Morpho-Semantic Features for Rule-based Tamil Enconversion

J Balaji
Junior Research Fellow
Dept. of CSE
College of Engineering
Guindy, Anna University,
Chennai-600025

T V Geetha
Professor
Dept. of CSE
College of Engineering
Guindy, Anna University,
Chennai-600025

Ranjani Parthasarathi
Professor, Dept. of IST
College of Engineering
Guindy, Anna University.
Chennai-600025

Madhan Karky
Asst. professor
Dept. of CSE
College of Engineering,
Guindy, Anna University
Chennai-600025

ABSTRACT
This paper discusses the UNL Enconversion of Tamil sentences. The rich morphology of Tamil enables the Enconversion process to be based on morpho-semantic features of the words and their preceding and succeeding context. The use of case relation indicating morphological suffixes, POS tag and word level semantics allows the rule based Enconversion to be independent of the syntactic structure of the sentence. These UNL graphs are used to build a conceptual level index.

General Terms
Natural Language Processing, Knowledge representation, Information Extraction

Keywords
Universal Networking Language, Enconversion, Morpho-Semantic features

1. INTRODUCTION
One of the objectives of Semantic processing of Natural language is the understanding of the language and the transformation of the natural language sentence into a semantic representation. Such representations are used for incorporating semantics into language technology applications such as machine translation, text summarization, information retrieval including cross lingual information retrieval. An Interlingua, language independent framework that encapsulates the meaning of sentences in terms of concepts and relations has been the focus of researchers working in areas that deal with multiple languages like machine translation and cross lingual information retrieval [8]. One such Interlingua framework called KANT [14], which has been targeted towards technical text in controlled sub-domain. UNITRANS [15], a bidirectional system, operates cross-linguistically but still accounting for knowledge that is specific to each language.

However, Universal Networking Language (UNL) is one such Interlingua framework which was originally designed to aid machine translation [8]. UNL deals with concepts, that are represented as Universal words (UWs) and defines a set of Relations that can exist between them. A sentence in the natural language is represented by UNL as a graph in which nodes represent concepts and links represents relations. In addition attributes are also associated with both concepts and relations of the UNL graph. Each UNL concept of the graph is obtained by referring an UNL knowledge base which provides concepts for terms/words of the natural language. Each concept is also placed in a UNL specified concept hierarchy represented as constraints in the UNL KB. The UNL consortium has at present defined a standard set of 46 UNL relations which essentially represent the semantic role of each word in the sentence. This UNL based semantic representation of the sentence has been used for a number of language technology applications. The most common application is the use of UNL representation as an Interlingua for machine translation [8]. UNL representation has also been used for semantic based text summarization [11]. The most recent use of UNL representation is in building concept based indexes for multilingual search engine [6] and in mono lingual advanced search [10].

This paper describes the UNL Enconversion of Tamil sentences without considering the syntactic parsing where the UNL graph produced by the Enconverter is represented as a multi-list structure for building conceptual indices. This paper concentrates on the use of morphological as well as word based morphological and semantic features for the Enconversion process.

This paper is organized as follows. Section 2 provides a review of related work. Section 3 describes the UNL Enconversion for Conceptual Search. Section 4 describes in detail the two passes of the UNL Enconversion process for Tamil. Finally section 5 describes the performance of the Enconversion process and its role and effectiveness in building a conceptual index.

2. RELATED WORK
The UNL Enconversion is basically a semantic analysis problem where various linguistic features have been used based on the characteristics of the input natural language. The Enconversion process can basically be classified based on the type and degree of usage of various linguistic features.

The usage of the level of linguistic analysis such as use of morphology, syntax and semantics for the Enconversion depends on how the various UNL relations are characterized by the specific natural language. The English UNL Enconversion is centered on the use of a standard syntactic parser with a limited amount of morphology and semantics [8]. The French UNL Enconverter generates UNL expressions using an incremental parser which converts the expressions into a semantic graph using a rule based approach [2]. Enconversion using a multi

functional linguistic processor ETAP-3 converts Natural Language into ETAP-3's internal representation that is essentially a normalized syntactic structure (Norm SS) which is then converted to the required UNL representation [3]. Again, this approach is centered on the use of syntactic structures to aid the Enconversion process. Enconversion process for Arabic, a highly inflected language and of relatively free word nature requires strong interaction between morphological and syntactic processing. Therefore, a rule based approach for Arabic Enconversion cannot use the normal pipeline model where morphological analysis is followed by syntactic analysis for resolving ambiguity [9].

