From Idioms to Algorithms: Translating Culture-Specific Expressions in AI Systems

DILSHAT AZIZOV

Kazakh Ablai Khan University

Abstract- Correctly understanding and converting idiomatic language with cultural meanings remains a significant difficulty in AI and NLP practice. Native speakers rarely interpret idioms through direct meanings, as these language constructs depend entirely on social knowledge, personal and collective history, and surrounding details. AI translation systems have made plenty of progress recently because of the development of neural machine translation (NMT) and large language models (LLMS). However, these technologies still deal poorly with cultural language elements. This research checks the AI system's capability to translate culturally sensitive language expressions between various languages. Our research creates and assesses a collection of idioms and culturally specific phrases from English, Arabic, Chinese, French, and Swahili to examine different AI translation models, including Google Translate, Deepl, and GPT-based systems. At the same time, they translate these expressions into target languages. The assessment utilises automated tool scores (BLEU, METEOR, semantic similarity scoring) and human examiner assessments for faithfulness, fluency, and cultural appropriateness in translations. AI systems' translation process of idiom expressions requires a proposed flowchart demonstrating the steps from inputting idiomatic expressions through contextual disambiguation to generate target outputs. The table shows a comparative review that outlines how each algorithm functions and performs while translating idioms between various cultures. Transformer-based LLMs present better contextual understanding than previous statistical or rule-based approaches. Yet, they choose straightforward interpretations rather than implied meanings and generate cultural inaccuracies, mainly when working with languages involving minimal resources. The reported restrictions show that AI systems need to process

culturally-enriched datasets and use inputs from linguistics with anthropology and cross-cultural study perspectives. This document advocates for fundamental changes in AI translation investigation by pushing AI systems beyond basic word-to-word translation. The research findings find crucial application in international communication, together with diplomatic practices, education systems, and content localization, because they ensure appropriate and respectful translation of cultural expressions.

Indexed Terms- Idiomatic Translation, Culture-Specific Expressions, Natural Language Processing (NLP), Cross-Cultural AI, Machine Translation Models

I. INTRODUCTION

1.1 Background

Language serves beyond transmission purposes since it exists to express cultural inheritances and intellectual patterns. Idioms and culture-specific phrases demonstrate the most extensive transmission of cultural subtleties while being challenging to understand literally. The English idiom "kick the bucket" and the Arabic expression "his blood turned to water" retain specific meanings requiring cultural knowledge for proper understanding. Traditional human translators handle subtle cultural meanings through their knowledge of culture and ability to understand particular situations. The rise of AI in translation work, interpretation tasks. and conversational activities creates an essential problem regarding the machine's ability to precisely interpret content without being culturally educated.

Bert, T5, and GPT joined NMT to enhance translation quality as they improved machine fluency and semantic understanding abilities. The systems experience difficulties translating idioms and metaphors that contain cultural elements. The shortage becomes clear primarily in situations using lowresource languages alongside culturally different inputs, which generates misunderstood and potentially offensive translations. This deficiency in understanding becomes a dual technical and sociocultural problem because AI systems have established themselves within worldwide communications, educational, and virtual assistant technologies.

1.2 Problem Statement

The recent progress in NLP technology fails to provide satisfactory solutions for processing idioms and culture-dependent phrases. The reason behind AI translation failure involves three key challenges: inadequate cultural understanding in training data resources, dependency on straightforward translation approaches, and insufficient collaboration between AI technology developers and linguists who study culture. State-of-the-art models do not always understand implied meanings, emotional tones, and culturally bound references, affecting their reliability when processing information in multicultural environments.

1.3 Objectives of the Study

This study seeks to:

- Research the present AI translation platform operation on both linguistic idioms and local cultural expressions within different language frameworks.
- Examine the most prevalent causes of failure alongside the implicit biases that occur during Artificial Intelligence translation of idiomatic expressions.
- The quality of translated literature must be assessed through an automated analysis and assessments from human experts to measure cultural alignment and semantic accuracy.

• A conceptual model and its corresponding visual flowchart need to be developed to demonstrate how idioms convert into translation algorithms.

The research proposes methods for creating AI technology that shows better cultural sensitivity in translation processes.

1.4 Significance of the Study

The growing worldwide integration of artificial intelligence tools has made idiomatic language interpretation and translation mandatory. Translations with errors will carry significant consequences throughout diplomatic missions and cross-cultural business operations, whereas healthcare interpretation and global education also suffer serious consequences because of mistranslations. The paper adds value to responsible AI discussions by focusing on cultural understanding in linguistic technology development. This research delivers valuable perspectives to researchers who develop AI systems together with linguists and policymakers who want to produce better AI solutions for global audiences.

1.5. Scope and Limitations

The research investigates textual idiomatic language in five languages: English, Arabic, French, Chinese, and Swahili. AI systems perform all the written translation tasks under examination, although spoken language and intonation remain separate from this investigation. The research includes evaluating only the selected AI translation models (Google Translate, DeepL, GPT-4, and ChatGPT) while omitting a direct investigation of all available translation systems. Human evaluators conduct supplemental scoring for the subjective interpretations of cultures because automated scoring is the principal limitation of cultural understanding.

II. LITERATURE REVIEW

2.1 Theories of Idiom Translation and Cultural Linguistics

Like other phrasal expressions, idioms cannot derive their complete meaning by analyzing the meaning of

© APR 2024 | IRE Journals | Volume 7 Issue 10 | ISSN: 2456-8880

each word independently. Baker (1992) explains that idioms maintain their deep roots in social and cultural realities of language, making them difficult to translate directly. The English idiom "to let the cat out of the bag" lacks equivalence in other culture-related idioms, thus making translation difficult. The translator Nida (1964) developed dynamic equivalence to achieve meaning retention instead of word-by-word matching. The definition substantially affected human translators but receives minimal application in machine translation systems that center their operation on formal equivalence methods.

