



RESEARCH ARTICLE

Inflection Rules for English to Marathi Translation

Charugatra Tidke¹, Shital Binayakya², Shivani Patil³, Rekha Sugandhi⁴

^{1,2,3,4} Computer Engineering Department, University Of Pune, MIT College of Engineering, Pune-38, India

¹charutidke@gmail.com; ²sjain464@gmail.com; ³patil.shivani@gmail.com; ⁴rekha.sugandhi@gmail.com

Abstract— *Machine Translation is one of the central areas of focus of Natural Language Processing where translation is done from Source Language to Target Language preserving the meaning of the sentence. Large amount of research is being done in this field. However, research in Machine Translation remains highly localized to the particular source and target languages due to the large variations in the syntactical construction of languages. Inflection is an important part to get the correct translation. Inflection is basically the adding of appropriate suffix to the word according to the sentence structure to obtain the meaningful form of the word. This paper presents the implementation of the Inflection for English to Marathi Translation. The inflection of Nouns, Pronouns, Verbs, Adjectives are done on the basis of the other words and their attributes in the sentence. This paper gives the rules for inflecting the above Parts-of-Speech.*

Key Terms: - *Natural Language Processing; Machine Translation; Parsing; Marathi; Parts-Of-Speech; Inflection; Vibhakti; Prataya; Adpositions; Preposition; Postposition; Penn Tags*

I. INTRODUCTION

Machine translation, an integral part of Natural Language Processing, is important for breaking the language barrier and facilitating the inter-lingual communication. Marathi, an Indo-Aryan language derived from Sanskrit, is spoken by 70 million people in India. The script currently used in Marathi is called “baalbodha” which is a modified version of Devnagri Script [1]. While translating one language to another changing of the word order and its form according to the grammar of the target language is very important. For the scope of this paper the Source Language is English and Target Language is Marathi.

II. WORD ORDERING IN LINGUISTICS

The syntactic structure of a language is determined by the word order. Words are classified into 8 parts-of-speech (POS) [noun, pronoun, adjective, verb, conjunction, preposition, interjection]. The arrangement of these POS in sentences is determined according to the structure the language follows. English follows Subject-Verb-Object (SVO) structure while Marathi follows Subject-Object-Verb (SOV) structure [2]. Along with Marathi other Indo-Aryan language like Hindi, also follow the SOV structure.

III. IMPORTANCE OF ADPOSITIONS IN LINGUISTICS

Adpositions are words which can occur before or after a phrase, word, or a clause that is necessary to complete the meaning of a given sentence. Adpositions are mainly categorized as:

- Prepositions
- Postpositions
- Circumpositions

A. Prepositions

Prepositions are defined as the words placed before the complement [3]. Prepositions are used in English.

Example:

I value my family **above** everything else.

B. Postpositions

Postpositions are words which come after the complement.[2] Postpositions are used in Marathi, Hindi, Urdu, Korean, Turkish, and Japanese.

Example:

मी बाकी प्रत्येक गोष्टी पेक्षा माझ-या कुटुंबाला जास्त महत्त्व देतो.

C. Circumpositions

Circumpositions are words that appear on both sides of the complement. They are used in English, Dutch, Swedish, and French.

Example:

I will exercise regularly **from** now **on**.

The languages which follow SOV Structure use postpositions. Hence, while translating an English sentence.

(SVO structure) to a Marathi sentence (SOV structure), we need to change the prepositions (of English) to postpositions (of Marathi). This is a major issue which needs to be resolved for inflecting the nouns, verbs and cases (Vibhakti).

IV. INFLECTION

Generating inflection of a word is important to retain the correct form of the word in Marathi. Words can be classified in two types depending on the Inflection as [1]:

Inflectional Words:

- Noun
- Pronoun
- Adjective
- Verb

Non-Inflectional Words:

- Adverb
- Preposition
- Interjection
- Conjunction

The words are inflected on the basis of changing Gender (Masculine, Feminine, Neuter), Multiplicity (Singular, Plural), Tense (Present, Past, Future), and Case (Nominative, Accusative, Instrumental, Dative, Ablative, Genitive, Locative, Vocative).