Many Indian languages do not have a rigid fixed word order structure and the amount of freedom allowed by a language differs based on the richness of the morphology. In Languages like Hindi and Bangla for which Enconversion has been done, isolated case markers along with local word grouping generally specify the case relations between the different components of the sentence. Since most UNL relations are based on case roles, the handling of this aspect differentiates the approach to Enconversion for Hindi [6].

Another approach for English to UNL Enconversion [5] uses a two stage process where the conceptual arguments are first identified in the form of semantically relatable sequences (SRS) which are potential candidates for being linked by semantic relations. These are then mapped to form a parsed output and then UNL expressions [5]. Most of the Enconverters discussed above use a rule based approach. A statistical approach using a parser with associated morphological and syntactic linguistic features has also been attempted [7].

An approach for bangla words to UNL Enconversion discussed in [8][12] dealt with the verb morphology and the use of syntax tree. Another work on bangla to UNL Enconversion proposed in [13] uses a set of morphological and semantic rules for enconversion and fits the rules in the enconverter framework obtained from UNDL foundation, for enconverting bangla sentences.

All the UNL Enconverters discussed above use a structural syntactic parser for fixed word order languages and dependency like parser for partial free word order languages.

In this work, Enconversion for Tamil which is a morphologically rich language and which has a relatively higher free word order has been described. The approach used in this work uses morphological features and word based semantic features with limited amount of context to obtain UNL expressions from Tamil sentences.

3. UNL ENCONVERSION FOR CONCEPTUAL SEARCH

In this work, the UNL Enconversion is the document processing component of a concept based search engine designed for Tamil. This enconversion process considers simple important sentence constituents obtained from complex and compound sentences. This essentially means that the UNL Enconversion process is working on part of sentences and not on the complete sentence itself. The UNL Enconverter transforms these sentence constituents into the associated set of UNL graphs. In order to facilitate the building of conceptual indices, the UNL graph is represented as a multi-list structure [11]. The document processing described here in addition to carrying out UNL Enconversion also provides additional data for the index building process. These include word based features such as POS category of UNL concept, actual word associated with the concept, its position within the document and also sentence based features such as term and concept frequency across documents, position of word within a sentence and word within a document. With the three types of indices that are built from the enconverted output, namely, the concept-relation-concept, concept-relation, and concept indices, this information helps in the concept based searching and ranking [11].

4. TAMIL TO UNL ENCONVERSION

The Tamil-UNL conversion process makes use of morphological features conveyed by the word, and in some situations, semantics of UNL concepts. In Tamil, the Enconversion process takes place in two passes. During the first pass, the possible UNL relations are identified using morphological or word features and stored in a relation vector. There are some UNL relations that can be unambiguously determined in this pass itself. However there are morphological suffixes which signal more than one UNL relation. For example, some morphological case suffixes such as ***ilirundhu*** signals multiple UNL relations such as "***frm", "plf"," tmf", "src",*** while the case suffix ***kku*** signals UNL relations such as" ***to", "plt", "tmt",*** and the case suffix ***il*** signals UNL relations" "***plc", "opl"*** [8]. In this context the relation vector will have multiple values. During the second pass, the concepts that are linked by the UNL relation are determined. The second pass also helps in resolving ambiguity among the UNL relations obtained in pass 1. The disambiguation process uses the relation vector along with co-occurrence, parts of speech, connectives and semantics of the corresponding UNL concept and the context.

As already mentioned, Tamil is a morphologically rich language. Here, case markers are conveyed by morphology of noun rather than separately as prepositions as in English or as karaks as in Hindi. Therefore, morphological analysis of the word, in addition to providing part of speech also provides case relation information about the nouns. Verb morphology provides the tense, aspect, mood, person, number, and gender of the verb. Normally the person, number and gender of the verb are expected to match with the noun acting as subject/agent of the sentence. The richness of the noun morphology allows noun phrases indicating different cases to occur in any position in the sentence. Normally, Tamil follows a subject-verb-object form.

Information conveyed in other languages by position of sentence constituents and syntax, is conveyed by morphological features in Tamil. In this paper, this characteristic of the Tamil language is exploited in the Tamil to UNL Enconversion process. The details of this are presented below.

4.1 UNL Related Information obtained during first pass

The first pass of the rule based Enconverter, uses word based features such as morphological suffixes or the presence of certain specific words to determine UNL relations. In this pass, word based features required for UNL enconversion are obtained. The presence of an UNL relation is signaled by morphological suffixes such as case suffixes, adjectival suffixes and adverbial suffixes or by the presence of certain standalone

connective words. However, morphological suffixes indicating case relations can specify more than one UNL relation which is indicated in the first pass by an associated relation vector. The details of this are presented below.

