IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Connecting People Through Language Translation using AI (Mobile App)

1.Anwesh Tayade , 2.Darshan Dhole, 3.Amit Vyawahare ,4.Ayush Shinde,5.Prof. Anuradha Kale
[1]Student Researcher, [2]Student Researcher, [3]Student Researcher, [4]Student Researcher, [5]Instructor
[1]Computer Science and Engineering, [1]PRMCEAM Badnera, Amravati, India, [2]Computer Science and Engineering, [2]PRMCEAM Badnera, Amravati, India, [3]Computer Science and Engineering, [3]PRMCEAM Badnera, Amravati, India, [4]Computer Science and Engineering, [4]PRMCEAM Badnera, Amravati, India, [5]Computer Science and Engineering, [5]PRMCEAM Badnera, Amravati, India

Abstract: India's linguistic diversity, featuring 22 recognized languages and numerous local dialects, poses a major obstacle to digital inclusion and intercultural communication. Although global translation tools excel in standardized languages, they often face challenges with the subtleties of regional expressions, local sayings, and culturally unique terms. This document suggests creating a localized translation app aimed at promoting genuine interaction among speakers of various Indian languages. Employing advanced Large Language Models (LLMs) and Natural Language Processing (NLP) methods, the study concentrates on recognizing and translating "non-literal" regional phrases that conventional systems frequently misread. The approach includes refining multilingual models using selected datasets of informal language and utilizing Retrieval-Augmented Generation (RAG) to deliver context-sensitive translations. Initial findings indicate that emphasizing phrase-level cultural context greatly enhances user understanding and learning results in contrast to literal translation. This study aids in the democratization of digital resources in India, providing a scalable approach to close the communication divide in education, commerce, and social engagement.

.

Index Terms **- Component, formatting, style, styling, insert.**

I. INTRODUCTION

In an increasingly globalized society, the ability to communicate across borders is no longer a luxury but a necessity. However, despite the physical world being more connected than ever through travel and digital infrastructure, a profound "language silo" persists. With over 7,000 languages spoken globally, millions of individuals remain excluded from global discourse, economic opportunities, and social integration due to linguistic barriers.

Traditional methods of translation—ranging from physical lexicons to rule-based digital translators—often lacked the nuance, speed, and portability required for real-time human connection. The emergence of Machine Translation (MT) marked a significant shift, but it was the advent of Artificial Intelligence (AI) and Neural Machine Translation (NMT) that truly revolutionized the field by allowing systems to understand context, idioms, and syntax rather than just word-for-word substitution.

While web-based translation tools exist, they often fail to address the spontaneity of face-to-face interaction. There is a critical need for a seamless, mobile-centric solution that leverages AI to provide instantaneous, high-fidelity translation. Current challenges include maintaining accuracy in diverse dialects, reducing latency in speech-to-speech communication, and ensuring the technology is accessible to non-technical users in everyday scenarios.

This research presents the development of a mobile application designed to bridge these gaps using advanced AI frameworks. By integrating **Natural Language Processing (NLP)** and real-time voice recognition, the proposed application aims to foster genuine human connection. This paper explores the technical architecture of the app, evaluates its translation accuracy across multiple language pairs, and discusses its potential impact on social and professional inclusivity.

Problem Statement:

1. The Contextual Accuracy Gap

Most conventional translation tools struggle with the nuances of human speech, such as idioms, cultural metaphors, and regional dialects. When AI lacks deep contextual understanding, the resulting translation is often "robotic" or factually incorrect, leading to misunderstandings that can hinder social bonding or professional trust.

2. High Latency in Real-Time Interaction

For a translation tool to truly "connect" people, it must operate at the speed of natural conversation. Existing web-based or heavy-duty software often suffers from processing delays (latency). In a face-to-face setting, a delay of even a few seconds breaks the flow of dialogue, making the technology a barrier rather than a bridge.

3. Lack of Portability and Accessibility

While powerful AI models exist on desktop platforms, they are often unavailable to users in the "moment of need"—such as a traveller at a local market or a healthcare worker in the field. There is a lack of optimized, mobile-first AI applications that provide high-performance translation without requiring high-end hardware or constant high-speed data.

4. The "Human" Element in UI/UX

Many current apps are designed as utility tools rather than connection tools. They focus on text-in/text-out functionality but ignore the user experience (UX) required for two people to share a single device or engage in a fluid, hands-free conversation.

Literature Review:

Name of the paper	Authors	Methodologies	Year published
A Semantic Context-Aware Automatic Quality Scoring Method for Machine Translation Based on Pretraining Language Model	FANGMIN TAN1AND HUAJU WANG	Pretrained Language Model (PLM) A **novel PLM** is established which is explicitly designed to **combine multiple key features and tasks**. This specialized architecture is used for encoding both the **large-scale initial (source) sentences and object (target/machine translation) sentences**.	2024
A Word Sense Disambiguation Method Applied to Natural Language Processing for the Portuguese Language	CLOVIS HOLANDA DO NASCIMENTO , VINICIUS CARDOSO GARCIA , AND RICARDO DE ANDRADE ARAÚJO	BERT model specifically pre-trained for the Portuguese language. **BERT is a contextual powerhouse** that revolutionized NLP by allowing models to grasp the full, simultaneous context	2024

