

Telugu to English Translation using Direct Machine Translation Approach

T. Venkateswara Prasad¹, G. Mayil Muthukumar²

¹ Dept. of Computer Sc. & Engg., C. M. J. University, Shillong, Meghalaya, India

² Technical Director, National Informatics Centre, Govt. of India, New Delhi, India
(¹tvprasad321@gmail.com, ²muthu@nic.in)

Abstract- The motivation behind working on a translation system from Telugu to English were based on the principles that

- There are many translation systems for translating from English to Indian languages but very few for vice versa. Telugu is a language that exhibits very strong phrasal, word and sentence structures next to Sanskrit, which makes the work organized on one hand but complex in handling on the other. This work demonstrates one such machine translation (MT) system for translating simple and moderately complex sentences from Telugu to English.
- Of the many MT approaches, the direct MT is used for translation between similar or nearly related languages. However, the direct MT has been used in this work for conversion from Telugu to English, which is quite complex compared to other Indian languages. The purpose of using direct MT for development of such a tool was to have the flexibility in usage, keeping it simple, look for rapid development and primarily to have better accuracy than all the known system.
- There are very large numbers of elisions/ inflection rules in Telugu requiring complex morphs, like those in Sanskrit. A large number of rules for handling inflections were to be developed along with the grammar rules.

The outcomes were compared with Google Translator, a publicly available translation web based system. The outcomes were found to be much better, as much as 90 percent more accurate. This work shall bring forth deeper insights into Telugu MT research.

Keywords- Machine translation (MT), direct MT, Telugu to English, natural language processing (NLP), elisions, inflections.

I. INTRODUCTION

Languages that are descent from Brahmi script are very good in grammar. The sentences are constructed strictly according to the norms laid out and there are very less chances of any deviation or violation. All Indian languages have

descended from Brahmi script. Telugu is said to have split from proto-Dravidian languages around 6th to 3rd century BCE [13].

Telugu language is a highly structured, disciplined, suave and rich in terms of expression, style and construction. It exhibits clear and structured implementation of grammar in the best possible manner while including present day corruptions (or vulgarity) and foreign words. There is a clear and specific purpose and meaning of each letter. Slight modification in the way a letter is written can change the meaning itself, e.g., kada and kaDa are two unique words having different meaning. Similarly, rama, rāma and ramā are three different usages. The language also provides large numbers of exceptions in usage thus making it more complex, beautiful and expressive [1].

The richness of Telugu language lies in the extremely large number of words representing different moods, expressions, contexts, etc. Ancient Telugu usage often known as “Grāndhika” had well defined grammar, classes of words, morphology, etc. Telugu language currently encompasses words of five categories, viz., a) of its own (purest form), b) of Sanskrit origin, c) of corrupt form of Sanskrit words, d) of colloquial usage and e) of other states/nations. Normally, the words of colloquial usage are not considered to be part of the Telugu grammar since it is considered as vulgar, was only prevalent with working class people [11].

Due to the modernization in the last century including serious impact of the media and cinema, the colloquial usage has taken centre stage of the grammar. When used in poetic sense, Telugu language exhibits very high level of grammatical usage. It is notable that each Telugu letter together with the consonants must be spoken very clearly with proper emphasis and intonation.

Tools for machine translation (MT) from English to certain Indian languages and from one Indian language to another are available; however, such tools for MT from Indian language to English are very few.

Indian languages are many in number but have a similar subject-object-verb (SOV) pattern of grammar, unlike the English that has SVO pattern or the VSO pattern of Arabic and Japanese. It is worth notable that translation from English

to any Indian language is a relatively easier process, whereas vice-versa is very complex.

This research work brings forth the process of converting Telugu sentences into its equivalent English sentences. Telugu grammar, vocabulary and style as documented by well known Telugu and British scholars during the British rule in India were studied in depth [1-2]. These books were selected since they were published during the mid 19th and early 20th century until when the Telugu language was relatively free from the heavy corruptions of the modern day literature.

II. MT SYSTEM APPROACH

Bernard Vauquois' pyramid is shown in Fig -1 depicting comparative depths of intermediary representation, interlingual machine translation at the peak, followed by transfer-based, then direct translation [3].