4.1.1 Signaling UNL relation by case suffixes of nouns

When case suffixes are attached to the noun, certain UNL relations are signaled unambiguously. Most often, the noun to which the case suffix is attached is one of the words taking part in the UNL relation. The UNL relation obtained through the case suffixes of nouns are "***pos", "ben",*** and ***"obj".*** This is illustrated with an example.

As shown in Fig.1, in the word pazhaththai (W_i) which has the case ending ***ai*** (Accusative case marker), the case ending indicates that W_i should be related to some other concept in the sentence through the UNL relation "***obj"***.

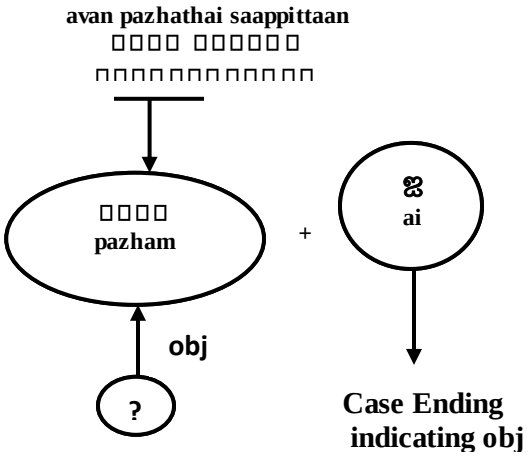

Fig. 1 Example for relation "*obj"*

Similarly, the UNL relation, " ***ben"*** can be obtained by the case ending ***ukkaaga*** associated with the word W_i and "***pos"*** relation can be obtained by the case endings ***udaya, in*** and ***athu*** (Genetive case marker) associated with the word W_i.

4.1.2 Signaling of UNL relations by Adjectival and Adverbial Suffixes

The semantics "***mod"*** and "***man"*** can be determined unambiguously by the Adjectival suffix (***aana, iya***) and Adverbial suffix (***aaga***). Again the respective adjective or adverb forms one word taking part in the UNL relations "***mod"*** and "***man"***.

Fig 2 shows the signaling of "***mod"*** relation using adjectival suffixes.

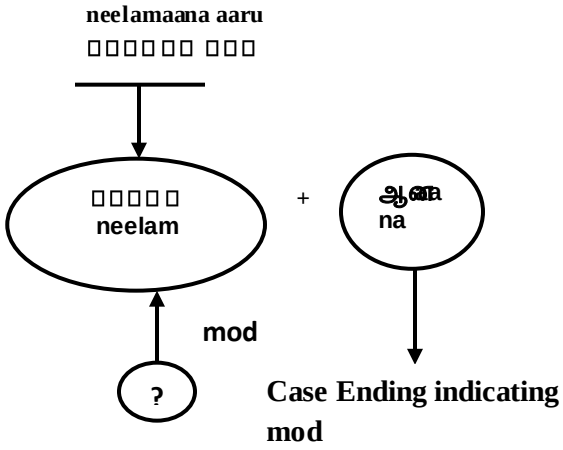

Fig. 2 Example for relation "***mod"***

Signaling of the "***man"*** relation by the adverbial suffixes is shown in Fig 3.

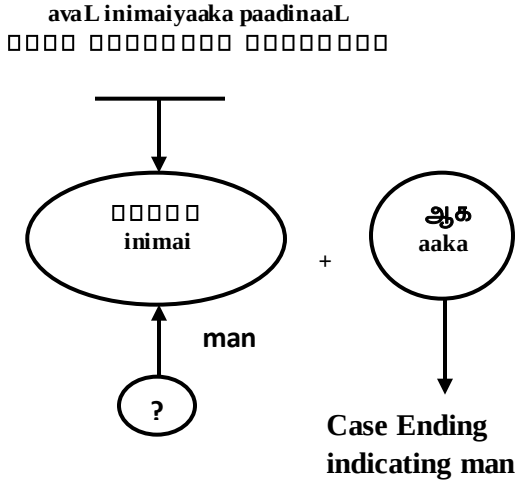

Fig. 3 Example for relation "***man"***

4.1.3 Signaling of UNL relations by Connectives

Similarly certain connectives signal certain semantic relations unambiguously. However, in these cases the connective word itself does not take part in the UNL relation. The connectives ***maRRum, mattumallaamal, aagiya, pondra, mudhaliya, vazhiyaaka, allathu, enpathu,*** and ***idaiye*** come under this category .

Let us consider the example given in Fig 4. The connective word ***maRRum*** indicates the UNL relation "***and"*** between two different concepts. The set of UNL relations that comes under this connective category is mentioned in Table-1.