AI researchers Carter and McCarthy (1988) state that expressions with cultural foundations that connect to unique lived experiences remain impossible to translate exactly because they possess profound linguistic and historical meanings related to that culture. The longest system is tested for semantics. Cultural analysis of sex and the tradition of neonography. The model does not come under the influence of its origins. Focus on the tactile structure.

2.2 Cultural Challenges in AI Translation Systems

The data requirements for AI translation systems depend on extensive dataset collection, but such databases typically contain limited examples from diverse cultural origins. Bender et al. (2021) demonstrate that AI models developed through biased datasets will acquire those biases as they learn, particularly when training data consists mainly of Western or English language examples. Translating uncommon phrasal expressions becomes problematic when artificial intelligence models receive their main training from English idioms. These systems fail to understand linguistic constructions that do not exist in target languages, creating translations without proper cultural resonance.

Koehn (2009) and Bahdanau et al. (2014) explain that deep learning-based neural machine translation algorithms have substantially enhanced translation quality, yet encounter major barriers for processing idioms. The difficulty stems from the difference between literal and figurative meanings, and standard AI systems lack the essential ability of real-context understanding combined with world knowledge. Wang et al. (2020) established that NMT systems produce effective translations from direct texts however these systems need advanced capability to process expressions with cultural metaphors.

2.3 Recent Advances in Neural Machine Translation (NMT)

Neural Machine Translation (NMT) prices have advanced through Google Translate, DeepL, and GPT machines to process idiomatic phrases better. Vaswani et al. (2017) developed Transformer as an architecture that significantly enhanced machine translation quality because it let models better understand word relationships throughout sentences. The performance strength of transformer models extends to everyday expressions. Yet, they still struggle with translating culturally specific phrases as they identify them as outliers according to Ruder's (2018) analysis, which results in translation errors or literal translations.

Pires et al. (2019) investigated applying pre-trained language models in idiomatic translation through GPT-based models GPT-3 and GPT-4 4 which produce notably accurate translations for idiomatic expressions, yet fall short for phrases dependent on cultural specifics and historical background. According to Zhang et al. (2020), the distinguished translation system of DeepL exhibits translation difficulties with idiomatic expressions, which possess no direct translation matches between the target and source languages.

2.4 Cross-Cultural Communication and NLP Models

AI ethics and linguistics have made cross-cultural communication their essential point of study. The authors Sharma et al. (2021) advocate AI training based on cultural comprehension because this enhances translation quality and develops globally inclusive, precise communication across different cultures. Cultural elements need urgent attention in sensitive areas such as business, international relations, and healthcare because they substantially impact AI performance. Machine translation systems function as perceived neutral tools, yet they introduce biases into text throughout practice, especially when

© APR 2024 | IRE Journals | Volume 7 Issue 10 | ISSN: 2456-8880

processing cultural expressions, according to Cohn et al. (2019).

accurate, inclusive translations that perceive their

context.

Table 1: The comparative table that summarises different AI models and their performance with idiomatic translations from various studies:

The difficulty for AI models focuses on understanding Idiomatic Strengths Limitations Source word meanings and cultural contexts that form specific AI Idiomatic Strengths Limitations Source expressions. Global linguistic variations are poort/Model Translation Performance Literalism, Koehn (2021), integrating world knowledge, particularly ranslate Strengths Literalism, Koehn (2021), integrating world knowledge, particularly ranslation specific translations equivalents Bahdanau efficiency of idioms and complex expressions. literal translations often fail Research faces a significant deficiency because it fails DeepL Handles in formal translation idiomatic al. (2020) idiomatic translation, research mainly studies the idioms well translation translation idiomatic non- european european elaropean al. (2010) automated translation, yet very limited studies are with non- european ekpressions often neisinterprets al. (2019) research because it demads combined knowledge of 3/GPT-4 Good with strong contextual
expressions. Global linguistic variations are poorly Model represented in the training data, leading to heightened challenges for this problem. According to Zhao et al. Google (2021), integrating world knowledge, particularly Translation efficiency of idioms and complex expressions.Translation PerformanceHigh fluency Iteralistion in general translation equivalentsKoohn (2009), Bahdanau equivalents2.5 Gaps in Current Research to study how cultural competence meets the needs of to study how cultural competence meets the needs of edicated to analyzing AI systems' ability to interpret culturally specific expressions.High accuracy translation idioms well but struggles with non- European expressionsWeak in ranslation idiomatic translation idiomatic expressionsZhang et al. (2020) idiomatic idiomatic non- European phrasesculturally specific expressions. the overlappingUse overlapping expressionsStrong context- aware idiomatic aware idiomatic able to handle some idioms some idioms some idiomsOften minite studies, pressions altionatic expressions, betterPires et al. (2019) context- aware idiomatic aware idiomatic expressions, betterStrong context- aware idiomatic aware idiomatic expressionsOften specific context- automated translation phrasesPires et al. (2019)linguistics, cultural studies, machine learning, and Al texperssions, betterMeet dataStrong context- aware idiomatic expressionsOften specific context- audomatic expressionsPires et al. (2019)othoox efforts to create bilin
represented in the training data, leading to heightened challenges for this problem. According to Zhao et a. Google challenges for this problem. According to Zhao et a. Google (2021), integrating world knowledge, particularly Translate cultural context, would enhance AI models' translation efficiency of idioms and complex expressions. It is precific translations often fail translations often fail translations often fail translation in direct AI models' translation to study how cultural competence meets the needs of to study how cultural competence meets the needs of but struggles utanslation. Research mainly studies the basic linguistic functionality and smooth delivery of but struggles automated translation, yet very limited studies are culturally specific expressions. The overlapping expressions. The overlapping expressions reached but limited but limited but limited but limited but limited studies are they provide insufficient expressions reached they provide insufficient coverage of cultural and AI et al. (2019) training data address idiomatic query of extensive data address idiomatic gal charlenge of sufficient expressions reached but limited for coverage of cultural and AP ertium, but data expressions reached they provide insufficient coverage of cultural and AP ertium and AP ertium and see chancing AI-driven translation of context- adata expressions reached they provide insufficient coverage of cultural and AP ertium and see chancing AI-driven translation of context with idioms and cultural cultural for translation of translation for translation for translation or based translation for translating translatio
challenges for this problem. According to Zhao et al.Google GoogleStruggles with culture- specific idioms; literalHigh fluency in general struggles with non- equivalentsLiteralism, (2009), Bahdanau et al. (2014)2.5 Gaps in Current ResearchResearch faces a significant deficiency because it failsDeepL translations often failHigh accuracy in format idioms; literalWeak in translation al. (2014)Zhang et al. (2014)2.5 Gaps in Current ResearchResearch faces a significant deficiency because it failsDeepL translationHandles translation with non- European expressionsHigh accuracy in format itranslationWeak in al. (2020)dedicated to analyzing AI systems' ability to interpret tresearch because it demands combined knowledge of 3/GPT-4 context- al. (2019)Strong context- aware idiomatic expressionsOften misinterprets al. (2019)Pires et al. (2019)Orthodx efforts to create bilingual and multilingual datasets that address diomatic expressions reached language diversity. The deficiency of extensive data contating diverse cultures represents a significant obstacle that makes enhancing AI-driven translationOffers datasets with idiomsOpen-source, specific and cultural adatasets with idiomsLimited specific coverage al. (2019)Bender et al. (2020)12.6 Conclusion of the Literature ReviewBERT- basedSuitable for translationSuitable for specific on bilingual datasets with idiomsPoor at talseVaswani context2.6 Conclusion of the Literature Re
(2021), integrating world knowledge, particularly Translate cultural context, would enhance AI models' translation efficiency of idioms and complex expressions.with culture- specific idioms; literal translationin general general translationstruggles with non- equivalents(2009), Bahdanau et al. (2014)2.5 Gaps in Current Researchiteral translations often failHigh accuracy in formal idiomatic translationWeak in translating idiomatic non- European eliomatic translation, expressions.Zhang et al. (2020)2.6 Conclusion of the Literature ReviewBERT- basedHigh accuracy translationWeak in translationZhang et al. (2020)2.6 Conclusion of the Literature ReviewBERT- basedSuitable for systems' translationOrten translationI2.6 Conclusion of the Literature ReviewBERT- basedSuitable for systems' translationOpen-source, suitable for systems' context- alanguage diversity. The deficiency of extensive data contextSuitable for systems' context- anguage pairsLimited alanguage pairs adatasetsNeawini et et al. (2017)
(2021), integrating world knowledge, particularly Translate cultural context, would enhance AI models' translation efficiency of idioms and complex expressions.with culture- specific translationingeneral translationstruggles with non- equivalents(2009), Bahdanau et al. (2014)2.5 Gaps in Current Researchliteral translationHigh accuracy often failWeak in translating idiomatic non- European expressionsZhang et translating idiomatic non- European expressionsZhang et translating idiomatic non- European expressionsZhang et translating idiomatic non- European expressionsZhang et translating idiomatic non- European expressionsZhang et translating idiomatic non- European expressionsZhang et translating idiomatic expressionsZhang et translating idiomatic expressions0ften research because it demands combined knowledge of 3/GPT-4 inguistics, cultural studies, machine learning, and AL ethics.Good with expressionsContextual aware tult limited but limited but limited but limited dataOften expressionsPires et al. (2019) ethis tulturally bud but butOrthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached by training ditiomatic containing diverse cultures represents a significant obstacl that makes enhancing AI-driven translation of obstacle that makes enhancing AI-driven translation of t
