International Journal of Engineering & Technology, 7 (4.36) (2018) 542-548

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

A Rule Based Approach for Translation of Causative Construction of English and Malayalam for the Development of Prototype for Malayalam to English and English To Malayalam Bilingual Machine Translation System

T.K. Bijimol[1*], John T. Abraham[2]

[1]Research Scholar, Bharathiar University, Coimbatore, TN.
[2]Asst. Professor, Bharath Math College, Trikkakkara, Cochin, Kerala.
E-mail:johntabraham@yahoo.com
**Corresponding author E-mail:annabi2003@gmail.com*

Abstract

Malayalam is one of the Indian languages and it is a highly agglutinative and morphologically rich. These linguistic specialties of Malayalam determine the quality of all kinds of Malayalam machine translation systems. Causative sentences translations in Malayalam to English and English to Malayalam were analysed using Google Translation System and identified that causative sentence translation in these languages is not up to the mark. This paper discusses the concept and method of causative sentence handling in Malayalam to English and English to Malayalam Machine Translation Systems. A Rule-based system is proposed here to handle the causative sentence in both languages.

Keywords: *Rule Based machine translation system, google NMT, causative sentences, impersonal causatives, interpersonal causative, malayalam to english machine translation systems, english to malayalam translation systems.*

1. Introduction

Machine translation (MT) is a translation of natural language content in one language into another by computer. Language is an art that must be mastered through hundreds and thousands of hours of practice, there are some great benefits that can be derived from the use of machine translation. Some of the major benefits that can be gained from machine translation are save time, lower costs, ability to translate in many languages, memorize key terms and web content and web page translation.

There are different methods for machine translation which are Rule Based Machine Translation (RBMT) systems, Statistical Machine translation, Example Based Machine Translation, Neural Machine Translation and Hybrid Machine Translation [6]. RBMT builds linguistic rules based on morphological, syntactic and semantic information related to source and target language that is divided into three types which are direct method, transfer based method and Interlingua based method [8]. A statistical translation model (TM) is a mathematical model, used to perform the language translation by statistical modeling [2]. Example-based machine translation is a method of machine translation often characterized by its use of a bilingual corpus with parallel texts as its main knowledge base at run-time [9]. The neural machine translation is one which reads a sentence and sends out a correct output translation by attempting to build and train single large neural networks [16]. Finally, hybrid method is the combined approach which merges the benefits of various machine translation techniques to achieve an overall better language translation [11]. Though a number of Machine Translation systems between Indian and non-Indian languages have already been developed there are only very few systems developed for south Indian languages [7]. The MT systems developed have many shortcomings in terms of rule set, dictionary and translation methodology. It is a fact that further work is needed in MT as a whole to produce meaningful translations. This paper discusses the concept of causative sentences and how this concept improves the accuracy of existing English to Malayalam and Malayalam to English MT systems.

2. The causative sentence in malayalam and english

Causative verb is a special category of verb which shows that someone or something caused something to occur [26]. In a causative sentence the real subject of the sentence caused someone else to do something or being in certain condition instead of doing by himself [23]. The way of expression of causation is vary from language to language. Most of the Indian language show morphological causation like Malayalam, Tamil, Urdu, Hindi and soon. English shows morphological lexical and analytical causation. Here this study considers only the analytical causatives in English and their translation issues. The analytical causative sentence is two verb sentences. The first verb is a functional verb and the second verb is the content verb. The functional verb expresses the causation, the tense and aspect of the verb [22]. The main verb comes in its root form or past form.

Example: He had/got his hair cut.

The syntax and semantics of Malayalam is very unique that is it does not show the verb agreements for person, number and gender. It follows the order SOV (Subject Object Verb)[1].

Malayalam has the feature morphological Causativisation which is constructed by Morphological affixations. Transitive or causative markers are represented by the syntactic importance of Malayalam Verb Phrase (VP). Verbs which necessitate or at least imply the presence of three nominal of which one is an Initiator, the other the Actor or performer, and the third an Object, may be labeled Causative verbs. Causative verbs always imply an Actor as well as an Initiator of the action performed by the Actor [10]. The usage of causative construction is very popular in Malayalam. The causative verbs in Malayalam follow the same case and Tense, Aspect and Modality (TAM) agreement of transitive verbs and verbs of causative alternation.

