Josuga: Jos Journal of Humanities Vol. 7, No. 2, December 2025

ARTIFICIAL INTELLIGENCE USE IN TRANSLATING NIGERIAN SLANG: PRAGMATIC EVALUATIONS AND PRACTICAL PATHWAYS FOR IMPROVEMENT

ZAYNAB BOLANLE RAJI-ELLAMS
Bayero University, Kano

ABSTRACT

Nigerian slang is often difficult to translate because it tends to be informal, idiomatic, and culturally salient. This study examines five artificial intelligence (AI) models with respect to human translators in their ability to translate the quality of Nigerian slang expressions, adapting three theoretical frameworks: Relevance, Equivalence, and Polysystem. 31 common words or phrases of Nigerian slang were translated by ChatGPT, Grok, Copilot, Gemini, DeepSeek, and professional human translators. To rate the translated words or phrases, bilingual experts rated each successful translation of the slang on a 5-point Likert scale across three theories rated on pragmatic intent, semantic meaning, and cultural appropriateness. Human translations rated higher consistently across all three dimensions of pragmatic intent, cultural fit, and semantic meaning. Human translations were able to best preserve the meaning, context, and relevant cultural fit of the Nigerian slang words or phrases being translated. From highest to lowest relevance, of AI models were Grok, ChatGPT, Gemini, Copilot, and then DeepSeek. Statistically significant differences were confirmed between human and AI model outputs, with greater differences between human translation outputs compared to AI model outputs in pragmatic and cultural dimensions. This study shows that AI models are rapidly improving but are currently incapable of fully considering the cultural and contextual dimensions of translating Nigerian slang expressions, pointing to the necessity of having human agents involved in informal language translation projects. To improve future translation quality, AI model outputs can be improved through a hybrid human-AI workflow and training datasets enriched with culturally relevant material.

KEYWORDS: Artificial Intelligence, Pragmatics, Relevance Theory, Polysystem Theory, Equivalence Theory, Nigerian Slang, and Machine Translation.

INTRODUCTION

Artificial Intelligence (AI) has very rapidly evolved from a technology initially utilised by a specialised group of experts to one integrated into numerous aspects of contemporary daily life. Most individuals now frequently utilise AI chatbots as their primary source of information, a position that was previously occupied by search engines. AI is now being used for tasks such as letter writing, text editing, and information gathering, and for advanced functions such as data analysis, automated customer service, and face recognition, amongst others. One aspect AI has revolutionised is translation. AI now enables users to translate languages with which they have no prior familiarity into a language they are familiar with, thereby facilitating seamless communication. However, while there have been studies on how AI has been leveraged for translation from one language to another, there has been limited scholarly attention on how AI can be leveraged to translate Nigerian slang accurately, considering factors such as tone, inflection, and nuanced meaning.

Jakobson (2000), whose work is highly regarded as the starting point for the definition of translation, divides translation into three distinct types. The first type is intersemiotic translation or 'transmutation which he defines as the translation between different codes, such as visual, auditory, and the word. The second type is intralinguistic translation or 'rewording', which he defined as the translation within a language, and thirdly, interlinguistic translation, which is the translation between different languages.

Language can be defined as a system of communication between a particular group of people. Language can further be divided into formal and informal language based on the audience and purpose. Formal language refers to the use of grammatically correct structures, complex sentence formations, and adherence to standard conventions of vocabulary and grammar (Zega et al., 2024). Informal language, however, is mostly spontaneous and casual because it is mostly used amongst individuals who share personal relationships and is often used in informal contexts such as social gatherings (Eble, 1996). It is also used by a group of people with a common characteristic or trait, like a generation or an age group, and it mostly involves the use of slang expressions.

Eble (1996) defines slang as 'an ever-changing set of colloquial words and phrases that speakers use to establish or reinforce social identity or cohesiveness within a group or with a trend or fashion in society at large'. Slang expressions serve

various communicative functions, such as fostering friendliness, giving a sense of intimacy or belonging, and sometimes to ensure the message passed is only understood by certain people and not everyone who comes across the message. Nigeria is a country with numerous languages, and each language has its own slang expressions. All of the slang expressions from these languages, alongside slang expressions in English, which is the official language of the country, combine to form the Nigerian slang, which is popularly known as "Naija slang".

