

# How to Handle Split Antecedents in Tamil?

Vijay Sundar Ram R. and Sobha Lalitha Devi  
AU-KBC Research Centre,  
MIT Campus of Anna University,  
Chennai, India  
{sundar, sobha}@au-kbc.org

## Abstract

Resolution of the anaphoric entities in natural language text is very much essential to extract the complete information from the text. In this paper, we present a methodology to resolve one of the difficult pronouns, plural pronouns with split antecedents in Tamil. We have used a salience measure based approach with salience factors obtained from sub-categorization information of nouns and selectional restriction rules of the verbs. We have evaluated our approach with Tamil novel corpus and the results are encouraging.

## 1 Introduction

Anaphoric expressions in natural language text help in bringing cohesion to the text. The resolution of these anaphoric expressions is vital in developing information extraction and understanding systems. Theoretically various anaphoric expressions such as pronominal, reflexives, reciprocals, distributors, one pronoun, definite descriptions, VP anaphora, and zero anaphora are well studied. Automatic resolution engines for various types of anaphors were presented from early 80's of the last century, starting with Hobb's (1978) naïve approach followed by knowledge rich approaches by Carter et al (1987), Carbonll and Brown (1988), and Rich and LuperFoy (1988). These approaches were followed by knowledge poor approaches by Lappin and Leass (1994), Kennedy and Baguraov, Mitkov (1998) etc. Centering theory based approach was introduced by Grosz, Joshi and Kuhn (1979, 1981). The task of anaphora resolution got boosted with various Machine Learning (ML) techniques. The first ML approach was presented

by Dagan and Itai (1980) and various ML techniques were later used.

Byron (2001) has mentioned difficult anaphors which are excluded in most of the systems and they are as follows. i) Constructions which are required to interpret pronouns with split antecedents or cataphora. ii) Pronouns with antecedents, different from NPs such as clauses. iii) Pronouns with no antecedents in the discourse such as deictic or generic pronouns. There are very less number of automatic resolution engines for these difficult anaphors. In this paper, we present an algorithm for automatic resolution of one of the difficult anaphors; plural pronouns with split antecedents. We have studied the split antecedents in Tamil, a morphologically rich and verb final South Dravidian language and came up with an algorithm to resolve it. Consider example 1 given below:

Ex 1:

- a) *nepaal pirathammar ke. pi. ooli*  
Nepal Prime minister K P Oli  
*inthiya pirathammar moodiyai*  
Indian Prime Minister Modi  
puthu dilliyil canthiththaar.  
New Delhi meet(V)+past+3h  
(Nepal Prime minister KP Oli met Indian Prime Minister Modi in New Delhi.)
- b) *avarkal ceythiyaalarkalai*  
They press-people(N)  
Ithirabath-aucil canththinar.  
Hyderabad-House(N)+loc meet(V)+pst+3p  
(They met the press people at Hyderabad House.)

In the above example 1, the plural pronoun 'avarkal' (they) in sentence 1.b, refers to *nepaal pirathammar ke. pi. ooli* (Nepal Prime minister K P

Oli) and *inthiya pirathamam moodiyai* (Indian Prime Minister Modi), where these two entities have occurred in the subject and object of sentence 1.a.

Split antecedents are well studied in the framework of computational model in various languages and the details are as follows. Kosuga (2014) has presented a study on Japanese reciprocal anaphor ‘otagai’ with split antecedents. Han et al. (2011) have presented a behavioral study of grammatical status of ‘caki’ in Korean, which takes split antecedents as referent. Split antecedents were considered for coreference annotation in various languages such as Spanish, Catalan, Italian, English, Polish etc. MATE, AnaCorra, and Polish Coreference annotation schema support annotation of split antecedents. There are no published works on Split Antecedents in Indian languages and particularly in Tamil and our work is first of its kind.

Split antecedents are well studied under different constructions. Following are the different constructions explored in English. Split antecedents occur with a relative clause construction as in example 2.