Malayalam shows only morphological causation, for creating a causative situation in Malayalam the main verb undergo morphological inflections [13]. When its need to express the causation the causative suffix are attached to the verb stem. In Malayalam two types of causative suffixes are found the 1st causative and the 2nd causative. The 1st causative suffix is (ഇക്കുക) 'ikkuka' and the 2nd causative suffix is (പ്പിക്കുക) 'ppikkuka'. When the real subject want to mention the performer of the action the verb takes the 2nd causative form and that kinds of causatives are termed as interpersonal causatives and if the subject did not wants to mention the real performer of the action the verb takes its 1st causative form that type of sentence are named as impersonal causatives **[**22].

3. Rule Based Machine Translation (RBMT)

RBMT is also called Knowledge Based Machine Translation that retrieves rules from bilingual dictionaries and grammars based on linguistic information about source and target languages [24]. Target sentences are generated on the basis of syntactic, semantic and morphological features of specific language. There are three types of RBMT systems that are direct translation, transfer-based translation and interlingua-based translation. Direct machine translation method was the first machine translation method and a system named IBM701 was designed in 1954 using it [27]. Literal translation method is another name of direct translation and it has only minimum structural analysis in order to translate Source Language (SL) into Target Language (TL). Morphological analyser, bilingual dictionary and reordering rules are main components in this method. Morphological analyser identifies the basic structure of SL words and bilingual dictionary helps to look up the SL text and its equivalent TL forms. Re ordering rules are used to perform simple grammatical rearrangement based on the TL rules. The quality of the system is based on the morphological analyzer, bilingual dictionary and re ordering rules[6].

4. Literature review

4.1. English to malayalam translations

The first English to Malayalam Translation was proposed in 2008 and it uses statistical approach for translation [19]. In 2009, R. Rajan et al. [20] have used the RBMT method to convert English sentences to Malayalam. Mary Priya Sebastian et al. [16], have also proposed an SMT system based approach in 2011 and Aneena George [2] proposed an SMT system based approach for English to Malayalam conversion and it was in 2013, Nithya B. et al. [18] have developed an approach in 2013 that was based on hybrid translation. J. Sangeetha et al. [11] have proposed a hybrid system based approach in 2014 that translates English to Indian languages such as Tamil, Malayalam and Hindi. Statistical and ruled based methods are used in this work.

4.2. Malayalam to english translations

In 2012, a transfer based RBMT approach is proposed by Latha R Nair et al. [14] for Malayalam to English translations and Rajesh. K. S et al. [21] have proposed a Hybrid approach which is the combination of word-aligned parallel corpus based and dictionary lookup methods that was in 2012. Another Malayalam to English translation system was proposed by Anju E.S and Manoj Kumar K. V [3] in 2014, using EBMT method.

4.3. Google NMT (english to malayalam and malayalam to english translations available)

Google Neural Machine Translation System (GNMTS) is bas a neural machine translation (NMT) system developed by Google [29]. Neural machine translation (NMT) is an approach to machine translation that uses a large artificial neural network [28]. It is able to translate 100 more languages including English to Malayalam and Malayalam to English. This system learns about the translation from millions of examples (corpus).

5. Relevance of model selection

According to literature survey on machine translation methods and study on causative construction, causative marker attached to the verb decides the suffixes to be attached in the objects. But phrasal chunks as the main element for constructing the sentence, the case marker selection will not be feasible in SMT method [12]. Still any existing Malayalam MT systems do not managed causative sentences in the perspective of Machine Translation. Unfortunately sufficient quantities of Malayalam corpus containing causative sentences with its different morphosyntactic forms are not available. For the translation of Malayalam causative sentences into English and vice versa is not possible through modern techniques like NMT (Google Translate) or Machine Learning approach because of unavailability of sufficient corpus. In this scenario this work treated causative sentences with a well-developed Rule-based direct approach. Direct approach is suitable here because simple causative sentences in both the Malayalam and English are only handled by this system.

6. Methodology

Figure 1 shows the general block diagram of bilingual Malayalam-English and English-Malayalam machine translation system. The five main modules in the systems are Preprocessing Module, Causative Sentence Processing Module, Malayalam-English Translation Module, English-Malayalam Translation Module and Bilingual Dictionary

