

International Multidisciplinary Research Journal

Golden Research Thoughts

Chief Editor
Dr.Tukaram Narayan Shinde

Publisher
Mrs.Laxmi Ashok Yakkaldevi

Associate Editor
Dr.Rajani Dalvi

Honorary
Mr.Ashok Yakkaldevi

Golden Research Thoughts Journal is a multidisciplinary research journal, published monthly in English, Hindi & Marathi Language. All research papers submitted to the journal will be double - blind peer reviewed referred by members of the editorial board. Readers will include investigator in universities, research institutes government and industry with research interest in the general subjects.

Regional Editor

Dr. T. Manichander

International Advisory Board

Kamani Perera
Regional Center For Strategic Studies, Sri Lanka

Mohammad Hailat
Dept. of Mathematical Sciences,
University of South Carolina Aiken

Hasan Baktir
English Language and Literature
Department, Kayseri

Janaki Sinnasamy
Librarian, University of Malaya

Abdullah Sabbagh
Engineering Studies, Sydney

Ghayoor Abbas Chotana
Dept of Chemistry, Lahore University of
Management Sciences[PK]

Romona Mihaila
Spiru Haret University, Romania

Ecaterina Patrascu
Spiru Haret University, Bucharest

Anna Maria Constantinovici
AL. I. Cuza University, Romania

Delia Serbescu
Spiru Haret University, Bucharest,
Romania

Loredana Bosca
Spiru Haret University, Romania

Ilie Pinteau,
Spiru Haret University, Romania

Anurag Misra
DBS College, Kanpur

Fabricio Moraes de Almeida
Federal University of Rondonia, Brazil

Xiaohua Yang
PhD, USA

Titus PopPhD, Partium Christian
University, Oradea, Romania

George - Calin SERITAN
Faculty of Philosophy and Socio-Political
Sciences Al. I. Cuza University, Iasi

.....More

Editorial Board

Pratap Vyamktrao Naikwade
ASP College Devrukh, Ratnagiri, MS India Ex - VC. Solapur University, Solapur

Iresh Swami

Rajendra Shendge
Director, B.C.U.D. Solapur University,
Solapur

R. R. Patil
Head Geology Department Solapur
University, Solapur

N.S. Dhaygude
Ex. Prin. Dayanand College, Solapur

R. R. Yaliker
Director Management Institute, Solapur

Rama Bhosale
Prin. and Jt. Director Higher Education,
Panvel

Narendra Kadu
Jt. Director Higher Education, Pune

Umesh Rajderkar
Head Humanities & Social Science
YCMOU, Nashik

Salve R. N.
Department of Sociology, Shivaji
University, Kolhapur

K. M. Bhandarkar
Praful Patel College of Education, Gondia

S. R. Pandya
Head Education Dept. Mumbai University,
Mumbai

Govind P. Shinde
Bharati Vidyapeeth School of Distance
Education Center, Navi Mumbai

G. P. Patankar
S. D. M. Degree College, Honavar, Karnataka

Alka Darshan Shrivastava
Shaskiya Snatkottar Mahavidyalaya, Dhar

Chakane Sanjay Dnyaneshwar
Arts, Science & Commerce College,
Indapur, Pune

Maj. S. Bakhtiar Choudhary
Director, Hyderabad AP India.

Rahul Shriram Sudke
Devi Ahilya Vishwavidyalaya, Indore

Awadhesh Kumar Shirotriya
Secretary, Play India Play, Meerut (U.P.)

S. Parvathi Devi
Ph.D.-University of Allahabad

S.KANNAN
Annamalai University, TN

Sonal Singh,
Vikram University, Ujjain

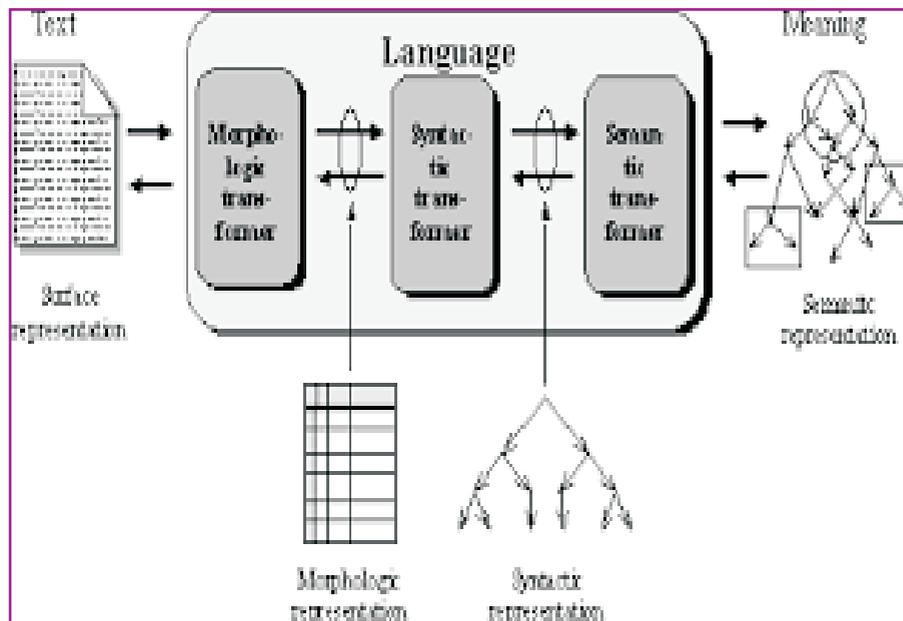
Satish Kumar Kalhotra
Maulana Azad National Urdu University