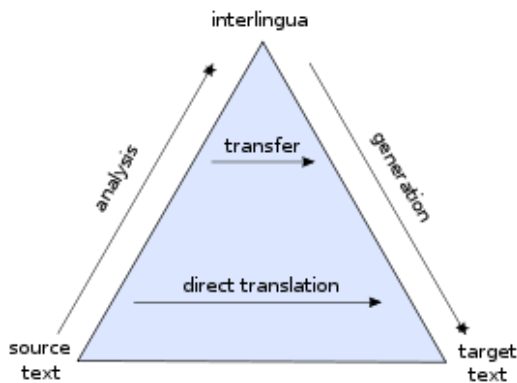


Figure 1. Bernard Vauquois' pyramid showing generalized model of MT

Machine translation can use a method based on linguistic rules, which means that words will be translated in a linguistic way — the most suitable (orally speaking) words of the target language will replace the ones in the source language. It is often argued that the success of machine translation requires the problem of natural language understanding to be solved first.

Rule-based methods parse a text, usually creating an intermediary, symbolic representation, from which the text in the target language is generated. According to the nature of the intermediary representation, an approach is described as interlingual MT or transfer-based MT. These methods require extensive lexicons with morphological, syntactic, and semantic information, and large sets of rules.

Given enough data, MT programs often work well enough for a native speaker of one language to get the approximate meaning of what is written by the other native speaker. The difficulty is getting sufficient data of right kind to support the particular method. For example, the large multilingual corpus of data needed for statistical methods to work is not necessary for the grammar-based methods. But then, the grammar

methods need a skilled linguist to carefully design the grammar that they use.

Following are the known approaches of MT:

- a) **Rule-based:** The rule-based MT paradigm includes transfer-based MT, interlingual MT and dictionary-based MT paradigms.
 - **Transfer-based machine translation:** To translate between closely related languages, a technique referred to as shallow-transfer machine translation may be used.
 - **Interlingual:** Interlingual MT is one instance of rule-based MT approaches. In this approach, the source language, i.e. the text to be translated, is transformed into an interlingual, i.e. source-/target-language-independent representation. The target language is then generated out of the interlingua.
 - **Dictionary-based:** MT can use a method based on dictionary entries, which means that the words will be translated as they are by a dictionary.
- b) **Statistical:** Statistical MT tries to generate translations using statistical methods based on bilingual text corpora, such as the Canadian Hansard corpus, the English-French record of the Canadian parliament and EUROPARL, the record of the European Parliament. Where such corpora are available, good results can be achieved translating similar texts, but such corpora are still rare for many language pairs.
- c) **Example-based:** Example-based MT (EBMT) approach was proposed by Makoto Nagao in 1984. It is often characterized by its use of a bilingual corpus as its main knowledge base, at run-time. It is essentially a translation by analogy and can be viewed as an implementation of case-based reasoning approach of machine learning.
- d) **Hybrid MT:** Hybrid MT (HMT) leverages the strengths of statistical and rule-based translation methodologies. Several MT organizations (such as Asia Online, LinguaSys, Systran, etc.) claim a hybrid approach that uses both rules and statistics. The approaches differ in a number of ways:
 - **Rules post-processed by statistics:** Translations are performed using a rules based engine. Statistics are then used in an attempt to adjust/correct the output from the rules engine.
 - **Statistics guided by rules:** Rules are used to pre-process data in an attempt to better guide the statistical engine. Rules are also used to post-process

the statistical output to perform functions such as normalization. This approach has a lot more power, flexibility and control when translating.

There has been debate on the suitability of statistical based MT on rule-based MT and vice versa for long; [19] concludes that it is purely dependent on the kind of applications and that these days a hybrid approach is being used more widely so as to combine the goodness of both approaches. Rule based NLP for demonstrating improvement in disease normalization in biomedical texts was also used [17]. The rule-based approach for MT of Arabic text was employed in [18]. Elaborated details on different approaches of MT and specific emphasis were put on Knowledge based MT (KBMT) are given in [16]. Latest views are also presented on the classification of different approaches in seminal work on English to Telugu MT [15].

In addition to the above classification of approaches, researchers have used various other methods like neural networks, fuzzy logic, genetic algorithms, hidden Markov models, etc. in different domains/languages for achieving better a) organization, b) rules and c) accuracy.