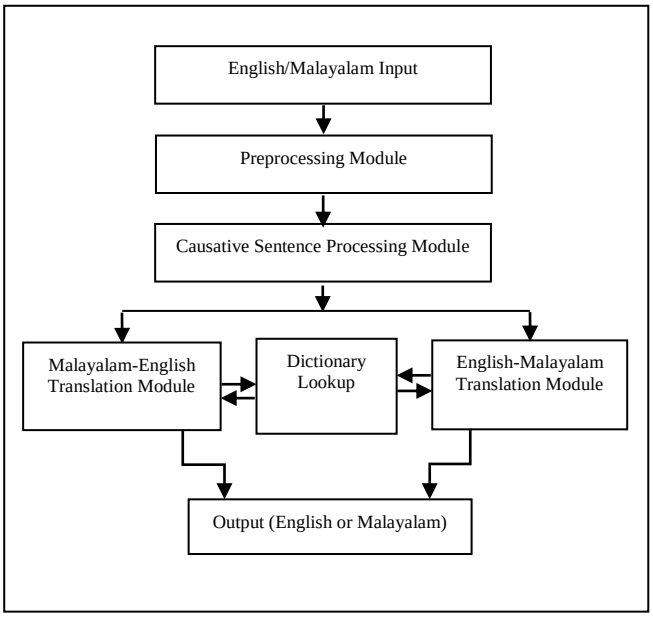

Figure 5.1: Block diagram of computational model

Figure 1. Block Diagram of Bilingual Malayalam-English and English-Malayalam Machine Translation System

Malayalam and English have completely different structures so it is required to design all modules separately for both the systems.

6.1. Preprocessing module

i) Tokenization ii) Suffix separation (for Malayalam) iii) POS tagging and iv) Word Order Identification are the four functions in this module. The input sentence is analysed and finds the set of morphemes which are used to check the type of sentence. Look up the POS (Parts of Speech) information and attach POS tags with each individual unit. These units can be used for word order identification and further causative sentence processing.

Tokenization is the basic process in NLP. In tokenization the input sentence is divided into set of small units called tokens which may be words or numbers or punctuation marks [4]. Word boundaries are identified and based on that the sentence is divided into tokens. English is space-delimited language so words can separate by white spaces as boundaries. But language like Malayalam is referred to as un segmented and there has no clear boundaries. So it is required to use lexical and morphological information for tokenization in Malayalam.

As discussed above suffix separation process is necessary to divide the inflected words into individual tokens. Suffix separation is the process of splitting the suffixes from inflected words. As Malayalam is inflected language it uses lexical and morphological information for suffix separation. Suffix separation process is done here using sandhi rules. Primary or secondary object is read from right to left, processes the check letter and identifies and removes the suffix with the help of sandhi rule look up table [15]. This process is not required when input text is in English

Example: Malayalam Input text - അവൾ കുട്ടിയെകൊണ്ട്റിപ്പോർട്ട്എഴുതിപ്പിച്ചു(avaL kuttiyekoNdu riportt ezhuthippicchu), the word 'കുട്ടിയെ' (kuttiye) is required to divide into കുട്ടി (kutti) +എ (e) for getting root word. 'എ' (e) is null in English translation.

English Text: She had the child write the report

Part of Speech (POS) Tags are attached to the individual units of a sentence in this function. It is also known as grammatical tagging and it is the processing of marking up sentence units with parts of speech information such as nouns, verbs, adverbs etc. based on its definition and context.

Malayalam sentence follows the order Subject-Object-Verb (SOV) but English follows Subject-Verb– Object (SVO) order. Word Order Identification function identifies the word order with the help of POS tagging and identify the tense of input sentence.

6.2. Causative sentence processing module

English make use the causative verbs for expressing causation. Each causative verb has their distinct semantics but they follow the same syntax. In English only the causative verb shows the person, number, gender (PNG), tense and aspect inflections. The main verb remains as its infinitive form of past tense form.

These two casuals have two distinct forms. For the making of causative verbs in Malayalam causative suffixes–'i'(ഇ), 'ppi'(പ്പി), and 'ththu'(ത്തു) are added to the end of the main verb according to the ending vowel of the main verb[13] Mainly three types of verb ending are found in Malayalam according to the ending vowel the same verb allows two types of causative suffixes and same are not allowed to take the two causative suffixes.

(1) Verb ends with 'a' the suffix 'i' is added with it [22].

