Critical editions of Tamil works: exploratory survey and future perspectives

Jean-Luc CHEVILLARD

(CNRS, Université Paris-Diderot)

<http://www.linguist.univ-paris-diderot.fr/~chevilla/>

Table of Contents

- Abstract
- Acknowledgement and background
- Digital *Tēvāram*, Project Madurai, TEI
- Tapping the ancient knowledge of Tamil scholars
- Metrical text (MT) and sandhi-split text (SST)
- Adapting the TEI markup to the *Tēvāram*
- Exploring the meters in *Tēvāram*
- Statistics for Campantar
- Towards electronic critical editions
- Short conclusion
- Bibliography

Abstract: This paper will deal with the possible representations of ancient Tamil texts in digital form. It does not offer a technical (software) solution for that problematic question. It is rather the exposition, through several examples, of recurring difficulties which require attention. Addressing them is important for a proper understanding of the history and transmission of those texts. Logical frameworks, such as the one provided by the Text Encoding Initiative (TEI), exist, but software tools for comfortably dealing with the situation remain a desideratum.

Acknowledgement and background (1)

TEI P5:

Guidelines for Electronic Text Encoding and Interchange
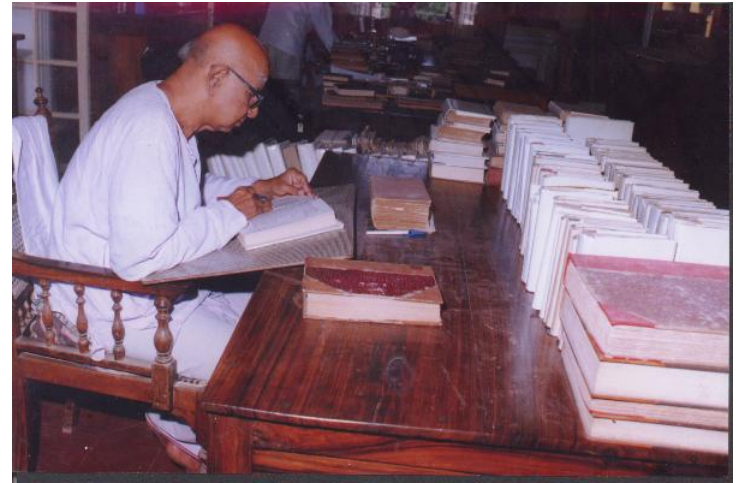
Acknowledgement and background (2a)

- This presentation is in some sense the continuation of an effort started in July 1997 which has resulted in the publication in 2007 of a CD-ROM, containing data in HTML format.

Acknowledgement and background (2b)

The CD-ROM description is as follows:

- Digital Tēvāram / கணினித் தேவாரம்
- with the complete English gloss of the late V. M. Subramanya Ayyar (IFP) and furnished with a full concordance of the Tamil text accompanied by 6 hours of MP3 audio recordings (illustrating all the 24 paṇ-s) various maps (showing all the 274 talam-s) and other related material.
- See link on French Institute of Pondicherry web site. (www.ifpindia.org)

Acknowledgement and background (3a)

I have often been helped and stimulated in my endeavours by the invaluable online resource called "Project Madurai", which has been prepared for a number of years by voluntaries called "Promads", under the leadership of Dr. K. Kalyanasundaram. Project Madurai is found at http://www.projectmadurai.org/ and the site hosted 338 files at the end of September 2009. Some of those texts are metrical texts (see for instance: PM0150: தேவாரம் [1A], etc.) and others are sandhi-split texts (see for instance PM0100: தொலகாப்பியம்)

Acknowledgement and background (3b)

While preparing the current presentation, I have also derived inspiration from another collective endeavour, called the TEI. See the title page of the 1345 pages PDF document available on their web site at the URL:

http://www.tei-c.org/index.xml

TEI P5:

Guidelines for Electronic Text Encoding and Interchange

by the TEI Consortium

Originally edited by C.M. Sperberg-McQueen and Lou Burnard for the ACH-ALLC-ACL Text Encoding Initiative

Now entirely revised and expanded under the supervision of the Technical Council of the TEI Consortium

