Online Bilingual Thesaurus for Subjects in the Humanities: A Case Study

A. Neelameghan and K. S. Raghavan

Documentation Research and Training Centre,
Indian Statistical Institute, Bangalore 560059, Karnataka, India
anm2002@vsnl.net, raghavan@drtc.isibang.ac.in

Abstract. A thesaurus is a useful knowledge organization tool. Growing globalization of information exchange calls for multi-lingual inter-disciplinary thesauri for use in information processing and retrieval. This paper discusses the authors' experiences of designing and developing online bilingual thesauri for subjects mainly in the humanities: Tamil-English thesaurus for 'classical Tamil studies' and compares with earlier work on Farsi-Vedic Sanskrit-English thesaurus for 'religious mysticism.' Building corpus of terms, selection of terms, identification of relationships among concepts, structure and hyper-linking for surfing, special problems of synonyms and homonyms, determining equivalence and near-equivalence of terms in Tamil and English, Indexing of Tamil and English terms and searching are discussed. The formation of dynamic webs of related terms through hyper-linking is mentioned.

Keywords: Bilingual Thesaurus; Online Thesaurus; Knowledge organization; Humanities

1   Introduction

For more than 150 years librarians have addressed issues of information storage and retrieval and have developed tools that are effective in handling traditional materials. There are deliberations on the future of traditional tools in the context of networked resources and the nature of information retrieval on the Web. Many projects are following standards-based approaches building upon terminology and knowledge organisation systems. Concurrently, within the Web community, there has been growing interest in vocabulary-based techniques, with the realization of the challenges posed by Web searching and retrieval applications. This is evident from the provision made in metadata standards including Dublin Core for descriptors from controlled vocabularies such as thesauri. Harpring (1999) gives an overview of the Getty's vocabularies with examples of their use in Web retrieval interfaces and collection management systems. Ontologies incorporating thesauri or related semantic models underpin ongoing projects in information retrieval in the digital environment. Several types of devices are being used to facilitate knowledge

organization (KO) and vocabulary management in information processing, organization, search and retrieval. Such devices include classification schemes, taxonomies, thesauri, ontologies, concept maps, semantic maps, lexicons (termnets, framenets), etc. In practice even a combination of such devices may be used.

An information retrieval thesaurus designed to assist KO in information systems, information indexing and retrieval differs from a conventional language thesaurus, such as Roget's *Thesaurus of the English Language*. The latter designed for a particular language is essentially a dictionary of synonyms and near-synonyms in that language. It helps in selecting an appropriate term for a given concept for which the user may have only a vague idea about a suitable term. A thesaurus for KO, on the other hand, is used for vocabulary management, e.g. in information systems: to enable the use of standard terms (for given concept/s) in indexing, and in user interfaces for formulating search expressions and search strategies by facilitating navigation between related terms (Narrower, Broader and other Related terms). Such a thesaurus can be used online or offline both at the time of data entry and indexing and for formulating search expressions for information retrieval by end user or a search intermediary. Use of such devices is known to result in improving recall and precision performance.

The growth of information resources in different languages and the resulting growth of multilingual databases have generated a need for tools that facilitate cross language retrieval. Multilingual tools for KO, such as, multilingual thesauri, are finding wider applications and becoming increasingly more useful as diverse groups "from different cultural and linguistic backgrounds seek access to diverse pieces of information" (Jorna and Davies, 2001) and information professionals have to work with information systems, sources and users with diverse language and cultural backgrounds. The proposed IFLA *Guidelines for Multilingual Thesauri* is a clear indication of this growing interest. All these developments characterize increasing globalization in almost all spheres of human activity.

2   Tamil-English Bi-Lingual Thesaurus

The Government of India recently accorded Tamil the status of a Classical language. The Central Institute of Indian Languages, Mysore has initiated a digital library project under the Central Plan Scheme for Classical Tamil (Sharada and Manjunath, 2006). In this context the need for a machine-readable bilingual thesaurus for vocabulary control, knowledge organization and related processes was felt. A project was therefore initiated to design and develop such a bilingual thesaurus (Tamil and English) covering the domain of ancient Tamil Studies. This paper is based on this ongoing project.

In an earlier paper we presented a report on the development of an online multi-lingual multi-faith thesaurus, F-THES, covering the subject Religious Mysticism (Neelameghan and Raghavan, 2005). In building F-THES, the approach adopted was to start with Sufi terms (in Farsi) and seek corresponding terms for these in English and Vedic Sanskrit. Briefly, the thesaurus record for each Farsi descriptor included the Scope Note (SN), any Synonyms, BT, NTs and RTs and the equivalent or near-equivalent terms in Vedic Sanskrit and English. The BTs, NTs, RTs and the equivalent terms in Vedic Sanskrit and English are hyper-linked to the respective thesaurus records. The structuring, linking, and display techniques adopted in F-THES were aimed at achieving an indirect integration of three language thesauri (Farsi, Vedic Sanskrit, and English) for Religious Mysticism.