Example:paRaya(പറയ)→transitive Form→paRayunnu(പറയുന്നു)+'i'→paRayippikkunnu(പറയിക്കുന്നു)→പറയിപ്പിക്കുന്നു pRayippikkunnu

(2) Verb ends with 'ka' the suffix 'ppi' is added with it

Example: eTukka(എടുക്ക)→transitive form→eTukkunnu (എടുക്കുന്നു) + 'ppi' →eTuppikkunnu (എടുപ്പിക്കുന്നു)(this types of verb posses only one causative fom)

Table 1: Forms Causative Sentences in English and Malayalam

Sl. No	Causative Type in English	Tense	Causative impersonal statement form in Malayalam	Causative interpersonal statement form in Malayalam
1	had	past indefinite	ഇച്ചു (iccu)	പ്പിച്ചു)ippiccu)
2	got	past indefinite	ഇച്ചു (iccu)	പ്പിച്ചു (ippiccu)
3	made	past indefinite	ഇച്ചു (iccu)	പ്പിച്ചു (ippiccu)
4	has/have	present indefinite	ഇക്കുന്നു (ikkunnu)	പ്പിക്കുന്നു (ippikkunnu)
5	get/gets	present indefinite	ഇക്കുന്നു (ikkunnu)	പ്പിക്കുന്നു (ippikkunnu)
6	make/makes	present indefinite	ഇക്കുന്നു (ikkunnu)	പ്പിക്കുന്നു (ippikkunnu)
7	will have	future indefinite	ഇക്കും (ikkum)	പ്പിക്കും (ippikkum)
8	will get	future	ഇക്കും (ikkum)	പ്പിക്കും (ippikkum)

		indefinite		
9	should have	future indefinite	ഇക്കേണ്ടതാകുന്നു (ikkendathakunnu)	പ്പിക്കേണ്ടതാകുന്നു (ippikkendathakunnu)
10	shall have	future indefinite	ഇക്കേണ്ടതാകുന്നു (ikkendathakunnu)	പ്പിക്കേണ്ടതാകുന്നു (ippikkendathakunnu)
11	must have	future indefinite	ഇക്കേണ്ടതാകുന്നു (ikkendathakunnu)	പ്പിക്കേണ്ടതാകുന്നു (ippikkendathakunnu)
12	should get	future indefinite	ഇക്കേണ്ടതാകുന്നു (ikkendathakunnu)	പ്പിക്കേണ്ടതാകുന്നു (ippikkendathakunnu)
13	shall get	future indefinite	ഇക്കേണ്ടതാകുന്നു (ikkendathakunnu)	പ്പിക്കേണ്ടതാകുന്നു (ippikkendathakunnu)
14	must get	future indefinite	ഇക്കേണ്ടതാകുന്നു (ikkendathakunnu)	പ്പിക്കേണ്ടതാകുന്നു (ippikkendathakunnu)
15	will make	future indefinite	ഇക്കും (ikkum)	പ്പിക്കും (ppikkum)
16	had had	past perfect	ഇച്ചിട്ടുണ്ടായിരുന്നു)iccittundayirunnu)	പ്പിച്ചിട്ടുണ്ടായിരുന്നു (ippiccittundayirunnu)
17	had got	past perfect	ഇച്ചിട്ടുണ്ടായിരുന്നു)iccittundayirunnu)	പ്പിച്ചിട്ടുണ്ടായിരുന്നു (ippiccittundayirunnu)
18	had made	past perfect	ഇച്ചിട്ടുണ്ടായിരുന്നു)iccittundayirunnu)	പ്പിച്ചിട്ടുണ്ടായിരുന്നു (ippiccittundayirunnu)
19	has/have had	present perfect	ഇച്ച് കഴിഞ്ഞിരിക്കുന്നു(iccu kazhinjirikkunnu)	പ്പിച്ച് കഴിഞ്ഞിരിക്കുന്നു(ippiccu kazhinjirikkunnu)
20	has/have got	present perfect	ഇച്ച് കഴിഞ്ഞിരിക്കുന്നു(iccu kazhinjirikkunnu)	പ്പിച്ച് കഴിഞ്ഞിരിക്കുന്നു(ippiccu kazhinjirikkunnu)
21	have/has made	present perfect	ഇച്ച് കഴിഞ്ഞിരിക്കുന്നു(iccu kazhinjirikkunnu)	പ്പിച്ച് കഴിഞ്ഞിരിക്കുന്നു(ippiccu kazhinjirikkunnu)
22	am/ is/are having	present continuous	ഇച്ചുകൊണ്ടിരിക്കുന്നു (iccukondirikkunnu)	പ്പിച്ചുകൊണ്ടിരിക്കുന്നു)ippiccukondirikkunnu)
23	am/is/are getting	present continuous	ഇച്ചുകൊണ്ടിരിക്കുന്നു (iccukondirikkunnu)	പ്പിച്ചുകൊണ്ടിരിക്കുന്നു(ippiccukondirikkunnu)
24	was/were having	past continous	ഇച്ചുകൊണ്ടിരുന്നു (iccukondirunnu)	പ്പിച്ചുകൊണ്ടിരുന്നു (ippiccukondirunnu)
25	was/were getting	past continous	ഇച്ചുകൊണ്ടിരുന്നു (iccukondirunnu)	പ്പിച്ചുകൊണ്ടിരുന്നു (ippiccukondirunnu)
26	is/am/are making	present continuous	ഇച്ചുകൊണ്ടിരിക്കുന്നു (iccukondirikkunnu)	പ്പിച്ചുകൊണ്ടിരിക്കുന്നു)ippiccukondirikkunnu)
27	was/were making	past continous	ഇച്ചുകൊണ്ടിരുന്നു (iccukondirunnu)	പ്പിച്ചുകൊണ്ടിരുന്നു (ippiccukondirunnu)
28	am/is/are going to have	planned future	ഇക്കുവാൻ പോവുകയാണ്)ikkuvan povukayanu)	പ്പിക്കുവാൻ പോവുകയാണ്)ippikkuvan povukayanu)
29	am/is/are going to make	planned future	ഇക്കുവാൻ പോവുകയാണ്)ikkuvan povukayanu)	പ്പിക്കുവാൻ പോവുകയാണ്)ippikkuvan povukayanu)