Ex 2:

Marry met *a man* and John met *a woman who know each other well*. (McKinney-Bock, 2013)

There are different theoretical solutions for this split antecedents in relative construction. McKinney-Bock et al (2013) have presented a head-external approach and Ning Zhan (2007) have proposed a syntactic derivation approach. Split antecedents are dealt with VP-ellipsis construction as in example 3.

Ex 3:

‘Sally want to sail around the world and Barbara wants to fly to South Africa and they will, if money is available’ (Webber 1978)

‘Sally will sail around the world and Barbara will fly to South Africa’

Gatt and van Deemter (2009) have studied the characteristics of plural pronouns with split antecedents in GNOME corpus. They have studied the similarity and distance between the plural pronoun and their antecedents.

Cristea et al. (2002) have presented a paper investigating the difficult problems that could arise in anaphora resolution and proposed some solutions within the frame work of a general anaphora resolver. They have discussed on the methodology to resolve the plural pronouns with split antecedents. Consider example 4.

Ex 4:

- a) *John* waited for *Maria*.
- b) *They* went for pizza.

During the interpretation of the above sentence, a new discourse entity (DE) must be proposed for the group [John, Maria] as soon as the referential expression ‘Maria’ is parsed. Cristea et al. (2002) came-up with a set of ideas.

- a) Groups should have a property of similarity of their elements and that group formation is triggered by a first referent to it.
- b) A group is considered only if it is verbalized as such in the text and it does not exist until it is referred to.
- c) World knowledge is needed for group identification. We should use similarity measures to identify members of the group.
- d) A new DE should be proposed when no match between the current entity and the preceding DE arise above a threshold.

With these introductions to split-antecedents, we continue the paper as follow. The following section describes about Tamil and anaphora resolution works in this language. In the third section, we present our approach to resolve split-antecedents in Tamil using selectional restriction rules, sub-categorization information and salience measure (Lappin and Leass, 1994). The fourth section has description on the experiments and evaluation. The paper concludes with a concluding section.

## 2 Pronoun Resolution in Tamil

Tamil is a morphologically rich and highly agglutinative language. It belongs to Dravidian family of languages. It is a verb final, nominative-accusative and relatively free-word order language. Subject and finite verb has person, number and gender (PNG) agreement. Similarly 3<sup>rd</sup> person pronoun’s

PNG has agreement with its antecedent. 1<sup>st</sup> person and 2<sup>nd</sup> person pronouns have number agreement with its antecedents. Among Indian languages, there are a few automatic anaphora resolution works done in languages such as Tamil, Hindi, Bengali, Punjabi and Malayalam. Similar to what was mentioned by Byron (2001), these resolution engines do not attempt the difficult anaphors. One of the earliest anaphora resolution works in Indian languages was ‘Vasisth’ presented by Sobha (2000, 2002) for Hindi and Malayalam. Considering anaphora resolution in Tamil, there are few works on resolution of third person pronouns. The details are as follows. Sobha (2007) using salience measure, Akilandeshwari et al. (2013) using Conditional Random Fields (CRFs), Balaji et al. (2012) using bootstrapping approach and Ram and Sobha (2013) using Tree-CRFs. Sobha et al.(2014) have presented a generic pronominal resolution engine for resolving pronouns in Indo-Aryan and Dravidian languages. Akilandeshwari et al. (2012) have studied a different construction in Tamil, where the 3<sup>rd</sup> person pronouns are agglutinated with relative-participle verbs and they have presented a CRFs based approach for resolving these pronouns.

### 3 Our Approach for Resolution of Plural Pronouns with Split Antecedents

We attempt to resolve the plural pronouns with split-antecedents using selectional restriction rules of the verb, categorizing the nouns based on its sub-categorization information and ranking the possible antecedents using salience factor weights. In the following sub-sections we explain Sub-categorization of nouns and Selectional restriction rules.

#### 3.1 Selectional Restriction Rules

The verbs describe the action or the process in the nature and this allow the verbs to take nouns with specific sub-categorization feature as its syntactic arguments. This is defined as the selectional restriction (SR) rules of a verb. Consider the sentence in example 5.