		of a sentence to understand the meaning of each word.	
Img2Vocab: Explore Words Tied to Your Life With LLMs and Social Media Images	KANTA YAMAOKA , KO WATANABE, KOICHI KISE, (Member, IEEE), ANDREAS DENGEL , AND SHOYA ISHIMARU 4,	**System Prototype:** An online vocabulary exploration system was prototyped. **Personalization:** The core mechanism was displaying a learner's selected Instagram photos alongside a GPT-3 generated sentence that utilized vocabulary tied to the context of the image.	2025
Framework for Deep Learning-Based Language Models Using Multi-Task Learning in Natural Language Understanding: A Systematic Literature Review and Future Directions	RAHUL MANOHAR SAMANT , MRINAL R. BACHUTE, SHILPA GITE, AND KETAN KOTECHA	**Systematic Search:** A defined process for searching the literature (between 2011 and 2021) to identify relevant language models used in NLU and NLP. **Identification and Investigation:** Identifying, investigating, and analyzing various language models. **Analysis and Synthesis:** The review points out that unsupervised learning method-based language models show potential but face challenges in designing a general-purpose framework. **Proposal/Conclusion:** Proposing building steps for a **conceptual framework** to enhance the performance of language models in NLU.	2022

Name of the paper	Authors	Methodologies	Year published
A Novel Approach to Continual Knowledge	SHAILASHREE K. SHESHADRI, DEEPA GUPTA, BISWAJIT PAUL, AND J. SIVA BHAVAN	**Continual Knowledge Transfer (CKT)** This is the central methodological	2025

Transfer in Multilingual Neural Machine Translation Using Autoregressive and Non-Autoregressive Models for Indic Languages		innovation. CKT is designed for efficient and scalable incremental adaptation to new languages while minimizing **catastrophic forgetting** (parameter interference) of previously learned languages.	
Implicit Cross-Lingual Word Embedding Alignment for Reference-Free Machine Translation Evaluation	MIN ZHANG, HAO YANG, YANQING ZHAO, XIAOSONG QIAO, SHIMIN TAO, SONG PENG, YING QIN, AND YANFEI JIANG	The methodologies used in this work center on using **Multilingual Knowledge Distillation** to achieve **cross-lingual word embedding alignment** implicitly, and then applying this resulting alignment to **reference-free machine translation (MT) evaluation** using established metrics.	2023
Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation	B. NATARAJAN, E. RAJALAKSHMI, R. ELAKKIYA, KETAN KOTECHA, AJITH ABRAHAM, LUBNA ABDELKAREIM GABRALLA , AND V. SUBRAMANIYASWAMY	The methodologies used in the proposed framework are divided into two main components: one for **Sign Language Recognition and Text Generation** and another for **Sign Gesture Video Production** from spoken sentences.	2022

Name of the paper	Authors	Methodologies	Year published
Toward Low-Resource Languages Machine Translation: A Language-Specific Fine-Tuning With LoRA for Specialized Large	XIAO LIANG, YEN-MIN JASMINA KHAW, SOUNG-YUE LIEW, TIEN-PING TAN AND DONGHONG QIN	**Language-Specific Fine-Tuning with Low-rank adaptation (LSFTL)** provides a targeted and computationally	2024

Language Models		lightweight "language injection" for Large Language Models, allowing them to rapidly and effectively learn the unique grammar and vocabulary of a low-resource language.	
An Evaluation of LLMs and Google Translate for Translation of Selected Indian Languages via Sentiment and Semantic Analyses	ROHITASH CHANDRA, ARYAN CHAUDHARI AND YESHWANTH RAYAVARAPU	Large Language Models (LLMs) an LLM is a powerful, highly-parameterized neural network that learns the statistical patterns of human language from massive datasets to become a versatile tool for nearly any text-based task.	2025
The Impact of Translating Resource-Rich Datasets to Low-Resource Languages Through Multi-Lingual Text Processing	ABDUL GHAFOOR, ALI SHARIQ IMRAN SHER MUHAMMAD DAUDPOTA, ZENUN KASTRATI, ABDULLAH, RAKHI BATRA AND MUDASIR AHMAD WAN	**Machine Translation (MT)** and **Natural Language Processing (NLP)** **NLP** is the **broad field** of teaching computers language comprehension, and **MT** is one of its **most complex and practical applications**.	2021
An Adapted Few-Shot Prompting Technique Using ChatGPT to Advance Low-Resource Languages Understanding	SAEDEH TAHERY AND SAEED FARZI	Two-step cross-lingual transfer approach this method **first uses machine translation/transfer to create a dataset**, and **then uses that synthetic data to provide highly relevant examples to a powerful LLM**, leveraging the LLM's superior reasoning for a final, high-quality output.	2025

Reference:

- YASIR ABDELGADIR MOHAMED (Member, IEEE) (2024) The Impact of Artificial Intelligence on Language
- MIHAI NADAŞ , LAURA DIOŞAN , AND ANDREEA TOMESCU. (2025). Synthetic Data Generation Using Large Language Models: Advances in Text and Code
- JI-BUM CHUNG1 AND TAEHYUN KIM (2025) Leveraging Large Language Models for Enhanced Back-Translation: Techniques and Applications