edited by Lou Burnard and Syd Bauman

Oxford — Providence — Charlottesville — Nancy 2008
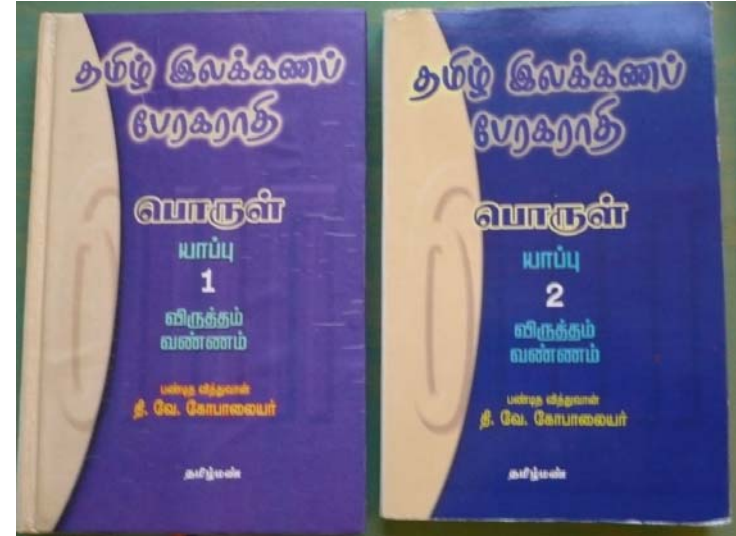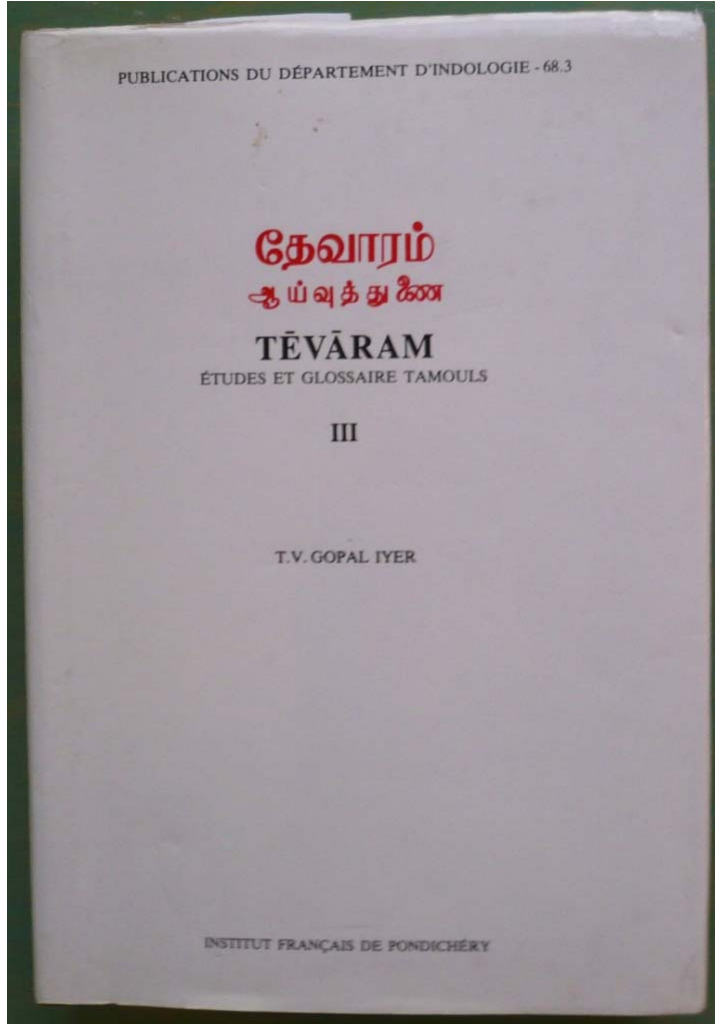
Acknowledgement and background (3c)

- The general goals of the TEI consortium are stated in the following way: "The Text Encoding Initiative (TEI) is a consortium which collectively develops and maintains a standard for the representation of texts in digital form. Its chief deliverable is a set of Guidelines which specify encoding methods for machine-readable texts, chiefly in the humanities, social sciences and linguistics."

- In the course of this presentation, I shall draw examples and ideas from two chapters in the TEI guidelines:
  - the chapter that deals with the representation of metrical text (TEI-P5, pp.187-202, Chapter 6: Verse)
  - the chapter that deals with tools for making critical editions (TEI-P5, pp.379-398, Chapter 12: Critical Apparatus)

- These chapters have inspired me in my search of methods for representing texts which are to be handled under two (or more) varieties

Tapping the ancient knowledge of Tamil scholars **(1)**

A powerful reason for exploring these fields is that an enormous amount of information on texts such as the *Tēvāram* is available in books prepared by traditional scholars from Tamil Nadu, among whom I would like to mention first of all the late pandit T.V. Gopal Iyer (1926-2007). It is important that the knowledge gathered by him and by other scholars is shared with the younger generations, and especially with those who can give it a wide circulation on the internet.