A. Noun Inflection

Noun inflection is performed on the basis of change in Gender, Multiplicity or Case (Vibhakti). The inflection of a word can be determined from the word endings. Following table describes the word endings and its inflections.

TABLE I
TERMINATING VOWEL OF ITS ROOT [2].

Terminating Vowel of the Root	Plural Inflection		
	Masculine	Feminine	Neuter
अ	No change	आ ई	ए
आ	ए	आ	...
इ	No change	No change	No change
ई	No change	आ या	ए
उ	No change	No change	No change
ऊ	No change	अ	ए
ए	...	आ	ई
ऐ	...	आ	...
ओ	No change	आ	...
औ	No change

The above table will be found helpful in determining the plural form of a noun by terminating vowel of its root. For instance, the plural form of “बायको” must be “आ” making up “बायका” as “आ” stands opposite to the vowel “ओ” in the column super scribed Feminine[2].

The word “गाय” must be inflected to “गायी” as “अ” would be replaced by “ई”. Another example of a word-ending in “ई”, “पायरी” would get inflected as “पायरीया” wherein “या” is the word ending.

Case is an inflected form of Noun by which its relation to other words in the sentence is indicated.

For example:

He drove the car.

“त्याने गाडी चालवली”

TABLE II
CASE TERMINATIONS [2].

Case (Vibhakti)	Singular	Plural
Nominative	---	---
Accusative	स , ला	स , ना
Instrumental	ने, भि	नी भि
Dative	स , ला	स , ना
Ablative	हून रून	हून रून
Genitive	चा,ची,चे	चा,ची,चे
Locative	त	त
Vocative	---	नो

B. Verb Inflection

When we translate the verb using Marathi Dictionary we get the gerundial form i.e it is given with the particle “णे”.

Example:

“खेळणे” – to play

For inflecting the verb, first we need to derive the verbal root (धातु) and then add personal endings to it, to indicate its relation to the noun. We can get verbal roots by dropping the particle “णे” from the gerundial form. Inflection of the verb depends upon the following particulars:

The Gender (लिंग): Masculine, Feminine and Neuter.

The Number (वचन): Singular, Plural.

Person (पुरुष): First, Second and Third.

Tenses (काल): Present, Past and Future.

Sometimes personal endings may also depend on moods (अर्थ), the constructions (प्रयोग), the participle and the verbal nouns (धातुसाधीति) [2].

The table for verb inflection is given below (Table no. III).

TABLE III
RULES FOR DHAATU TO VERB [1].

पुरुष (Person)	पुल्लिङ्गी (Masculine)				स्त्रीलिङ्गी (Feminine)				नपुंसक (Neutral)			
	अनेकवचन Singular		अनेकवचन Plural		एकवचन Singular		अनेकवचन Plural		एकवचन Singular		अनेकवचन Plural	
	काल (Tense)	व	काल (Tense)	व	काल (Tense)	व	काल (Tense)	व	काल (Tense)	व	काल (Tense)	व
प्रथम (First)	Pre	तो	तो	तो	तो	Pre	ते	ते	ते	Pre	ते	ते
	Pt	लो	लो	लो	लो	Pt	ले	ले	ले	Pt	ले	ले
	F	इन	इन	इन	इन	F	ई	ई	ई	F	ई	ई
		इन	इन	इन	इन		इ	इ	इ		इ	इ
		इन	इन	इन	इन		इ	इ	इ		इ	इ
		इन	इन	इन	इन		इ	इ	इ		इ	इ
द्वितीय (Second)	Pre	तोस	तोस	तोस	तोस	Pre	तेस	तेस	तेस	Pre	तेस	तेस
	Pt	लास	लास	लास	लास	Pt	लेस	लेस	लेस	Pt	लेस	लेस
	F	शील	शील	शील	शील	F	लीस	लीस	लीस	F	लीस	लीस
		शील	शील	शील	शील		लीस	लीस	लीस		लीस	लीस
		शील	शील	शील	शील		लीस	लीस	लीस		लीस	लीस
		शील	शील	शील	शील		लीस	लीस	लीस		लीस	लीस
तृतीय (Third)	Pre	तो	तो	तो	तो	Pre	ते	ते	ते	Pre	ते	ते
	Pt	लो	लो	लो	लो	Pt	ले	ले	ले	Pt	ले	ले
	F	ईल	ईल	ईल	ईल	F	लील	लील	लील	F	लील	लील
		ईल	ईल	ईल	ईल		लील	लील	लील		लील	लील
		ईल	ईल	ईल	ईल		लील	लील	लील		लील	लील
		ईल	ईल	ईल	ईल		लील	लील	लील		लील	लील