NOUN PHRASE PROCESSING IN COMPUTATIONAL TRANSLATION

Mr. Gadekar Manoj Bhagwat

Head, Department of English, Shriman Bhausaheb Zadbuke Mahavidyalaya, Barshi,
Dist- Solapur .



English, Marathi, Computation, translation, database, processing.

INTRODUCTION

The modern life is engrossed number of technological applications, computational processes and so many other things which have created and brought about great changes in every walk of life, but the very basis of all these developments is nothing but it's only the communication that has made all the difference. This difference is in every walk of life, however, the cultural, anthropological, social, geographical, historical and many more other dimensions have added their own role to the development of the human civilization. The linguistics as a branch of knowledge has its own significance for in the course of time it also has changed itself by adding something to its study. The varieties of different angles have blurred the boundaries of specialization and have their own effect on linguistics also. The fundamental studies in the linguistics have been furthering their development with the help of

ABSTRACT

The paper runs over the fundamental problem of machine translation, particularly, it is in regard to English-Marathi machine translation. The number of differences, divergence and structural variations directly or indirectly affect the semantic representation of a particular language into another language. In this paper the processing of Noun Phrase in English for translating it into Marathi is considered. It presents the problem of Noun Phrase structure in English and Noun Phrase structure in Marathi. Both these Noun Phrases are brought at the equal level and the semantic phenomenon of both of

them is considered. The most common issue regarding the nouns with processing is related to number, gender and case. The concepts of English pre-modifiers and post modifiers are not altogether the same. They are based on English linguistic requirements and Marathi language has altogether different structural patterns which has created the problems at this basic level. If the problem is resolved at the phrase level itself, it can facilitate the translation process. And as the Noun Phrase is among the highly important phrases, it needs to be resolved at the very basic stage.

KEYWORDS: phrases, Noun Phrase,

changing needs and demands of time. The computational linguistics is such a branch which has facilitated understanding of human cognition and other perspectives have been added to the study of language. The Information retrieval system; the information extraction; summarization; punctuation; Spell checker; and machine translations have been developing with rapid speed. The unbelievable communication technological developments have enhanced the possibilities of human understanding and sped up the development in the area of communication. However, the machine translation which is still the centre of inadequacies and deficiencies has been the major focus of discussion of the scholars in modern world. Especially, the languages coming under the same group of language family are easy to translate from one language to another language, but basically the languages that come from different language families encounter problems at variety of levels. The present paper is an attempt to present one of such obstacles which is the major concern of latest developments in translation technologies especially the mechanical translation.

The present paper is concerned with the nouns only. The noun phrases in English in comparison with the noun phrases in Marathi are the focal point of the present paper. The phrases in English are altogether different in their structure and organization, as far as there is the concern of noun phrases in Marathi. To understand the differences in the noun phrases in these two languages, it is necessary to systematically identify the basic features of the Noun Phrases in English and the Noun Phrases in Marathi. The structure of these phrases is completely different and basically the linguists study the structural aspects of variety of linguistic units and later on the semantic phenomenon are considered. The semantic phenomena are more important, but the less attention is paid at them in the natural language studies, because it is very abstract and intricately intelligible. In the present context of study, it consists of the morphological elements, logical elements as well as the structural elements, the immediate constituents are used to analyse the phrases. Especially, the effect and influence of structural issues is tremendous on linguistic analysis.

The noun phrases systematically consist of two different issues 'pre' and 'post' modifiers. The modifiers are pre modifier and post modifiers and head of noun phrase can be a noun, pronoun, determiner, adjective, adverb, genitive phrase, etc. and pre-modifiers in the noun phrase can be adjectives, pronouns, adverbs, enumerators, determiners etc. and the post-modifiers in noun phrase can be prepositional phrases, relative clauses, adjectives, nouns, adverbs etc.