Tapping the ancient knowledge of Tamil scholars **(2)**

PUBLICATIONS DU DÉPARTEMENT D'INDOLOGIE - 68.3

தேவாரம்
ஆய்வுத்துணை

TĒVĀRAM
ÉTUDES ET GLOSSAIRE TAMOULS

III

T.V. GOPAL IYER

INSTITUT FRANÇAIS DE PONDICHÉRY

Especially important for this presentation is his 1991 *Tēvāram Āyvut Tuṇai,* but also his in 2004 தமிழ் இலக்கணப் பேரகராதி in 17 volumes

Tapping the ancient knowledge of Tamil scholars **(3)**

PUBLICATIONS DE L'INSTITUT FRANÇAIS D'INDOLOGIE No. 68,1

TĒVĀRAM

HYMNES ŚIVAÏTES DU PAYS TAMOUL

தேவாரம்

பண்முறை

édition établie par T. V. GOPAL IYER
sous la direction de FRANÇOIS GROS

volume I Ñāṉacampantar

ஞானசம்பந்தர்

publié avec le concours du Fonds International
pour la Promotion de la Culture

INSTITUT FRANÇAIS D'INDOLOGIE
PONDICHÉRY
1984
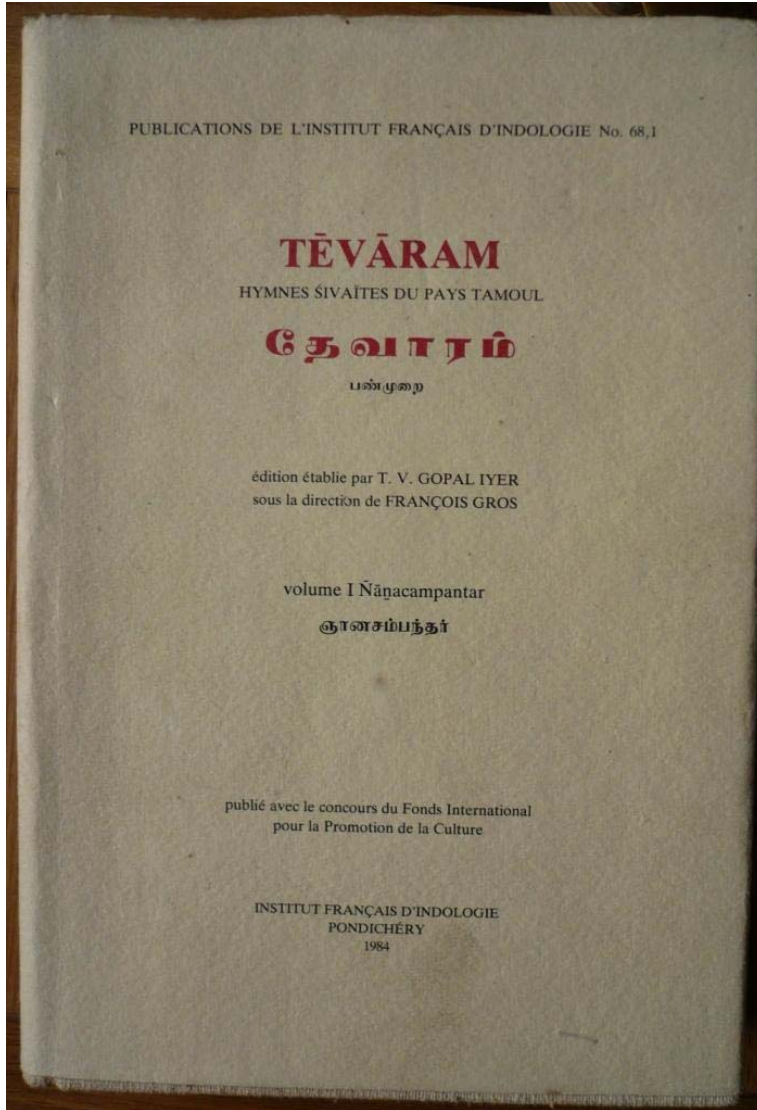
The *Tēvāram Āyvut Tuṇai* was preceded by a two volume edition of the *Tēvāram*: the sandhi split text which it contains is available on the Digital Tēvāram CD.

Metrical text (MT) and sandhi-split text (SST) (1)

The right-hand side on this screen shows the text of one stanza in the **PM** (Project Madurai) and in the 1984 volume of the Tēvāram edition by T.V. Gopal Iyer **(TVG)**

செந்நெல் அம் கழனிப் பழனத்து அயலே செழும் புன்னை வெண் கிழியில் பவளம் புரை பூந்தராய் துன்னி, நல் இமையோர் முடி தோய் கழலீர்! சொலீர்---
பின்னுசெஞ்சடையில் பிறை பாம்புஉடன் வைத்ததே? (TĒV. 2-1_1) [**TVG** edition]

செந்நெ லங்கழ னிப்பழ னத்தய லேசெழும் புன்னை வெண்கிழி யிற்பவ எம்புரை பூந்தராய் துன்னி, நல்லிமை யோர்முடி தோய்கழ லீர்சொலீர்
பின்னு செஞ்சடை யிற்பிறை பாம்புடன் வைத்ததே? (TĒV. 2-1_1) [**PM** text]

Metrical text (MT) and sandhi-split text (SST) (2)

- The question will then naturally arise: "which of those two texts, MT or SST, is to be considered as the base text?"