III. DIRECT MACHINE TRANSLATION

The direct MT system is considered to be the most primitive approaches of all carrying out replacement of the words in the source language with words in the target language. This is carried out in the same sequence and without much linguistic analysis or processing. The only resource direct MT uses is a bilingual dictionary, and that is why it is also known as dictionary-driven MT.

While certain researchers consider it to be quite unsophisticated approach and obsolete for many years, while some believe that direct MT has been considered useful for translation between two similar or near related languages. Systems falling under such approach are used for translation between Sanskrit and Hindi, Punjabi and Hindi, and so on. Description of evaluation of direct MT approach between Punjabi and Hindi is given in [21]. Earlier, [20] used the direct MT for English to Swedish translation.

Rule-based translation is one of the forms of MT, the rule-based MT paradigm includes transfer-based MT, interlingual MT and dictionary-based MT paradigms. Some experts call direct MT approach as part of the rule-based MT and consider it to be different from dictionary based MT approach. There is also a scope of combining the features of two or more approaches together for bringing out better translation results.

Of all these approaches, the direct MT approach was chosen for the proposed research on Telugu to English MT, keeping in view that the aspects of a) rapid software application development, b) higher accuracy, c) customizable MT, and d) provisioning of very simple and easily understandable design.

It is strongly believed that direct MT still has a place in today's automated translation tools. Such approaches are used where both vocabulary and syntax are standardized, in domains like weather reports, financial profiles, and many e-commerce applications. For implementation of such approach, word-for-word or phrase-for-phrase substitution is all that is needed.

Records reveal that human translation projects provided an unacceptably high level of error rates. The direct MT has proved to be very useful where initial tests had shown that both translation memories and rules-based machine translation systems produced poor results with text that has little or no repetition on the sentence level; or even high repetition on the word/phrase level.

Since direct MT does not require human post-editing in most of the cases, using MT in this kind is highly welcomed by translators and buyers needing very quick, cheap and moderately good quality of translation.

Many of the words are formed by combining two or more related words. Sandhis are actually conjugations of two or more words and elisions are reverse of sandhi, i.e. splitting of a word into two or more components. The more is the usage of elisions in Telugu, the structure of the sentence is considered the better [12].

For Telugu, certain work has been done on MT to/from Telugu related to handling of corpora and building of tree bank [6-7]. Most of the work has been built around Hindi language and generalized to all Indian languages as they follow the same SVO structure [6] with slight variations in placement of articles, pre/post-positions, etc. Morphological synthesis of English – Telugu MT was done [8]. Very less is available for MT from Indian languages to English. One recent attempt has been documented for Malayalam to English [10]. A lucid account of various useful works done on MT on Indian languages is given in [5].

Currently, there is only one known web based Telugu MT system available in the form of Google Translator [4]. A large number of experiments were conducted on the Google Translator to obtain the translation of various simple and moderately complex statements. Google Translator could not provide good translation of many words since the elision section was not handled adequately.

IV. EXPERIMENTAL WORK

Due to the vastness of the subject, the scope was limited to important portions of language translation. The assumptions/initial boundaries made for the purpose are (a) translation for simple Telugu statements are to be undertaken, (b) more focus to be given on word morphology that forms the most complex part of the research.

With these premises, a comprehensive software tool by the name "Telugu to English Translation Suite" was developed in Access Basic on Windows platform. A limited dictionary of Telugu to English database comprising of over 2000 words

was developed. As the Telugu language comprises extremely large number of conjunctions/ elisions/ inflections or sandhi forms, over 650 of them were analyzed, grouped in 222 paradigms and incorporated in the software suite, Table I.

TABLE I. TELUGU – ENGLISH DICTIONARY

Description	Qty
Telugu Verbs	399
Telugu Nouns	908
Telugu Pronouns	2
Telugu Adverbs	247
Telugu Adjectives	125
Telugu Prepositions	299
Telugu Ordinals	40
English Irregular verbs	362
Verb forms	276
Pronoun forms	109
Elision rules	649

Broadly, the system has been divided into five parts or modules, Figure II, viz.