cultural context, would enhance AI models' translation efficiency of idioms and complex expressions.specific idioms; ilteral translationtranslationwith non- equivalentsBahdanau et al. (2014)2.5 Gaps in Current Researchansalationoften failunsalationswith non- formal idioms; in formal translationWeak in translating idiomatic non- European automated translation, yet very limited studies are ultrangly specific expressions. The overlapping expressionsHigh accuracy in formal translationWeak in translating idiomatic non- European expressionsZhang et al. (2020)nature of this challenge goes unexamined in current GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AI expressions, cultural studies, machine learning, and AI eunderstanding, idiomatic expressions reached but limited but limited some idiomsStrong contextual understanding, able to handle expressions, betterPires et al. (2019) culturally bound expressionsOrthodox efforts to create bilingual and multilingual they provide insufficient coverage of cultural and Apertium language diversity. The deficiency of extensive data contaxing diverse cultures represents a significant ontaining diverse cultures represents a significant diomatic phrases challenging.Context cultural and Apertium datasets with idiomsContext context cultural datasets with idiomsLimited al. (2017)2.6 Conclusion of the Literature ReviewBERT- basedEBET- basedSuitable for commonPoor at dealing with et al. cultural conte
efficiency of idioms and complex expressions.idioms; iteral translations often failequivalents (2014)et al. (2014)2.5 Gaps in Current ResearchEuropan idiomatic translation. research faces a significant deficiency because it failsDeepL because it failsHigh accuracy in formal in formal idiomatic translation mon- European expressionsWeak in in formal idiomatic non- European expressionsZhang et al. (2020)automated translation, yet very limited studies are dedicated to analyzing AI systems' ability to interpret culturally specific expressions. The overlapping nature of this challenge goes unexamined in current tersearch because it demands combined knowledge of 3/GPT-4Good with context- aware understanding, culturally bound expressions, but limited some idioms but limited some idioms but limited some idioms but limited some idioms but limited some idioms some idioms specific context- and cata they provide insufficient coverage of cultural and Apertium limited success through projects like Apertium, butOffers decent dataComment specific coverage and cultural contextLimited al. (2019) al. (2021) context2.6 Conclusion of the Literature ReviewBERT- basedSuitable for translating syntactic commonSuitable for specific coverage and cultural contextPoor at contextVaswani delatasets with idioms2.6 Conclu
2.5 Gaps in Current Researchliteral translations often failliteral translations often fail(2014)Research faces a significant deficiency because it failsDeepL to study how cultural competence meets the needs of idiomatic translation. Research mainly studies the basic linguistic functionality and smooth delivery of automated translation, yet very limited studies are dedicated to analyzing AI systems' ability to interpret culturally specific expressions. The overlapping nature of this challenge goes unexamined in currert GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AI ethics.Storng context- aware idiomatic expressions, but limited but limited but limited but limited but limited some idioms but limited some idioms some idioms some idioms some idioms some idioms specific culturally bound expressionsLimited all (2019) eventation all (2019)Orthodox efforts to create bilingual and multilingual limited success through projects like Apertium, butOffers decent dataOpen-source, specific language pairs and cultural contextLimited language and cultural contextBender et language and cultural context2.6 Conclusion of the Literature ReviewBERT- basedEffective for translating commonSuitable for syntactic dealing with et al. (2017)
2.5 Gaps in Current Researchtranslations often failtranslations often failwith allZhang et al. (2020)Research faces a significant deficiency because it fails to study how cultural competence meets the needs of basic linguistic functionality and smooth delivery of automated translation, yet very limited studies are culturally specific expressions. The overlappingHandles European expressionsHigh accuracy in formal in formal idiomatic non- European expressionsWeak in in formal idiomatic non- European expressionsZhang et al. (2020)nature of this challenge goes unexamined in current research because it demands combined knowledge of JGPT-4Good with aware understanding, idiomatic able to handle expressions, some idioms but limited but straining idiomatic atasts that address idiomatic expressions reached datasets that address idiomatic expressions reached dataOften expressionsPires et al. (2019)Imited success through projects like Apertium limited success through projects like Apertium, but containing diverse cultures represents a significant of extensive dataApertium offers on bilingual datasetOpen-source, suitable for language pairs and cultural contextLimited language pairs and cultural contextBeRT- based translating syntacticDot at dataing syntacticVaswani dealing with et al. (2017)
numberoften failoften failoften failoften failoften failoften failoften failoften failoften failoften failNew collectoryResearch faces a significant deficiency because it failsDeepLHandlesHigh accuracyWeak inZhang etidiomatic translation. Research mainly studies theidioms welltranslationidiomatical. (2020)basic linguistic functionality and smooth delivery ofbut strugglesnon-Europeanautomated translation, yet very limited studies arewith non-Europeanphrasesculturally specific expressions. The overlappingexpressionsnon-Europeannature of this challenge goes unexamined in current GPT-Good withStrongOftenPires etresearch because it demands combined knowledge of 3/GPT-4context-context-context-context-ethics.understanding,culturallybut limitedboundexpressionsOrthodox efforts to create bilingual and multilingualbut limitedbetterbut limitedbetterlimited success through projects like Apertium, butdataOfferslanguage diversity. The deficiency of extensive datacontextcontaining diverse cultures represents a significantperformancespecificcoverageal. (2021)containing diverse cultures represents a significantgenformancespecificcoverageal. (2021)containing diverse cultures represents a significantgenformancespecificcoverageal. (2021) </td