The intricacies lay not in the structural organisation of these phrase elements but the constituents of the phrases and their properties, because the nouns have different semantic properties like number: singular and plural, or the gender like masculine, feminine, and neuter. These properties affect the translation process because according to the number of noun in the noun phrase there is the deviation in the basic meaning of the modifiers as well as there is the effect of the gender of modifier phrase on the preceding pre modifier. It may be in terms of gender like masculine, feminine and neuter. However, it is in terms of the meaning of the succeeding nouns. This is all about the structure of the noun phrase and its pre- modifiers, but it is not the only issue of the pre modifiers but also it is the issue of post modifiers. The same thing is true to the post modifiers. Unless and until the phrases are properly translated they cannot meet exactly the same meaning because their meaning would not be clear if they are not properly perceived and communicated. There are number of factors that contribute to the ambiguity, but still it can be resolved to the larger extent with the help of the systematic classification of these adjectives and their relations to the noun or head of the noun phrase.

Simply, the ambiguities can be resolved only by understanding the semantic properties and other such features which are useful to systematically organise the meaning and relate them. In terms of the Target Language (TL) in which we want to convert concepts and their meanings, this type of analysis and organization is useful to the larger extent to translate the basic meaning of the Source Language (SL) to be converted into target language. Once this primary phrase level analysis and conversion is in effect, it can help to the larger extent to bridge the gap between different phrases and relate them in terms of the target language.

While translating the phrases, it is attempted to create the digital data-bases which are the part of creating and processing of dictionary. Dictionaries are useful to relate and match the actual meaning of the sentences, however in relation to the phrases in terms of the sentences it is found that they are completely irrelevant and unmatchable to each other. Hence there is ambiguity or lack of clear-cut sense. Sometimes it is

found that the meaning communicated by the source language is altogether immature in the target language and in relation to translation. So the application becomes useless. In this paper, it is attempted to classify the words, and create such database (DB) which is useful to the translation of exact notions from source language to target language. It becomes very useful in the long run of the translation and effectively translating the source language into target language.

While bringing English and Marathi languages on the equal platforms, it is found that these are the major hurdles in the way of translating phrases. Because the structure of Marathi Phrases and the structure of English phrases is totally different, and it cannot be directly translated with the word-to-word matching, but the problem can be resolved by creating separate databases for English and Marathi and vice versa process in both of these languages, it is quite challenging. Still efforts can be made in this regard by creating English database separately and Marathi database separately. The same database cannot be effectively used as an alternative because of the cultural differences in both these languages. For example one polysemic word cannot be the issue of consideration in another language, so the present research paper makes an attempt to perceive the noun phrases in English in terms of the noun phrases in Marathi. These noun phrases when they are put into computational program, they look for the exact meaning in terms of the properties of that individual. However the hindrances cannot be dismissed entirely but they can be minimized to the larger extent and the barrier of high level ambiguity can be reduced to the low level ambiguity, as this is a natural language. It may rest with certain inadequacies and deficiencies but still the challenges remain the challenges. These challenges in the present technological system are difficult to entirely overcome because still there is time to develop the artificial intelligence in its fuller extent.

The following details regarding the noun phrase and its structure are useful to understand the limitations and problems that all linguists, especially, a computational linguists encounter during the processing of natural languages while translating the source language into target language. The following details about the structure of the noun phrase tells us that the head of the noun phrase can be either of the open class words, and the same is true to the modifiers also. But the meaning is mostly controlled by the different properties of the noun it is not the discussion of the ambiguity. But of the plain translation which faces the problems of structural differences and semantic differences.

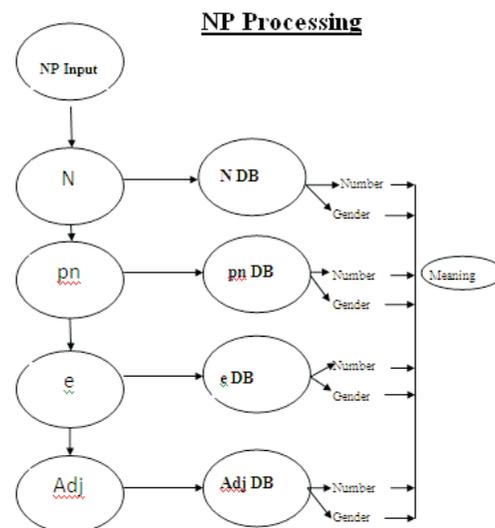


Fig.1

The above analysis of the noun phrases is useful to identify the basic structure of the phrase. With the help of this analysis one thing is quite clear that the head of the noun phrase can be a noun(N), pronoun(pn), adjective(Adj), enumerator(e) and some other. However, the modifier also is controlled by the noun. All the heads of the noun phrase, especially, when it is the matter of translating a noun phrase into Marathi, the most common issue will challenge is the choice of the meaning shade which is appropriate in the particular context is