- Conversion to Roman Telugu form (by transliteration)
- Application of Telugu morphology on the words
- Application of machine translation by replacing each Telugu word by equivalent English word
- Maintaining word order
- Application of English morphology (called here as reverse morphology)

There were 450 Telugu sentences categorized into five groups as listed in Table II, were taken from [1] and [14]. The TETS system was tested basically for the first two categories.

The developed software suite was rigorously experimented with large number of different types/structures of sentences. The outcomes of the software suite were also compared with the Google Translator (currently the only known publicly available translation site). The results were very encouraging as the accuracy of the developed software was very much higher.

TABLE II. CATEGORIZATION OF TELUGU TEST SENTENCES

Group	Description of test/example sentence	Number
I	Very Simple Telugu Sentences	346
II	Simple Telugu Sentences	65
III	Complex Telugu Sentences	29
IV	Very Complex Telugu Sentences	15
V	Free Flowing Telugu Paragraphs	Many

The test sentences/corpora were put into the MT system developed for MT from Telugu to English and were found comparatively to be very successful.

V. RESULTS AND DISCUSSIONS

Telugu being a free word-order structure language, MT from English to Telugu can be easy. However, the vice-versa is very complex keeping in view the complexity of English language structure.

Handling of two elisions in Telugu text were successfully implemented with accuracy of translation as high as 90 percent over the given test statements. Though the translation of idioms, style, feelings, handling synonyms of a word, etc. aspects have not been touched at this stage, the translation results were over 60 percent better than the web based Google Translator.

Sample outcomes of the MT to English as well as comparison with the outputs of Google Translator are tabulated in Table III. Some of the outcomes resulting translation specific to tenses have also been detailed in Table III. Some examples of poor or bad translation are given in Table IV.

The TETS system was also tested using free flowing sentences from various websites of newspaper companies. The parsing of lexicon, splitting or stripping of suffices, and their translation to English was very much satisfactory. Only those words could not be translated accurately that form very complex elisions/ inflections, or those not available in the dictionary or those having many synonyms.

It is most notable that the dictionary for Telugu to English MT should be populated with words that are spoken/used as they are. This means, there can be more words in the dictionary than predicted. For example, the Telugu equivalent for December is represented commonly in day-to-day usage by the words *DiseMbaru డిసెంబరు* as well as *Dishambar డిశంబర్*, however, if the dictionary is built only with the standard version, it is sure that the accuracy of translation will drastically reduce.

VI. CONCLUSION

With the present work, it was brought out that for successful translation of Indian languages, special emphasis has to be done on handling inflections/ elisions. There are large numbers of words that have three or more elisions.

For the first time, successful implementation of direct MT on two dissimilar languages was demonstrated through this work.

Addition of more linguistic rules related to handling of elisions/inflections and the word ordering system would enhance the accuracy of the proposed translation system.