2.1 Subject and Languages

This paper reports work related to the development of a bi-lingual (Tamil-English) machine-readable thesaurus for ancient classical Tamil studies (TAMTH). The time-line for the thesaurus is terminology of the Sangam period and earlier. In reality, however, this time period did not restrict the coverage of the thesaurus, as there are indeed a large number of works of later periods that comment on, analyze, evaluate or make comparisons of, works of the classical periods. In terms of subject coverage practically all subjects are covered. The following table summarizes the principal differences between the present project and F-THES.

	Subject Coverage	*Time / Period*	*Structure & Presentation*
F-THES	Religious Mysticism	No period restriction	Structure defined to generate independent language thesauri, if required; Context specifying elements used only occasionally
TAMTH	Entire universe of subjects	Sangam Period	Structure based on Tamil terms as the base / source (descriptor) with corresponding terms in English language; Context specifying elements used for every Descriptor

Table1. Differences Between F-THES & TAMTH

2.2   Sources of Terms

Indexing of literature using the thesaurus was not a part of this project; it was decided to adopt the deductive method. Given the fact that there is no information retrieval thesaurus for the domain in English or Tamil, the thesaurus had to be built using a '*bottom up*' approach starting with Tamil

terms. The approach was to collect a broad set of relevant candidate terms in Tamil using several print-on-paper and electronic sources:

1. *Tamil lexicon. Published under the authority of the University of Madras. Madras: University of Madras; 1924-1939. Reprint 1982. v. I-VI + Supplement.* [An online version is also available]
2. *Cologne online Tamil lexicon. [Based on Tamil Lexicon and supplement, 1924-1939]. http://webapps.uni-koeln.de/tamil/ (COTL)*
3. *Commemorative bibliography of the first 1008 books published by the South India Saiva Siddhanta Works Publishing Society, Tinnevelly Limited / By S.R. Ranganathan and R. Muthukumaraswamy. Tirunelveli: The Society; 1961.*
4. *Periya puranam: a Tamil classic on the great Saiva Saints of South India / By Sekkizhaar. Condensed English version by G. Vanmikanathan and N. Mahalingam. Madras: Sri Ramakrishna Math; [1985].*
5. *A catalogue of vaisnava literature (on microfilms in the Adyar Library, the Bodleian Library and the American University Library) / By Charles S.J. White.*
6. *Sub-forms of Tamil poetry and their classification / By S.R. Ranganathan and V.Thillainayagam. Annals of Library Science, 10(3-4); 1963; 175-185*
7. *WordNet 2.1 (online)*

An important issue relates to script and transliteration. At present Tamil terms transliterated into Roman script are being used in the thesaurus. The conversion of transliterated terms to the Tamil script is important and will be taken up in the later stages of the thesaurus project. It was, however, important to decide on a standard transliteration scheme in view of the fact that there are several schemes in use. It was decided to use the transliteration scheme employed in COTL for the descriptor field. This standard is based on the transliteration used by the *Tamil Lexicon* (University of Madras) but without diacritical marks. An examination of COTL indicates that two types of transliteration are used; one involving use of only Roman alphabets and the other involving the use of Roman alphabets as well as Arabic numerals.

Example:
(1) pOrkkezuvajnci (2) po1rkkel6u-van1ci

The first type involving use of only Roman alphabets is used in the Descriptor field; however, the second form may occur in the SN field. To facilitate search by both forms the Descriptor field as well as the SN field have been indexed. An examination of the conventions with regard to transliteration of Tamil revealed that at least one other scheme is being used. The University of

Chicago has made available on the web a number of lexical tools for South Asian languages. Among these is the earlier edition of the University of Madras Tamil Lexicon. The transliteration used by the University of Chicago avoids use of upper case letters and is, in many respects different from the one adopted by COTL. This form of transliteration has also been used in the present project as an additional data field so as to facilitate linking to the online lexicon. (See also section 5 on linking transliterated descriptor to corresponding term in Tamil script).