- In what follows, I shall mention a few arguments for preferring one or the other.

- I shall conclude that, if we want to try to preserve (or to reconstruct) the totality of the historical knowledge associated with a text such as the *Tēvāram*, we must have both texts available and find a way to digitally encode the **dual textual reality** which they represent.

Metrical text (MT) and sandhi-split text (SST) (3)

- Some of the arguments given by those who consider the sandhi-split text (SST) as the base text are as follows:

- It is easier to understand

- It seems usually possible to reconstruct the metrical text (MT) on the basis of the SST if we know the sandhi rules, given by grammar, and if we know the name of the meter (and possibly the name of the *paṇ* [பண்])

Metrical text (MT) and sandhi-split text (SST) (4)

- Some of the arguments given by those who consider the metrical text (MT) as the base text are as follows:

- It is the **transmitted text** : if we use older printed books and palm-leaf manuscripts, this is what we find.

- The sandhi-split text (SST) is an interpretation and different scholars might have (or have had in the past) different interpretations. If we provide only the SST, this is equivalent to eliminating possible meanings.

Metrical text (MT) and sandhi-split text (SST) (5)

- However, whatever our own opinion concerning the primacy of the MT or the SST, one can probably say that they are as indissolubly linked as form and meaning. In the case of the *Tēvāram*, those who traditionally transmitted it, the ஒதுவார், recited a metrical text, but they also understood that text and it seems natural to think that they could perceive the presence of words inside that text, even when the words were somehow hidden by the metrical structure.

- Therefore, it would be highly desirable to define a standard for providing to modern users, in digital form, dual texts containing, as parallel components, a MT and a SST. And several types of parsers could be developped for **validating the alignment** of the two components.

Adapting the TEI markup to the Tēvāram (1)

- However, before tackling dual texts, it appears that a preliminary necessary step is to explore the existing available frameworks for dealing with verse.

- The Text Encoding Initiative (TEI) provides a number of modules which can be applied (or adapted) to the Tamil case.

- In the following screens, I shall present extracts from the 6th chapter of the (P5) TEI guidelines (pp.187-202).

- Those extracts will be followed by possible adaptations to the *Tēvāram* corpus.

Adapting the TEI markup to the Tēvāram (2)

A short poem by Emily Dickinson (TEI guidelines, p.187)

```
<text>
<front>
<head>1755</head>
</front>
<body>
<l>To make a prairie it takes a clover and one bee,</l>
<l>One clover, and a bee,</l>
<l>And revery.</l>
<l>The revery alone will do,</l>
<l>If bees are few.</l>
</body>
</text>
```

Adapting the TEI markup to the Tēvāram (3)

Grouping of lines [Allen Ginsberg] (TEI guidelines, p.188)

```
<text><body>
<head>My Alba</head>
<lg>
<l>Now that I've wasted</l>
<l>five years in Manhattan</l>
<l>life decaying</l>
<l>talent a blank</l>
</lg>
<lg>
<l>talking disconnected</l>
<l>patient and mental</l>
<l>sliderule and number</l>
<l>machine on a desk</l>
</lg>
</body></text>
```

Adapting the TEI markup to the Tēvāram (4)

<text><body>

<head>TĒVĀRAM 2-1</head>

<lg>

<l>செந்நெ லங்கழ னிப்பழ னத்தய லேசெழும்</l>

<l>புன்னை வெண்கிழி யிற்பவ எம்புரை பூந்தராய்</l>

<l>துன்னி, நல்லிமை யோர்முடி தோய்கழ லீர்சொலீர்</l>

<l>பின்னு செஞ்சடை யிற்பிறை பாம்புடன் வைத்ததே?</l>

</lg>

<lg>

<l> எற்று தெண்டிரை யேறிய சங்கினொ டிப்பிகள்</l>

<l>பொற்றி கழ்கம லப்பழ னம்புகு பூந்தராய்ச்</l>

<l>சுற்றி நல்லிமை யோர்தொழு பொன்கழ லீர்சொலீர்</l>

<l>பெற்ற மேறுதல் பெற்றிமை யோபெரு மானிரே</l>

</lg>

</body></text>

Adapting the TEI markup to the Tēvāram (5)