Chennai maRRum maduraiyil koyilkaL uLLana

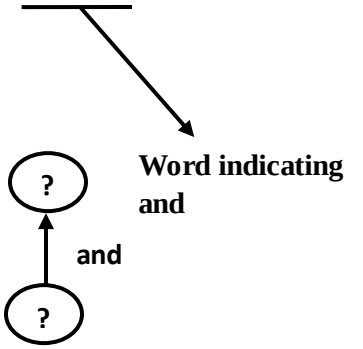

Fig. 4 Example for relation "*and"*

Certain UNL relations such as "***iof"*** and "***nam"*** connect more than two concepts. The connective ***aagiya*** identifies such cases. The example in Fig 5 illustrates such an UNL relation.

kaveri vaigai ponni paalaaRu aakiya nadhikaL

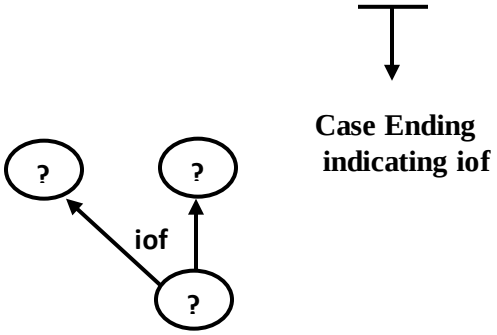

Fig. 5 Example for relation "***iof"***

4.1.4 Signaling of multiple UNL relations

Certain case suffixes attached to nouns signal more than one UNL relation. Suffixes of this category are given below.

The case suffix ***ilirundhu*** associated with word $\mathbf{W_i}$ indicates multiple UNL relations such as "***frm", "plf", "tmf", "src".*** These are stored in the relation vector. An example of this case is shown in Fig.6 below. The ambiguity will be resolved during the second pass using the semantics of the word $\mathbf{W_i}$. If the semantics of $\mathbf{W_i}$ is ***icl>time***, then the UNL relation would be "***tmf"***. If the semantics is ***icl>place,*** then the UNL relation would be "***plf"*** and so on.

kaalayilirundhu poojai nadanthukondirunthathu

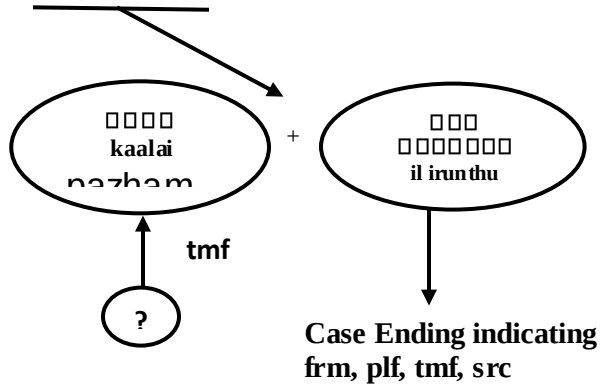

Fig. 6 Example for case suffix "***ilirundhu"***

Similarly, the case suffix ***udan*** or ***odu*** associated with word $\mathbf{W_i}$ indicates multiple UNL relations such as "***cag", "cob"*** and "***ptn".*** Fig. 7 shows an example

raman seethaiyudan kaattukku sendRaan

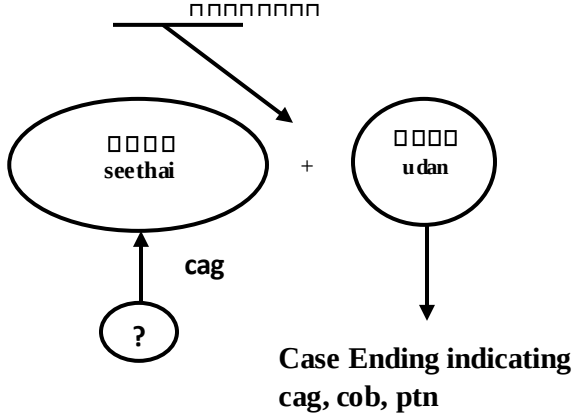

Fig. 7 Example for case suffix "***udan"*** or ***" odu"***

Fig. 8 illustrates the signaling of the UNL relation "***agt", "ins"*** and "***con"*** for the case suffix ***aal*** associated with the word $\mathbf{W_i}$.