2.3 Data / Term selection

Initial examination indicated that over 1,25,000 (one hundred and twenty-five thousand) Tamil terms may constitute the corpus. In the actual creation of the thesaurus records, two distinct approaches were possible: To create records by entering terms in 'A to Z' sequence using a source such as the *Tamil Lexicon*. This was found to be inconvenient for identifying BT, NT, RT, USE, and UF term(s) for each descriptor. A subject-wise approach was found more helpful. Based on an examination of the corpus of terms, the major disciplines / basic classes have been identified and defined according to the schedule of Basic classes of Colon Classification. The examination further revealed that Tamil terms in the physical and life sciences were relatively fewer compared to those in the Humanities and Social Sciences. Currently there are about 50000 records in the thesaurus. The major disciplines and sub-disciplines are listed below. It is possible that a more detailed structuring of subjects may be necessary when new records are added to the thesaurus.
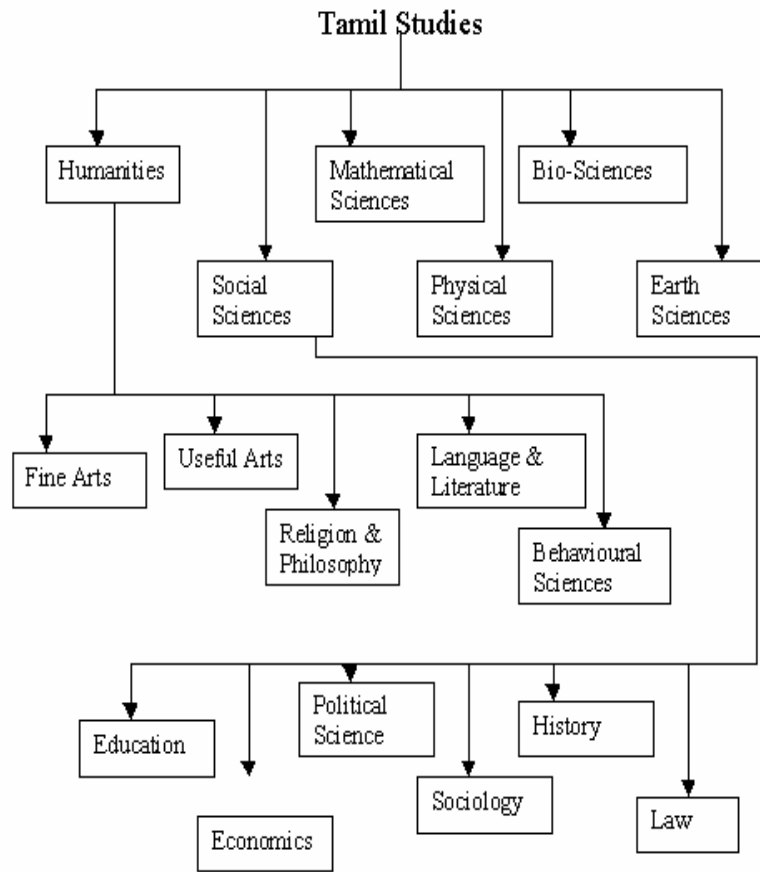
Fig. 1: An Overview of the Disciplines / Sub-Disciplines

2.4

Structure

Most of the terms denoted concepts in Humanities and Social sciences. In this circumstance the problems usually encountered in vocabulary management in culture-specific domains need to be handled. In practice a user may use a non-Descriptor; such a term is cross-referred to the appropriate Descriptor.

Example:
ArEcikam (plantain)   vAzai (plantain)
   US vAzai (plantain)   UF ArEcikam (plantain)
Where US = USE   where UF = Used For

The Scope Note (SN) provides the meaning(s) of the term, the context(s) of its use(s), alternative transliteration form, etc.

Example:
aNukkirakam (divine grace)
  SN: Grace; mercy; show grace; to liberate souls from bondage. A
              paNca-kiruttiyam
  BT: paNcakiruttiyam (God's functions)
  RT: aNukkirakittal (bestow grace)
  RT: aruL (benevolence)
  RT: tiruvaruL (divine grace)
    Etc.

The thesaurus, in many cases links to more than one level of hierarchy.
Examples:
vAzai (plantain)
  SN Plantain tree
  UF Agkucam (plantain)
    Arampai (plantain)
    ArEcikam (plantain)
  BT tAparavastu (plant)
   BT2 tAparanUl (botany)
  NT aTiccIppu (plantain vartiety)
    cami (plantain variety)
    ilaivAzai (plantain variety)
ilaivAzai
  SN A kind of plantain that does not yield fruit; a kind of plantain that
     yields stony fruits
  BT vAzai (plantain)
   BT2 tAparavastu (plant)
    BT3 tAparanUl (botany)

The standard guidelines for determining hierarchical relationships: genus-species, parent-sibling, organizational hierarchy (master-servant), whole-part, etc have been adopted. However, it was found that non-hierarchical associative relationships are more numerous and more difficult to determine. Nevertheless, non-hierarchically related terms are useful in enhancing retrieval effectiveness. The typology of RT relationships developed in the early 1970s at the Documentation Research and Training Centre, ISI, Bangalore, deemed to be a pioneering work, and updated recently (Neelameghan and Raghavan, 2005; 2006) has been used.