Ex 5:

raam aappil caappittaana.  
 Ram(N) apple(N) eat(V)+past+3sn  
 ‘Ram ate an apple’.

Here ‘raam’ (Ram) has the sub-categorization feature [+animate, +human] and ‘aappil’ (apple) with [+edible]. The SR features required by the verb ‘caapitu’ (eat) for selecting its subject and object are [+animate] and [+edible] respectively. If there is a violation in SR rules, the sentence can be syntactically correct but it will not be semantically correct (Arulmozhi 2006). Verb has the right to select its arguments. We have grouped the verbs according to the sub-categorization information of the subject and object nouns. A group of commonly used 1500 verb senses are analyzed and 500 SR rules are derived from these verbs in-house. The SR rules do not cover figurative usage of language. The sub-categorization features of a noun are explained in the next section. A sample rule is shown in Figure 1.

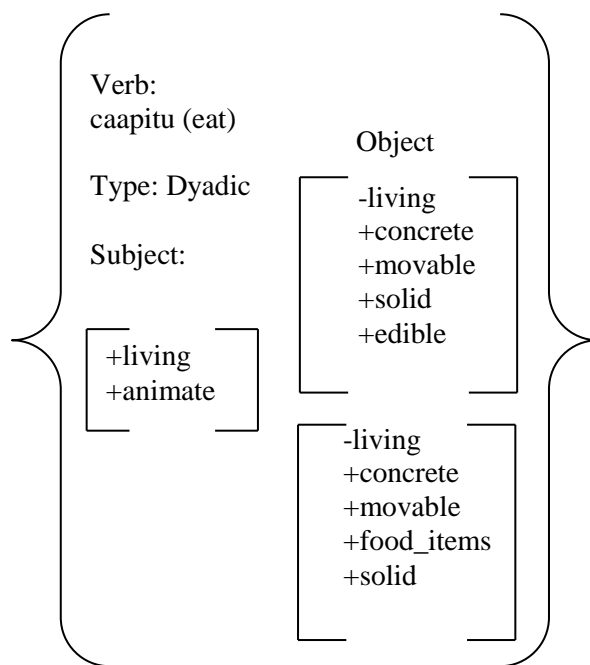
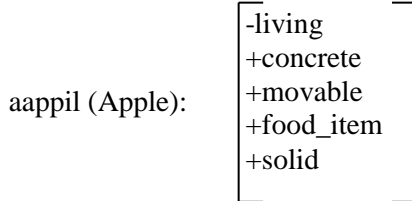


Figure 1: Selectional restriction rule for ‘caapitu’ (eat).

#### 3.2 Sub-Categorization

Sub-categorization features explain the nature of a noun. Essentially, the arguments of the verb, subject and object are analyzed using these features. These features may include the type of noun, its characteristics, state etc. Sub-categorization information include the features such as [+animate], [+concrete], [+edible] etc (Arulmozhi 2006).

There are totally 104 sub-categorization features. Using the sub-categorization features, which are related to the nouns, the SR features of the verb selects the nouns as its syntactic arguments. We have categorised 4500 frequently occurring nouns in Tamil. The Sub-categorization feature for the noun ‘aappil’ (apple) is presented in Figure 2.



**Figure 2:** Sub-categorization features of the noun ‘aappil’ (apple)

These sub-categorization features are used as nodes in building a language ontology. This language ontology is built with respect to the usage of language. Due to this, it deviates substantially from the taxonomy of nature. The sub-categorization features for the nouns can be obtained easily by traversing through various nodes (Arulmozhi, 2006). The nouns are grouped under each node, so we get a coarse to fine grained information of each noun. The ontology starts with [+entity] as the head noun and it divides into [+living] and [-living].

### 3.3 Resolution of Plural Pronouns

Using the SR rules and the sub-categorization information of nouns we try to resolve the plural pronouns in a two-step process. In the first step we try to group the noun phrases to form groups which can be possible split-antecedents. The nouns are grouped based on the sub-categorization information and following the verb’s SR rule restriction rule. Consider examples 6.