mostly the part of the number and gender of the head of the noun phrase. Not all the adjectives are equally responsible for the controlling of meaning but the controlling of the meaning is mostly related to the adjectives, especially the adjectives of size, colour and volume. These three adjectives according to the gender and number of a head of noun phrase are to be translated into Marathi. For example if there is 'flowers' as the head of noun phrase the colour adjective is according to the head and the number in the same head is plural is used automatically. The meaning shade changes in Marathi but in English phrase there is no such change. By affixing certain elements to the adjective, we cannot do any processing in machine translation. The same thing becomes a challenge for processing the noun phrase in translation. Rests of the objects are actually in the same order which is needed for the noun phrase in English and in Marathi. This is about the pre-modifiers but the case with post-modifiers is altogether different. The post-modifiers alter after the head of the noun phrase but in machine translation for Marathi, they are considered in the beginning. It results into complication. This complication poses the biggest challenge before the machine translation system.

There are number of factors that affect the unity of the structure and then conversion into a target language meaning. The target language meaning is not exactly when there is the change in syntactic items. There are several hurdles in translation on the basis of the properties of words. It can assist to the larger extent to resolve the rest of the ambiguity is the specification in the database of all the adjectives according to their properties and their association with a singular or plural noun as well as their relation to the masculine, feminine or neuter gender that affects the unity of structure and the unity of meaning. Because of this very phenomenon the natural language processing makes a big difference. This research paper is based on a problem that superficially looks simple but at the bottom, it has tremendous effect on the translation of phrases and consequently clauses, sentences paragraphs and the discourse.

In the conclusion it is necessary to note that the noun phrase while processing with the help of computational tools faces number of problems, but this is the problem that lies within. It is internal structure, especially; it rests with the adjectives related to volume, colour and size. There are a number of other factors that contribute to the ambiguity problem but this is usually general problem that has the effect on the process of generating meaning in the target language. This research paper has focus on the problem in machine translation and its effects on the entire development of translation environment with the help of computational tools or computer itself. This problem if it is resolved at the initial stage while creating the database can assist the developers of machine translation systems to maximize the success rate and minimize the errors which will in turn produce better translation systems. Superficially it looks to be a simple issue but while considering the essence of noun phrase in the entire discursive practices, it plays pivotal role and contributes to the entire understanding of natural languages, because the noun phrase is the only phrase which has maximum number of functions performed in the clauses. It stands as a subject, object, complement and adverbial of all the five functions of phrases in a clause this is the only phrase that has maximum number of functions. In brief with this discussion it is quite clear that the problem can be solved by creating separate parts of speech tags for such adjectives which are responsible for ambiguity. In turn it will help to generate better system for future.

REFERENCES

1. Croft, Bruce W., ed. (2003), *Language modeling for information retrieval*, Dordrecht ; Boston : Kluwer Academic Publishers.
2. Delmonte, Rodolfo, (2007), *Computational linguistic text processing : logical form, semantic interpretation, discourse relations and question answering*, New York: Nova Science Publishers.
3. Gelbukh, Alexander, (2002), *Computational linguistics and intelligent text processing*, Berlin: Springer Verlag.
4. Roark, Brian, (2007), *Computational approaches to morphology and syntax*, Oxford : Oxford University Press.
5. Sikkel, Klaas, (1997), *Parsing schemata*, Berlin Springer Verlag.
6. Teubert, W, (2007), *Corpus linguistics : critical concepts in linguistics*, London ; New York : Routledge.
7. *Computational Linguistics Special Report on Seminar on "Concept Dictionary and Automatic Translation"*, New Delhi, 18th March, 1985. New Delhi Universal Digital Communication Research Institute 1985

Publish Research Article

International Level Multidisciplinary Research Journal For All Subjects

Dear Sir/Mam,

We invite unpublished Research Paper, Summary of Research Project, Theses, Books and Book Review for publication, you will be pleased to know that our journals are

Associated and Indexed, India

- * International Scientific Journal Consortium
- * OPEN J-GATE

Associated and Indexed, USA

- EBSCO
- Index Copernicus
- Publication Index
- Academic Journal Database
- Contemporary Research Index
- Academic Paper Database
- Digital Journals Database
- Current Index to Scholarly Journals
- Elite Scientific Journal Archive
- Directory Of Academic Resources
- Scholar Journal Index
- Recent Science Index
- Scientific Resources Database
- Directory Of Research Journal Indexing

Golden Research Thoughts
258/34 Raviwar Peth Solapur-413005, Maharashtra
Contact-9595359435
E-Mail-ayisrj@yahoo.in/ayisrj2011@gmail.com
Website : www.aygrt.isrj.org