2.5 Database and Software
TAMTH was built using WINISIS 1.5, considering its features and capabilities for database building, indexing, retrieval and hypertext linking and the facility

to export records to XML format. Some large multilingual thesauri, such as, AGROVOC of FAO, MACROTHERSAURUS (originally developed by OECD) and others have been developed using the software. The fields and tags in the TAMTH are given below. The fields in italics are for future use.

Field Tag	Field Name	Repeatability	Sub-field Delimiters
1	Descriptor		a
11	Descriptor1		a
2	SN	Repeatable	
3	US	Repeatable	
4	UF	Repeatable	
5	BT	Repeatable	
6	NT	Repeatable	
7	RT	Repeatable	
71	EQE	Repeatable	
72	EQT	Repeatable	
73	EQK	Repeatable	
90	BS	Repeatable	
91	Type	Repeatable	
92	IN	Repeatable	
93	CN	Repeatable	
99	Remarks	Repeatable	
900	Links		abcdefg

Table 2. Data Fields

The descriptor field contains the Tamil term in the transliterated form. The process of identifying equivalent terms in English is in progress and the BT, NT and RT terms in Tamil are being added to the records as necessary or as they become available.

3   Issues in Bilingual Thesaurus

The major problems and issues encountered in the design and development of TAMTH can be grouped into two broad categories:

   a) Issues related to concepts and their representation
   b) Issues related to relationship between concepts

3.1   Lack of Terms

Often concepts in Humanities are unique to one or a few communities – normally defined by language, religion, geography, occupation, etc, characteristics. A frequent problem in building this bilingual thesaurus, therefore related to non-availability of appropriate terms in English. Several concepts that were subjects of discourses in Tamil Studies there were no corresponding terms in the English language. Often, therefore, such concepts

represented by a single word or phrase in Tamil had to be represented using a phrase in English (or even a sentence)

Example:
cAttunARRu = Young plants planted in place of the dead ones

3.2 Synonyms

Another issue related to the use of several synonyms or near-synonyms in Tamil to denote a concept. The decision relating to the choice of descriptor was not always easy.

Example:
Elam (cardamom))
SN Cardamom plant, elettaria cardamomum; cardamom; a
   spice
     UF ilAjncali (cardamom)
     UF ilAjnci (cardamom))
     UF kALintam (cardamom))
     UF kaNmali (cardamom)
     BT tAparavastu (plant)
        BT2 tAparanUl (botany)
IlAjncali (cardamom))
     US Elam (cardamom))
IlAjnci (cardamom)
     US Elam (cardamom)

Similarly, from each of the other UF terms.

3.3 Homonyms

Many Tamil terms were homographs with two or more meanings.

Examples:
Etu = Petal. Flower. Eyelid. Strip of Palmyra leaf used for writing. Book of Palmyra leaves. Leaf of a book. Section of a plantain leaf cut near the base for use as plate. Cream. Body. Greatness. Excellence. Fault. Blemish.
iTimpam = Baby, small child; misery; spleen; egg of birds; castor plant).
This required extensive use of context-specifying term(s) (individualizing elements) in different contexts of the Descriptor.

Examples:
ETu (plant), ETu (cream), ETu (quality), ETu (leaf), ETu (greatness), etc.
iTimpam (baby); iTimpam (misery); iTimpam (spleen); iTimpam (egg); iTimpam (castor)

3.4 Relationships

The handling of relationships between concepts presented several problems. Here we discuss only the predominant types and issues occurring in the present thesaurus.

3.4.1 Multiple RTs and NTs

There were several situations in which a large number of terms related to a descriptor had to be handled.

