmei n^U
222	DE	Icelandic Thorn	tamil uyirmei pU
223	DF	German sharp s	tamil uyirmei mU
224	E0	a grave	tamil uyirmei yU
225	E1	a acute	tamil uyirmei rU
226	E2	a circumflex	tamil uyirmei lU
227	E3	a tilde	tamil uyirmei vU
228	E4	a diaeresis	tamil uyirmei zhU
229	E5	a ring	tamil uyirmei LU
230	E6	ae ligature	tamil uyirmei RU
231	E7	c cedilla	tamil uyirmei nU
232	E8	e grave	tamil vowel k (ik)
233	E9	e acute	tamil vowel ng (ing)
234	EA	e circumflex	tamil vowel c (ikc)
235	EB	e diaeresis	tamil vowel ny (iny)
236	EC	i grave	tamil vowel t (it)
237	ED	i acute	tamil vowel N (iN)
238	EE	i circumflex	tamil vowel th (ith)
239	EF	i diaeresis	tamil vowel n (in^)
240	F0	Icelandic eth	tamil vowel p (ip)
241	F1	n tilde	tamil vowel m (im)
242	F2	o grave	tamil vowel y (i<)
243	F3	o acute	tamil vowel r (ir)
244	F4	o circumflex	tamil vowel l (il)

245	F5		o tilde		tamil vowel v (iv)
246	F6		o diaeresis		tamil vowel zh (izh)
247	F7		divide symbol		tamil vowel L (iL)
248	F8		o with oblique stroke		tamil vowel R (iR)
249	F9		u grave		tamil vowel n (in)
250	FA		u acute		grantha vowel j (ij)
251	FB		u circumflex		grantha vowel sh (ish)
252	FC		u diaeresis		grantha vowel s (is)
253	FD		y acute		grantha vowel h (ih)
254	FE		Icelandic thorn		grantha vowel ksh (iksh)
255	FF		y diaeresis		(vacant)
----- ----- -----					