C. Adjective inflection

Adjective is a verb which is joined to a noun to qualify it. Inflection of adjective depends upon gender, multiplicity, attachment of postpositions to the noun modified by such objective. When genitive case makers or some prepositions are attached to nouns, it produces adjective [4].

TABLE IV
ADJECTIVE TERMINATION FOR “अण्” [4]

Changing part in masculine form	Feminine	Neuter	Oblique form
अण्	अण्य	ए	या

D. Pronoun inflection

A pronoun is a word that can be substituted for a noun or a noun phrase. Pronoun inflection is similar to noun inflection but there are some special cases which need to handle separately. For the verbs such as “like”, “want”, “will”, “need” and “would”, the pronoun inflection are different than general cases. Some sentences have the same structure along with the parse tree, gender, multiplicity, cases but have different translation of pronoun. These sentences will typically include the above mentioned verbs.

Following table determines the translation of pronouns if the sentence has any of the verbs “like”, “want”, “will”, “need” or “would”.

TABLE V
EXCEPTIONAL PRONOUN INFLECTION

English word	Multiplicity	Marathi
I	S	मला
We	P	आम्हाला
You	S	तुला
You	P	तुम्हाला
He	S	त्याला
She	S	तिला
They	P	त्यांना

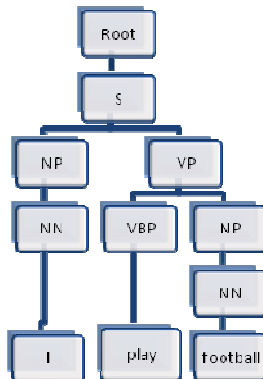
Following cases gives a comparison of different cases where the same pronoun is used but has different inflections depending upon the verbs used in the sentence.

Case 1 : (First person singular)

Example 1:

English sentence: I play football.

Parse Tree:



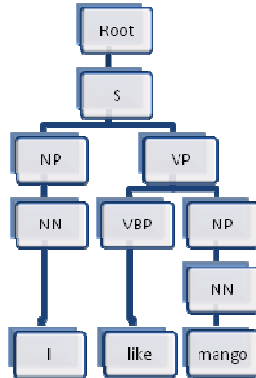
Marathi translation (After inflection):

मी खेळतो फुटबॉल.

Example 2:

English sentence: I like mango.

Parse Tree:



Marathi translation (After Inflection):

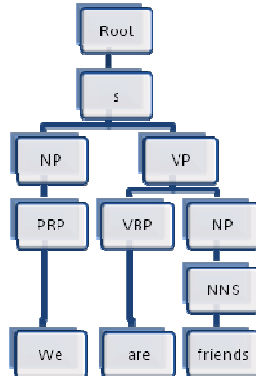
मला आवडतो आंबा.

Case 2:(first person plural)

Example 1:

English sentence: We are friends.