REFERENCES

- [1] Arden A. H., "A Progressive Grammar of the Telugu Language with Copious Examples and Exercises", 2ed, Society for Promoting Christian Knowledge, Madras, 1905
- [2] Brown C. P., The Grammar of the Telugu Language, W. H. Allen & Co., London, 1857
- [3] Machine Translation Approaches, Wikipedia, available at www.wikipedia.com
- [4] Google Translator, Google, available at <http://translate.google.com>
- [5] Sitender, Seema Bawa, "Survey of Indian Machine Translation Systems", Int. J. of Computer Sc. & Tech., Vol. 3, Issue 1, Jan. - March 2012
- [6] Kolachina, P., Kolachina, S., Singh, A.K., Naidu, V., Husain, S., Sangal, R., Bharati, A., "Grammar Extraction from Treebanks for Hindi and Telugu", 2009 available at <http://researchweb.iiit.ac.in/~samar/data/grammarextraction-LREC10.pdf>
- [7] Chaitanya Vempaty, Viswanatha Naidu, Samar Husain, Ravi Kiran, Lakshmi Bai, Dipti M Sharma, and Rajeev Sangal, "Issues in Analyzing Telugu Sentences towards Building a Telugu Treebank", Proc. of 12th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2010), Lect. Notes in CS 6008, pp. 50-59, 2010
- [8] Suryakanthi T., Prasad S. V. A. V. and Sharma Kamlesh, "Morphological Synthesis in English to Telugu Machine Translation System", Proc. of International Conference on System Modeling & Advancement in Research Trends (SMART 2012), 2012
- [9] Anusaaraka, Indian Institute of Information Technology, Hyderabad, available at <http://nlp.iiit.net/~anusaaraka>
- [10] Latha R Nair, David Peter & Renjith P Ravindran, "Design and Development of a Malayalam to English Translator – A Transfer Based Approach", Int. J. of Computational Ling. (IJCL), Volume 3, Issue 1, 2012
- [11] Telugu Language, Wikipedia, available at en.wikipedia.org/wiki/Telugu_language
- [12] Venkateswara Prasad T. and Mayil Muthukumar G., "Handling Elisions in Telugu to English Machine Translation", Journal of Natural Language Processing, 2012 (submitted)
- [13] Wikipedia [Brahmic family of scripts], 2012, available at en.wikipedia.org/wiki/Brahmic_family_of_scripts
- [14] Learn Telugu, MyLanguages.org, available at mylanguages.org/learn_telugu.php
- [15] Sri Badri Narayanan R., English – Telugu Rule Based Machine Translation System, Master's Thesis, Amrita Vishwa Vidyapeetham University, Coimbatore, TN, India, 2012, available at nlp.amrita.edu:8080/project/mhrd/ms/Final_Thesis.pdf
- [16] Tripathi Sneha and Sarkhel J. K., Approaches to Machine Translation, Annals of Library and Info. Sc., Vol. 57, pp. 388-393, 2010, available at nopr.niscair.res.in/bitstream/123456789/11057/1/ALIS%2057%284%29%20388-393.pdf
- [17] Ning Kang, Bharat Singh, Zubair Afzal, Erik M van Mulligen, and Jan A Kors, Using rule-based natural language processing to improve disease normalization in biomedical text, *J Am Med Inform Assoc*, 2012, doi:10.1136/amiajnl-2012-001173
- [18] Khaled Shaalan, Rule-based Approach in Arabic Natural Language Processing, Int. J. of Info. & Comm. Tech., Vol. 2, No. 3, pp. 11-19, 2010, available at www.ieee.ma/IJICT/IJICT-SI-Bouzoubaa-3.3/3%20%20Khaledl.pdf
- [19] Linguistics Beta, Stack Exchange.com, available at linguistics.stackexchange.com/questions/350/machine-translation-rule-based-and-statistics-based-approaches
- [20] Ahrenberg Lars and Holmqvist Maria, Back to the Future? The Case for English-Swedish Direct Machine Translation, 2004, DOI 10.1.1.125.6062
- [21] Josan G. S. and Lehal G. S., Evaluation of Direct Machine Translation System For Punjabi To Hindi, 2009, available at www.advancedcentrepunjabi.org/pdf/directtrans.pdf

Mr. T. Venkateswara Prasad is a researcher working in the field of natural language processing for last few years. He is currently pursuing his post graduate studies at CMJ University, Shillong, Meghalaya, India. His research areas include machine translation, human computer interaction, discourse analysis and translation.

Mr. G. Mayil Muthukumar is associated with National Informatics Centre, Ministry of Communications & IT, Govt. of India in the capacity of Technical Director. He is currently directing major e-governance projects of the Central and State Governments. His research areas include data mining and data warehousing, design and analysis of algorithms, supercomputing, cloud computing, etc.