Naija slang portrays the rich and diverse cultural heritage that the country has to offer. It draws from the various local languages, dialects, and even sometimes international linguistic trends. It is also worth mentioning that the same slang might mean different things depending on the tone, context, and the identity of both the speaker and the listener. All these factors make it a hard task for AI systems to translate slang expressions correctly, as it often involves a deep understanding of not only Nigerian culture and language but also the situation that led to the use of the slang. However, a pragmatic analysis which involves the study of how language is used in context to achieve specific communication aims and objectives, can help train AI systems on how to not only translate slang expressions based on the syntax and semantics of the source language but also on the tone, style, and intent of the speaker, thereby facilitating translations that are more accurate and contextually appropriate.

There are several AI systems specifically built for translations, as translating with AI systems is significantly cheaper than the traditional manual processes. This is due to the fact that systems such as DeepL, Smartcat, and Quillbot can translate large volumes of text speedily, saving time, labour, and financial costs that usually come with human translators. Also, the ability to translate a significant number of languages with a high degree of accuracy contributes to the demand for AI translation systems. For example, DeepL offers translation services for over 30 languages, while Quillbot for 45 languages, and Google Translate for over 100 languages (Sujatha, 2024).

However, AI systems are often only effective for translating formal languages; the challenges of tone, style, and intent make it harder for them to translate informal languages such as slang expressions, proverbs, and idioms. Whereas several articles have been written on the use of AI for translation of formal languages, there has been limited scholarly attention on the use of AI for translation of informal languages, and even fewer on translating country-specific slang expressions.

The selected AI chatbots used in this study include: ChatGPT 4o from OpenAI, Grok AIfrom X, Gemini from Google, Copilot from Microsoft, and DeepSeek Chat from Deepseek. These chatbots are recognised as the top AI systems from some of the biggest companies in the world.

Several studies have been conducted to measure the effectiveness of AI systems in interlinguistic translations. A study conducted by Mohammed Mohsen published in 2024 compared the use of Google Translate, ChatGPT 3.5, and ChatGPT 4 in the translation of academic abstracts between English and Arabic. The results from the study showed that ChatGPT 3.5 and ChatGPT 4 outperformed Google Translate in translating from Arabic to English and in translating from English to Arabic. The study also found that ChatGPT 4 had a significant advantage over Chat 3.5 in Arabic-English translation, while there was no statistically significant advantage when translating from English to Arabic, and Google Translate displayed limitations in translating some contextual aspects and translated some terms literally. However, these abstracts were in academic contexts and were written in formal language, and not in informal language like slang expressions. The study also only employed three independent raters to carry out assessments of the translation outputs and just tested the effectiveness and precision of machine translation represented by Google Translate against that of two Large Language Models (Mohsen, 2024).

Another study is by Moneus & Sahari (2024). They researched to examine the contrasts between human and AI translations of legal texts to evaluate the possibility that there is no difference between the two, and to also examine rising concerns that the need for human translators will decline in the face of AI development. A collection of legal texts was chosen from various contracts, allocated to human translators and also AI translation systems, and the differences in translation were then examined using a contrastive methodology. The study aimed to prove the hypotheses that there are no statistical differences between human and AI translations in Arabic and that there are no statistical differences between human and AI translations in English. Ten professional human translators took part in translating six legal texts, and three AI models were also used in translating the same texts. The study identified that while speed, scalability, and consistency were the weaknesses of human translators, they served as strengths of the AI models, and while human translators were able to understand and interpret context, cultural nuances, and idiomatic expressions, and also handled ambiguous phrases and sentences correctly, AI models had problems replicating the same feat. The study

also mentioned that there has been a growing interest in combining the advantages of human and AI translation systems to create hybrid systems that can leverage the speed of AI translation with the accuracy and expertise of human translators. However, while this study was applied to legal documents that have their own terms and jargon, it is still in the context of formal language, and the documents translated were also in Arabic.

Yet another study conducted by Yan et al. (2024) made a comprehensive evaluation of the translation abilities of GPT-4 by assessing translations across three language pairs: Chinese-English, Russian-English, and Chinese-Hindi. The selected texts also cut across three domains. News, Technology, and Biomedical. Through qualitative research, they were able to identify distinctive patterns in the translation approaches of GPT-4 and human translators. They found that GPT-4 exhibited lexical consistency and tended towards giving translations that were overly literal and achieved performance success comparable to junior translators while falling behind senior translators.