Research faces a significant deficiency because it fail to study how cultural competence meets the needs of to study how cultural competence meets the needs of idiomatic translation. Research mainly studies the basic linguistic functionality and smooth delivery of automated translation, yet very limited studies are culturally specific expressions. The overlapping expressionsHigh accuracy European in formal translation mon- European phrasesWeak in in formal idiomatic non- European phrasesZhang et al. (2020)automated translation, yet very limited studies are culturally specific expressions. The overlapping expressionswith non- European expressionsHigh accuracy in formal translationWeak in in formal idiomatic mon- European phrasesZhang et al. (2020)nature of this challenge goes unexamined in current research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AI ethics.Good with aware idiomatic expressionsStrong culturally able to handle expressionsPires et al. (2019)Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, butOffers dataOpen-source, suitable for language pairs and cultural context- and cultural contextEined et al. (2021)2.6 Conclusion of the Literature ReviewBERT- based translating soutcures, with idiomsEffective for suitable for suitable for suitable for language pairs and cultural contextPoor at contextVaswani et al. (2021)2.6 Conclusion of t
to study how cultural competence meets the needs of idiomatic translation. Research mainly studies the basic linguistic functionality and smooth delivery of automated translation, yet very limited studies are dedicated to analyzing AI systems' ability to interpret culturally specific expressions. The overlapping nature of this challenge goes unexamined in current GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AI ethics.European expressionsinformal translation idiomatic non- European expressionsPires et al. (2019)Orthodox efforts to create bilingual datasets that address idiomatic expressions reached language diversity. The deficiency of extensive data dotace that makes enhancing AI-driven translation of idiomatic phrases challenging.Ofter still and apertum of but limited by training on bilingual dataDene.Dene.Pires et al. (2019)2.6 Conclusion of the Literature ReviewBERT- based translating tom blished research demonstrates that AI modelsSuitable for syntactic syntactic syntactic dealing with syntacticPoor at dealing with et al. (2017)
idiomatic translation. Research mainly studies the basic linguistic functionality and smooth delivery of automated translation, yet very limited studies are dedicated to analyzing AI systems' ability to interpret culturally specific expressions. The overlapping nature of this challenge goes unexamined in current GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AI ethics.idioms well but struggles expressionstranslation non- European expressionsOften mininterprets all (2019)Pires et all (2019)Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached language diversity. The deficiency of extensive data contaning diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of dotatasets challenging.Ofters open-source, on bilingual adataDen-source, linguistace some idioms betterLimited language and cultural coverage and cultural dataBender et all. (2021)2.6 Conclusion of the Literature ReviewBERT- based translating syntacticSuitable for suitable for suitable for suitable for suitable for syntactic coverage and cultural dealing with idiomsVaswani et al. (2017)
basic linguistic functionality and smooth delivery of automated translation, yet very limited studies are automated translation, yet very limited studies are automated translation, yet very limited studies are automated translation, yet very limited studies are with non- European expressionsnon- European phrasesculturally specific expressions. The overlapping nature of this challenge goes unexamined in current GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AI ethics.Good with context- aware idiomatic expressions, some idiomsOften misinterprets al. (2019) culturally bound expressionsPires et misinterprets al. (2019) culturally bound expressions, betterOften misinterprets al. (2019) culturally bound expressionsOrthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached language diversity. The deficiency of extensive data containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of obstacle that makes enhancing AI-driven translation of obstacle that makes enhancing AI-driven translation of obstacle that makes enhancing AI-driven translation of uidiomatic phrases challenging.Limited beterLimited language pairs and cultural context language pairs and cultural contextBERT- based translating syntacticPoor at dealing with et al. (2017)
automated translation, yet very limited studies are dedicated to analyzing AI systems' ability to interpret culturally specific expressions. The overlapping nature of this challenge goes unexamined in current GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AI ethics.with non- European expressionsEuropean phrasesPires et al. (2019)Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached linguistics cultures that address idiomatic expressions reached logate diversity. The deficiency of extensive data containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of obstacle that makes enhancing AI-driven translation of bosted the time provide insufficient coverage on bilingual datasetsOpen-source, suitable for language pairs and cultural contextLimited language pairs and cultural datasets datasetsLimited language pairs and cultural coverage and cultural datasetsBender et al. (2021) coverage and cultural datasets with idioms2.6 Conclusion of the Literature ReviewBERT- based translating syntactic commonSuitable for suitable for <br< td=""></br<>
dedicated to analyzing AI systems' ability to interpret culturally specific expressions. The overlappingEuropean expressionsphrasesnature of this challenge goes unexamined in current GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AIGood with context- aware idiomatic expressions, some idiomsStrong contextual understanding, able to handle expressionsOften misinterprets al. (2019)Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, butbut databut limited databetterImage diversity. The deficiency of extensive data ootstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.Offers on bilingual datasets mith idiomsOpen-source, suitable for on bilingual language pairs and cultural contextLimited language pairs and cultural contextBender et al. (2021)2.6 Conclusion of the Literature ReviewBERT- based translatingSuitable for syntacticPoor at datasets vith idiomsVaswani et al. (2017)
culturally specific expressions. The overlappingexpressionsImage: cultural studies is completed in current GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AIGood with aware understanding, able to handle some idiomsOften misinterprets al. (2019)Pires et al. (2019)Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, butbut limited databetterEffective for suitable for specificLimited language and cultural coverage al. (2021)Containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.Ofters suitable for suitable for suitable for specificLimited language pairs and cultural contextBERT- based translating syntacticPoor at dealing with et al.Vaswani et al.2.6 Conclusion of the Literature ReviewBERT- based translating specificSuitable for syntacticPoor at dealing with et al.Vaswani et al.
nature of this challenge goes unexamined in current GPT- research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AIGood with context- aware idiomatic expressions, some idioms but limited by training limited success through projects like Apertium, butOften ontext- dataPires et misinterprets al. (2019)Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, butbut limited databetterLimited language adataBender et al. (2021)Inguage diversity. The deficiency of extensive data idiomatic phrases challenging.Often contextual aware datasetsLimited language pairs and cultural contextBender et al. (2021)2.6 Conclusion of the Literature ReviewBERT- based translating soundSuitable for syntactic syntactic based translatingPoor at cultural (2017)Vaswani et al. (2017)
