# A literature review on Interactive Telugu-Language Dialogue System

SADARAM VIDYA SAGAR[1], DR. SRINIVASU BADUGU[2]
[1] *Research Scholar, Department of Computer Science and Engineering (IT), University College of Engineering (A), Osmania University, Hyderabad.*
[2] *Professor, Department of Computer Science and Engineering, Stanley college of Engineering & Technology for women (A), Abids, Hyderabad.*

***Abstract*— In an increasingly interconnected world, the increase in demand for versatile and effective multilingual Natural Language Processing (NLP) systems is being driven. This demand is seen as a significant challenge in regions characterized by rich linguistic diversity, such as Telugu, where NLP applications are currently underrepresented. The endeavour to pioneer an innovative solution addressing this challenge is the focus of this research, with emphasis on the design and development of an interactive Telugu-language dialogue system. This system is equipped with capabilities for keyword extraction, question validation, and multilingual sentiment analysis. Current NLP systems often fall short when applied to low resourced languages like Telugu. The inherent complexities of language, coupled with the limited availability of training data and linguistic resources, pose substantial hurdles. Furthermore, the critical requirement of complete and accurate information for effective dialogue systems remains largely unaddressed. A holistic approach to tackle these challenges is proposed by this research. A novel interactive dialogue system is introduced, specifically designed for the Telugu language, where the integration of advanced keyword extraction and question validation mechanisms ensures that user queries can be effectively understood and processed by the system. Additionally, state-of-the-art Zero-Shot Learning techniques are leveraged to develop a multilingual sentiment analysis module tailored to Telugu language, thereby enhancing the system's capability to comprehend and respond to user sentiment. As an additional innovation, the incorporation of English vocabulary models is explored to improve the overall performance of the Telugu language dialogue system. The aim is to enhance the system's responsiveness, comprehension, and user experience by bridging the linguistic gap between Telugu and English.***

***Index Terms- Telugu-Language Text Dialogue System, Natural Language Processing, Tokenization, Part-of-Speech Tagging, Keyword Extraction.***

## I. INTRODUCTION

The topic of interactive dialogue system encompass a broad spectrum of human interaction, encompassing language diversity, technology, and the evolving nature of communication. In today's globalized society, an astonishing array of languages is spoken, reflecting the rich cultural tapestry of our world. This linguistic diversity, while a source of cultural richness, also poses formidable challenges for effective communication. Language barriers can impede various aspects of human interaction, from basic conversations to accessing critical information. These barriers can affect daily activities, ranging from seeking medical care to participating in educational endeavors. Bridging these gaps is paramount for fostering inclusivity, accessibility, and societal progress.

The phenomenon of language diversity is a testament to the intricate tapestry of cultures and histories that comprise our global community [1]. Thousands of languages are spoken worldwide, each representing a unique reservoir of cultural heritage and identity. However, this diversity also gives rise to substantial communication challenges. In multilingual societies, individuals often encounter situations where they need to navigate between languages for various aspects of their lives, be it education, healthcare, or civic engagement. For instance, in a healthcare setting, effective communication between healthcare providers and patients is paramount for accurate diagnosis and treatment. Language barriers can obstruct this process, potentially leading to misunderstandings, misdiagnoses, or inadequate care. In addition, businesses and organizations operating in a globalized landscape face the imperative of effective

communication with stakeholders from diverse linguistic backgrounds [2]. From multinational corporations conducting international business transactions to non-profit organizations implementing humanitarian projects in remote regions, the ability to bridge language divides is instrumental for success. The impact of language diversity is not limited to physical interactions; it extends to the digital realm as well. In an era dominated by online content and social media, ensuring accessibility and inclusivity for speakers of various languages is pivotal for reaching a wide audience and fostering an inclusive digital environment. Therefore, addressing these communication challenges necessitates innovative solutions within the realm of Natural Language Processing (NLP). This encompasses the development of robust translation systems, intelligent dialogue interfaces, and technologies that facilitate seamless multilingual interactions [3].

