

Performance of AI Translation of Simple, Compound, and Complex Sentences of a Modern Short Fiction

Prof. Kolakaluri Madhu Jyoti*1, Dr. Revathi Kudumula*2

¹ Former Vice Chancellor of Dravidian University, Kuppam. Head, Department of Telugu Studies Sri Padmavathi Mahila Visvavidyalayam, Tirupati-517502. Andhra Pradesh

² Assistant Professor, Department of English Language and Literature, Sri Padmavathi Mahila Visvavidyalayam, Tirupati-517502. Andhra Pradesh.

Abstract

The advancement of machine translation (MT) technologies has significantly impacted cross-linguistic accessibility, particularly in literary domains. Sentence structure plays a vital role in shaping clarity, coherence, and cohesion in written and spoken communication when we translate. This study examines the performance of machine translation tools in translating Telugu short fiction into English, with a specific focus on the treatment of simple, compound, and complex sentence structures. By conducting a comparative syntactic analysis of a selected Telugu short story and its English translation generated through the online MT tool, 'Microsoft'. The study identifies patterns of structural transformation, syntactic distortion, and morphological inconsistencies. Emphasis is placed on the fidelity of sentence type preservation and the semantic implications of syntactic shifts. The research reveals that while simple sentences are largely translated with minimal error, compound and complex sentences often suffer from misinterpretation, clause fragmentation, or unnatural rendering. These findings highlight the limitations of current MT systems in processing literary syntax and suggest the need for more linguistically informed models. This study contributes to the growing field of computational linguistics and translation studies, offering insights into the intersection of sentence structure and machine-generated literary translation.

Key Words: Machine Translation, Simple, Compound, and Complex Sentences, Literary Translation, Short Fiction, Syntax.

I. Introduction:

Machine Translation (MT) has revolutionized language accessibility in the digital era, playing an essential role in bridging linguistic and cultural divides. However, while MT systems have improved significantly in translating factual or technical content, translating literary texts, particularly on short fiction remains a complex task. Literary language is rich in figurative expressions, emotional undertones, idiomatic expressions and varied sentence structures. This paper investigates how sentence types, particularly simple, compound, and complex constructions, are handled when short fiction is translated from Telugu to English using MT tools. The study also discussed two more sentences, compound-complex and fragment which are traced in the selected prose work.

II. Objectives of the Study:

This research article aims to investigate selected text by a hybrid approach. By this, study focuses on to examine the effectiveness of Microsoft AI translation in preserving sentence structures from Telugu short fiction to English with RBMT model, to identify errors and inconsistencies in translating simple, compound, and complex sentences by the help of clauses and to understand how a sentence type influences translation of a literary prose.

III. Review of Literature:

Research studies in MT have long acknowledged challenges in translating structural differences in languages. Telugu, a Dravidian language, follows a subject+object+verb (SOV) order,

10.48047/jocaaa.2025.34.11.25

while English typically follows a subject+verb+object (SVO) order. This syntactic mismatch often results in structural shifts in MT output also. Researchers such as Pushpak Bhattacharya in *Machine Translation*, Makoto Nagao in *How Fat can It Go? Machine Translation* emphasized the importance of syntactic parsing and sentence segmentation in improving MT accuracy. In this book Bhattacharya introduces Rule Based Machine Translation concepts. This research outlines the information about the different methods to machine translation, as well as related works. However, limited work has been done specifically on Telugu to English literary translation and sentence type analysis.

IV. Methodology:

A qualitative analysis is being conducted using short excerpts from Telugu short fiction “Pagilina Gali Budaga”(1937) which is written by prominent author, Peddi Botla Subbaramaiah. Many methods were formulated for analyzing sentences. Here the translated English data is analysed with the Rule Based Machine Translation model (RBMT). Sentences were categorized as simple, compound, and complex based on standard grammatical definitions. The translated outputs were analysed for syntactic fidelity, semantic accuracy, and naturalness of expression.