Parse Tree:



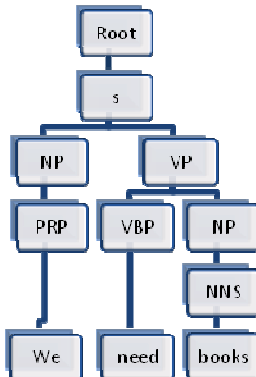
Marathi translation (After inflection):

आम्ही आहोत मित्र.

Example 2:

English sentence: We need books.

Parse Tree:



Marathi translation (After inflection):

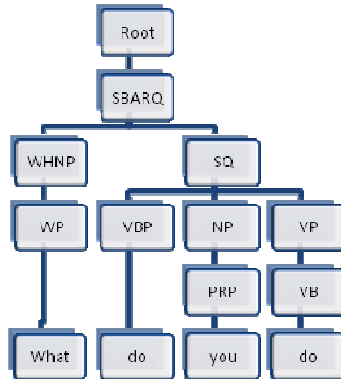
आम्हाला पाहिजे पुस्तककृ

Case 3: (Second Person Singular)

Example 1:

English sentence: What do you do?

Parse Tree:



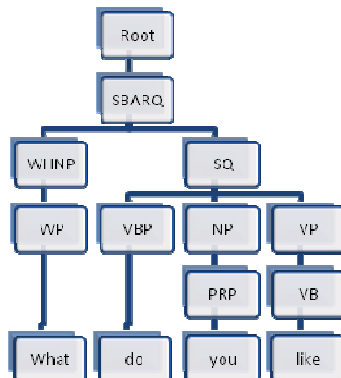
Marathi translation (After inflection):

काय तु करत आहे?

Example 2:

English sentence: What do you like?

Parse Tree:



Marathi translation (After inflection):

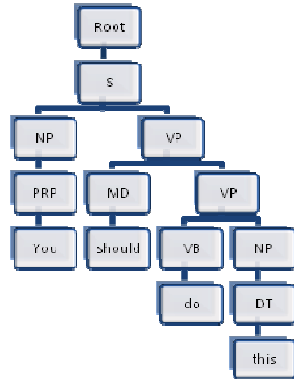
काय तुला आवडतं?

Case 4: (Second person plural)

Example 1:

English sentence: You should do this.

Parse Tree:

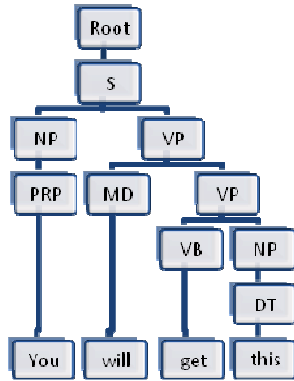


Marathi translation(After inflection):
तुम्ही पाहीजे करायला हे.

Example 2:

English sentence: You will get this.

Parse Tree:



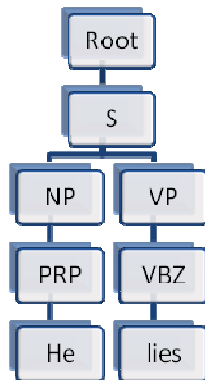
Marathi translation (After Inflection):
तुम्हाला मिळेल हे.

Case 5: (Third person singular)

Example 1:

English sentence: He lies.

Parse Tree:

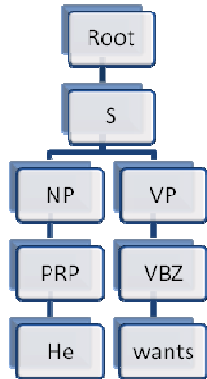


Marathi translation (After inflection):
तो खोटं बोलतो.

Example 2:

English sentence: He wants.

Parse Tree:



Marathi translation (After inflection):

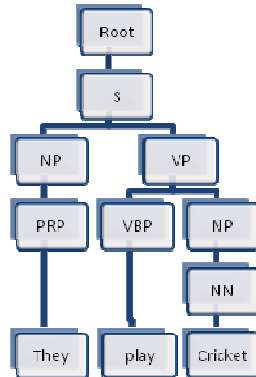
त्याला पाहीजे.