Charles-Kenechi (2024) also made a scoping review of existing literature that extensively discussed the benefits and challenges that AI faces in translation studies. The review revealed that while AI tools added numerous benefits like efficiency, speed, cost-effectiveness, and scalability to translation, they also have deficiencies and technical challenges that originate from natural language processing and overly literal translation of texts. The study concluded by stating that incorporating AI tools into translation offers great potential and barriers and while AI translation helps make international communication more efficient, constant development and refinement are necessary to enhance AI translation capabilities and as it is improved through the joint efforts of developers, linguists and cultural sociologists, AI systems could give more precise and bias free translations that can help connect the world.

Fu & Liu (2024) noted that while AI translation brings convenience, it also imposes severe challenges on the dissemination of knowledge. They conducted a study that investigated the linguistic features of ChatGPT 3.5 and 19 Master of Translation and Interpreting students in China from lexical and syntactic levels. The data gathered showed that at the lexical level, human translators use more words but repeat more of the same vocabulary, while ChatGPT is better at translating technical or domain-specific terms correctly. At the syntactic level, human translators use more sentences to translate a passage, but the said individual sentences tend to be shorter in terms of the number of words used. Also, human translators often transform sentences from passive voice into active voice more than ChatGPT does,

and they also tend to deconstruct long and complex sentences into shorter clauses more skilfully.

However, all these studies conducted research that was focused on formal and technical texts with standardised terminologies, while there is a lack of research examining the performance of AI systems in translating language forms that are informal, culturally embedded, and rely heavily on context. Also, most available resources are centred around popular languages such as English, Chinese, Russian, Hindi, and Arabic, neglecting languages that are mostly fluid, nonstandard, and deeply rooted in cultural contexts that are constantly evolving, particularly in underrepresented linguistic varieties such as Nigerian Pidgin English. Furthermore, few studies apply theoretical frameworks such as Relevance Theory, Equivalence Theory, and Polysystem Theory to evaluate AI translation output. This study, however, fills these gaps by extending research into the realm of Naija slang expressions that were previously unexplored, applying robust theories to informal language contexts, and providing pragmatic and cultural analysis of AI and human translations.

The present study aims to evaluate AI effectiveness in translating Nigerian slang expressions by comparing AI-generated translations with human translations using Relevance, Equivalence, and Polysystem theories. Overall, this study seeks to answer the question: to what extent do AI systems accurately interpret and translate Nigerian slang expressions compared to human translators?

THEORETICAL FRAMEWORK

The responses from the human translators and the AI chatbots are analysed using the following theories: Relevance Theory, Polysystem Theory, and Equivalence Theory.

The Relevance theory

The Relevance theory was proposed by Dan Sperber and Deirdre Wilson in 1986 and is used within the fields of cognitive linguistics and pragmatics as a framework for understanding the interpretation of utterances. The main assumption of the theory is that “humans are endowed with a biologically rooted ability to maximise the relevance of incoming stimuli (including linguistic utterances and other communicative behaviour)” (Sperber & Wilson 1986). This study utilises the

Relevance theory to analyse how effectively AI and human translators capture the intended contextual meanings of the slang expressions in their translations.

The Polysystem theory

The Polysystem theory was proposed by Itamar Even-Zohar in the year 1978 as a dynamic model that conceptualises that language, literature, and culture all function together as interconnected subsystems within a dynamic system. And depending on prevailing sociocultural norms, the translated texts occupy either a central or a peripheral position within a given dynamic system (Even-Zohar, 1990). This helps to explain how translated texts are influenced by the systems of language, literature, and culture, especially for slang expressions that are constantly shifting and evolving. In this study, this theory is adopted to analyse how translations by AI chatbots are placed within this dynamic system. This study adopted Eugene Nida's Equivalence theory, which separates equivalence into two types: Formal equivalence, which focuses on staying close to the original structure and wording of the source text, and Dynamic equivalence, which focuses on conveying the intended meaning of the source text as naturally as possible in the target language (Nida, 1964). This theory is used to compare AI-generated slang translations to human translations to determine whether AI prioritises formal or dynamic equivalence.

These theories are used to dissect the qualitative, cultural, and contextual relevance of AI and human translations of Naija slang, to evaluate the adequacy and cultural sensitivity of AI-generated translations in comparison to human translations.

METHODS

Study Design

In this comparative, quantitative study, we evaluated the quality of translations of Nigerian slang using three theoretical frameworks of interest in translation studies: Relevance Theory (Sperber & Wilson, 1986), Equivalence Theory (Nida, 1964), and Polysystem Theory (Even-Zohar, 1990). These frameworks were employed to assess the semantic fidelity, pragmatic relevance, and cultural appropriateness of translations of the slang expressions.