V. NMT Vs. RBMT:

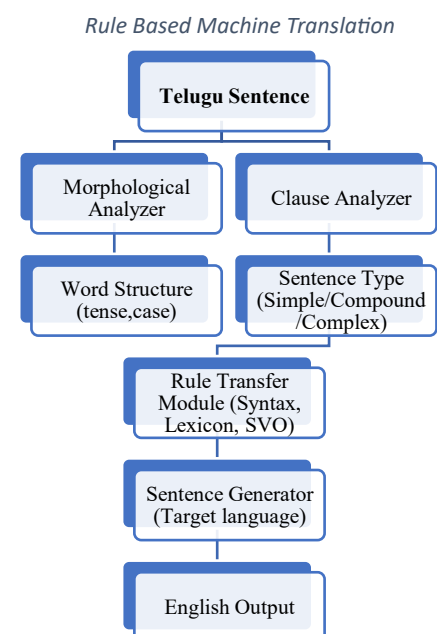
The difference between NMT (Neural Language Model Translation) and RBMT (Rule Based Machine Translation) is working with their approach to language processing and translation methodology. NMT works with deep learning models which is typically based on neural networks e from large parallel corpora to translate any language. Mostly, it works on neural model, like an encoder-decoder with particular attention. It is trained on millions of sentence pairs. It learns patterns, grammar, context and vocabulary from the data. Whereas, RBMT works on manually written linguistic rules and bilingual dictionaries to translate text from one language to another. When it comes to creative writings translation, both NMT and RBMT are not successful. RBMT is rule-based, transparent and grammar focused tool but not flexible. NLM/NMT is data-driven, context based and more fluent, but less explainable. Though NMT is advanced it has many technical hitches when it translates a literary text from one language to another language.

VI. Rule-Based Machine Translation (RBMT):

RBMT works by applying a set of linguistic rules (grammar, syntax, morphology, etc.) to translate sentences from a source language (Telugu) to a target language (English). The RBMT utilizes the linguistic information to decode the source and target languages. RBMT is superior in grammar and it is more predictable when we compared to other methods. When it comes to sentence type (simple, compound, and complex) RBMT handles them through structured, rule-governed steps. The present research article is following RBMT based MT for the select text. The flow chart figure shows how RBMT works in the MT.

VII. Introduction to Simple, Compound and Complex Sentences:

Sentence structure plays a dynamic role in shaping clarity, coherence, and association in written and spoken communication. This article explores the foundational types of sentences in English i.e., simple, compound, and complex which are



identified in the short fiction to demonstrate how each contributes to effective in language use. A simple sentence contains a single independent clause is delivering straightforward ideas with clarity. A compound sentence, made up of two or more independent clauses joined by coordinating conjunctions, which helps to link related ideas and maintain flow. A complex sentence combines one independent clause with one or more dependent clauses, allowing for nuanced expression of relationships such as cause, contrast, time, and condition. The study also discusses about other sentences of compound-complex sentences and fragment sentences. The compound-complex sentence is forms with two independent clauses and at least one dependent clause which expresses more ideas in one sentence. Whereas fragment sentences are not full sentences but they give expressive information and adds efficacy to the discourse. The analysis of the sentences has shown as Source Language (SL) and Target Language (TL). For SL text is given transliterate script also.

VIII. Analysis of Sentences:

1. Simple Sentences:

According to the structures, there are three kinds of sentences in English, i.e., Simple, Compound and Complex sentences. Simple Sentence Contains one independent clause which should has one subject and one finite verb. It expresses a single thought that can stand on its own.

Rule Based Machine Translation works on clause, subject+verb+object (SVO) structure. Hence it matches source sentence with a corresponding simple sentence pattern in the target language. It works on morphological rules to handle verb conjugation, case markers, tense, etc.

Example:

1. SL: పని చూసుకుందామని పైలు ముందుకు లాక్కున్నాడు మురారి.

(Pani cūsukundāmani phailu munduku lākkunnāḍu murāri)

TL : Murari pulled the file forward to look after the work.

Source data from Telugu to target language, English is translated by direct one-to-one mapping using grammar rules. MT tools were largely successful in translating simple sentences.