Telugu, one of the classical languages of India, holds a venerable position in the linguistic landscape of the Indian subcontinent. Spoken by millions of people primarily in the states of Andhra Pradesh and Telangana, Telugu boasts a rich literary heritage that spans centuries. The development of Telugu-language dialogue systems carries profound significance in the broader context of linguistic inclusivity and technological accessibility [4]. By enabling seamless interaction with technology in their native language, Telugu speakers can access a wide array of services, ranging from information retrieval to entertainment, with ease and efficiency.

Moreover, Telugu-language dialogue systems play a pivotal role in democratizing access to information and services. In domains such as education and public services, where accessibility and inclusivity are paramount, the availability of technology in one's native language is a fundamental right. These systems have the potential to empower individuals who may have limited proficiency in other languages, particularly English. In education, for instance, students can benefit immensely from learning materials and educational resources presented in their mother tongue [5]. This not only enhances comprehension but also fosters a deeper connection with the subject matter. Additionally, in public services, such as government portals or helplines, providing information in Telugu can bridge the gap between citizens and essential services, ensuring equitable access for all.

Technological advancements have emerged as a critical solution to these challenges. Natural Language Processing (NLP), a subfield of artificial intelligence, is dedicated to empowering machines to comprehend and process human language. Its ambit encompasses a wide array of tasks, including language translation, sentiment analysis, and dialogue systems. Recent years have witnessed a remarkable surge in the capabilities of NLP, with systems like ChatGPT showcasing the potential to facilitate multilingual communication and provide support across diverse linguistic landscapes.

Furthermore, the evolution of NLP is intricately entwined with the broader digital revolution. The proliferation of the internet and the rise of digital communication platforms have propelled the need for sophisticated language processing technologies. From online customer support chatbots to automated content creation, NLP is at the forefront of enabling machines to navigate the intricacies of human language, thereby enhancing accessibility and user experience in the digital realm. In essence, the background and context of this topic encapsulate a dynamic interplay between language diversity, technological innovation, and the imperative of effective communication in our interconnected world.

- Natural Language Processing (NLP)

Natural Language Processing (NLP) is a multidisciplinary field at the intersection of computer science, artificial intelligence, and linguistics [6]. It focuses on enabling computers to understand, interpret, generate, and respond to human language in a manner that is both meaningful and contextually relevant. NLP seeks to bridge the gap between human communication and computational understanding, allowing machines to interact with humans in a way that feels natural and intuitive.

- Language Understanding:

At the core of NLP lies the task of language understanding. This involves breaking down complex linguistic structures into smaller, manageable units. These units can range from individual characters and

words to phrases, sentences, and even entire documents. Techniques like tokenization and syntactic analysis are used to parse and analyse text, extracting essential information about its structure, grammar, and semantics.

- Semantics and Meaning Extraction:

Understanding the meaning of language is a central challenge in NLP. This involves discerning the intent behind a piece of text, identifying entities (such as names, places, and dates), and extracting relationships between them. Techniques like named entity recognition (NER) and dependency parsing are employed to achieve this. Additionally, sentiment analysis aims to determine the emotional tone conveyed in a text, which can be crucial in applications like social media monitoring, customer feedback analysis, and market sentiment tracking.

- Language Generation:

In addition to understanding language, NLP is concerned with generating coherent and contextually appropriate responses. This aspect of NLP involves tasks like text summarization, machine translation, and text generation. Cutting-edge models like the transformer architecture, which powers models like GPT-3, have revolutionized the ability to generate human-like text by leveraging techniques such as self-attention and deep learning.

1. Research gap

Researching the gap in developing an interactive dialogue system for the Telugu language using machine learning (ML) and deep learning (DL) models involves understanding the current state of technology and identifying areas that need improvement or lack sufficient research. Here are some key research gaps that could be explored:

Data Scarcity and Quality

- Limited Annotated Data: There is a significant lack of high-quality annotated datasets for training and evaluating dialogue systems in Telugu.
- Domain-Specific Data: Few domain-specific datasets exist for specialized applications like healthcare, customer service, and education.

Linguistic Challenges

- Morphological Richness: Telugu, being a morphologically rich language, poses challenges in tokenization, lemmatization, and part-of-speech tagging.
- Dialectal Variations: Variations in dialects and colloquial usage can impact the performance of dialogue systems.