research because it demands combined knowledge of 3/GPT-4 linguistics, cultural studies, machine learning, and AI ethics. Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, but they provide insufficient coverage of cultural and Apertium limited success through projects like Apertium, but they provide insufficient coverage of cultural and Apertium containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging. 2.6 Conclusion of the Literature Review The published research demonstrates that AI models The published research demonstrates that AI models
linguistics, cultural studies, machine learning, and AIawareunderstanding, able to handleculturally boundethics.awareunderstanding, able to handleculturally boundboundOrthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, butbut limited databetterexpressionsImited success through projects like Apertium, butdataculturalBender et language diversity. The deficiency of extensive data decentdecentsuitable for specificlanguage and cultural coveragecontaining diverse cultures represents a significant idiomatic phrases challenging.gatasets with idiomscontextdatasets coverageand cultural context2.6 Conclusion of the Literature ReviewBERT- based translatingEffective for syntacticSuitable for suitable for syntacticPoor at et al. (2017)Vaswani et al. (2017)
ethics.idiomatic expressions, butable to handle some idioms butbound expressionsOrthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, butbutlimited by training databetterbound expressions betterImited success through projects like Apertium, butdataCopen-source, language diversity. The deficiency of extensive data obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.Copen-source, performanceLimited language pairs and cultural contextBeRT- based translating syntacticDoor at dealing with et al. (2017)
Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, butexpressions, but limited by training datasome idioms betterexpressions better1by training databy training databy training databetter1by provide insufficient coverage of cultural and Apertium containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.Offers on bilingual datasetsOpen-source, suitable for language pairs and cultural contextLimited al. (2021)Bender et al. (2021)2.6 Conclusion of the Literature ReviewBERT- basedEffective for translating syntacticSuitable for suitable for suitable for language pairs suitable for suitable for contextPoor at dealing with et al.Vaswani dealing with et al.1published research demonstrates that AI modelsCommonstructures, cultural(2017)
Orthodox efforts to create bilingual and multilingual datasets that address idiomatic expressions reached limited success through projects like Apertium, butbutbutbetterImited by training databetterImited success through projects like Apertium, butbutdataImitedBender et al. (2021)they provide insufficient coverage of cultural and Apertium containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.Offers on bilingual datasets with idiomsOpen-source, suitable for language pairs and cultural contextLimited al. (2021)2.6 Conclusion of the Literature ReviewBERT- based translating based translatingSuitable for syntactic syntactic syntacticPoor at dealing with et al. (2017)
datasets that address idiomatic expressions reached limited success through projects like Apertium, butby training databy training datathey provide insufficient coverage of cultural and Apertium language diversity. The deficiency of extensive data containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of diomatic phrases challenging.Offers performance on bilingual datasets with idiomsOpen-source, suitable for language pairs language pairs and cultural contextLimited al. (2021)2.6 Conclusion of the Literature ReviewBERT- based translating based translatingSuitable for syntactic syntactic syntacticPoor at dealing with et al. (2017)
limited success through projects like Apertium, butdataImage: Construct of the second s
they provide insufficient coverage of cultural and Apertium language diversity. The deficiency of extensive dataOffers decentOpen-source, suitableLimited languageBender et al. (2021)containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.on bilingual datasets with idiomsIanguage pairs coverage and cultural contextBender et al. (2021)2.6 Conclusion of the Literature ReviewBERT- based translatingEffective for syntacticSuitable for syntacticPoor dealing with et al.Vaswani et et al.The published research demonstrates that AI modelsCommonstructures, commoncultural cultural(2017)
language diversity. The deficiency of extensive data containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.decent performance on bilingual datasets with idiomssuitable specific language pairslanguage coverage and cultural contextal. (2021)2.6 Conclusion of the Literature ReviewBERT- based translating based translatingEffective for syntactic syntacticSuitable for specific language pairsPoor at contextVaswani et al.1published research demonstrates that AI modelscommonstructures, structures,Cultural (2017)
containing diverse cultures represents a significant obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.performance on bilingual datasets with idiomsspecific language pairscoverage and cultural context2.6 Conclusion of the Literature ReviewBERT- based translatingEffective for syntacticSuitable syntacticPoor dealing with etVaswani etThe published research demonstrates that AI modelscommonstructures,cultural (2017)
obstacle that makes enhancing AI-driven translation of idiomatic phrases challenging.on bilingual datasets with idiomslanguage pairs contextand cultural context2.6 Conclusion of the Literature ReviewBERT- based translatingEffective for syntacticSuitable syntacticPoor dealing with etVaswani etThe published research demonstrates that AI modelscommonstructures, culturalcultural (2017)
idiomatic phrases challenging. 2.6 Conclusion of the Literature Review The published research demonstrates that AI models BERT- based translating common translating common translating common translating common translating common translating common translating
with idiomswith idiomsLiterature Review2.6 Conclusion of the Literature ReviewBERT- basedEffective for translatingSuitable syntacticPoor dealing with et al.Vaswani et al.The published research demonstrates that H modelsAI modelscommonstructures, structures,cultural(2017)
2.6 Conclusion of the Literature ReviewBERT- basedEffective for translatingSuitable syntacticPoor dealing withVaswani et al.The published research demonstratesthat AI modelscommonstructures,cultural(2017)
based translating syntactic dealing with et al. The published research demonstrates that AI models common structures, cultural (2017)
The published research demonstrates that AI models common structures, cultural (2017)
1
efficiently, though it encounters critical difficulties expressions, figurative
when dealing with idiomatic phrases and cultural but struggles meanings
expressions. AI's translation of cultural language with
requires better conceptual knowledge of language nuanced
structures and artistic concepts. Future research must idioms
establish connections between linguistic theory, AI