Components in line: Langland's Piers Plowman (TEI Guidelines, p.191)

```
<l>
<seg>In a somer seson,</seg>
<seg>whan softe was the sonne,</seg>
</l>
<l>
<seg>I shoop me into shroudes</seg>
<seg>as I a sheep were,</seg>
</l>
<l>
<seg>In habite as an heremite </seg>
<seg>unholy of werkes,</seg>
</l>
<l>
<seg>Went wide in this world </seg>
<seg>wondres to here.</seg>
</l>
```

```
<அடி>
<சீர்> செந்நெ </சீர்>
<சீர்> லங்கழ </சீர்>
<சீர்> னிப்பழ </சீர்>
<சீர்> னத்தய </சீர்>
<சீர்> லேசெழும் </சீர்>
</அடி>
<அடி>
<சீர்> புன்னை </சீர்>
<சீர்> வெண்கிழி </சீர்>
<சீர்> யிற்பவ </சீர்>
<சீர்> எம்புரை </சீர்>
<சீர்> பூந்தராய் </சீர்>
</அடி>
.................................................
```

Adapting the TEI markup to the Tēvāram (7)

*Pope's Essay on criticism **[iambic pentameter]** (TEI Guidelines, p.194)*

```
<div
type="book"
n="1"
met="-+|-+|-+|-+|-+/"
rhyme="aa">
<lg n="1" type="paragraph">
<l>'Tis hard to say, if greater Want of Skill</l>
<l>Appear in <hi>Writing</hi> or in <hi>Judging</hi> ill;</l>
<l>But, of the two, less dang'rous is th'Offence,</l>
<l>To tire our <hi>Patience</hi>, than mis-lead our
<hi>Sense</hi>:</l>
</lg>
</div>
```

Adapting the TEI markup to the Tēvāram (8)

Inserted alexandrine in Pope's Essay on criticism (Guidelines, p.196)

<l n="356"> A needless alexandrine ends the song, </l>

<l n="357" met="-+|-+|-+|-+|-+|-+" real="++|-+|-+|+-|++|-+"> That, like a wounded snake, drags its slow length along.

</l>

<l n="356"> A need<seg type="foot" n="2" real="--">less

a</seg>lexandrine ends the song,</l>

<l n="357" met="-+|-+|-+|-+|-+|-+">

<seg n="1" real="++"> That, like </seg> a wounded snake,

<seg n="4" real="+-"> drags its </seg>

<seg n="5" real="++"> slow length </seg>

along.

</l>

<அடி>
<சீர் met="தேமா"> செந்நெ </சீர்>
<சீர் met="கூவிளம்"> லங்கழ </சீர்>
<சீர் met="கூவிளம்"> னிப்பழ </சீர்>
<சீர் met="கூவிளம்"> னத்தய </சீர்>
<சீர் met="கூவிளம்"> லேசெழும் </சீர்>
</அடி>
<அடி>
<சீர் met="தேமா"> புன்னை </சீர்>
<சீர் met="கூவிளம்"> வெண்கிழி </சீர்>
<சீர் met="கூவிளம்"> யிற்பவ </சீர்>
<சீர் met="கூவிளம்"> எம்புரை </சீர்>
<சீர் met="கூவிளம்"> பூந்தராய் </சீர்>
</அடி>
.... (upto the end of *Tēvāram* 2-1)

Adapting the TEI markup to the Tēvāram (10)

- The next step forward is to create a dual text, including both the MT and the SST, as in the following screen.

- Such a tagged text could be created automatically on the basis of existing texts, namely the MT, freely available the Project Madurai web site (http://www.projectmadurai.org/) and the SST prepared by T.V. Gopal Iyer, freely available on the French Institute of Pondicherry web site (www.ifpindia.org), which appears here in red.

- It is of course **a challenge for gifted programmers**, to find out whether, after performing automatic alignment of the two texts, they could create the metrical labels which appear in green on the following screen.

<அடி பொருள்="செந்நெல் அம் கழனிப் பழனத்து அயலே செழும்">
<சீர் met="தேமா"> செந்நெ </சீர்>
<சீர் met="கூவிளம்"> லங்கழ </சீர்>
<சீர் met="கூவிளம்"> னிப்பழ </சீர்>
<சீர் met="கூவிளம்"> னத்தய </சீர்>
<சீர் met="கூவிளம்"> லேசெழும் </சீர்>
</அடி>
<அடி பொருள்="புன்னை வெண் கிழியில் பவளம் புரை பூந்தராய்">
<சீர் met="தேமா"> புன்னை </சீர்>
<சீர் met="கூவிளம்"> வெண்கிழி </சீர்>
<சீர் met="கூவிளம்"> யிற்பவ </சீர்>
<சீர் met="கூவிளம்"> எம்புரை </சீர்>
<சீர் met="கூவிளம்"> பூந்தராய் </சீர்>
</அடி>
.... (upto the end of தேவாரம் 2-1)

Exploring the meters in Tēvāram (1)

- We could continue our examination of possible ways for adapting the strategies developed in the TEI guideline for inserting markup in ancient Tamil poetical compositions, provided time were illimited.

- However, this is a long task and it is advisable at this stage to ponder on the benefits which can be obtained if the task is accomplished, one among them being a reconquest of the traditional field of knowlege called யாப்பு.