Model Adaptation

- Transfer Learning: There is a need for research on effective transfer learning techniques to adapt pre-trained models (typically in English) to Telugu.
- Multilingual Models: Exploring the potential of multilingual models like mBERT and XLM-R for Telugu dialogue systems.

Context Understanding and Management

- Contextual Understanding: Enhancing the ability of dialogue systems to understand and maintain context over long conversations in Telugu.
- Code-Switching: Handling code-switching between Telugu and English, which is common in real-world conversations.

Evaluation Metrics

- Standardized Benchmarks: Lack of standardized benchmarks and evaluation metrics for Telugu dialogue systems.
- User-Centric Evaluation: More focus on user-centric evaluation methods to assess system performance from an end-user perspective.

Semantic and Pragmatic Understanding

- Semantic Parsing: Developing robust semantic parsing techniques tailored for Telugu.
- Pragmatic Understanding: Addressing the nuances of pragmatics, such as implicature and speech acts, in Telugu dialogue.

Dialogue Management Strategies

- Reinforcement Learning: Exploration of reinforcement learning approaches for dialogue management in Telugu.
- Hybrid Models: Combining rule-based and data-driven approaches for more effective dialogue management.

Cultural and Societal Context

- Cultural Sensitivity: Ensuring that dialogue systems are culturally sensitive and contextually relevant to Telugu-speaking users.
- Ethical Considerations: Addressing ethical issues such as bias, fairness, and user privacy in dialogue systems.

Real-Time Processing

- Latency: Reducing latency in real-time interactive dialogue systems.
- Scalability: Ensuring scalability of dialogue systems to handle large volumes of interactions.

Application-Specific Challenges

- Education: Designing dialogue systems that can cater to educational needs, including tutoring and language learning in Telugu.
- Healthcare: Developing dialogue systems for healthcare applications that can effectively understand and respond to medical queries in Telugu.

Focusing on these gaps can guide researchers and developers in creating more effective and robust interactive dialogue systems for the Telugu language using machine learning and deep learning models.

## II. LITERATURE

Telugu-language interactive dialogue system (chatbot) Rambabu Banothu et al [9] focused to elucidate the Speech-to-Speech connection between farmers and the government by using a speech recognition system for the Telugu language. This system will serve as the interface for a website that disseminates information about government initiatives and commodities pricing via voice-based communication. he user's text is already academic.

Arijit Das et al [10] presented a system for the automated process of question-answering in natural language, using a vast reservoir of unstructured text corpora. Traditionally, computer systems have shown greater efficiency in extracting information from big datasets or repositories via the use of keyword matching. Web search operates based on the fundamental concept of keyword matching. In order to enhance the ranking of the gathered list of URLs, prominent search engines use many factors such as the page-rank algorithm, user's web cache, location, preference, and other relevant considerations. The fundamental concept behind the quality assurance (QA) system differs from that of syntactic search algorithms. In this context, the desired response is anticipated regardless of the quantity of corresponding phrases included in both the query and the answer. The difficulty of developing a QA system in English and other languages with rich resources is addressed by the use of machine-readable dictionaries, such as WordNet, as well as ontologies. The primary obstacles in developing a question answering (QA) system for low resource languages such as Bengali are the lack of annotated text, inadequate WordNet, and limited ontology.

*Birgit Popp et al [11] proposed The Chatbot Language (CBL) framework as a means to efficiently create and implement Conversational User Interfaces (CUI) using crowd sourcing platforms, eliminating the need for specialized technical knowledge. The CBL library is designed to provide specific Command-Line User Interface (CUI) capabilities, using the JavaScript programming language at a high-level.*

*Hao-Yung Chan et al [12] designed and implemented a conversation system that addresses the challenges faced by decision makers at emergency operation centers (EOC) in effectively and accurately using information. The exponential increase in data volume has posed challenges in decision-making processes within the field of catastrophe management.*

Ryuichiro Higashinaka et al [13] organized a competition named "The Dialogue System Live Competition", whereby the primary audience included mostly of researchers within the conversation community. The event involved the observation and assessment of a real-time discussion between users and dialogue systems. The primary objective of the event was to foster the development of cutting-edge methodologies in dialogue systems and facilitate the exchange of challenges encountered in existing discussion systems within the dialogue community. The competition consists of two distinct components: preliminary selection and a live event. During the first phase of the selection process, a total of eleven

systems were assessed via the use of crowd-sourcing. Three systems were present at the live event, engaging in conversations with selected speakers and being assessed by the audience.