Data Collection

Selection of Slang Expressions

We identified 31 Nigerian slang expressions that represented forms of informal, idiomatic, and culturally specific Nigerian English and Nigerian Pidgin.

Slang expressions were gathered from the existing literature, social media platforms, and from informal conversational contexts to ensure they were representative of contemporary usage.

Translation Sources

We collected translations for the slang expressions from six sources:

AI Models: Five contemporary AI language models were chosen to produce English translations of the slang expressions.

Human Translators: A group of 10 professional human translators with experience in translating between Nigerian English and Pidgin, who produced human reference translations, was used. Each AI model and human translator independently produced a corresponding English translation of every slang expression. The human reference translations were then evaluated, while the translations provided were not the same in most cases, the meanings were close, and the most frequently occurring translations were picked.

Scoring Method

Each dimension was evaluated using a 5-point Likert scale:

- 1 = Poor preservation of meaning/context/cultural fit
- 3 = Moderate preservation
- 5 = Excellent preservation, near-native equivalence

Raters and Procedure

A panel consisting of three bilingual experts fluent in Nigerian English, Nigerian Pidgin, and Standard English evaluated each translation based on the three theoretical dimensions. Raters were trained in the theoretical frameworks and scoring criteria in advance to allow for consistency in their ratings.

The final score for each translation on each dimension was calculated as the mean of the three raters' ratings.

Data Analysis

Data Preparation

The scores were arranged in a structured dataset of 31 slang items × 6 translation sources × 3 theoretical dimensions, which produced a total of 558 observations.

Statistical Methods

Descriptive statistics (mean and standard deviation) were calculated for each translation source for each of the theoretical dimensions. One-way analysis of variance (ANOVA) tests were used to evaluate whether statistically significant differences existed between translation sources for each of the three theories. To assess the translations of AI models compared to human translations, pairwise independent-samples t-tests (with adjustments for multiple comparisons using the False Discovery Rate (FDR) method) were then performed. All analyses were completed using Python's SciPy and statsmodels libraries.

Ethical Considerations

As this study involved the analysis of textual data and did not utilise human subjects, ethical approval was not required. Human translators provided informed consent for their participation and their translations to be used.

RESULTS AND FINDINGS

The study assessed the translation quality of Nigerian slang using three theories: Equivalence, Polysystem, and Relevance. The translations were compared on five AI models, ChatGPT, Grok, Copilot, Gemini, and DeepSeek, with translations produced by human translators. The translations were assessed using a 5-point Likert Scale, with higher values indicating more accurate and culturally relevant translations.

Statistical Analysis

Descriptive statistics indicated that human translators obtained the highest mean scores in all three dimensions: Relevance (mean = 4.75, SD = 0.53), Equivalence (mean = 4.82, SD = 0.55), and Polysystem (mean = 4.69, SD = 0.76). Among AI models, the Grok model also received the highest average scores (Relevance mean = 4.41, Equivalence mean = 4.11, Polysystem mean = 4.19), with ChatGPT, Copilot, Gemini, and DeepSeek following with lower averages on all three dimensions.

One-way ANOVA tests indicated there were statistically significant differences between the translation sources for all three theories (Relevance: $F = 22.94$, $p < 0.0001$; Equivalence: $F = 21.98$, $p < 0.0001$; Polysystem: $F = 23.81$, $p <$

0.0001), confirming that the quality of translation can differ statistically between AI and human translators.

Pairwise t-tests comparing AI models with human translations indicated that relative to human output, Grok's performance did not differ significantly after adjusting for multiple comparisons in the Relevance and Equivalence dimensions, while ChatGPT was close, with only Relevance showing a statistically significant difference (adjusted $p = 0.045$). In the other AI models, Copilot, Gemini, and DeepSeek, there were statistically significant gaps from humans' performance in all three dimensions, with Polysystem scores reflecting persistent challenges in achieving cultural adaptation and contextual appropriateness.

Overall, while some models display a relatively close approximation to human levels of semantic and pragmatic accuracy, statistically significant differences persist between AI models and human translators, particularly in terms of cultural appropriateness with respect to Polysystem theory.