Examples:
The Tamil descriptor pirapantam (anthology) has over 96 NTs.
The Tamil descriptor aRul (benevolence) has over 40 RTs.
The standard practice in thesaurus construction is to list all NTs and RTs. Further, NTs to the same descriptor are normally treated as RTs to one another. Let us consider an example:
maturaikkANci (anthology)

   SN A collection of ten ancient Tamil poems
   NT tirumurukARRuppaTai (poem)
   NT porunARRuppaTai (poem)
   NT perumpANARRuppaTai (poem)
   NT ciRupANARRuppaTai (poem)
   NT mullaippATTu (poem)
   NT maturaikkAjnci (poem)
   NT neTunalvATai (poem)
   NT kuRijncippATTu (poem)
   NT paTTiNappAlai (poem)
   NT malaipaTukaTAm (poem)

In the conventional approach to building thesauri, the standard practice is to create a record for each NT that will take the following form:
tirumurukARRuppaTai (poem)

SN One of the ten ancient Tamil poems in the anthology
  maturaikkANci
   BT maturaikkANci (anthology)
   RT porunARRuppaTai (poem)
   RT perumpANARRuppaTai (poem)
   RT ciRupANARRuppaTai (poem)
   RT mullaippATTu (poem)
   RT maturaikkAjnci (poem)
   RT neTunalvATai (poem)
   RT kuRijncippATTu (poem)
   RT paTTiNappAlai (poem)
   RT malaipaTukaTAm (poem)

Similar records will normally be created for each of the other NTs of maturaikkANci (anthology). In the context of the present thesaurus, this resulted in a very large number of linked entries. With a view to reducing the size of the thesaurus without compromising on the linking, in all cases involving five or more NTs and / or RTs to a descriptor, the full list of NTs and / or RTs as the case may be, is given in one schedule under the descriptor and references are given to this schedule from each of the NTs or RTs for that descriptor. This was necessary considering the fact that the thesaurus is growing and in future more terms will get added to the schedules of NTs and RTs for a descriptor. For example the above schedule for tirumurukARRuppaTai (poem) will now be:

tirumurukARRuppaTai (poem)
For a list of related poems
SEE maturaikkANci (anthology)

Clicking on maturaikkANci (anthology) will link the user to the schedule under this term.

3.4.2 Hierarchical Relationships
Several kinds of hierarchical relationships were encountered.
Original/Variant-Relationship
kavitai (poem)
  NT catakam (poem of 100 stanzas)
   kaviyam (epic poem)
   kuRavajnci (dramatic poem)
   mAlai (chain / garland type)
   paLLu (descriptive / expository)
   and others
Whole/Part-Relationship
kavitai (poem)
  NT muNNaTi (first line of poem)
   muri (curitakam at close of poem)

3.4.3 Lateral Relationship (RT) (Non-hierarchical Associative Relationship)
Examples of some of the lateral relationships encountered including some that had not been defined in the taxonomy of lateral relationships referred to above, are given below:
  Creator/Created entity-Relationship
  CivappirakAcamuNivar (author)
   SN Author of a didactic poem of 40 stanzas, naNNeri
   RT naNNeRi (poem)
Entity/Example/case-Relationship
  paLLu (descriptive poem)

RT <u>mukkUTaRpaLLu (poem)</u>
Entity/Content/theme-Relationship
  nanti III (theme)
    SN Theme of the poem nantikkalampakam, 9th century
    RT <u>nantikkalampakam (poem)</u>
Entity/Attribute-Relationship
  yAppANantam (defect)
    SN A defect in composition in which the name of hero of a
     poem is clumsily set in the midst of attributes
    RT kavitai (poem)
Entity/Derived entity-Relationship
  ativaravu (mnemonic)
   SN Mnemonic formed of the initial syllables of stanzas of a poem
   RT <u>aTivaraiyaRai (index)</u>
   RT <u>kavitai (poem)</u>
  kavitai (poem)
    RT <u>ativaravu (mnemonic)</u>
    RT <u>aTivaraiyaRai (index)</u>

4 Index and Search

Recognizing the importance of providing for multiple approaches, both the Descriptor and SN fields are indexed. Tamil terms in COTL transliteration are arranged in a single alphabetical sequence. In the SN all words other than those enumerated in the stop-word list are indexed. The SN contains English words and Tamil words in alternative transliteration. A word extracted from the SN field is prefixed with ZZ= such that all English words and words in alternative transliteration will be filed in a single alphabetical sequence, following the Tamil terms. In addition the transliterated form of the descriptor based on the convention used by the University of Chicago (*Online Tamil Lexicon*) (Field 11 in the database) is also indexed and linked to the *Online Tamil Lexicon* (see section 5 below).

5 Linking Online Lexicon

In view of the availability of several lexical tools on the Web, an attempt was made to link the terms in the thesaurus to an appropriate online tool. The University of Chicago has made available several lexical tools especially for languages of South Asia on the Web. One such tool is an edition of the University of Madras Tamil Lexicon. It was thought that linking the terms in the thesaurus to such a tool would serve two important purposes:

- It would allow users of the thesaurus to view the term in Tamil script;
- It would also provide more detailed information about the term including its various meanings, etc.