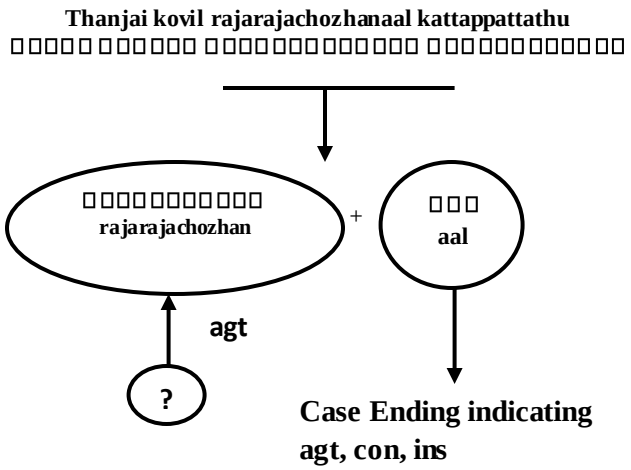

Fig. 8 Example for case suffix "aal"

4.2 UNL Relation Ambiguity Resolution and Construction of UNL graph

The second pass of the UNL enconverter performs two tasks, using additional information such as POS and semantic category of the UNL relation signaling word $\mathbf{W_i}$, and/or its context $\mathbf{W_{i\pm k}}$. In the context where UNL relation has been unambiguously determined by the first pass, there is a necessity to find the two concepts that are involved in the UNL relation in order to build the UNL graph. However, when there is an ambiguity regarding the UNL relation, this ambiguity has to be resolved before the UNL graph is constructed.

4.2.1 UNL graph construction in the unambiguous context

The UNL relation of nouns with the case suffixes are usually with the corresponding main verb. The verb normally occurs at the end of the sentence. Therefore if POS of $\mathbf{W_{i+k}}$ where (k>=1) word is verb then the $\mathbf{W_{i+k}}$ word is connected to the case suffixed noun $\mathbf{W_i}$ by the indicated UNL relation. This is illustrated in the Fig. 9 below for ***obj*** relation.

avan pazhathai saappittaan
Wi-1 Wi Wi+1
He ate the fruit

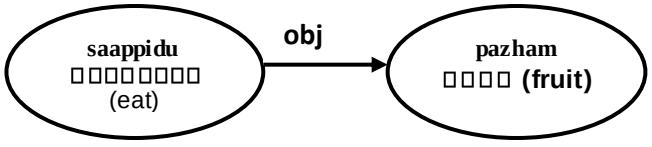

Fig. 9 Identifying concept based on case suffix "ai"

However, in some cases like "*avan pazham saappittaan*", more level of semantics is needed to obtain the semantic relations. The semantics of the verb as well as the other concepts are used to obtain the relations between the concepts.

UNL relation specified by adjectival and adverbial suffixes requires the determination of the concept to which this adjective or adverb is connected. The adjective or the adverb $\mathbf{W_i}$ is normally associated with the succeeding $\mathbf{W_{i+k}}$ whose POS is noun or verb respectively. Fig. 10 illustrates the "***mod***" relation obtained between the concepts for the example used in Fig. 2. Here, the word $\mathbf{W_i}$ with the adjectival suffix ***aana*** is connected with the corresponding noun $\mathbf{W_{i+1}}$.

neelamaana aaru
Wi Wi+1
Long river

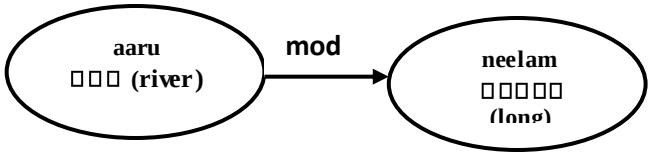

Fig. 10 Identifying concept based on the case suffix "aana"

For the example shown in Fig. 3, Fig. 11 illustrates that the word $\mathbf{W_i}$ with the adverbial suffix ***aaga*** is related to the succeeding main verb $\mathbf{W_{i+1}}$ with the UNL relation "***man***".

avaL inimaiyaaka paadinaaL
Wi-1 Wi Wi+1
She sung sweety

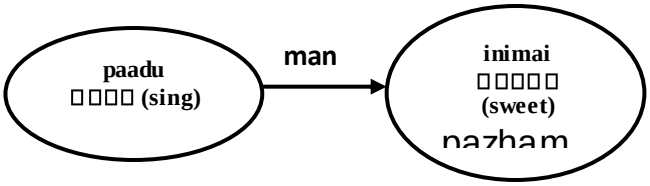

Fig. 11 Identifying concept based on case suffix "aaga"

In the case of connective indicating UNL relation, the connective itself is not part of the UNL Graph. Hence in these cases, the UNL graph construction requires the determination of all the associated concepts. In this context, for the connectives ***maRRum*** and ***mattumallaamal,*** the $\mathbf{W_{i-1}}$ and $\mathbf{W_{i+1}}$ concepts are connected by the UNL relation specified by $\mathbf{W_i}$. Fig. 12 illustrates this for the example shown in Fig. 4.