DISCUSSION

This study utilised three theoretical frameworks (Relevance, Equivalence, and Polysystem) to compare the performance and translation quality of Nigerian slang expressions by five contemporary AI language models against human translators. Our findings reveal the strengths as well as limitations of what is currently available through AI translation when working with culturally embedded, contextually reliant language, specifically, informal contexts and idiomatic Nigerian slang.

As has been found in previous studies on machine translation (e.g., Lo et al., 2021; Bawden et al., 2020), human translators generally demonstrated superiority in translation performance in all dimensions when compared to AI models. In fact, humans demonstrated elaborate understandings of the sociolinguistic and pragmatic phenomenon of Nigerian slang as evidenced by their near-perfect scores on the Likert scale. This is consistent with the notion that human translators are traditionally far better suited to address contextual subtleties, pragmatic implicatures, and cultural references in the source language that AI has not yet fully grasped (Munday, 2016).

The qualitative analysis further underscores the challenges faced by AI models when translating Nigerian slang expressions. The slang “Opueh!” is an informal term generally used to refer to sex, and this translation was accurately

captured by human translators. AI models, however, produced literally and contextually incorrect translations for this slang, with three models (Copilot, DeepSeek, and ChatGPT) translating it as an expression of surprise, Grok translated it to mean "Plenty," and Gemini translated it to mean "Trouble". Another example is the slang "Oblee!" used by X users to refer to enjoyment and partying was accurately translated by human translators. Of all the AI models, Grok provided a somewhat accurate translation (Living it up), two AI models (Copilot and ChatGPT) translated the slang as a term used for police and law enforcement agents, Gemini translated it as someone stylish or well dressed, and DeepSeek translated it to mean nonsense. Through the lens of Relevance Theory, these AI translations lacked the contextual assumptions needed to derive the intended cognitive effect and the intended meaning of the speaker, which leads to communication breakdowns where pragmatic adequacy is compromised. (Sperber & Wilson, 1986)

"Send your aza" is a slang used to ask for a person's account details with the intention of crediting their account. Human translators and all AI models accurately translated this slang, successfully achieving both lexical and dynamic equivalence. However, the slang "Idan mi" is used as a term to hype people up and was accurately translated by human translators as "My boss" or "My idol". ChatGPT and Grok provided their translations as "My G! The real MVP!" and "My boss" respectively, while Copilot, Gemini, and DeepSeek translated it as a term used to hype one's self up. From an Equivalence theory perspective, humans were able to reflect not only the lexical equivalence but also the emotional and social context of the source expression, thereby achieving dynamic equivalence as well (Nida, 1964). This suggests that while the selected AI models were also able to sometimes replicate this achievement, it largely depended on their familiarity with the slang expressions in the source language, and there is no underlying architecture to detect and attempt to replicate contextual meanings and implications in their translations. This goes further to prove the limitations AI has when it comes to metaphoricity and culture. (Koehn, 2020)

While scoring high for Relevance and Equivalence, ChatGPT, Grok, and Gemini AI models maintain lower scores for Polysystem, in other words, not gaining sufficient cultural adaptation or sociolinguistic appropriateness in terms of distinctions between aspects of informal speech (Even-Zohar, 1990). This can be seen from slang expressions such as "O ti lo" and "Pepper don rest". In this way, we can argue that these AI systems maintain at least a form of literal and functional correctness, but there are still inadequate cultural norms and informality that factor

into their translations. Copilot and DeepSeek, however, fell significantly short, showing challenges with idiomatic and context-enhanced phrases. This is consistent with previous observations that many AI translation models, especially those that were not specifically trained on common or low-resource languages, tend to provide literal or formal renditions that lack cultural fit (Tiedemann, 2020). This matters because slang is significant to identity and group belonging (Eble, 1996), and effective translational work requires products that are culturally sensitive to preserve the social significance of slang.

The practical implications of these conclusions are that while AI might assist in generating rough translations or help to moderate content, human judgment is essential when developing translation products that are contextually and culturally sensitive, particularly when translating informal discourse rich in idioms and slang. This challenge dovetails with claims that hybrid human and AI workflows have been proposed in translation practice to achieve a balance between efficiency and cultural fidelity (Way, 2018).

LIMITATIONS

The research also has several limitations that must be taken into consideration when interpreting the findings.

Firstly, the dataset used, while rich in cultural variation, only had 31 slang expressions and was limited to a very specific context of language use. Although we were able to identify important trends in our sample dataset, findings may not allow for full generalisability to other cultural slang systems or contexts of language use with differing socio-pragmatic systems. Future research using larger datasets that reflect a greater variety of language use will enhance analyses and interpretation.

