

# DIFFICULTIES IN LEARNING TAMIL : IRREGULARITIES IN THE TAMIL SCRIPT

N. ANBARASAN

APPLESOFT

No 39, Ist Cross, Ist Main, Shiva nagar, West of chord road  
Bangalore - 560 010, India. <email: aplesoft@vsnl.com>

---

## Abstract

Computers and TVs have become easy source of Tamil Language and culture in today's modern world of technology. These perform the role of carriers of Language and culture. Computers play a dominant service on Language learning while TVs help in Language acquisition and supplements Language learning. As the content preparation is centered around computers, the mistakes in developing fonts, difficulties in presenting multimedia contents compounds the problems in learning Tamil scripts through distance / online education and classroom learning.

Now a days, some multimedia software are available to augment the requirement of teaching/learning materials for Tamil. However, they are not able to eliminate the difficulties of the learner in recognising all varieties of scripts used in the society.

With the experience in developing multimedia software for teaching / learning Tamil as First Language and Second Language, the author highlights certain issues / problems related to learning Tamil script in general and CALL (Computer Assisted Language Learning) in particular.

## Introduction

Tamil Language is rich, old and it has its own heritage. The script for Tamil Language was conceived since time immemorial and the present form is after undergoing various changes at various periods of time. During its long cherished existence, it under-went various changes, corrections, modifications to form the present form based on its applicability and adoptability in the ever changing writing method and devices from time to time. The present and in the recent past (It is considered because of its continued usage by some Sections of the Society) form of the Tamil script is believed to have been evolved based on its "adoptability" for hand writing mechanisms. However, due to passage of time and driven by the lust for new innovations and to go for new methods to suit the needs, methods and devices got changed from time to time. The method devised for handwriting, obviously was found to be having certain problems and irregularities in terms of irregular hand movements, non-uniform allographs of vowels in terms of shape, position and conjugation.

## Tamil Orthography

Tamil writing system (Letters) consists of vowels, consonants and syllables (Vowel consonants). A syllable is formed when a consonant is followed by a vowel. When vowel joins with the previous consonant, it appears as vowel signs (allographs of vowels). The allograph

may appear on top, bottom, left, right or both sides of the 'a' consonant. The shape of allograph of vowels such as i, ii, u, uu etc., are normally vary when joined with different consonants.

1. Graphemic variation of secondary symbol: Even though the allographic variations are found for vowels இ, ஈ, உ, ஊ. It is predominant in the case of உ and ஊ. Please note that the allographic variation of ஆ for certain archaic letters is not taken into consideration.

i) Allograph of இ :

It appears in 5 different forms:

1. with க, ச, ஞ, ண, த, ந, ர, ள, ன
2. with க, ச, த as in ' , ' , "
3. with ம, ழ, ற
4. with ங, ப, ய, வ
5. with ல

ii) Allograph of ஈ : Similarly as for இ.

iii) Allograph of உ : It appears in 12 different forms and they are classified based on the shape of Secondary symbol and shape (allograph) of the consonant changed, as in:

1. கு : The end portion of the consonant க is slightly reduced and makes half circle from bottom and moves backwards.
2. ங, ப, ய, வ : A small vertical line is attached to the right bottom of the 'a' consonants.
3. சு : A small vertical line is attached to the end of the ending horizontal line.
4. னு, து, நு : The end portion of the respective consonants are reduced and the allograph து is attached to the remaining portion of the consonants at the end.
5. ணு, று, னு : Here also, the letter formation is similar to above.
6. லு : The 'a' consonant is not changed, but the allograph து just attached to the bottom.
7. மு : The allograph is just a continuation of the stroke after writing 'a' consonant.
8. டு : The Horizontal stroke of the letter ட is dropped and allograph is added to the remaining portion of the letter.
9. ரு : The last slant stroke is dropped and allograph " is added to the remaining portion of the letter.
10. ளு : To the end of the 'a' consonant, allograph " is added to the remaining portion of the letter.
11. ழு : The last portion of the 'a' consonant is changed and allograph is added to the remaining portion of the letter.
12. ஜு : The 'a' consonant is not changed and allograph ழ is written next to the 'a' consonant from right top.

iv) Allograph of ஊ : It appears in 12 different forms with an extended sign to 2 consonants as in:

1. கூ
2. ஙூ, பூ, யூ, வூ
3. சூ
4. னூ, தூ, நூ
5. ணூ, றூ, னூ
6. லூ
7. மூ
8. டூ
9. ரூ
10. ளூ

11. ೀ
12. ೀ

## Teaching Tamil Script Methodology

For any one interested in learning more than conversational Tamil, learning the complex Tamil script is an inevitable task. Now a days Tamil script teaching is based on pattern perception, shape similarity and contrastive observation. In this method, Tamil letters are grouped into the following 11 groups.