□□□□□□ □□□□□□□ □□□□□□□□□ □□□□□□□□□
□□□□□
Chennai maRRum maduraiyil koyilkaL uLLana
Wi-1 Wi Wi+1 Wi+2 Wi+3
Temples are located in Chennai and Madurai

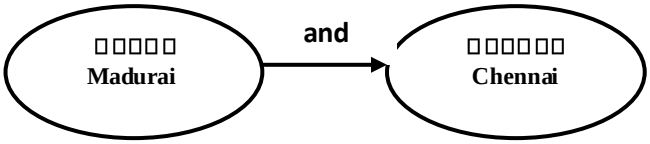

Fig. 12 Identifying concepts based on the connectives "maRRum"

For the connectives ***idaiye, idaiyil***, W_{i-1} and W_{i-2} take part in the UNL relation specified by W_i. Fig. 13 shows an example for ***idaiye identify*** the concepts for the relation "***int"***

□□□□□□ □□□□□□□□□ □□□□□ □□□□□ □□□□□□
Cauvery koLLidam idaiye paalam uLLathu
Wi-2 Wi-1 Wi Wi+1 Wi+2
Bridge is between Cauvery and koLLidam

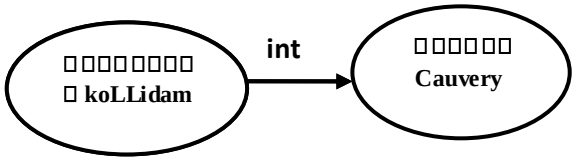

Fig. 13 Identifying concepts based on the connectives "idaiye"

In the UNL relation specified by ***vazhiyaaka,*** indicating UNL relation ***via,*** W_{i-1} and the succeeding W_{i+k} which have POS tag as verb will be connected by ***via*** relation.

□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□
□□□□□□
Tamilnattilirundhu rameswaram vazhiyaaga
Wi-2 Wi-1 Wi
□□□□□□□□□□ □□□□□□ □□□□□□□□□
Ilangaikku raaman sendraan
Wi+1 Wi+2 Wi+3
From Tamilnadu raman went to srilanka via rameswaram

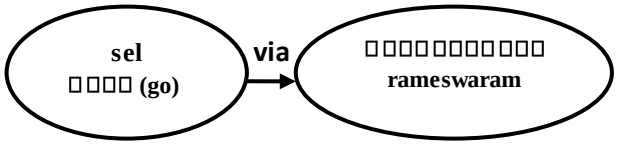

Fig. 14 Identifying concepts based on the word "vazhiyaaga"

The connectives ***aagiya, pondra*** and ***mudhaliya*** specifiy the UNL relation "***iof".*** Here the UNL relation is between W_{i+1} and the preceding W_{i-k} concepts whose semantic category matches with the semantic category of W_{i+1} An example for aagiya is shown in

□□□□□□ □□□□ □□□□□□ □□□□□□ □□□□ □□□□□□
kaveri vaigai ponni paalaaRu aakiya nadhikaL
Wi-4 Wi-3 Wi-2 Wi-1 Wi Wi+1
Rivers like cauvery vaigai ponni paalaaru

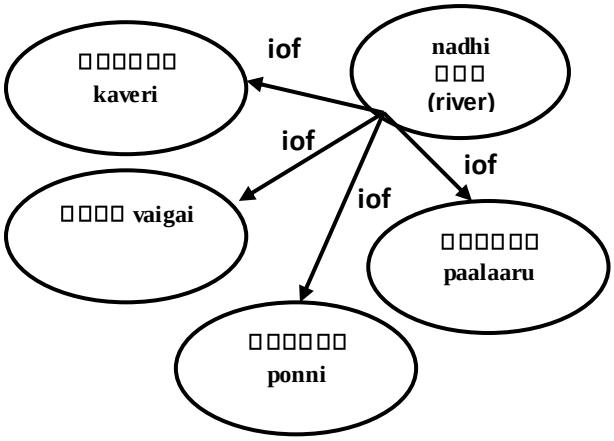

Fig. 15 Identifying concepts based on the word "aagiya"

4.2.2 UNL graph construction in the ambiguous context

In this section, the use of linguistic features that are used to resolve ambiguity of UNL relations are described.

As discussed in section 4.1.4, the possible UNL relations such as [***frm, plf, tmf, src***] are stored in a relation vector. The semantics of the word W_i (icl>time) is checked for obtaining the correct ***relation*** to the main verb of the sentence. If the semantics of W_i would be "***tmf".***The other UNL relations such as "***plf"*** can be determined using the semantics such as ***icl>place, icl>city, icl>country, iof>city*** etc. Similarly, the relations "***frm", "src"*** can be disambiguated with associated semantics.