- For instance, we can recognize in the metrical pattern seen in the previous screen a specimen of that auxiliary meter (பாவினம்) which is called கலித்துறை.

- Each line contains 5 cīr-s and the first cīr always ends in "மா" while the 4 following cīr-s all end in "விளம்".

- There is a regular pattern: "மா விளம் விளம் விளம் விளம்".

Exploring the meters in Tēvāram (2)

- In addition to the auxiliary meter (பாவினம்) called கலித்துறை which we have just mentioned, with one of its many possible patterns, there exist other auxiliary meters.

- The names of the canonical 12 pāviṉam-s, as seen in the *Yāpparuṅkalak kārikai*, are formed by combining the 4 words *āciriyam, vañci, veṇpā* and *kali* (which are the name of the 4 classical meters) with the 3 words *tāḻicai, tuṟai* and *viruttam.*

- The result is: *āciriyat tāḻicai, vañcit tāḻicai, veṇtāḻicai, kalit tāḻicai, āciriyat tuṟai, vañcit tuṟai, veṇtuṟai, kalit tuṟai, āciriya viruttam, vañci viruttam, veḷi viruttam, kali viruttam*

- A great number of those are found in the *Tēvāram*

- Precise explanations about those meters can be found for instance in the volumes 14 and 15 of T.V. Gopal Iyer's 2004 தமிழ் இலக்கணப் பேரகராதி already mentioned.

Exploring the meters in Tēvāram (3)

Additionally, the volume *Tēvāram, Āyvut tuṇai*, already mentionned, contains on pp.375-421 information on the all the meters used in the Tēvāram hymns. In order to further illustrate the metrical variety inside the Tēvāram, I now take another hymn as example, composed in a shorter meter, called **vañci viruttam** (வஞ்சி விருத்தம்)

- அருத்தனை யறவனை யமுதனைநீர்
- விருத்தனைப் பாலனை வினவுதிரேல்
- ஒருத்தனை யல்லதிங் குலகமேத்தும்
- கருத்தவன் வளநகர் கடைமுடியே.

Tēvāram, 1-111

Exploring the meters in Tēvāram (4)
(T.V. Gopal Iyer, *Tēvāram, Āyvut tuṇai*, pp.384)

104) ஆசிரிய இணைக்குறட்டுறை:

    'ஆட லரவசைத்தான் அருமாமறை தான்விரித்தான் கொன்றை'   *1-1*

    'சூடிய செஞ்சடையான் சுடுகா டமர்ந்தபிரான்'   *1-2*

    'ஏடவிழ் மாமலையா ளொருபாக மமர்ந்தடியா ரேத்த'   *1-3*

    'ஆடிய எம்மிறையூர் புகலிப் பதியாமே'   *1-4*

முதலடியு மூன்றுமடியும் ஐஞ்சீர் பெற, ஏனையடிகள் ஒருசீர் குறைந்தன.

105) —— 108) இவையுமது.

109) வஞ்சி விருத்தம்: அடிதோறும் முதலிருசீர் விளச்சீர்; மூன்றுஞ் சீர் காய்ச்சீர்.

110) —— 115) இவையுமது. மூன்றுஞ்சீராம் காய்ச்சீர் தண்பூவாக வரினும் இழுக் காது.

116) கட்டளைக் கலித்துறை:

    'அவ்வினைக் கிவ்வினை யாமென்று சொல்லு மஃதறிவீர்'   *1-1*

(இறுதிச்சீர்க்கண் ஆய்தம் அலகு பெற்றது.)

    'பிறந்த பிறவியிற் பேணியெஞ் செல்வன் கழலடைவான்'   *10-1*

Exploring the meters in Tēvāram (5)

- அருத்தனை யறவனை யமுதனைநீர்
- **[viḷac cīr – viḷac cīr – kāyccīr]**
- விருத்தனைப் பாலனை வினவுதிரேல்
- **[viḷac cīr – viḷac cīr – kāyccīr]**
- ஒருத்தனை யல்லதிங் குலகமேத்தும்
- **[viḷac cīr – viḷac cīr – kāyccīr]**
- கருத்தவன் வளநகர் கடைமுடியே.
- **[viḷac cīr – viḷac cīr – kāyccīr]**

Tēvāram, 1-111
(வஞ்சி விருத்தம், வியாழக்குறிஞ்சி)

Exploring the meters in Tēvāram (6)