Case 6:- (Third person plural)

Example 1:

English sentence: They play cricket.

Parse Tree:



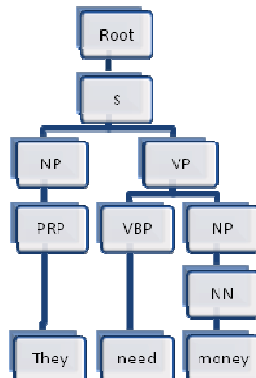
Marathi translation (After Inflection):

ते खेळतात क्रिकेट

Example 2:

English sentence: They need money.

Parse Tree:



Marathi translation (After Inflection):

त्यांना पाहिजे पैसे.

In all the above cases, the first example gives the inflection of pronouns according to the general inflection rules and the second examples give the inflection of pronoun according to the rules defined in Table 5.

V. IMPLEMENTATION OF INFLECTION

For implementation of the inflection we need to store following information to the database.

TABLE VI
DICTIONARY FORMAT

English word	POS Tag	Gender	Tense	Multiplicity	Degree	Case
--------------	---------	--------	-------	--------------	--------	------

POS tags are identified from parser.

Gender tags are identified from word endings or the gender of the word stored in a dictionary.

Tense of the sentence can be determined from Penn tags that we get from the parser. Following table gives PENN tags and its tense:

TABLE VII
GETTING THE TENSE FROM THE PENN TAGS

Penn_Tag	Tense
VB,VBG,VBP,VBZ	Present
VBD,VCN	Past
MD	Future

Multiplicity can also be determined from Penn tags. Following table gives PENN Tags and its multiplicity:

TABLE VIII
GETTING THE MULTIPLICITY FROM THE PENN TAGS

Penn_Tag	Multiplicity
NN,NNP,VBD,VB,VBZ	S
NNS,NNP	P

The **case** of the noun is determined on the basis of subject, object or prepositions.

The **degree** of the person can be determined from the following table:

TABLE IX
GETTING THE DEGREE DEPENDING UPON THE PERSON

Person	Degree
I, Me, My, Mine, We, Our, Us.	1
You, Your, Yours	2
He, She, It, They, His, Him, Her, Them, Their.	3

Using the above retrieved information, we can apply various Inflection rules given in the above sections to get the correct inflection. The inflected words can then be mapped to the SVO structure of Marathi to get the correct translation.

Example:

English Sentence:

He came to my house to meet Maya but she was not there.

Parse Tree:

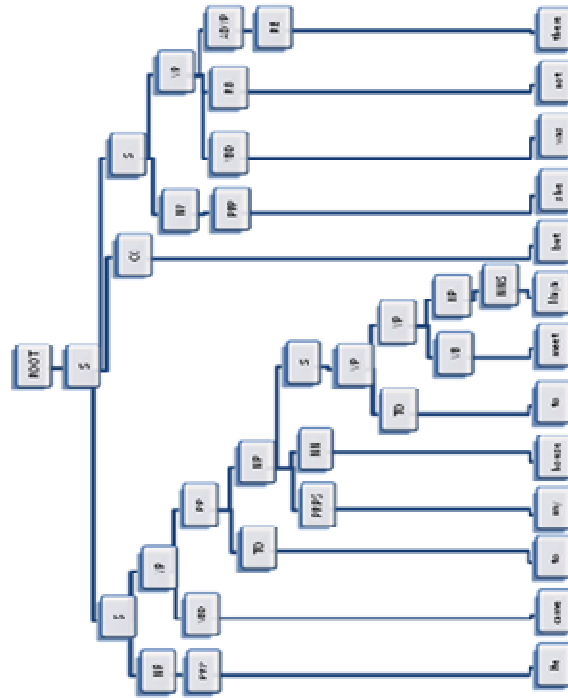


TABLE X
RULES FOR INFLECTION

Eng_Word	Marathi_Lemma	Inflected_Word	Rules to get Inflected
He	तो	तो	_____
came	येणे	आला	Verb table, 3 rd person, Masculine, Singular, Past Tense.
to	ला		
my	माझं		
house	घर	घरी	Case table, Locative Singular.
to	ला	_____	_____
meet	भेटणे	भेटायला	Case table, Dative Singular.
Maya	माया	मायाला	Case table, Accusative Singular.
but	पण	पण	_____
she	ती	ती	_____
was	होती	नव्हती	_____
not	नाही		_____
there	तिथे	तिथे	_____

