

TamilNet Transcription: version 1.0 launched

[TamilNet, Friday, 01 June 2007, 21:49 GMT]

The TamilNet Transcription has been improved and modified since its launch for trial on April 28, 2007. It is now simplified further by minimising the interference of punctuation marks in transcription and enhanced by the introduction of context-sensitive algorithm in automation. It is also expanded to accommodate phonological variations in Roman to Tamil automation and in some instances, Tamil to Roman, without touching ambiguities. We invite thorough scrutiny and feedback from readers before releasing the algorithm for public use. We also invite proposals from interested well-wishers to develop conducive software, confining to the specifications of the system.

Changes and Improvements:



With the help of feedback from the editorial staff of TamilNet and

readers, the following changes and improvements have been made in the TamilNet Transcription.

1. Colon is removed before *n* because of its interference with reading. By doing so, we are equating the pronunciation of ந with ன in transcription. Please note that Tamil grammatical texts are specific about the subtle difference in pronunciation between the two *n*'s. However, this is not felt in ordinary speech today and may not affect the transcription seriously.
2. For automation of the two *n*'s, from Roman to Tamil, a context-sensitive algorithm is introduced. In the beginning of words and when precedes *th*, the *n* will be automated in Tamil as ந. In all other instances, it will become ன.
3. Error may arise in automating compound words where the second word begins with *n* and in some sanskrit words which some people prefer to write with ந. Examples: சைவந்தி, வெந்தீர், அந்நியன், ஆநந்தகுமாரசுவாமி, விபுலாநந்தர் etc. To automate such words from Roman to Tamil, type a colon before *n* as shown here: *chaiva:neethi*, *ve:n:neer*, *a:n:niyan*, *Aa:nanthakumaarasuvaami*, and *Vipulaa:nanthar*. However, in Tamil to Roman transcription, the words will appear as *chaivaneethi*, *venneer*, *anniyar*, *Aananthakumaarasuvaami*, *Vipulaananthar* etc., without causing interference to readers.
4. The instances where *ch* is pronounced as *s* without any ambiguity, are accommodated in the transcription. Eg.: பசி (*pasi*) அம்சம் (*amsam*). They can be reverted back to Tamil script in automation.
5. Some readers have brought to our notice the pronunciation shifts in hard consonants (*vallinam*). We do agree with the phenomenon. But, there are ambiguities and dialectal differences. In such instances, we prefer to stick to the alphabet. However, we have accommodated those phonological deviations in reverse automation ie. from transcription to Tamil script, for people who prefer to write in those ways in English.
6. In transcription, the Grantha ஸ will become *s*. But in reverse transcription, only a single consonant *s* will be automated to ஸ. In other instances it will show ச. In reverse transcription, type 's (with an apostrophe preceding *s*) to produce ஸ on any required occasion.
7. The Grantha ஹ will become *h* in transcription. In reverse transcription, only in the following instances *h* will be automated to ஹ: (a) in the beginning of a word; (b) a single consonant. In other instances *h* will be automated to க. In reverse transcription, type 'h (apostrophe preceding *h*) to produce ஹ on any required occasion.
8. The *aaytham* ஃ is transcribed by a colon added to *h* (:h).

9. The vowel consonants of *r* (ர, ரா etc.) following consonant *n* (ன்) is transcribed to *t*.

See the PDF document for revised [Transcription Tables](#) and use the [web-tool](#) provided by TamilNet for online transcription and automation.

Find this article at:

<http://www.tamilnet.com/art.html?catid=13&artid=22339>

Reproduction of this news item is allowed when used without any alterations to the contents and the source, TamilNet, is mentioned

Email

Close