- Multilingual sentiment analysis for interactive dialogue system

Mohammad AL-Smadi et al [14] included the creation and implementation of a deep learning model using Gated Recurrent Units (GRU) and characteristics derived from the Multilingual Universal Sentence Encoder (MUSE). The Pooled-GRU model, which has been suggested, was trained using Arabic evaluations of hotels in order to tackle two ABSA tasks: (1) the extraction of aspects, and (2) the categorization of aspect polarity.

Mauajama Firdaus et al [15] proposed a multilingual multitask methodology to integrate intent detection and slot filling, which are the two major tasks in spoken language understanding (SLU), across three distinct languages. Intent detection involves the identification of the user's objective or intention, whereas slot filling entails the collection of relevant information from the user's utterance in the form of slots. Given the strong correlation between these two objectives, we suggest using a multitask approach to address both simultaneously. A transformer model is used as a common phrase encoder for three languages, namely English, Hindi, and Bengali.

Haiyun Peng et al [16] suggest the incorporation of two efficacious attributes for encoding phonetic information, hence facilitating its integration with textual information. In this study, we present the hypothesis of Disambiguate Intonation for Sentiment Analysis (DISA), a neural network model that is constructed using the principles of reinforcement learning. The DISA system disambiguates the intonations associated with each Chinese letter (pinyin), therefore acquiring accurate phonetic representations.

Anas El-Ansari et al [17] suggested the incorporation of two efficacious attributes for encoding phonetic information, hence facilitating its integration with textual information. In this study, we present the hypothesis of Disambiguate Intonation for Sentiment Analysis (DISA), a neural network model that is constructed using the principles of reinforcement learning. The DISA system disambiguates the intonations associated with each Chinese letter (pinyin), therefore acquiring accurate phonetic representations.

Ahlam Fuad et al et al [18] presents AraConv, the first Arabic end-to-end generative model for task-oriented dialogue systems. AraConv utilizes the multi-lingual transformer model mT5 with various configurations. In addition, we provide an Arabic conversation dataset, namely Arabic-TOD, which was used for training and evaluating the proposed AraConv algorithm.

- Multi-lingual vocabulary models

Sebastián E. Rodríguez et al [19] focused on identifying hate speech in environments that are monolingual, multilingual, and cross-lingual. In order to do this, we employed a pre-trained language model called Language Agnostic BERT Sentence Embeddings (LabSE), which we used for both feature extraction and end-to-end classification. We examined a variety of representations, including bags of words, bags of characters, and sentence embeddings taken from the Multi-lingual BERT, as well as models like Support Vector Machines and tree-based models.

Sahil Manoj Bhatt et al [20] proposed Five NLG tasks for the tourism area in the benchmark TOURISMNLG, and associated datasets with conventional train, validation, and test splits are made available. Additionally, earlier data science proposals for tourist issues did not take use of transfer learning's most current advantages.

Faisal Alshargi et al [21] presented Xword, an online multi-lingual framework designed for the purpose of automated word extension. The expansion job in Xword is supported by a combination of pre-trained ad hoc word embedding models and n-gram models. Xword now incorporates two linguistic systems, namely Arabic and German. The Xword platform shows the outcomes of each model, both in isolation and as a whole.

Ayaz H. Khan et al [22] presented a model for the categorization of tweets in several languages, namely English and Roman Urdu, over a varied spectrum of

classes. The model is designed using a non-hierarchical architecture. Previous research in the field of tweet categorization has shown a limited scope, mostly concentrating on either a single language or a uniform number of classifications, often including positive, very positive, negative, and extremely negative sentiments. The model being presented is grounded on the principles of semi-supervised learning. The suggested feature selection strategy enhances the model's autonomy and adaptability in identifying and incorporating current and popular phrases.