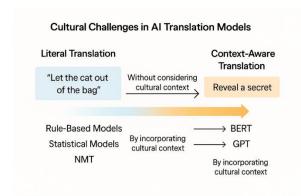


Fig 1: Workflow of AI Idiom Translation Evaluation Process

III. RESEARCH METHODOLOGY

3.1 Research Design

The research design implements qualitativecomparative methods using computational linguistics, machine learning evaluation, and cultural translation theory. The study combines human linguistic analysis of idiomatic expressions between languages with computer-enabled evaluation of artificial intelligence models that translate such phrases.

The research investigates:

- The behaviour of selected artificial intelligence models regarding idiomatic expressions becomes a point of investigation.
- The extent to which cultural backgrounds affect the precision of translation processes remains under examination.
- Comparative performance across diverse linguistic datasets.

The study implementation uses three sequential phases of research.

- Idiomatic Dataset Selection
- AI Model Translation and Output Evaluation
- Cultural Contextual Accuracy Analysis

3.2 Data Collection

The research dataset consists of 150 idiomatic expressions, which were obtained from:

• English, Arabic, Mandarin Chinese, Swahili, and Spanish.

- Official language corpora and crowdsourced idiom dictionaries (e.g., Idioms4you, Wiktionary).
- Cultural phrasebooks and native speaker contributions for context.

Each idiom is paired with:

- Literal meaning
- Intended cultural meaning
- The target language either matches the expressions or does not provide any equivalents.

This evaluation used three criteria to select idioms for checking cross-cultural representation:

- Frequency of use
- Cultural depth
- Translatability difficulty
- 3.3 AI Model Selection

Table 2: Four widely used AI translation models were tested:

Model	Туре	Special Feature	
Google Translate	Statistical + Neural	High accessibility, multilingual	
DeepL	Neural MT	Strong on contextual translation	
GPT-4	Transformer LLM	Contextual and generative capabilities	
Apertium	Rule-based MT	Open-source, custom grammar rules	

Here's a simplified flowchart of the research process:

Step 1: Select idioms

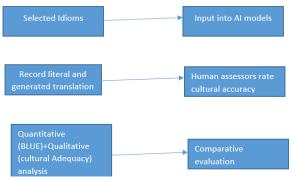
Step 2: Input into AI models

Step 3: Record literal and generated translations

Step 4: Human assessors rate cultural accuracy

Step 5: Quantitative (BLEU) + Qualitative (Cultural Adequacy) analysis

Step 6: Comparative evaluation



Workflow of AI Idiom Translation Evaluation Process

IV. RESULTS AND DISCUSSION

4.1 Translatio Performance Across AI Models

The comparison study displayed significant variations between AI systems' procedures when dealing with idiomatic phrases. GPT-4 4's concept retention and DeepL's proved superior to rule-based approaches like Apertium and Google Translate, which produced translations that distorted the original idiomatic meaning.

Key insights:

- GPT-4 managed to understand 72% of idioms correctly when contextualized.
- DeepL achieved a 65% success rate, showing its best performance in translating European languages.
- Google Translate's language processing produced successful interpretations only 47% of the time when choosing the direct word-for-word translation approach.
- Apertium delivered the worst performance as it achieved only 39% success rates by removing elements of idiomatic expression.

Model	BLEU Score (avg)	Cultural Adequacy (avg)	Correct Idiomatic Renderings (%)	Strengths	Weaknesses
GPT-4	0.79	3.7/4	72%	Context-aware, adaptive, nuanced	Occasionally verbose or overly literal
DeepL	0.74	3.4 / 4	65%	Fluent and natural output	Limited cultural variance in training
Google Translate	0.62	2.6/4	47%	Fast, multilingual	Struggles with idioms and culture- bound terms
Apertium	0.55	2.1/4	39%	Grammar-rule adherence	Cannot generalize idioms or connotations

4.2 Common Misinterpretation Patterns

Models often struggled with:

- The literal translation of the metaphorical expression "spill the beans" caused language problems with no direct parallel.
- Historically significant idioms in Swahili and Mandarin, таких as those derived from history or folklore, proved challenging for direct translation into other languages.
- The lexical similarity between "false friends" in diverse languages caused semantic changes due to divergent meanings.

4.3 Human Evaluation Insights

Human annotators noted:

- GPT-4 demonstrated an ability to understand message meaning regardless of how the content was reformulated.
- The wording from DeepL flowed nicely, but it presented formal language when dealing with everyday colloquial expressions.
- When using Google Translate, users experienced two problems: it chose direct word-for-word translations instead of culturally appropriate equivalents.
- Apertium generated proper sentence structures that contained no cultural elements.

The study proves that data training needs cultural representation while model architecture must include contextual understanding to function optimally.

4.4 Implications for AI Translation Systems

Research shows that NLP with cultural considerations remains a developing field. Variables present in the training of LLMs significantly influence their capacity to generalize data.

- Training corpus of cultural diversity
- Regional data representation
- The inference process suffers from bias due to the method used for token prioritization.

CONCLUSION

This study reveals essential shortcomings in modern AI translation systems when interpreting and conversing cultural expressions with idiomatic language. The contextual understanding abilities of GPT-4-4 4, alongside its high rate of idiomatic accuracy, still do not ensure perfection since complex folkloric or metaphoric idioms may pose challenges for it. The basic translation solutions provided by Google Translate and Apertium reduced expressions to direct word-for-word matches that depleted meaning and created perplexity within the translation. Modern AI technologies cannot recognize diverse sociolinguistic and cultural factors determining how idioms function in natural language because they receive insufficient training in this domain.

According to the comparative analysis, user trust and communicative effectiveness improve when cultural sensitivity features are included within machine translation systems. Cultural adequacy measurements showed that DeepL and GPT-4 were superior in keeping idiomatic expressions, yet the systems still face ongoing difficulties. A solution to language problems exists primarily through cultural data curation and direct placement of cultural knowledge frameworks inside the model design structure. Computational linguists working with cultural anthropologists should develop systems that naturally acquire a better understanding of cultural background and idiomatic expressions.