(3) Verbs ends with 'la', 'L', 'zha','ra' the suffix is 'ththu' added at the verb end

Example: para(പര)→transitive verb form→parakkunnu(പരക്കുന്നു)→ 'ththu'→parathunnu(പരത്തുന്നു)→പരത്തിക്കുന്നു

Like other transitive verbs all Malayalam causative verbs shows tens inflections also [10].

Example as like 'cheyyippicchu' (ചെയ്യിപ്പിച്ചു),'cheyyippikkunnu'(ചെയ്യിപ്പികുന്നു) 'cheyyippikkum'(ചെയ്യിപ്പിക്കും) in past-present-future correspondingly.

Example: Malayalam Text in impersonal form:

അവൻ അവന്റെ കാർ കഴുകിച്ചു

(avaNavantecarkazhukiccu)

English Text : He had his car washed

The ending suffix 'icchu' indicate the past tense of the main verb 'wash'

Example: Malayalam Text in interpersonal form:

അവൻ വേലക്കാരനെ കൊണ്ട് അവന്റെ കാർ കഴുകിപ്പിച്ചു

(avaNvelakkaranekoNduavantecarkazhukippiccu)

English Text: He had the servant wash his car

The ending suffix 'പ്പിച്ചു'(ippicchu) indicate the past tense of the main verb 'write' and word 'കൊണ്ട്'(kondu) represents the form is causative interpersonal.

This work handles twenty nine various forms of causative sentences in English and Malayalam which is shown in table 1.

Causative Sentence Processing Module is the main module in this work and have are two functions which are i) Causative Sentence Identification and Impersonal-Interpersonal Causative Sentence Processing. Malayalam and English Causative sentences have different structures. In Malayalam if the sentence follows the order 'sub + object + main verb +'iccuka' or 'ippikkuka' form then sentence is considered as causative. The sentence is in the form of 'have/get/make' + object +main verb then sentence is English Causative. As mentioned in section 2, a causative sentence contains Initiator, actor or performer and a causative verb. If the sentence contains a performer then it is in interpersonal form. Select impersonal causative sentence or interpersonal causative form based on the performer exists or not. In this module the proposed system identify the sense and type of causative sentences with help of impersonal/interpersonal causative sentence rules.

6.2.1. English causative verb identification

After receiving input, initially system identifies the sentence is in the form of 'have/get/make'+object+main verb, and select impersonal causative sentence or interpersonal causative form based on the performer exist or not.

Algorithm

begin

If the sentence follows the order 'sub + have/get/make + object + main verb' then

Sentence is causative. Performs the steps 1 to 5

1.Scan the sentence and identify the Causer (Subject), Causee (Secondary obect) Patient

(primary object) and main verb tense in the input sentence with the help of POS tagging.

a. Use table 1 to identify the tense of the main verb.

2.Look up the animate_inanimate table to identify the Causee and Patent are animate or inanimate

3.If Causee and Patient exist in the sentence then sentence is interpersonal causative. So perform steps b to e.

b. Look up the bilingual dictionary to choose the Malayalam equivalent of Causee

c. If Causee is animate then

Add Dative +instrumental case marker (‘kondu’) at the end of Causee

d. Look up the bilingual dictionary to choose the Malayalam equivalent of patient

e. If patient is animate then add dative case marker at the end of patient

Else add null case marker at the end of patient

Else if sentence contains only patient then sentence is impersonal causative. So perform steps f and g

f. Look up the bilingual dictionary to choose the Malayalam equivalent of patient

g. If patient is animate then add dative case marker atthe end of patient

Else add null case marker atthe end of patient

4.Look up the verb dictionary to choose the causative verb depends upon the tense and causative type.