Ex 6:

- a) raam siitavai vakuppil  
Ram(N) Sita(N)+acc class(N)+loc  
canthithaan.  
meet(V)+past+3sm  
(Ram met Sita in the class.)
- b) avarkal unavu\_vituthikku  
They (plural PRP) hotel (N)+dative  
cenranar.

go(V)+past+3p  
(They went to the hotel)

Ex 7:

- a) puunai pullil  
Cat(N) grass(N)+loc  
vilaiyadi-koNtirunthathu.  
play(V)+progressive  
(The cat was playing on the grass.)
- b) nay manalil  
Dog(N) sand(N)+loc  
vilaiyadi-koNtirunthathu.  
play(V)+progressive  
(The dog was playing in the sand.)
- c) unavai kaNtathum avai  
Food(N) on\_seeing(V) they(plural-PRP)  
ooti\_vanthana.  
came\_running  
(On seeing the food they came running)

In example 6, there are three nouns before the plural pronoun ‘avarkal’ (they). The sub-categorization of these nouns are as follows:

- a) raam (Ram):[+living; +animate; +vertebrate; +mammal; +human; -female]
- b) siita (Sita):[+living; +animate; +vertebrate; +mammal; +human; +female]
- c) vakkupu (class):[-living; +concrete; -movable; +artifact; +building]

The verb in sentence 6.a and 6.b are ‘canthippu’ (meet) and ‘cel’ (go). The SR rules for these verbs are as follows;

- a) Verb: ‘canthippu’ (meet)  
Subj: [+living; +animate; +vertebrate; +mammal; +human;]  
Obj: [+living; +animate; +vertebrate; +mammal; +human;]
- b) Verb: ‘cel’ (go)  
Subj: [+living; +animate]  
Obj: [-living; +concrete; -movable]

And the plural pronoun ‘avarkal’ (they) has occurred in the second sentence 6.b. The antecedent of “avarkal” (they) occurs as two separate NPs [raam, siita].

In example 7, there are three nouns before the occurrence of the plural pronoun ‘avai’ (they). The sub-categorization of these nouns are as follows:

- a) puunai (Cat): [+living; +animate; +vertebrate; +mammal; -human; avion; +carnivorous; +cat\_family]
- b) naay (Dog): [+living; +animate; +vertebrate; +mammal; -human; -avion; +carnivorous; +dog\_family]
- c) unavu (food): [-living; +concrete; +food\_item]

The verbs in sentence 7.a, 7.b and 7.c are ‘villaiyatu’ and ‘ooti\_va’, where both the verbs take [+living; +animate] as subject argument. And the sentence 8.c has ‘avai’ (they), plural pronoun. Here the split antecedents are [puunai, naay].

We describe the methodology to perform the resolution of plural pronouns, which do not refer to a plural noun phrase, on text preprocessed with syntactic information such as morphological analysis (Ram et al, 2010), POS tag (Sobha et al, 2010), chunk information, clause boundary (Ram et al, 2012) and named entity (Malarkodi et al, 2012). The morphological analyser gives an indepth analysis of each word, such as root word, suffixes and its labels and person, number and gender (PNG) information. The clause boundary identifier marks the matrix clause and sub-ordinate clause boundaries, which helps in adding positional constraint features.

Following are the steps involved in resolving the plural pronoun. In the first step, we enrich the nouns and the verbs with their sub-categorization information, and SR rules respectively. The named entities (NEs) are mapped to the sub-categorization features, so we get the sub-categorization information using the NE information as described in the example 8.

Ex 8:

- a) Person: [+living; +animate; +vertebrate; +mammal; +human;]
- b) Location: [-living; -moveable; +landscape]

In the second step, when a plural pronoun is encountered in the sentence, the preceding portion of the sentence and two preceding sentences are considered for analysis, as Gatt et al. (2009) have shown that the distance between plural pronouns and their antecedent are very few sentences away.

The noun phrases in the preceding sentences are analysed and grouped to form the possible antecedents. For grouping the NPs, the NPs need to satisfy the following matching conditions.