Example:

2. SL: కళ్ళు పైల్ లోకి చూస్తున్నాయి

(Kaḷḷu phail lōki cūstunnāyi)

TL: Eyes are looking into the file

3. SL : నేను కళ్లారా చూశాను

(Nēnu kaḷḷārā cūśānu)

TL: I saw it with my own eyes

4. SL: రోడ్లన్నీ జనంతో రద్దీగా కోలాహలంగా కిటకిటలాడుతున్నాయి.

(Rōḍḷannī janantō raddīgā kōlāhalaṅgā kiṭakiṭalāḍutunnāyi.)

TL: All the roads were crowded with people.

5. SL: బస్సు పక్కన ఇద్దరు ముగ్గురు అడుక్కునే వాళ్ళు కనిపించారు.

(Bas'su pakkana iddaru mugguru aḍukkunē vāḷḷu kanipiñcāru.)

TL: Two or three beggars were seen next to the bus.

Although there were no structural corrections regarding the sentence type, but some vocabulary related errors are observed. In the fourth example, SL sentence “కోలాహలంగా”, “కిటకిటలాడుతు” words are not translated. Hence, the true literariness of the creative writing is missing. The intensity of the action will not be conveyed appropriately. This is quite common in the creative works. The study has observed that there are errors in terms of TL words in the translated discourse.

2. Compound Sentences:

Compound Sentence contains two or more independent clauses joined by coordinating conjunctions. A sentence is made of two simple sentences joined by a coordinating conjunctions (For, and, nor, but, or, yet, still etc.)

In RBMT component it identifies conjunctions (and, but, yet etc.), splits the source sentence into clauses and translates each clause separately using the rules. Then finally machine rejoins clauses using English coordinators.

Example:

1. SL: మురారి కొంచెం ఆశ్చర్యపోయి అంతలోనే నవ్వుతూ 'బాగానే ఉన్నామండి అన్నాడు. (SL)

(Murāri koñcem āścaryapōyi antalōnē navvutū'bāgānē unnāmaṇḍi annāḍu.)

TL: Murari was a little surprised and smiled and said, 'We're fine. (TL)

Each clause is translated using rule sets, then joined with a coordinating conjunction. In the first example sentence co-ordinating conjunctions are joined with three independent clauses by the 'and'. "Murari was a little surprised", "smiled" and "we are fine" are independent clauses which are joined with 'and'. So, the sentence has three independent clauses by a single speaker.

2. SL: మురారి లేచి 'ధ్యాంక్స్' అంటూ ఇంకేమీ మాట్లాడలేక లేచి వచ్చేశాడు.

(Murāri lēci'thyāñks' aṅṭū inḱēmī māṭlāḍalēka lēci vaccēśāḍu)

TL: Murari got up and said, 'Thank you,' and, unable to say anything else, got up.

3. SL: కష్టించి పనిచేసి కడుపు కట్టుకుని తనను కాలేజీలో చేర్పించింది.

(Kaṣṭiñci panicēsi kaḍupu kaṭṭukuni tananu kālējīlō cērpīñcindi.)

TL: English: She worked hard and got her stomach shut and enrolled herself in college.

It often failed to join clauses correctly, misused conjunctions, or produced run-on sentences. In the second example multiple independent actions are joined by 'and' conjunction. If we observe, the third example MT has identified type of the SL sentence and changed into compound sentence. But meaning of the translation is completely wrong. "కడుపు కట్టుకుని" is translated as "stomach shut", this is not correct. Moreover, the sentence is about narrator's mother, her struggle for raising her children up. To translate this thought meaningfully into TL the MT has not successful. As a results appropriate meaning is not conveyed accurately. For the second examples also MT has done the same.

3. Complex Sentences:

Complex Sentence contains one independent/main clause and one or more dependent clauses. A Complex sentence always has a subordinate conjunction (as, because, since, after, though, when, where, if, unless etc.) or a relative pronoun (who, which, what, that etc.). Therefore, a complex sentence combines one independent clause with one or more dependent clauses, allowing for nuanced expression of relationships such as cause, contrast, time, and condition. Understanding the structure and function of these sentence types enables writers and speakers to convey meaning more precisely and adapt their style to different contexts. MT recognizes subordinators of clauses and identifies main clause and subordinate clauses. It applies specific rules for subordinate clause transformation on the discourse. Then it reconstructs the English sentence using English subordinators and proper word order.