5.Next go to the Word reordering module

Else print message ‘Sentence is not causative’ and continue with next input

End

6.2.2. Malayalam causative verb identification

This system follows the conventional Rule-based MT system.

Algorithm

Begin

If the sentence follows the order ‘sub + object + main verb +’iccuka’ or ‘ippikkuka’ form then

Sentence is causative. Performs the steps 1 to 5

1.Scan the sentence and identify the Causer (Subject), Causee (Secondary obect), Patient (primary object) and main verb form in the input sentence if any. Perform steps a to c.

a. Perform the suffix separation of Causee and patient if exist.

b. Use instrumental case marker ‘kondu’ to identify the Causee

c. Use table. 1 to identify the tense of the main verb

2.If Causee and patient exist in the sentence then sentence is interpersonal causative. So perform steps d

d. Look up the bilingual dictionary to choose the English equivalent of Causee and patient

3.Else if sentence contains only patient then sentence is impersonal causative. So perform step e.

e. Look up the bilingual dictionary to choose the English equivalent of patient

4.Look up the verb dictionary to choose the English causative verb depends upon the tense and causative type.

5.Next go to the Word reordering module

Else print message ‘Sentence is not causative’ and continue with next input

End

6.3. Malayalam-english translation module

Malayalam-English Translation Module has two functions which are i) Source Word Translation and ii) Output Sentence Generation. In target sentence generation module the system re-arrange the transformed linguistic units according to the word-order pattern of the English text with the help of rules. A list of English words in prescribed order is combined and gets the translation output here.

6.4. English-malayalam translation module

English-Malayalam Translation Module has three functions which are i) suffix addition ii) Source Word Translation and iii) target sentence generation.

When English is the input there is no suffixes with noun which can seen in the above sentence but while its Malayalam output is generated the suffix need to be added with noun. Sandhi rules are required for generating the suffix. For example, ‘ഏ’ (e) has to add with noun കുട്ടി (kutti). As per sandhi rules all the words end with ‘ഏ’ (e) are attached with suffix ‘യെ’(ye) and become ‘കുട്ടിയെ’ (kuttiye).

Malayalam words are translated into English words in Source Word Translation function. Target sentence generation function is used to re-arrange the transformed linguistic units as per the word-order pattern of the Malayalam text with the help of grammar rules. Malayalam words are combined and get the Malayalam sentence here.

6.5. Bilingual dictionary

Bilingual dictionary is the most important part of machine translation system. It is required to keep a dictionary for storing root words and with various inflectional forms of causative verbs of Malayalam separately and the main verbs and its various tens forms of Malayalam verbs and different forms of causatives. In this stage the system check into the dictionary of Source and target language and if the verb is available, the process is going into the next stage. Dictionary contains nouns, verbs, pronouns, adjectives, Part of Speech tagging information, case markers and causative forms in past/present/future tense. The root words stored along with the verbs in different tense.

English-Malayalam bilingual dictionary is also required to translate input text from English to Malayalam. Dictionary contains different forms of causative ‘have/get/make’ in past/present/future tense. Also it contains nouns, verbs, pronouns, adjectives, case markers, and the root words.

7. Translation evaluation

Human evaluation, Round trip translation evaluation and automatic evaluation and are available methods. BLEU, NIST, Word Error Rate, METEOR and LEPOR are used for automatic evaluation [6]. Human evaluation is the most reliable method to compare quality of various machine translation systems output [25]. A language expert is required to correctly evaluate the translated sentences. Adequacy and Fluency are two human evaluation criteria for the evaluation of MT output [5].

8. Experiment and result

Implementation of the proposed system is in Python 2.7. English tokenization and POS tagging is done by NLTK module. In Malayalam, tokenization is performed with the help of coding and POS tagging is done by dictionary processing. Input can give directly through the keyboard. Grammar rules and structure

transfer rules were implemented in the source code of the system. The system can work in Windows and Linux operating systems.
The proposed system is compared with Google Translator. The newly developed rule-based system performs better than the Google translator. More than 1000 English-Malayalam and Malayalam-English parallel sentences were utilized to train and test the system and same source sentences gave as the input of Google translator and compared the output of both systems. The Google translator translated all the main verb in their transitive form only it means that Google translator failed to identify the causative sense of in the source English or Malayalam sentence. For testing we included the impersonal and interpersonal causative sentences. The system gave quality output for English-Malayalam simple sentences with nearly an accuracy of 80%. The Malayalam-English system shows 75% accuracy for the human evaluation of simple sentences. In the context of long sentences our system failed to recognize the exact sentence patterns of source and target languages.