1. ೀ, ೀ, ೀ, ೀ, ೀ
2. ೀ, ೀ, ೀ, ೀ, ೀ
3. ೀ, ೀ, ೀ
4. ೀ, ೀ
5. ೀ, ೀ, ೀ
6. ೀ, ೀ
7. ೀ, ೀ
8. ೀ, ೀ, ೀ
9. ೀ, ೀ, ೀ
10. ೀ, ೀ
11. ೀ, ೀ, ೀ, ೀ, ೀ, ೀ

The allographs or secondary symbols of the vowels are introduced alongwith the vowels. While teaching/writing a vowel consonant (syllables), they pose lot of difficulties and they are grouped into six categories, such as

1. Secondary symbols which follow the 'a' consonant.  
Ex: ೀ, ೀ, ೀ ...
2. Secondary symbols which precede the 'a' consonant  
Ex : ೀ, ೀ, ೀ ...
3. Secondary symbols written on both sides of the 'a' consonant  
Ex : ೀ, ೀ, ೀ ...
4. Secondary symbols written on top of the 'a' consonant  
Ex : ೀ, ೀ ...
5. Secondary symbols written at the bottom of the 'a' consonant depending on the starting points and position of the secondary symbol. This type of syllables are written in three ways.
  - i) ೀ, ೀ, ೀ, ೀ, ೀ
  - ii) ೀ. ೀ. ೀ. ೀ. ೀ, ೀ
  - iii) ೀ, ೀ, ೀ, ೀ, ೀ, ೀ, ೀ

The secondary symbols (ೀ, ೀ) of long vowel ೀ, are extended sign to the short 2 consonant in the order shown, with exception of ೀ. ೀ formed differently.

6. Since the grantha letters are being taught at the end separately, the secondary symbol used exclusively for these letters are taught separately as in: ೀ, ೀ, ೀ, ೀ etc.

## Font Development

There are a variety of True type fonts available which are developed for use on graphics based systems like MS windows. The available Tamil fonts can be classified and grouped as follows:

1. Glyphs mapped directly on to ASCII code corresponding to key positions of a glyph in typewriter. These type of fonts do not require any additional keyboard managers to input Tamil text. They are cheap.
2. Glyphs designed for aesthetic requirements of DTP, titling etc, which needs a keyboard manager to input Tamil text. They normally make use of all possible code positions in the Microsoft Windows encoding standard.
3. Glyphs mapped on to upper ASCII (Latin -1) part of the ANSI encoding standard. These type of fonts can be used with the help of keyboard drivers only. They are normally used for database applications, bi-lingual user interface or Internet based applications.

In the first and third type of fonts, quality suffers due to slicing of letters. Sometimes these fonts donot produce acceptable quality under certain settings.

#### Internet and TV

Internet and satellite Television have become main source of Language learning and acquisition for heritage learners or Foreigners. As the contents are produced/created using computers for Internet and satellite Television, the errors in fonts in terms of wrong settings/design also reaches the potential learner hampering his learning activity.

#### Learning Difficulties

The learning difficulties of Tamil script could be summarised as follows:

1. Learners face problem about the place and positions of the secondary symbols. The position of secondary symbol vary as in டு, கி, கா, சு, கெ, கொ etc.
2. When the syllables are composed using category 1 or 3 type of fonts, some of the vowel consonanta are produced as:

பு	=	பி
டி	=	பி
பு	=	பு, ப,
ழு	=	ழு
யீ	=	யீ

These type of variations are observed on the popular satellite TVs.

3. Allograph variation of 'a' consonants when forms a vowel consonant.

ர =>ர; ன =>ன; ண =>ண; ழ =>

4. Variation in letter form.

சூ = சூ

ஐ	=	ஐ
ஐா	=	ஐ ஈ
ஐை	=	ஐ
ஐஐ	=	ஐஐ ஈ
ஐா	=	ஐ ஈ
ஐ	=	ஐ or -
ஐ	=	ஐ
ஐ்	=	ஐ்
ஐி	=	ஐி

5. For the learners or an occasional language user where Tamil is not being used in social contexts, letter identification often becomes a problem with certain letters having similar shapes as in :

க	ச
மு	மு
ல	வ
ஓ	ஓ
ஐ	ஐ

ironically these letters are also ambiguous for computers in identifying these in applications such as OCR.

#### Developing Multimedia based Tamil learning software

By making use of the facilities like animations, colour and interaction, the hand movements while writing Tamil script and contrast in shape in similar letters can be depicted fairly well. The learner could be able to understand the process in a better way. It also helps in retaining more of the learning.

For example while teaching vowel consonants of ஊ. It could be shown with an incremental hand movement in different colours as

ம மு மு
ப ப ப

The animation facility could be used for repetitive writing of the same letter to reinforce the writing hand movement and for better grasping and understanding.

#### Problems in developing language learning software

Due to the complexities of script, it is difficult to develop a game/exercise for vocabulary developments, letter identification by making use of keyboard as input. As the graphemes of allograph in isolation do not have a meaning, it can't be used in word building or in any vocabulary development exercise or games. This situation forces to use either the basic vowels and consonants- which is not the way the standard keyboard is based on.

#### Conclusion

I have attempted to highlight the complexities and irregularities of Tamil Script, the non-standard

fonts usage which further makes the Tamil script worst. Eventhough, Tamil Script was not developed by way of planned and organised activity, it is high time to arrive at a standard Tamil script as well. Let us try to simplify the Tamil Script and Tamil keyboard in order to make the learning Tamil more interested and a funfilled experience with the help of computers and more simple to use for all purpose even by a common man.