□□□□□□□□□□□□□□□ □□□□
□□□□□□□□□□□□□□□□□□□
kaalayilirundhu poojai nadanthukondirunthathu
Wi Wi+1 Wi+2
Worship happened from morning

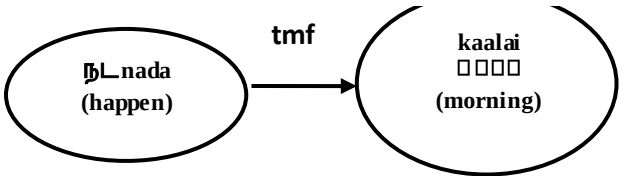

Fig. 16 Identifying concepts based on the case suffix "ilirundhu"

Similarly, the UNL relations ***"to", "plt", "tmt", "gol"*** can be disambiguated with the associated semantics.

Further cag The UNL relation "***cag"*** can be obtained using the semantics of the word $\mathbf{W_i}$ as (icl>person). Similarly, the UNL relations "***cob", "ptn"*** can be obtained with the semantics of the word $\mathbf{W_i}$.

□□□□□□ □□□□□□□□□ □□□□□□□□□□□□
□□□□□□□□
raman seethaiyudan kaattukku sendRaan
Wi-2 Wi Wi+1 Wi+2
Raman went to forest with seethai

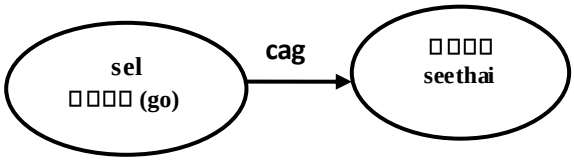

Fig. 17 Identifying concept based on the case suffix "udan"

For the ambiguity associated with the case suffix aal (shown in fig 8) the semantics of the word $\mathbf{W_i}$ (icl>person) and the POS tag of $\mathbf{W_{i+k}}$ th word (passive verb) are checked to connect $\mathbf{W_i}$ and $\mathbf{W_{i+k}}$. this is illustrated in fig .18

□□□□□ □□□□□□ □□□□□□□□□□□□ □□□□□□□□□□□
Thanjai kovil rajarajachozhanaal kattappattathu
Wi-1 Wi Wi+1
Thanjai temple was built by rajarajachozhan

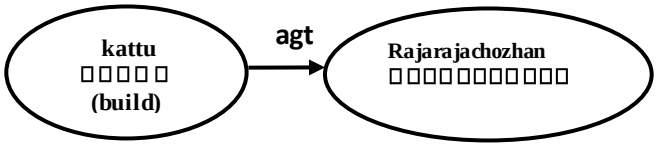

Fig. 18 Identifying concepts based on the case suffix "aal"

it can be observed from the above discussion, that word based morphological suffixes and context are sufficient for enconverting Tamil sentence to UNL expressions and no syntactic structure information is necessary.

5. RESULTS

Enconversion process consists of 53 rules encompassing the various categories and UNL relations.

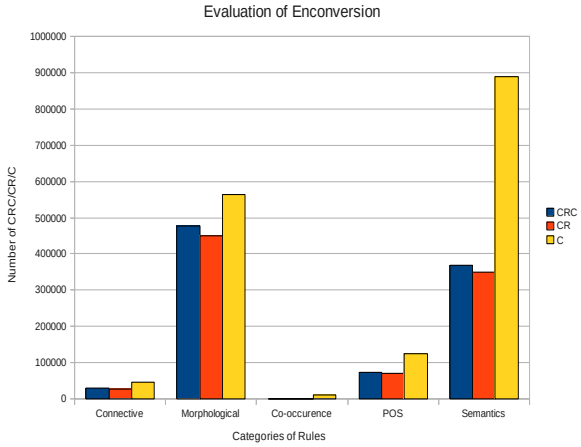

Fig. 19

The performance of this Enconversion process has been evaluated using a tourism domain corpus of 33000 documents. Fig. 19 shows the number of Concepts (C), Concept-Relation (CR) and Concept-Relation-Concept (CRC) obtained by the Enconversion process in each category. It can be seen that maximum number of C and CRC are obtained from morphological information and semantics. 90% of the Enconversion is achieved using morphological rules and semantic rules. Again there is not much difference in the number of C and CRCs in the morphological category. However in the semantics category, the number of C is comparatively more than the number of CRCs**.** This can be explained by the fact that in the test set of tourism documents obtained from the web, a number of isolated words exist in many documents. Such isolated words are identified as concepts based on the semantics. The difference between the number of C and CRC shows the efficiency of the Enconversion in connecting the concepts using UNL relations. Similarly, Fig. 20 shows the comparison of UNL relations obtained for the concepts.