<அடி பொருள்="அருத்தனை, அறவனை, அமுதனை,---நீர்---">
<சீர் met="விளம்"> அருத்தனை </சீர்>
<சீர் met="விளம்"> யறவனை </சீர்>
<சீர் met="காய்"> யமுதனைநீர் </சீர்></அடி>
<அடி பொருள்="விருத்தனை, பாலனை, வினவுதிரேல்,">
<சீர் met="விளம்"> விருத்தனைப் </சீர்>
<சீர் met="விளம்"> பாலனை </சீர்>
<சீர் met="காய்"> வினவுதிரே</சீர்> </அடி>
<அடி பொருள்="ஒருத்தனை, அல்லது இங்கு உலகம் ஏத்தும்">
<சீர் met="விளம்"> லொருத்தனை </சீர்>
<சீர் met="விளம்"> யல்லதிங் </சீர்>
<சீர் met="காய்"> குலகமேத்துங்</சீர்></அடி>
<அடி பொருள்="கருத்தவன், வள நகர்---கடைமுடியே.">
<சீர் met="விளம்"> கருத்தவன் </சீர்>
<சீர் met="விளம்"> வளநகர் </சீர்>
<சீர் met="காய்"> கடைமுடியே </சீர்> </அடி> (தேவாரம் 1-111)

Exploring the meters in Tēvāram (7)

We could continue providing examples to illustrate the possible format for the tagged text which could be produced on the basis of the data already available online, but the கலித்துறை (தேவாரம், 2-1) and the வஞ்சி விருத்தம் (தேவாரம், 1-111) already given should be sufficient. It seems now preferable to point towards the wider perspective, which is to accumulate precise information on the metrical structure of global corpora (such as the தேவாரம், and other collections). And it is not necessary to emphasize how useful this kind of information can be for the editors of ancient texts.

Exploring the meters in Tēvāram (8): Statistics for Campantar

In order to hint at possible targets for those who would plan to study systematically those corpora, I include on this screen, a few lines extracted from a statistical chart which I have prepared on the basis of the 1991 volume தேவாரம் – ஆய்வுத் துணை, already mentioned. These lines indicate the total counts, inside the part which is attributed to சம்பந்தர், for some of the most frequents meters.

- kali viruttam: 144 hymns
- āciriya viruttam: 79 hymns
  - aṟucīr āciriya viruttam: 56 hymns
  - eḻucīr āciriya viruttam: 15 hymns
  - eṇcīr āciriya viruttam: 8 hymns
- kalit tuṟai: 55 hymns
- taravu koccakam: 28 hymns
- āciriya iṇaikkuṟaṭṭuṟai: 27 hymns
- kalinilaittuṟai: 19 hymns
- vañci viruttam: 13 hymns
- Vañcit tāḻicai: 9 hymns

Exploring the meters in Tēvāram (9)

- This is of course incomplete, and I have omitted in this enumeration the meters that are the least frequent. I have also not compiled the indications given in தேவாரம் ஆய்வுத் துணை concerning the patterns, such as "மா விளம் விளம் விளம் விளம்" (தேவாரம் 2-1) or "விளம் விளம் காய்" (தேவாரம் 1-111).

- A reasonable target would be to try to create a parser which could automatically detect such patterns, and to compare its output with the descriptions found in தேவாரம் ஆய்வுத் துணை.

- It should also be noted that there does not exist complete unanimity among traditional scholars concerning the meters of all the hymns. See for instance the difference of opinion between T.V. Gopal Iyer and other scholars concerning தேவாரம் 1-133

Exploring the meters in Tēvāram (10) [conflicting interpretations]

Project Madurai

(& Thevaaram.org)

<அடி **meter**="கலித் துறை">
<சீர்>வெந்தவெண்</சீர்>
<சீர்>பொடிப்பூசு</சீர்>
<சீர்>மார்பின்விரி</சீர்>
<சீர் > நூலொருபால்</சீர்>
<சீர் >பொருந்தக்</சீர்>
</அடி>

Tēvāram Āyvut tuṇai (p.386)

<அடி **meter**="தரவு கொச்சகம்">
<சீர்>வெந்தவெண்பொடிப்</சீர்>
<சீர்> பூசுமார்பின்</சீர்>
<சீர் > விரிநூலொரு</சீர்>
<சீர்>பால்பொருந்தக்</சீர்>
</அடி>

Towards electronic critical editions (1)

- If we suppose fulfilled the program which has just been sketched, what would remain to be done, in order to have an edition of the தேவாரம் which would incorporate other components of what is known about its text, and which can be found for instance inside the தேவாரம் ஆய்வு துணை volume?

- The next step would be to try to give a full presentation of all the **variant readings** which have been found inside தேவாரம் manuscripts and printed editions and which are listed on pages 425-536 of the தேவாரம் ஆய்வு துணை volume. I have not made a very precise count, but the number of variants listed is probable not far from 4000.

Towards electronic critical editions (2)

- For instance, the two verses that we have used as examples for illustrating the use of metrical structure tags both have variant readings (பாடபேதங்கள்).

- In the case of தேவாரம் 1-111, there is one variant (see தேவாரம், ஆயவுத் துணை, p.439, item 1196): some sources read அமுதனை where other sources read அமுதினை.