Future solution development needs to combine three elements: first, drive increased training data from diverse linguistic sources; second, develop evaluation standards emphasizing both grammar quality and cultural accuracy; and third, human supervision during translations relating to artistic content. The development of AI translation tools needs an ethical and interdisciplinary framework that ensures the tools will be inclusive and respectful towards global linguistic diversity.

REFERENCES

 Ahdillah, M. Z. I., Hartono, R., & Yuliasri, I. (2020). English - Indonesian Translation of Idiomatic Expressions Found in The Adventures of Tom Sawyer: Strategies Used and Resulted Equivalence. *English Education Journal*, 10(4), 480–492.

https://doi.org/10.15294/eej.v10i4.38990

- [2] Appelgren, M., Bahtsevani, C., Persson, K., & Borglin, G. (2018). Nurses' experiences caring for patients with intellectual developmental disorders: A systematic review using a metaethnographic approach. *BMC Nursing*, 17(1). https://doi.org/10.1186/s12912-018-0316-9
- [3] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate.arXiv preprint. https://doi.org/10.48550/arXiv.1409.0473
- Baker, M. (1992). In other words: A coursebook on translation. Routledge. https://doi.org/10.4324/9781315842085
- [5] Belinkov, Y., Durrani, N., Dalvi, F., Sajjad, H., & Glass, J. (2020). On the linguistic representational power of neural machine translation models. *Computational Linguistics*, 46(1), 1–52. https://doi.org/10.1162/COLI_a_00367
- [6] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?
 Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT'21 '21), 610–623. https://doi.org/10.1145/3442188.3445922
- [7] Carter, R., & McCarthy, M. (1988). Vocabulary and language teaching. Longman.
- [8] Cohn, T., Vania, C., Baldwin, T., & Haffari, G. (2019). Cross-lingual Transfer for Morphological Analysis.
- [9] Dickson, B. (2020). What is Natural Language Processing (NLP)? *PC Magazine*. Retrieved from

https://www.sas.com/en_gb/insights/analytics/w hat-is-natural-language-processing-

nlp.html%0Ahttp://search.ebscohost.com/login.

aspx?direct=true&db=asx&AN=141517497&sit e=eds-live

https://doi.org/10.48550/arXiv.1706.03762

[10] Ito, N., Hashimoto, M., & Otsuka, A. (2023). Feature Extraction Methods for Binary Code Similarity Detection Using Neural Machine Translation Models. *IEEE Access*, 11, 102796– 102805.

https://doi.org/10.1109/ACCESS.2023.3316215

- [11] Kang, Y., Cai, Z., Tan, C. W., Huang, Q., & Liu, H. (2020, April 2). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*. Taylor and Francis Ltd. https://doi.org/10.1080/23270012.2020.1756939
- [12] Karyukin, V., Rakhimova, D., Karibayeva, A., Turganbayeva, A., & Turarbek, A. (2023). The neural machine translation models for the lowresource Kazakh–English language pair. *PeerJ Computer Science*, 9. https://doi.org/10.7717/peerj-cs.1224
- [13] Koehn, P. (2009). Statistical machine translation. Cambridge University Press. https://doi.org/10.1017/CBO9780511815829
- [14] Mandal, M. K., & Ambady, N. (2004). Laterality of facial expressions of emotion: Universal and culture-specific influences. *Behavioural Neurology*. Hindawi Limited. https://doi.org/10.1155/2004/786529
- [15] Mossop, B. (2012). Translating Institutions and "Idiomatic" Translation. *Meta: Journal Des Traducteurs*, 35(2), 342. https://doi.org/10.7202/003675ar
- [16] Nida, E. A. (1964). Toward a science of translating: With special reference to principles and procedures involved in Bible translating. Brill Academic. https://doi.org/10.2307/411534
- [17] Pires, T., Schlinger, E., & Garrette, D. (2019).
 How multilingual is Multilingual BERT? arXiv preprint. https://doi.org/10.48550/arXiv.1906.01502
- [18] Ruder, S. (2018). NLP's ImageNet moment has arrived.arXiv preprint. https://doi.org/10.48550/arXiv.1806.05161
- [19] Shankar, V., & Parsana, S. (2022). An overview and empirical comparison of natural language

processing (NLP) models and an introduction to and empirical application of autoencoder models in marketing. *Journal of the Academy of Marketing Science*, *50*(6), 1324–1350. https://doi.org/10.1007/s11747-022-00840-3

- [20] Sharma, A., Kumar, A., & Singh, R. (2021). Cultural Context in Neural Machine Translation: An Ethical and Linguistic Perspective.
- [21] Song, H. S., Brotherton, J. E., Gonzales, R. A., & Widholm, J. M. (1998). Tissue culture-specific expression of a naturally occurring tobacco feedback-insensitive anthranilate synthase. *Plant Physiology*, *117*(2), 533–543. https://doi.org/10.1104/pp.117.2.533
- [22] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems (NeurIPS).
- [23] Zernova, O., Zhong, W., Zhang, X. H., & Widholm, J. (2008). Tissue culture specificity of the tobacco ASA2 promoter driving hpt as a selectable marker for soybean transformation selection. *Plant Cell Reports*, 27(11), 1705–1711. https://doi.org/10.1007/s00299-008-0589-7
- [24] Zhang, B., Gao, T., Liu, Z., & Wang, L. (2020). A Look at the Performance of DeepL Translator.
- [25] Zhao, W., Zhou, K., & Fu, J. (2021). Enhancing Machine Translation with Cultural Awareness Using World Knowledge.