8.1. Sample output

Table.3 shows the output generated by both Google Translate and the newly developed rule-based system.

Table 1: Sample English to Malayalam Output of Google Translator and Proposed System

Sl. No	Input (English)	Google Translate Output (Malayalam)	System Output (Malayalam)
1	John will have his house painted	*യോഹന്നാൻതന്റെഭവനംനിറച്ചെടുക്കും.(John will fill his home) (5th August 2018)	ജോൺഅവന്റെവീട് ചായംഅടിപ്പിക്കും
2	I am getting my car washed	*എന്റെകാർകഴുകുന്നത്ഞാൻകഴുകുന്നു(My car wash I am wash) (5th August 2018)	ഞാൻ എന്റെ കാർ കഴുകിച്ചുകൊണ്ടിരിക്കുന്നു
3	Girls made him cry	പെൺകുട്ടികൾഅദ്ദേഹത്തെകരയിപ്പിച്ചു (girls made him cry) (5th August 2018)	പെൺകുട്ടികൾ അവനെ കരയിച്ചു
4	I am going to have the mechanic repair my car	*ഞാൻമെക്കാനിക്കായിഎന്റെകാർറിപ്പയർചെയ്യാൻപോകുകയാണ് (I as a mechanic I am going to repair my car) (5th August 2018)	ഞാൻ മെക്കാനിക്കിനെ കൊണ്ട് എൻറെ കാർ നന്നാക്കിപ്പിക്കുവാൻ പോവുകയാണ്

Table 2: Malayalam-English Translation Output of Google Translator and the Newly Developed System

Sl No	Input (Malayalam)	Google Translator Output (English)	Output of newly developed system (English)
1	അവൾ അവളുടെ മുടി വെട്ടിപ്പിച്ചു	*She cut off her hair	She had her hair cut
2	അവൻ അവൻറെ തുണി കഴുകിപ്പിച്ചു	*He washed his cloth (5th August 2018)	He had his dress washed
3	അവൻ എന്നെ കൊണ്ട് ഒരു കത്ത് എഴുതിപ്പിക്കില്ല	*He does not write a letter to me(5th August 2018)	He will not make me write a letter
4	അവൻവാതിൽ തുറപ്പിക്കും	*He will open the door (5th August 2018)	He will have opened the door
5	അവൻ അവന്റെ മുടി വെട്ടിച്ചോ?	*Did he cut off his hair? (5th August 2018)	Did he has his hair cut?

The table.1 and 2 show that the output of Google's MT system is completely wrong; in few cases GT (Google Translate) is able to produce correct translation (third example in table. 1). It has been noted that the GT is failed to identify the morphological causative construction in Malayalam and the analytical causative constructions of English.

9. Conclusion and future work

In this study we are proposing a Rule -based method for the translation of causative sentence in Malayalam and English and English to Malayalam causative sentences. Research reveals that this proposed system performs better than that of Google translator for the translation of causative sentences. It had been observed that sometimes our system did not correctly identify the correct POS information; it affects the quality of translation. Sense disambiguation is not handled in this work that is another limitation of the current system which can be overcome in future work. The proposed system is able to produce quality output for short sentences. In future it can be extend to lengthy causative sentences.

References

[1] Akshar B, Vineet C & Rajeev S, *Natural Language Processing A Paninian Perspective*, Prentice-Hall of India, New Delhi, (1996).

[2] Aneena G, "English To Malayalam Statistical Machine Translation System", *Article of International Journal of Engineering Research & Technology (*IJERT), Vol. 2, No. 7, (2013).

[3] Anju ES & Manoj Kumar KV, "Malayalam to English Machine Translation: An EBMT System", *IOSR Journal of Engineering (IOSRJEN)*, www.iosrjen.org, Vol.4, No.1, (2014).

[4] Steven B, Edward L & Ewan K, *Natural Language Processing with Python, O'Reilly Media Inc*., USA, (2009).

[5] Banchs RE, D'Haro LF & Li H, "Adequacy - Fluency Metrics: Evaluating MT in the Continuous Space Model Framework", *IEEE/ACM Transactions on Audio Speech and Language Processing*, Vol.23, No.3, (2015), pp.472-482.