Amandeep Singh Dhanjal et al [23] addressed the existing gap by proposing the development of an automated system, referred to as the Speech-to-Indian Sign Language Avatar (SISLA). This system aims to facilitate the translation of spoken language into Indian Sign Language. The whole system operates via a three-phase process. (i) The first stage encompasses the use of speech recognition (SR) technology to identify isolated words in English, Hindi, and Punjabi, within a speaker-independent setting. (ii) The subsequent stage involves the translation of the source language into Indian Sign Language (ISL). The ISL motions are represented by a 3D avatar that is built on HamNoSys. The application of SISLA has four primary components, namely requirement analysis, data collecting, technological development, and assessment.

## CONCLUSION

The core problem addressed by this research is the lack of sophisticated Natural Language Processing solutions tailored specifically for the Telugu language. This leads to a significant gap in digital communication tools for a large population of Telugu speakers. Additionally, the challenge of sentiment analysis in multilingual contexts further exacerbates this issue, as existing methods often struggle to provide accurate insights from non-English text. Furthermore, the problem of integrating English vocabulary models in a way that enhances, rather than hinders, the performance of a Telugu dialogue system represents a critical area of investigation. Balancing the nuances of two distinct languages within a single system requires careful formulation and testing.

Overall, the research seeks to address these multifaceted challenges in order to create more inclusive, accurate, and efficient NLP solutions for the Telugu-speaking community and beyond. To achieve this integrating English vocabulary models aims to enhance the system's adaptability to a bilingual user base, a multilingual sentiment analysis-based interactive dialogue system in Telugu through Zero-Shot Learning apart from preparing and aligning the English vocabulary with the Telugu dialogue system. Techniques such as translation, transliteration, or creating bilingual dictionaries may be employed. Additionally, preprocessing steps like tokenization and normalization ensure seamless integration.

## REFERENCES

[1] Kumar, Sunil, Arpan Kumar Kar, and P. Vigneswara Ilavarasan. "Applications of text mining in services management: A systematic literature review." *International Journal of Information Management Data Insights* 1, no. 1 (2021): 100008.

[2] Sivarajah, Uthayasankar, Zahir Irani, Suraksha Gupta, and Kamran Mahroof. "Role of big data and social media analytics for business to business sustainability: A participatory web context." *Industrial Marketing Management* 86 (2020): 163-179.

[3] Kusal, Sheetal, Shruti Patil, Jyoti Choudrie, Ketan Kotecha, Deepali Vora, and Ilias Pappas. "A systematic review of applications of natural language processing and future challenges with special emphasis in text-based emotion detection." *Artificial Intelligence Review* (2023): 1-87.

[4] Ma, Yukun, Khanh Linh Nguyen, Frank Z. Xing, and Erik Cambria. "A survey on empathetic dialogue systems." *Information Fusion* 64 (2020): 50-70.

[5] Zhang, Zheng, Ryuichi Takanobu, Qi Zhu, MinLie Huang, and XiaoYan Zhu. "Recent advances and challenges in task-oriented dialog systems." *Science China Technological Sciences* 63, no. 10 (2020): 2011-2027.

[6] Rahman, Mahbubur, Sara Nowakowski, Ritwick Agrawal, Aanand Naik, Amir Sharafkhaneh, and Javad Razjouyan. "Validation of a natural

language processing algorithm for the extraction of the sleep parameters from the polysomnography reports." In *Healthcare*, vol. 10, no. 10, p. 1837. MDPI, 2022.

[7] Li, Xinlu, Lexuan Zhang, Liangkuan Fang, and Pei Cao. "Multi-Intent Natural Language Understanding Framework for Automotive Applications: A Heterogeneous Parallel Approach." *Applied Sciences* 13, no. 17 (2023): 9919.

[8] Khan, Wahab, Ali Daud, Khairullah Khan, Shakoor Muhammad, and Rafiul Haq. "Exploring the frontiers of deep learning and natural language processing: A comprehensive overview of key challenges and emerging trends." *Natural Language Processing Journal* (2023): 100026.