- a) The NPs can be grouped together if they have same sub-categorization information or till the last but one node in the ontology is same. Example [+living; +animate; +vertebrate; +mammal; +human; +female] and [+living; +animate; +vertebrate; +mammal; +human; -female] are considered to be same since both are same till last but one node.
- b) Exceptions are as follows: In the case of NPs with sub-categorization [+living] and do not have [+human], we look for sub-categorization match between the NPs only till [+living; +animate] and such NPs are grouped together.

Following are the steps involved to form possible candidates by grouping the NPs.

- a) Identify the plural pronoun in  $n^{\text{th}}$  sentence.
- b) Consider sentence  $n-2^{\text{th}}$ ,  $n-1^{\text{th}}$  and in  $n^{\text{th}}$  sentence consider the portion preceding to the plural pronoun to form a candidate sentence set.
- c) For each sentence in the candidate sentence set; Noun Phrases in the sentence with conjunct suffix ‘um’ or conjunct word ‘maRRum’ (and) are united to form conjunct NPs. From now onwards the term NPs refers to both NPs and conjunct NPs.
- d) For each sentence in sentence set; if there exists NPs satisfying the matching condition, then the NPs are grouped together.
- e) Group the NPs that occur in same syntactic argument position and satisfy the matching condition across  $n^{\text{th}}$ ,  $n-1^{\text{th}}$  and  $n-2^{\text{th}}$  sentences.

S.No	Salience Factors	Weights
1	Same Ontology Nodes	30
2	NPs with following verbs	30
3	NPs with same syntactic argument position	20
4	NPs with different syntactic argument position	10

5	NPs are syntactic argument for verbs having same SR rules	30
6	NPs are syntactic argument for verbs with different SR rules	10
7	NPs in current n <sup>th</sup> sentence	30
8	NPs in n-1th sentence	20
9	NPs in n-2th sentence	10

**Table 1:** Saliency Factors and their Weights

In the third step, when the possible antecedents are formed by grouping the NPs, they are ranked based on the saliency factors derived from the features of NPs such as the sub-categorization information of NPs, the SR rules of verbs followed by the NPs and the syntactic argument position of the NPs in the sentences. The saliency factor weights (Lappin and Leass, 1994) are described in table 1. The weights for the saliency factors are initially manually assigned based on linguistic considerations and fine-tuned through experiments.

#### 4 Experiment, Results and Discussion

To analyse the plural pronouns, we choose a Tamil novel, ‘Ponniyin Selvan’ which was authored by Kalki, a well-known writer. As mentioned in Section 3.3, we processed the corpus with morphological analyser, POS tagger, chunker, pruner, clause boundary identifier and named entity recognizer. The corpus is made into a column format, where the information from each preprocessing module is added as a column. In the corpus, we considered the first 1000, plural pronouns, ‘avarkal’ and ‘avai’. These pronouns had four different types of antecedents such as plural noun phrase, conjunct NPs, split antecedents and the pronoun ‘avarkal’ also refers to honorific NP. The distribution of the pronouns with respect to their antecedents is presented in table 2.

S.No	Type of antecedent	Number of occurrence
1	Plural NP	789
2	Conjunct NPs	147
3	Split Antecedents	51
4	Honorific NP	18

**Table 2:** Distribution of plural pronouns based on their antecedents

In this experiment, we focus on plural pronouns with split antecedents. We considered the sentence having this plural pronoun and its preceding two sentences. In this set of sentences, as mentioned in

Section 3.3, we first tag the sub-categorization information for the nouns and SR rules of the verbs. After forming the possible antecedents by grouping NPs, we rank the possible antecedents with the saliency factor weights mentioned in Section 3.3 to find the antecedent. The performance evaluation is done with accuracy as the measure. The results are presented in table 3.

S.No	Total number of pronouns with split antecedents	Correctly tagged	Accuracy%
1	51	30	58.82

**Table 3:** Performance of resolution of plural pronouns with split antecedents

On analyzing the output, we found errors, when the preceding two sentences have similar NPs in the subject position. Consider the following example 9.