- In the case of தேவாரம் 2-1, there are three variants (see தேவாரம், ஆயவுத் துணை, p.442, item 1470):
  - some sources read « செந்நெல் » where other sources read « செந்நல் »
  - some sources read « முடிதோய் கழலீர் » where other sources read « தொழு பொற் கழலீர் »
  - some sources read « பின்னு சஞ்சடை » where other sources read « மின்னு சஞ்சடை »

Towards electronic critical editions (3)

The TEI P5 (Chap.12, « Critical Apparatus ») guidelines provide a module for dealing with such situations, but I shall not deal extensively with it here and shall give only one example (drawn from TEI-P5, p.393) illustrating the variation among sources for Chaucer's *Wife of Bath's Prologue*.

```
<l n="1">Experience though noon Auctoritee
<app>
    <lem wit="#El">Experience</lem>
    <rdg wit="#La">Experiment</rdg>
    <rdg wit="#Ra2">Eryment</rdg>
</app>
</l>
<l>Were in this world ...</l>
```

Towards electronic critical editions (4)

- The <**rdg**></**rdg**> tags used in the preceding screen indicate variant readings, which differ from the main reading, preferred by the editor and indicated by the <**lem**></**lem**> tag.

- The readings are attributed to witnesses, indicated by the attribute "wit", identified by a short designation, such as "E1", "La", etc.

- The information provided is that there are three different beginnings found for this poem in the sources:
  - "Experience though noon Auctoritee" (acc. to 1st witness)
  - "Experiment though noon Auctoritee" (acc. to 2nd witness)
  - "Eryment though noon Auctoritee" (acc. to 3rd witness)

- A similar strategy can of course be adopted for presenting the variants in தேவாரம் 2-1 (See next screen):

Towards electronic critical editions (5)

<**lg**>
<**l**>செந்நெ லங்கழ னிப்பழ னத்தய லேசெழும்
<**app**>
<**lem** wit="#1 #3 #4 #7 #9 #10 #11 #12"> செந்நெ</**lem**>
<**rdg** wit="#5 #58 #67"> செந்ந</**rdg**>
</**app**></**l**>
<**l**>புன்னை வெண்கிழி யிற்பவ எம்புரை பூந்தராய்</**l**>
<**l**>துன்னி, நல்லிமை யோர்முடி தோய்கழ லீர்சொலீர்</**l**>
<**app**>
<**lem** wit="#1 #3 #4 #5 #7 #9 #10 #11 #12"> செந்நெ</**lem**>
<**rdg** wit="#7">செந்ந</rdg>
</**app**></**l**>
<**l**>பின்னு செஞ்சடை யிற்பிறை பாம்புடன் வைத்ததே?</**l**>
<**app**>
<**lem** wit="="#1 #3 #4 #5 #7 #9 #10 #11 #12"> பின்னு </**lem**>
<**rdg** wit="#7"> மின்னு </**rdg**>
</**app**></**l**></**lg**>

Short conclusion

- We have now reached a stage where the natural continuation is to implement the suggestion put forward in the presentation. As I indicated in the initial section, one of my goals is to improve the presentation of the data which is found in the *Digital Tēvāram* CD-ROM (see Bibliography). The first improvement should consist in adding the tagging for the metrical structure. The second improvement should deal with the problem of variants.

- However, this type of work naturally leads to the desire for collaboration with all those who are interested in the handling and tagging of natural language digital corpora, and notably in the building of parsers for validating different aspects of the tagging.

- I would be glad to receive feedback and suggestions and welcome all proposals for collaborative work. Thanks for your attention.

Bibliography

- 1984, T.V. Gopal Iyer, *Mūvar Tēvāram (vol. 1, Ñāṉacampantar),* Publications de l'Institut Français d'Indologie (PIFI) 68-1, Pondicherry

- 1985, T.V. Gopal Iyer, *Mūvar Tēvāram (vol. 2, Appar and Cuntarar),* Publications de l'Institut Français d'Indologie (PIFI) 68-2, Pondicherry

- 1991, T.V. Gopal Iyer, தேவாரம், ஆய்வுத்துணை *(Tēvāram, Etudes et glossaires tamouls),* Publications de l'Institut Français d'Indologie (PIFI) 68-3, Pondicherry

- 2005, Gopal Iyer, T.V., [Kōpālaiyar, Ti. Vē.], *Tamiḻ Ilakkaṇap Pērakarāti*, 17 volumes. Tamiḻmaṇpatippakam, Chennai. [See especially vol. 14 & 15: Poruḷ: yāppu]

- 2007, Digital Tēvāram [KAṆIṈIT TĒVĀRAM], with the complete English gloss of the late V. M. Subramanya Ayyar, Chevillard Jean-Luc (General editor) and S.A.S. Sarma (Associate editor), Collection Indologie – 103, Institut Français de Pondichéry & École Française d'Extrême-Orient. [CD-ROM].