Example:

1. SL: అది వినీ వినగానే మురారికి ఒంటినిండా తేళ్లు, జెర్రులూ పాకినట్టయింది.

(Adi vinī vinagānē murāriki oṅṅiniṅṅā tēllu, jerrulū pākinattayindi.)

TL: When Murari heard it, he felt as if his body was full of scorpions and jerks.

If we observe RBMT proces, it works as three steps. As the first step the machine identifies clauses of the sentence (dependent and independent) then in the second step, it separates dependent and main clauses with the suitable conjunctions and matches with English structures. In the third step it generates translated sentence in the Target Language. In the above example, ‘అది వినీ వినగానే’ is a dependent clause. ‘మురారికి ఒంటినిండా తేళ్లు, జెర్రులూ పాకినట్టయింది’ is a main (independent) clause. Here in TL the translated word ‘jerks’ is not suitable to the context.

The most significant issues were observed in translating complex sentences by MT subordinate clauses were often mistranslated or omitted. As a result, there was a loss of meaning for a sentence.

2. SL: శంకరి పొద్దున లేచినప్పటి నుండి వంచిన నడుం ఎత్తకుండా చాకిరి చేస్తుంటుంది.

(Śaṅkari podduna lēcinappaṭi nuṅṅi vaṅcina naḍuṅṅ ettakuṅṅā cākiri cēstuṅṅundi.)

TL: Ever since Shankari got up in the morning, she has been doing chakiri without lifting her bent waist.

3. SL: ఎన్నికలు దగ్గర పడటంతో నినాదాలూ అరుపులతో ఊరంతా అక్షరాల సంతలాగుంది.

(Ennikalu daggara paḍaṅṅtō ninādālū arupulatō ūraṅṅā akṣarāla santalāṅṅundi.)

TL: As the elections drew to a close, the whole town was filled with slogans and shouts.

RBMT has recognized one independent clause and two subordinate clauses in the SL sentences of the second and third examples. So that it has identified as complex sentences and translation has been done. According to the syntactical identification, MT is working efficiently. But when it comes at meaning and connectivity it is highly uncertain. The study has identified clearly that the part of the meanings of the sentences do not seem correct. This has been demonstrated by the above TL sentences.

Similarly, in the second and third examples MT identified the ST as complex sentences but not meaningfully translated into TL. Consequently, it is clear that MT is working effectively for the practical work, but not in the morphological work. If we observe the words in the second example SL sentence, “నడుం ఎత్తకుండా”, and “చాకిరి” are not shown in the TL, “without lifting her bent waist” and “chakiri” as a part of the meanings, they do not seem to have worked that effectively. In the third example, the word in SL “సంతలాగుంది.” is not translated in the TL.

All over, compound sentences presented moderate challenges. Whereas the most significant issues were observed by MT in translating complex sentences. Subordinate clauses where often mistranslated or omitted, resulting in loss of meaning.

4. Compound-complex Sentences:

In the discourse, some sentences are formed with compound-complex sentences. They are discussed by the below example.

1. SL: మురారి క్యాంటీన్ లో ఒక మూల కూర్చుని టీ తాగుతూ ఉండగా పని కట్టుకుని దూరం నుంచి వచ్చి మురారి పక్కనే కూర్చుని పలకరించాడు శంకరావు 'ఎలా ఉన్నారు సార్' అంటూ.

(Murāri kyāṅṅīṅ lō oka mūla kūrcuni ṅī tāgutunṅṅagā, śaṅkar rāvu dūraṅṅ nuṅṅi vacci murāri pakkana kūrcuni, 'elā unnāru sār?' Ani palakariṅcāḍu.)

10.48047/jocaaa.2025.34.11.25

TL: While Murari was sitting in a corner of the canteen and drinking tea, Shankar Rao came from a distance and sat down next to Murari and greeted him, 'How are you, sir?'

In this compound-complex sentence there is dependent clause which is combined with multiple co-ordinated actions. This is also come under complex sentence. MT has done from target source language, Telugu to target language, English with appropriate meaning. In this sentence dependent clause is “While Murari was sitting... drinking tea” and multiple coordinated main actions, “came”, “sat”, “greeted” are used. But it is clearly observed that the creative literary essence is missing in the TL sentence.