Ex 9:

- a) vanthiyathevan maNtapatthai  
Vanthiyathevan(N) hall(N)+acc  
atainthaan.  
reach(V)+past+3sm  
(Vanthiyathevan reached the hall.)
- b) kunthavai vaanathiyai  
Kundhavai(N) Vanathi(N)+acc  
azaiththaal.  
call(V)+past+3sf  
(Kundhavai called Vanathi)
- c) avarkal vanthiyathevanai kaNtu  
they(PN) Vanthiyathevan(N)+acc see(V)  
ciriththanar.  
laugh(V)+past+3pl  
(They on seeing Vanthiyathevan laughed.)

In the above example the possible antecedents for the pronoun ‘avarkal’ are [vanthiyathevan, kunthavai] and [kunthavai, vaanathiyai]. Here ‘vanthiyathevan’ and ‘kunthavai’ have occurred in the subject position and this group of NPs got higher saliency score. But [kunthavai, vaanathiyai] is the correct antecedent. This shows the saliency weights have to be altered further after analyzing more plural pronouns.

## Conclusion

We have presented a methodology to resolve plural pronouns which refer to split antecedents in Tamil. Automatic resolution of split antecedents is less attempted and it is first of its kind in Tamil. Our algorithm works on salience measures, the salience factors for scoring are obtained from the sub-categorization information of the noun phrases and the SR rules of the verbs. We have tested the algorithm on plural pronouns occurred in a Tamil novel. The results are encouraging. We need to test this methodology on a corpus from other domains.