An example is given below to illustrate how this works.
Clicking on the term (Descriptor1) in the thesaurus will open the home page of the *Tamil Online Lexicon*. (See Fig. 2)
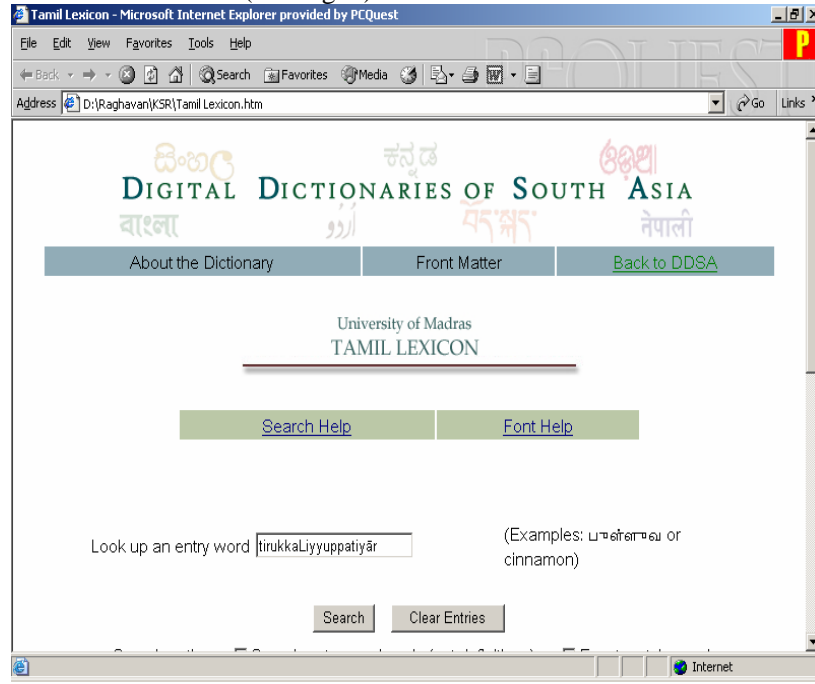

Figure 2

The search term can be entered in the search box and a search executed to display the relevant portion of the Lexicon. A search for '***tirukkaliyyuppatiyar***' will retrieve all the records in the lexicon in which the term occurs either as an entry word or in the description (text) (see Figure 3 below).
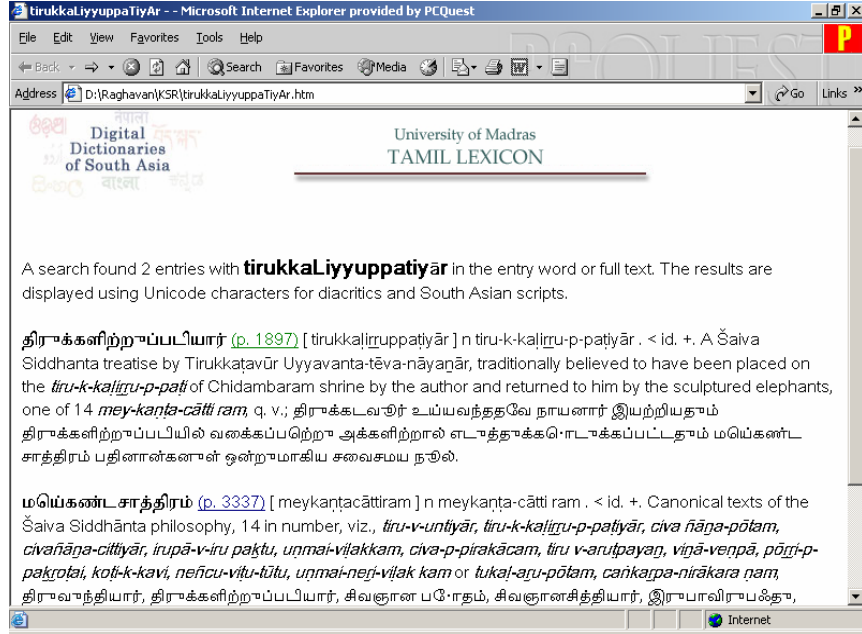
Figure 3