2. SL: శంకరావు ఉన్నట్టుండి పని కట్టుకొని వచ్చి పక్కనే కూర్చుని ఈ తరహా సోది ఎందుకు మొదలుపెట్టాడో అర్థం కాలేదు
(Śankarrāvu unnattunḍi pani kaṭṭukoni vacchi pakkane kūrcuni ī tarahā sōdi enduku modalupeṭṭāḍō arthaṁ kālēdu)
TL: Shankar Rao suddenly came to work and sat next to him and did not understand why he had started this kind of sodhi.

In the above example, MT has identified SL sentence as compound-complex sentence. Because here co-ordinated main actions are combined with dependent clause, “why”. Consequently, sentence type identification is done correctly but not successful in the terms of meanings. “సోది” in SL is not shown (translated) in TL. It is showing as it is in TL “sodhi”. Although there are no syntactical errors in the sentence but observed that there are errors in the terms of words.

5. Fragment Sentences:

The study also identified a few fragment sentences. Fragment sentences are the sentences which do not have any syntax but gives meanings.

1. SL: చాలా నెమ్మదిగా, సౌమ్యంగా.
(Cālā nem'madigā, saumyaṅgā)

TL: “Very slowly, very gently.”

Fragmented sentences are not full syntactical sentences, but present in the text and make the discourse more effective. Exclamatory fragment sentences are functioned as rhetorically. Such sentences are incomplete and give phrasal meaning only.

2. SL: 'నిజం.'
('Nijam.')

TL: It's true

3. SL: 'ఏం జీవితాలో ఏమిటో!
('Ēm jīvitālō ēmitō!)

TL: “What a life is all about!

Here, the third example is used as an exclamatory fragment sentence which is functioning as rhetorical and exclamatory. Usually, fragment sentences do not follow any fixed structures but they express strong feelings and emotions pertaining to the context. MT is successful in the first and second examples, but coming to the third example, the study has observed MT is not meaningfully done. Expression is missing when MT has translated from Telugu to English. Hence the fragmentation of emotional touch is vanished.

IX. Research Findings:

- With the analysis of the above example sentences of simple, compound and complex sentence by the RBMT rules, MT has been translating simple sentences by the help of direct transfer via lexical and grammar rules which are followed by morphology and word order.
- In the compound sentences, MT is identifying clauses and joining with conjunctions. It worked as clause segmentation and with co-ordinator rules.
- Complex sentences followed the strategies of identifying main and subordinating clauses with subordinator handling and clause ordering.
- Coming to the discussion, the analysis demonstrates that MT tools perform reasonably well with simple sentence structures but struggling with the syntactic and semantic complexities of compound and complex sentences.
- In literary translation, where nuance and connections are essential, such breakdowns significantly affect the quality and interpretability of the text. The tendency to oversimplify or fragment clauses results in a mechanical and sometimes incoherent output.
- The study examined translations of sentence types from Telugu to English by the RBMT system's accuracy, particularly in preservation of clause structure, conversion of sentence types, meaning of translated text and handling of conjunctions and subordinate elements which are correctly used.
- MT is implemented to translate simple, compound and complex sentences from the discourse formed using connectives a test suite of the sentences.

X. Conclusion:

This study highlights both the capabilities and limitations of machine translation in handling modern literary prose from Telugu to English. While MT can support basic translation tasks, it remains insufficient for capturing the stylistic and structural nuances of complex sentence forms in modern short fiction. The findings suggest a need for MT systems that are better attuned to syntactic structures and literary contexts of a generated text of a regional language. Further research involving larger corpora and advanced linguistic models may enhance the reliability of machine-assisted literary translation. Analysed syntactic richness and stylistic variability of short fiction, translating of texts poses unique challenges, especially when source and target languages differ significantly in structure and meanings. To get accuracy for creative literature human intervention is required.

XI. Appendix:

This table Include side-by-side examples of Telugu source, MT output, and reason of the sentences. It shows Simple, Compound and Complex sentences which are found from the first five paragraphs from the modern short fiction “Pagilinina Gali Budaga”.