[6] Bijimol TK & John TA, "A Study of Machine Translation Methods", *LAP LAMBERT Academic Publishing is managed by: SIA OmniScriptum Publishing*, (2018).

[7] Bijimol TK & John TA, "An Overview of Computational Linguistics and Malayalam", *International Conference on Recent trends in Engineering Science and Management*, (2015).

[8] Dove C, Loskutova O & Fuente R, *What's Your Pick: RbMT, SMT or Hybrid*?, (2012).

[9] Sumita E & Iida H, "Experiments and prospects of example-based machine translation", *Proceedings of the 29th annual meeting on Association for Computational Linguistics*, (1991), pp.185-192.

[10] Gopinathan NB, *Collected Papers on Malayalam Language and Linguistics*, International School of Dravidian Linguistics, (2008).

[11] Sangeetha J, Jothilakshmi S & Kumar RD, "An Efficient Machine Translation System for English to Indian Languages Using Hybrid Mechanism", *International Journal of Engineering and Technology (IJET)*, Vol.6, No.4, (2014), pp.1909-1919.

[12] Jayan V & Bhadran VK, "Difficulties in Processing Malayalam Verbs for Statistical Machine Translation", *International Journal of Artificial Intelligence & Applications (IJAIA)*, Vol.6, No.3, (2015), pp. 13-24.

[13] Narayanapillai KS, *Adhunika Malayalam vyakaraNam*, Kerala Bhasha Institute, (1995).

[14] Latha RN, David PS &Renjith PR, "Design and Development of a Malayalam to English Translator- A Transfer Based Approach", *International Journal of Computational Linguistics (IJCL)*, Vol.3, No.1, (2012).

[15] Mary Priya Sebastian, Sheena Kurian K and G. Santhosh Kumar, "A Classification of Sandhi Rules for Suffix Separation in Malayalam", *38th All India Conference of Dravidian Linguists*, (2010).

[16] Mary PS, Sheena KKB & Santhosh KG, "A Framework for Translating English Text Into Malayalam Using Statistical

Models", *2nd International Conference on Communication, Computing & Security*, (2011), pp.1-12.
[17] Mariana N, *Neural Machine Translation*, HPI Uni. Potsdam, BfR, (2017).
[18] Nithya B & Shibily J, "A Hybrid Approach to English to Malayalam Machine Translation", *International Journal of Computer Applications*, Vol. 81, No.8, (2013), pp.11-15.
[19] Preeti D & Preeti D, *Study and Development of Machine Translation System from Hindi Language to Dogri Language An Important Tool to Bridge the Digital Divide*, Ph.D Thesis, (2008).
[20] Rajan R, Sivan R, Ravindran R & Soman KP, "Rule Based Machine Translation from English to Malayalam", I*nternational Conference on Advances in Computing Control, & Telecommunication Technologies*, (2009).
[21] Rajesh KS, Veena AK & Dayakar Reddy CH, "Building a Bilingual Corpus based on Hybrid Approach for Malayalam-English Machine Translation", *Special Issue of International Journal of Computer Science & Informatics (IJCSI)*, Vol.2, No.1&2, (2012).
[22] Ram KV, *Ramkumarinte Sampourna Malayala vyakaranam*. Siso Books, Pattam, (2001).
[23] Randolph Q, Sidney G, Geoffery L and Jan S, *A Comprehensive Grammar of the English Language*, Dorling Kindersley, (2010).
[24] Tripathi S & Sarjgek JK, "Approaches to Machine Translation", *Annals of Library and Information Studies*, Vol.57, (2010), pp.388-393.
[25] Bremin S, Hu H, Karlsson J, Lillkull AP, Wester M, Danielsson H & Stymne S, "Methods for human evaluation of machine translation", *Small*, Vol.14, (2010), pp.55-67.
[26] Sven L, *Periphrastic Causative Verbs in English: What Do They Mean? The Expression of Causal Necessity and Causal Sufficiency in Ordinary English*, Department of Linguistics, Stanford University, (2010).
[27] Prasad TV & Muthukumaran GM, "Telugu to English Translation using Direct Machine Translation Approach", *International Journal of Science and Engineering Investigations*, Vol.2, No.2, (2013), pp.25-35.
[28] Jalaj T, *Python Natural Language Processing*, *Packt Publishing*, (2017).
[29] Wu Y, Schuster M & Chen Z, "*Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*, CoRR, (2016).