Clicking on the page number of the appropriate record will retrieve the complete relevant page of the *Lexicon* (Fig. 4.)
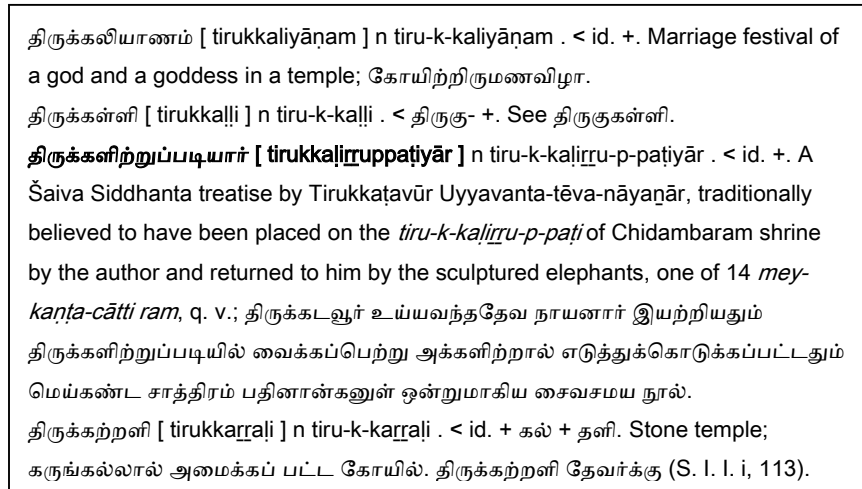
திருக்கலியாணம் [tirukkaliyāṇam] n tiru-k-kaliyāṇam . < id. +. Marriage festival of a god and a goddess in a temple; கோயிற்றிருமணவிழா.

திருக்கள்ளி [tirukkaḷḷi] n tiru-k-kaḷḷi . < திருகு- +. See திருகுகள்ளி.

திருக்களிற்றுப்படியார் [tirukkaḷiṟṟuppaṭiyār] n tiru-k-kaḷiṟṟu-p-paṭiyār . < id. +. A Šaiva Siddhanta treatise by Tirukkaṭavūr Uyyavanta-tēva-nāyaṉār, traditionally believed to have been placed on the *tiru-k-kaḷiṟṟu-p-paṭi* of Chidambaram shrine by the author and returned to him by the sculptured elephants, one of 14 *mey-kaṇṭa-cātti ram*, q. v.; திருக்கடவூர் உய்யவந்ததேவ நாயனார் இயற்றியதும் திருக்களிற்றுப்படியில் வைக்கப்பெற்று அக்களிற்றால் எடுத்துக்கொடுக்கப்பட்டதும் மெய்கண்ட சாத்திரம் பதினான்கனுள் ஒன்றுமாகிய சைவசமய நூல்.

திருக்கற்றளி [tirukkaṟṟaḷi] n tiru-k-kaṟṟaḷi . < id. + கல் + தளி. Stone temple; கருங்கல்லால் அமைக்கப் பட்ட கோயில். திருக்கற்றளி தேவர்க்கு (S. I. I. i, 113).

Figure 4

6   Concluding Remarks and Future Work

Bilingual thesauri in culture-specific domains such as the TAMTH for classical Tamil studies discussed above present certain problems.

a) First and foremost is the impracticality of explicitly displaying, under every descriptor, the complete range of RTs and NTs.

b) The need to handle a large number synonyms and homonyms (in Tamil, in this thesaurus) is another important issue that has to be handled.

c) It is inevitable to adopt a '*bottom-up*' approach for building such a thesaurus as this. It is difficult to realize a bilingual thesaurus as this by a merger of a Tamil thesaurus and an English thesaurus even if such thesauri were to exist.

d) The contribution that a web of RTs (see Fig 5 and Annex 1) could make in enhancing a search has already been emphasized. The facility to hyperlink makes it possible to employ appropriate tools to create maps of term clusters (and even link terms in a cluster to digital objects). These clusters are dynamic and their composition will change with the introduction of new related terms (BTs, NTs, RTs). In the Thesaurus itself these have been hyper-linked to facilitate surfing across the schedules of terms. In general RTs as they are used in thesauri represent a class of non-hierarchical relationships that have not been clearly understood. Construction of thesauri based on a clearer understanding and more standardized application of RTs should enhance retrieval. A more organized approach to application of RTs in thesaurus design has potential for search and information retrieval. It is possible that the inclusion of such semantic relationships can be made to serve in developing knowledge-based approaches to information retrieval such as suggesting search terms and various forms of query expansion query refinement.

e) Linking to appropriate online lexical tools can substantially enhance the utility and value of a thesaurus.