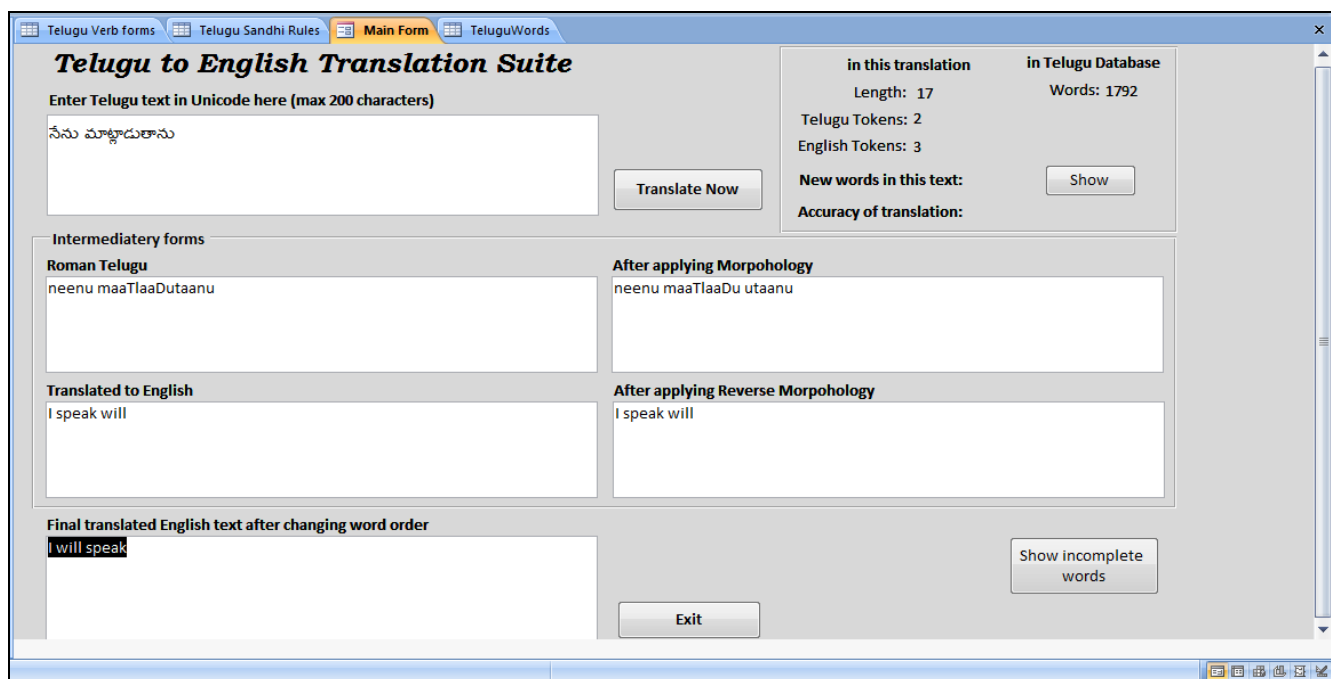


Figure 2. Screen shot of the TETS system

TABLE III. TRANSLATION OF SIMPLE SENTENCES FROM TELUGU TO ENGLISH

S. No.	Telugu text	Roman Telugu	English text	Translated by Google	Translated by our Suite*
1	ఆమె చూచుచున్నది	aame chuuchuchunnadi	She is seeing	She cucucunnadi	She is seeing
2	వారు వెళ్ళుచున్నారు	vaaru veLLuchunnaaru	they are going	They vellucunnaru	They are going
3	మీరు నడుచుచున్నారు	miiru naDuchuchunnaaru	You are walking	You naducucunnaru	you are walking
4	ఆమె వారిని తిట్టచున్నది	aame vaarini tiTTuchunnadi	She is abusing him/them	She tittucunnadi them	She is abusing them
5	మీరు టెలిఫోను వరకు నడుచిరి	miiru Telifoonu varaku naDuchiri	you walked till the telephone	You can telephone to naduciri	You till have walked telephone
6	నేను రేపు వ్రాయగలను	neenu reepu vraayagalanu	I will (or can) write tomorrow	I will write	I tomorrow can write
7	నేను మాట్లాడుతాను	nenu matlaaDutaanu	I will speak	I am talking	I will speak
8	మీరు మాట్లాడుతారు	meeru matlaaDutaaru	you will speak	You speak	You will speak
9	నేను మాట్లాడను	nEnu maTlaDanu	I don't speak	I do not talk	I do not speak
10	వారు అతనికి సలహా ఇచ్చారు	vaaru ataniki salaha ichchaaru	they advised him	They counseled him	They suggestion gave to him