[9] Banothu, Rambabu, S. Sadiq Basha, Nagamani Molakatala, Veerendra Kumar Gautam, and Suryakanth V. Gangashetty. "Speech Based Access of Kisan Information System in Telugu Language." In *Intelligent Human Computer Interaction: 12th International Conference, IHCI 2020, Daegu, South Korea, November 24–26, 2020, Proceedings, Part I 12*, pp. 287-298. Springer International Publishing, 2021.

[10] Das, Arijit, and Diganta Saha. "Question Answering System Using Deep Learning in the Low Resource Language Bengali." *Convergence of Deep Learning In Cyber-IoT Systems and Security* (2022): 207-230.

[11] Popp, Birgit, Philip Lalone, and Anna Leschanowsky. "Chatbot Language-crowdsource perceptions and reactions to dialogue systems to inform dialogue design decisions." *Behavior Research Methods* 55, no. 4 (2023): 1601-1623.

[12] Chan, Hao-Yung, and Meng-Han Tsai. "Question-answering dialogue system for emergency operations." *International Journal of Disaster Risk Reduction* 41 (2019): 101313.

[13] Higashinaka, Ryuichiro, Kotaro Funakoshi, Michimasa Inaba, Yuiko Tsunomori, Tetsuro Takahashi, and Reina Akama. "Dialogue system live competition: identifying problems with dialogue systems through live event." In *Increasing Naturalness and Flexibility in Spoken Dialogue Interaction: 10th International Workshop on Spoken Dialogue Systems*, pp. 185-199. Springer Singapore, 2021.

[14] Mohammad, AL-Smadi, Mahmoud M. Hammad, A. Sa'ad, AL-Tawalbeh Saja, and Erik Cambria. "Gated recurrent unit with multilingual universal sentence encoder for Arabic aspect-based sentiment analysis." *Knowledge-Based Systems* (2021): 107540.

[15] Firdaus, Mauajama, Asif Ekbal, and Erik Cambria. "Multitask learning for multilingual intent detection and slot filling in dialogue systems." *Information Fusion* 91 (2023): 299-315.

[16] Peng, Haiyun, Yukun Ma, Soujanya Poria, Yang Li, and Erik Cambria. "Phonetic-enriched text representation for Chinese sentiment analysis with reinforcement learning." *Information Fusion* 70 (2021): 88-99.

[17] El-Ansari, Anas, and Abderrahim Beni-Hssane. "Sentiment Analysis for Personalized Chatbots in E-Commerce Applications." *Wireless Personal Communications* 129, no. 3 (2023): 1623-1644.

[18] Fuad, Ahlam, and Maha Al-Yahya. "AraConv: Developing an Arabic task-oriented dialogue system using multi-lingual transformer model mT5." *Applied Sciences* 12, no. 4 (2022): 1881.

[19] Rodríguez, Sebastián E., Héctor Allende-Cid, and Héctor Allende. "Detecting hate speech in cross-lingual and multi-lingual settings using language agnostic representations." In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 25th Iberoamerican Congress, CIARP 2021, Porto, Portugal, May 10–13, 2021, Revised Selected Papers 25*, pp. 77-87. Springer International Publishing, 2021.

[20] Bhatt, Sahil Manoj, Sahaj Agarwal, Omkar Gurjar, Manish Gupta, and Manish Shrivastava. "TourismNLG: A Multi-lingual Generative Benchmark for the Tourism Domain." In *European Conference on Information Retrieval*, pp. 150-166. Cham: Springer Nature Switzerland, 2023.

[21] Alshargi, Faisal, Saeedeh Shekarpour, and Waseem Alromema. "Xword: A Multi-lingual

Framework for Expanding Words." In *Emerging Trends in Intelligent Computing and Informatics: Data Science, Intelligent Information Systems and Smart Computing 4*, pp. 162-175. Springer International Publishing, 2020.

[22] Khan, Ayaz H., and Muhammad Zubair. "Classification of multi-lingual tweets, into multi-class model using Naïve Bayes and semi-supervised learning." *Multimedia Tools and Applications* 79 (2020): 32749-32767.

[23] Dhanjal, Amandeep Singh, and Williamjeet Singh."An automatic machine translation system for multi-lingual speech to Indian sign language." *multimedia Tools and Applications* (2022): 1-39.