An issue that has often been discussed relates to the possibilities of enhancing the capabilities of a thesaurus by converting it to an ontology. Typically thesauri are based on recognition of equivalence, hierarchical and lateral

relations between concepts. The semantics of these relations are not explicitly defined. The semantics of BT/NT relations within a thesaurus is fairly straight forward and usually include the following types of relations:

- A variety / sub species (***Is-A*** relation)
- Whole-part (***Component of*** relation)

However, RT relations are more complex and include a range of different kinds of relations (see Fig. 5). With a view to
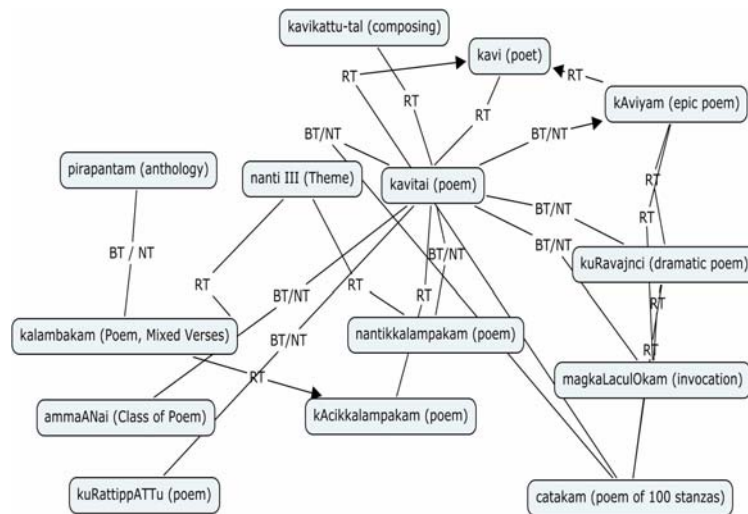

Figure 5. Web of Related Terms

define more explicitly the different kinds of RTs, a taxonomy of lateral relations has been developed (Neelameghan and Raghavan, 2005; 2006). Ontologies are believed to provide better semantic representation and machine understandable representation of knowledge. The feasibility of employing the taxonomy of lateral relations to convert the proposed bilingual thesaurus into an ontology will be explored.

References

[1] Harpring, P. (1999) "How forcible are the right words: overview of applications and interfaces incorporating the Getty vocabularies". *Proceedings of Museums and the Web 1999*, Archives and Museum Informatics.
http://www.archimuse.com/mw99/papers/harpring/harpring.html

[2] Jorna, Kerstin and Davies, Sylvie (2001). Multilingual thesauri for the modern world – no ideal solution? *Journal of Documentation*, 57(2); 284-295

[3] Neelameghan, A, and Raghavan, K.S. (2005). An online multi-lingual, multi-faith thesaurus: a progress report on F-THES. *Webology*, 2(4). http://www.webology.ir/2005]v2n4/a19.html

[4] Neelameghan, A, and Raghavan, K.S. (2005). Semantics of relationships in knowledge organization: lateral relationships. *SRELS Journal of Information Management,* 42(4); 361-382

[5] Neelameghan, A, and Raghavan, K.S. (2006). Semantics of relationships in knowledge organization – lateral relationships: update 1. *Information Studies*, 12 (2); 115-122

[6] Tudhope, Douglas [et al](2001). Augmenting Thesaurus Relationships: Possibilities for Retrieval. *Journal of Digital information*, 1. http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Tudhope/

ANNEX 1

Selecting NANTI III (THEME) from the Index will retrieve:
nanti III (theme)
     SN Theme of the poem
nantikkalampakam, 9th century
      RT kalampakam (poem, mixed verses)
      RT nantikkalampakam (poem)
Clicking on kalampakam (poem, mixed verses) will link to:
kalampakam (poem, mixed verses)
     SN A kind of poem composed of different kinds of stanzas
      BT pirapantam (anthology)
      RT ammANai (class of poem)
      RT kAcikkalampakam (poem)
      RT nantikkalampakam (poem)
Clicking on nantikkalampakam (poem) will link to:
nantikkalampakam (poem)
SN A poem on the Pallava king Nandi III, 9th century, in kilampakam
   Metre
BT kavitai (poem)
RT kalampakam (poem, mixed verses)
RT nanti III (theme)
Clicking on Kavitai will link to:
kavitai (poem)
	SN Poem, stanza, verse. For a specific poemcheck the Index by the
	    name / title of the poem, e.g.
	alagkarapajncakam (poem);
	    camuttiravilAcam (poem) For a list of all poems included in the
	    database check the Index under
	ZZ=poem
	RT catakam (poem, 100 stanzas)
	RT kavi (poet)
	RT kavikattutal (compose)
	RT kAviyam (epic poem)
	RT kogkuvENmAkkatai (epic poem)
	RT kuRattippATTu (poem)
	RT kuRavajnci (dramatic poem)
	RT mAlai (chain/garland type)
	RT magkaLaculOkam (invocation)
	RT paLLu (descriptive poem)