## References

- Akilandeswari A., Sobha Lalitha Devi. 2013. Conditional Random Fields Based Pronominal Resolution in Tamil. *International Journal on Computer Science and Engineering*, Vol. 5 Issue 6 pp 601 - 610
- Akilandeswari A, Bakiyavathi T and Sobha Lalitha Devi, 2012. at Difficult Pronominal in Tamil. In: *1<sup>st</sup> Workshop on Indian Language Data: Resources and Evaluation, Organized under LREC 2012*, Istanbul
- Arulmozhi P. 2006. Semantic Tagging for Language Processing. *Thesis submitted to Anna University Chennai, India*
- Aone C., and McKee D. 1993. A Language-Independent Anaphora Resolution System for Understanding Multilingual Texts. In *proceeding of ACL 1993*, pp 156-163.
- Balaji J., Geetha T.V., Ranjani Parthasarathi R., Karky M. 2012. Two-Stage Bootstrapping for Anaphora Resolution In: *Proceedings of COLING 2012*, pp 507-516
- Byron, D. 2001. A proposal for consistent evaluation of pronoun resolution algorithms.
- Carbonell J. G., and Brown R. D. 1988. Anaphora resolution: A multi-strategy approach. In: *12<sup>th</sup> International Conference on Computational Linguistics*, 1988, pp. 96-101
- Carter D. 1987. Interpreting anaphors in natural language texts. Chisester: *Ellis Horwood Ltd.*
- Cristea, D.; Dima, G. E.; Postolache, O. & Mitkov, R. 2002. Handling complex cases of anaphora resolution, In: *4<sup>th</sup> Discourse Anaphora and Anaphora Resolution Colloquium (DAARC2002)*, University of Lisbon, Portugal, pp. 7-12.
- Dagan I., and Itai. A. 1990. Automatic processing of large corpora for the resolution of anaphora references. In: *13th conference on Computational linguistics*, Vol. 3, Helsinki, Finland, pp.330-332.
- Gatt A and Kees van Deemter. 2009. Generating plural NPs in discourse: Evidence from theGNOME corpus. In: *Procs. of COGSCI workshop \Production of Referring Expressions*, Amsterdam
- Han, C., Dennis R. S., and Calen W. 2011. An Experimental Study of the Grammatical Status of Caki in Korean. *Japanese/Korean Linguistics 19* (JK 19). Eds. Ho-min S., Haruko C., William O', Leon S. and Sang Y. C., CSLI, Stanford, 81-94
- Hobbs J. (1978). Resolving pronoun references. *Lingua* 44, pp. 339-352.
- Joshi A. K., and Kuhn S. 1979. Centered logic: The role of entity centered sentence representation in natural language inferencing. In: *International Joint Conference on Artificial Intelligence*.
- Joshi A. K., and Weinstein S. (1981). Control of inference: Role of some aspects of discourse structure – centering. In: *International Joint Conference on Artificial Intelligence*, pp. 385-387.
- Kosuge, T. (2014). Japanese Reciprocal Anaphor *Otagai* with Split Antecedents in Disguise and Multi-Dominant Syntactic Structure. *JELS* 31, 319-324.
- Lappin S., and Leass H. J. (1994). An algorithm for pronominal anaphora resolution. *Computational Linguistics* 20 (4), pp. 535-561.
- Malarkodi C.S and Sobha Lalitha Devi . 2012. A Deeper Look into Features for NE Resolution in Indian Languages. In: *1<sup>st</sup> Workshop on Indian Language Data: Resources and Evaluation, Organized under LREC 2012*, Istanbul
- McKinney-Bock, K. 2013. Deriving Split-Antecedent Relative Clauses. *University of Pennsylvania Working Papers in Linguistics*: Vol. 19: Iss. 1, Article 14.
- Mitkov R. 1998. Robust pronoun resolution with limited knowledge. In: *17th International Conference on Computational Linguistics (COLING' 98/ACL'98)*, Montreal, Canada, pp. 869-875.
- Ning Zhang N. 2007. The Syntactic Derivations of Split Antecedent Relative Clause Constructions. *Taiwan Journal of Linguistics* Vol. 5.1, 19-47
- Ram R.V.S, Menaka S and Sobha Lalitha Devi. 2010. Tamil Morphological Analyser. In: “Morphological Analysers and Generators”, (ed.) Mona Parakh, LDC-IL, Mysore, pp. 1 –18.
- Ram R.V.S, Bakiyavathi T, Sindhujagopalan, Amudha K and Sobha Lalitha Devi. 2012. Tamil Clause Boundary Identification: Annotation and Evaluation. In: *1<sup>st</sup> Workshop on Indian Language Data: Resources and Evaluation, Organized under LREC 2012*, Istanbul
- Ram, R.V.S and Sobha Lalitha Devi. 2013. Pronominal Resolution in Tamil Using Tree CRFs. In *Proceedings of 6th Language and Technology Conference, Human Language Technologies as a challenge for Computer Science and Linguistics - 2013*, Poznan, Polan

- Rich, E. LuperFoy, S. 1988. Anaphora architecture for anaphora resolution In: *Second Conference on Applied Natural Language Processing (ANLP'88)*, pp. 18-24 Austin, Texas (USA).
- Sobha L. and Patnaik B. N. 2000. Vasisth: An Anaphora Resolution System for Indian Languages. In *Proceedings of International Conference on Artificial and Computational Intelligence for Decision, Control and Automation in Engineering and Industrial Applications*, Monastir, Tunisia.
- Sobha L. and Patnaik, B.N. 2002. Vasisth: An anaphora resolution system for Malayalam and Hindi. In *Proceedings of Symposium on Translation Support Systems*.
- Sobha L. 2007. Resolution of Pronominals in Tamil. Computing Theory and Application, *The IEEE Computer Society Press*, Los Alamitos, CA, pp. 475-79.
- Sobha Lalitha Devi and Pattabhi R K Rao T. (2010). "Hybrid Approach for POS Tagging for Relatively Free Word Order Languages", in the Proceedings of Knowledge Sharing Event on Part-Of-Speech Tagging, 25-26th March, 2010, LDC-IL, CIIL, Mysore.
- Sobha L., Sivaji B., Ram R.V.S., and Akilandeswari A. 2011. NLP Tool Contest @ICON2011 on Anaphora Resolution in Indian Languages. In: *Proceedings of ICON 2011*.
- Webber, B. L. 1978. A Formal Approach to Discourse Anaphora, *Ph.D. thesis, Harvard University*.