Telugu Sentence	English Sentence	Type	Reason
పని చూసుకుందామని పైలు ముందుకు లాక్కున్నాడు మురారి	Murari pulled the file forward to look after the work.	Simple	One independent clause.
కళ్ళు పైల్ లోకి చూస్తున్నాయి	Eyes are looking into the file.	Simple	Single subject and verb.
కానీ అతని బుర్రలోకి ఏమీ ఎక్కడంలేదు	But nothing climbed into his head.	Simple	One independent clause.

మెదడు అంతా మొద్దు బారిపోయి ఒక రకమైన గ్లారిని ఆవహించింది.	The whole brain became numb and covered with a kind of glare.	Simple	One subject, compound predicate.
పైలు అవతలికి తోసి చిరాగ్గా చేతి వేళ్ళు విరుచుకొని ఒక చేత్తో జుట్టు పైకి తోసుకొని కుర్చీలో సాగిలపడ్డాడు.	Pushing the file to the other side, he broke his fingers in frustration, pushed his hair up with one hand, and leaned back in his chair.	Compound	Coordinated actions with "and."
ఒక్క క్షణం కళ్ళు మూసుకున్నాడు.	He closed his eyes for a moment.	Simple	One independent clause.
మాసిన చేతి రుమాలుతో ముఖం తుడుచుకున్నాడు.	He wiped his face with a faded handkerchief.	Simple	Single action.
లంచ్ టైంలో క్యాంటీన్లో జరిగిన వ్యవహారం తలుచుకుంటే ఒళ్ళు జలదరిస్తుంది.	When you think of what happened in the canteen during lunch time, it makes me tingle.	Complex	Dependent clause starting with "when."
మెదడుకు కరెంట్ షాక్ తగిలినట్లవుతుంది.	The brain is hit by an electric shock.	Simple	Single clause.
శంకరావు తన తోటి గుమస్తా మితభాషి	Shankar Rao is his fellow clerk Mitabhasi.	Simple	One subject, one predicate.
అనవసరమైన విషయాల్లో జోక్యం పెట్టుకునే తరహా కాదు.	It is not the kind of interference in unnecessary things.	Simple	One independent clause.
తన పనేమిటో తానేమిటో తప్ప చవకబారు ప్రసంగాలు చేయడం ఇతరుల గురించి మాట్లాడటం అతనికి అలవాటు లేదు.	He is not used to making cheap speeches and talking about others except what he is doing and what he is.	Complex	Contains relative/dependent clauses.
నలుగురిలో పై కారణాల చేతనే పెద్ద మనిషిగా చలామణి అవుతున్నాడు	Of the four, it is for the above reasons that he circulates as a big man.	Complex	Contains a dependent phrase explaining a cause.
మురారి క్యాంటీన్ లో ఒక మూల కూర్చుని టీ తాగుతూ ఉండగా పని కట్టుకుని దూరం నుంచి	While Murari was sitting in a corner of the canteen and drinking tea, Shankar Rao came	Compound-Complex	Dependent "while" clause + multiple coordinated actions.

వచ్చి మురారి పక్కనే కూర్చుని పలకరించాడు శంకరావు 'ఎలా ఉన్నారు సార్' అంటూ.	from a distance and sat down next to Murari and greeted him, 'How are you, sir?'		
--	--	--	--

This research article is a part of the project, entitled, "Prospects and Challenges of Machine Translation of Telugu Literature with reference to Modern Short Fiction and Modern Poetry". This research project is financially supported by PM-USHA.

Works Cited:

- Bhattacharyya, Pushpak. Machine Translation. CRC Press, New York. 2015.
- Abrams, M. H. 1993. A Glossary of Literary Terms. Orlando, FL: Harcourt Brace College Publishers.
- Sri, Sri. *Mahaprasthanam*, "Chedu Paata", 1937. Vishalandhra Publication, August 2013.
- Bharati, A., Chaitanya, V., & Sangal, R. (2000). Natural Language Processing: A Paninian Perspective.
- Bojar, O., et al.. "Machine Translation" (WMT16).2016.
- Venkat Rao, C. "Translation and Transformation: Essays on South Indian Literature". 2014.