S. No.	Telugu text	Roman Telugu	English text	Translated by Google	Translated by our Suite*
11	నేను దాన్ని తేగలను	neenu daanni teegalanu	I can bring it	I tegalanu	I that can bring
12	ఆమె అతనిని యెన్నుకుంటుంది	aame atanini yennukuMTuMdi	she chooses him	She made him yennukuntundi	She him will select
13	అతడు దాన్ని గురించి వెద్దాడు	ataDu daanni guriMchi yeDchaaDu	he cried about that	He was about to cry	He that about wept
14	మేం దాన్ని గురించి మాట్లాడుతాం	meeM daanni guriMchi maaTlaaDutaam	we talk about it	We're about it matladutam	We that will speak about
15	ఈ చిప్ప బాగా కడుగు	ee chippa baagaa kadugu	Wash this plate well	This shell is wash	this plate well wash
16	గొడుగు అక్కడ పెట్టు	godugu akkada peTTu	Put the umbrella there	Umbrella, where it	umbrella keep there
17	ఆ పెట్టె అక్కడ పెట్టవద్దు	aa peTTe akkada peTTavaddu	Don't put that box there	The box where the pettavaddu	that box keep there don't
18	ఆ గుర్రము త్వరగా పరుగెత్తుతున్నది	aa gurramu tvaragaa parugettutunnadi	That horse is galloping quickly	The horse quickly parugettutunnadi	that horse quickly running is
19	నే(ను) మాట్లాడాను	ne(nu) matladaanu	I spoke	I talked to	I spoke
20	నే(ను) వ్రాశాను	ne(nu) vrasaanu	I wrote	I write	I wrote
21	నే తీసుకున్నాను	nee tiisukunna nee tiisukunnaanu	I took (or) I have taken	I took the	I have taken
22	అతను తీసుకున్నాడు	atanu tiisukunnaaDu	he took (or) he has taken	He took	he has taken
23	మేం మాట్లాడాం	meeM maaTlaDaaM	we spoke	We matladam	We spoke
24	అతడు తీసుకుంటాడు	ataDu tiisukuntaaDu	he will take	He was taking	He will take
25	మేము మాట్లాడుతాము	memu maaTlaDutaam	we will speak	We matladutamu	We will speak
26	నేను మాట్లాడుతున్నాను	neenu maaTlaaDutunnaanu	I am speaking	I was talking	I am speaking
27	నేను వ్రాస్తున్నాను	neenu vraastunnaanu	I am writing	I am writing	I am writing
28	అతను కారు నడుపుతున్నాడు	atanu kaaru naDuputunnaaDu	he is driving car	He ran the car	he car is driving
29	మేము ప్రేమిస్తున్నాము	meemu premistunnaamu	we are loving	We love	We are loving
30	మేము తీసుకుంటున్నాము	meemu tiisukuMtunnaamu	we are taking	We take	We are taking

TABLE IV. EXAMPLES OF BAD TRANSLATION TO ENGLISH

Sl. No.	Telugu Questions	Roman Telugu	English Questions	Translated by Google	Translated by our Suite*
1	వాడు ఎక్కడ? అతడు ఎక్కడ?	vaaDu ekkaDa? atadu ekkaDa?	Where is he?	Where he	exploit where? He where?
2	మీకు నేను తెలుసా?	miiku neenu telusaa?	Do you know me?	If you know me?	
3	నా పుస్తకం నీ(మీ) దగ్గర ఉన్నదా?	naa pustakaM nee(mii) daggara unnadaa?	Do you have my book?	Napustakam near you?	
4	అది ఎంత పెద్దది?	adi eMta peddadi?	How big is it?	It was so big?	that how much
5	మీకు నా సహాయం కావాలా?	miiku naa sahaayaM kaavaalaa?	Can I help you?	If you want me to help you?	to you my support
6	మీరు ఆంగ్లములో మాట్లాడగలరా?	miiru aaMglamuLO maTlaDagalaraa?	Do you speak English?	Do you speak English?	You in English speak
7	ఇది ఎంత దూరం?	idi eMta dooraM?	How far is this?	This is how far?	this how much far
8	టైమెంత? టైము ఎంత?	TaimeMta? Taimu eMta?	What time is it?	Taimenta? Time at how much?	- time how much
9	ఇదెంత? ఇది ఎంత?	ideMta? idi eMta?	How much is this?	Identa? This is how much?	- this how much
10	మీ(నీ) పేరేమి? నీ పేరు ఏమి?	mii(nee) pErEmi? nee pEru eemi?	What is your name?	Miperemi?/ Niperemi? What is your name?	- you name what