On doing human verification or correctness of UNL, 87% of the Enconversion is satisfactory. On analyzing the 13% that is not enconverted correctly, it was found that the absence of case suffixes and handling of nested graphs limits the accuracy of the Enconversion. Similarly the absence of anaphora resolution also affects the accuracy of Enconversion. These are the aspects that need to explore.

5. CONCLUSION

The details regarding the type of information used for each UNL relation in the Tamil Enconversion has been presented. While a few UNL relations could be determined by morphological suffixes, POS tag or semantics of current word most require the features of preceding or succeeding words in the context. This usage of word based features rather than structure based features is possible due to the rich morphology of Tamil, with morphological suffixes conveying case relations, adjectival and adverbial relations. This in turn, has resulted in the morphological information rather than the syntactic structure playing a main role in Tamil Enconversion process.

Future work focuses on the use of UNL framework for Tamil to UNL enconversion by adapting the enconversion rules described in this paper. Another enhancement is to handle relations for words without case suffixes and nested graphs of Tamil sentences.

6. REFERENCES

[1] Ali, M.N.Y. Al-Mamun, S.M.A. Das, J.K. Nurannabi, A.M. 2008 Dept. of CSE, East West Univ., Dhaka. Morphological analysis of Bangla words for Universal Networking Language Digital Information Management. ICDIM 2008. Third International Conference

[2] Gala, N. 2004. Using an incremental robust parser to automatically generate semantic graph. Proceedings of the 3rd workshop on Robust Methods of Analysis of Natural Language Data.

[3] Igor M. Boguslavsky, Leonid L. Iomdin , Victor G. Sizov 2003 Interactive enconversion by means of the ETAP-3 system, Proceedings of the International Conference on the Convergence of Knowledge, Culture, Language and Information Technologies

[4] Jain, M. and Damani, O. P. 2008. English to UNL (Interlingua) Enconversion. Indian Institute of Technology Bombay, India.

[5] RajatKumar Mohanty, Anupama Dutta, P. B. 2005. Semantically relatable sets: Building blocks for representing semantics. Machine Translation Summit.

[6] Mrugank Surve, Satish Kagathara, Pushpak Bhattacharyya, 2004, Agro Explorer Group, Agro Explorer: a Meaning Based Multilingual Search Engine, In Proceedings of the International Conference on Digital Libraries (ICDL), Volume 2, New Delhi, India.

[7] Nguyen, D.P.T. Ishizuka, M.A 2006 A Statistical approach for universal networking language-based relation extraction Research, Innovation and Vision for the Future, 2006 IEEE International Conference

[8] UNDL. 2009. Universal networking digital language. http: // www. undl. org/ Online; accessed 28 September 2009 .

[9] Sameh Alansary, Magdy Nagi and Noha Adly, 2009, A Library Information System (LIS) based on UNL knowledge infrastructure, Seventh International Conference on Computer Science and Information Technologies, 28 September - 2 October, Yerevan, Armenia

[10] Subalalitha, T.V.Geetha, Parthasarathi, R., and Karky, M. 2008. CoReX: A Concept Based Semantic Indexing Technique. SWM-08

[11] Virach Sornlertlamvanich, Tanapong Potipiti and Thatsanee Charoenporn, 2001 NECTEC, Thailand, UNL Document Summarization , The First International Workshop on MultiMedia Annotation, 30-31 January 2001, Tokyo, Japan

[12] Firoz Mridha, Zakir Hossain, Shahid AI Noor, 2010, Development of Morphological Rules for Bangla Words for Universal Networking Language, IJCSNS International Journal of Computer Science and Network Security, vol. 10 No. 10, October 2010

[13] Nawab Yousuf Ali and Mohammad Zakir Hossain Sarker and Jugal Krishna Das, 2011, Analysis and Generation of Bengali Case Structure Constructs for Universal Networking Language, International Journal of Computer Applications, Vol. 18, March 2011, page no. 34-41

[14] Deryle Lonsdale and Er M. Franz and John R. R. Leavitt, 1994, Large-Scale Machine Translation: An Interlingua Approach, In Proceedings of IEA/AIE-94, page no. 525-530 ACM

[15] Bonnie Dorr, Parameterization of the Interlingua in Machine Translation, In Proceedings of Fourteenth International Conference on Computational Linguistics, page no. 624-630, MIT Press