VI. CONCLUSION

In the field of Machine Translation for Indian Languages a great amount of work has been done bur for Marathi the research is limited. This paper focuses on the issue of English to Marathi Translation with proper Inflection. The main objective of this paper is to give the detailed description of the rules required for inflecting the words.

ACKNOWLEDGEMENT

The authors would like to thank MIT College of Engineering, Pune and Board of College and University Development (BCUD), University of Pune for the grant for conducting the research work for Multi-Lingual Machine Translation System under which the paper has been written.

REFERENCES

- [1] Abhijeet R. Joshi, M. Sasikumar, "Constructive approach to teach inflections in Marathi language", www.cdacmumbai.in/design/corporate_site/.../pdf.../CATIML1.pdf
- [2] Navalkar, Ganpatrao, "The Student's Marathi Grammar", Asian Educational Services, 2001.
- [3] Wren and Martin, "New Edition High School English Grammar & Composition", S. Chand,
- [4] Veena Dixit, Satish Dethe, Rushikesh K. Joshi, "Design and Implementation of a Morphology-based spell-checker for Marathi, an Indian Language", In special issue on Human Language Technologies as a Challenge for Computer Science and Linguistics. Part I.15, Pages 309-316. Archives of Control Sciences, 2006.
- [5] Gowilkar, Leela, Marathiche Vyakran, Mehta Publishing House, 2006.
- [6] Walambe, M.R, Sugam Marathi Vyakaran Lekha, Nitin Prakashan, 2005.
- [7] Mugdha Bapat, Harshada Gune, Pushpak Bhattacharyya, "A paradigm-based Finite State Morphological Analyzer for Marathi", The 23rd International Conference on Computational Linguistics (Coling), Beijing, August 2010.
- [8] Ljiljana Dolamic, "Influence of Language Morphological Complexity on Information Retrieval", La Faculte des Sciences de l'Universite de Neuchatel, 2010.
- [9] Rekha Sugandhi, Devika Pisharoty, Priya Sidhaye, Hrishikesh Utpat, Sayali Wandkar and, Rajendra Khope, Managing Tokens for Machine Translation from English to Marathi, International Journal of Engineering, Science and Research (IJESR), Volume 1, Issue 3, October 2011
- [10] Rekha Sugandhi, Devika Pisharoty, Priya Sidhaye, Hrishikesh Utpat, Sayali Wandkar, "Extending Capabilities of English to Marathi Translator", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 3, May 2012.
- [11] Trevor Cohn and Mirella Lapata, "Machine Translation by Triangulation: Making Effective Use of Multi-Parallel Corpora", Human Computer Research Centre, School of Informatics, University of Edinburgh.
- [12] G. Singh and D. Lobiyal, "A Computational Grammar For Hindi Verb Phrase", Proc. International Conference on Expert Systems for Development, 1994, , 28-31 Mar 1994, pp: 244 – 249, doi: 10.1109/ICESD.1994.302273.
- [13] Sayali Charhate, Rekha Sugandhi, Anurag Dani, Varsha Patil, "Adding Intelligence to Non-corpus based Word Sense Disambiguation", HIS 2012 Conference technical program, doi: 5 Dec 2012.
- [14] Rekha S. Sugandhi, Ritika Shekhar, Tarun Agarwal, Rajneesh K. Bedi, Vijay M. Wadhai, "Issues in Parsing for Machine Aided Translation from English to Hindi", Information and Communication Technologies (WICT), 2011, doi:11-14 Dec.2011.