Secondly, the AI systems used in this study were the commercially available models that were not fine-tuned to Nigerian linguistic corpora, and the performance of the systems reflects the general limitations of these globalised AI models when applied to localised, informal linguistic systems. This highlights the current capabilities and urgent need for more domain-specific AI models.

Thirdly, while this study makes use of sound theoretical frameworks, the subjective coding of translation adequacy introduces a potential reduction in reliability due to minor evaluator bias despite efforts at rater reliability. Further research utilising automated and pragmatic scoring tools or a larger panel of experts may increase reliability.

CONCLUSION

As seen from the foregoing, while they are highly effective with formal languages, AI systems in most cases struggle with culturally grounded and non-standard expressions due to limitations in pragmatic decoding, inferencing, and cultural familiarity. Unlike prior studies that focus primarily on formal language domains, this study contributes findings from culturally dense language varieties and emphasises the relationship between culture, language, and AI translation capabilities. Future research should further prioritise the development of training datasets heavily influenced by culture, translation algorithms that are sensitive to context, and translation workflows that draw on human cultural expertise.

REFERENCES

Bawden, R., Birch, A., Reddy, S., &Blunsom, P. (2020). Disentangling language and knowledge in task-oriented dialogue systems. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 1899–1905). https://doi.org/10.18653/v1/2020.acl-main.173.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems,* 33, 1877–1901. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html.

Charles-Kenechi, S. (2024). Artificial intelligence in translation studies: Benefits and challenges. *Cascades*: *Journal of the Department of French & International Studies*, 2(1), 5–15.

Eble, C. (1996). *Slang and sociability: In-group language among college students.* University of North Carolina Press.

Even-Zohar, I. (1990). Polysystem studies. *Poetics Today*, 11(1), 1–97. https://doi.org/10.2307/1773176.

Fu, L., & Liu, L. (2024). What are the differences? A comparative study of generative artificial intelligence translation and human translation of scientific texts. *Humanities and Social Sciences Communications*, 11, Article 1236. https://doi.org/10.1057/s41599-024-03726-7.

Jakobson, R. (2000). On linguistic aspects of translation. In L. Venuti (Ed.), *The translation studies reader* (pp. 113–118). Routledge.

Koehn, P. (2020). *Neural machine translation*. Cambridge University Press. https://doi.org/10.1017/9781108569275.

Lo, C.K., Tiedemann, J., & Hovy, E. (2021). Machine translation quality estimation: From black-box to glass-box evaluation. *Computational Linguistics*, 47(1), 131–174. https://doi.org/10.1162/coli_a_00395.

Mohsen, M. (2024). Artificial intelligence in academic translation: A comparative study of large language models and Google Translate. *Psycholinguistics*, 35(2), 134–156.

Moneus, A.M., & Sahari, Y. (2024). Artificial intelligence and human translation: A contrastive study based on legal texts. *Heliyon*, 10, e28106. https://doi.org/10.1016/j.heliyon.2024.e28106.

Munday, J. (2016). *Introducing translation studies: Theories and applications* (4th ed.). Routledge.

Nida, E.A. (1964). *Toward a science of translating: With special reference to principles and procedures involved in Bible translating.* Brill.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Blackwell.

Tiedemann, J. (2020). Machine translation for low-resource languages. In P. Koehn (Ed.), *Recent advances in NLP* (pp. 233–250). Springer. https://doi.org/10.1007/978-3-030-32308-1_12.

Way, A. (2018). Quality expectations of machine translation. *Translation Spaces*, 7(1), 90–111. https://doi.org/10.1075/ts.00005.way.

Yan, J., Yan, P., Chen, Y., Li, J., Zhu, X., & Zhang, Y. (2024). GPT-4 against human translators: A comprehensive evaluation of translation quality across languages, domains, and expertise levels. arXiv. https://doi.org/10.48550/arXiv.2407.03658.

Zega, R., Daeli, R., Zebua, E., Harefa, Y., Zai, F., Zendrato, M., & Hulu, J. (2024). Formal language use in conversations among second-semester students at English Education Study Program of Universitas Nias. *JELITA: Journal of English Language Teaching and Literature,* 5(2), 445–458. https://doi.org/10.56185/jelita.v5i2.794.

TABLES

Table 1. Slang expressions and their AI and human translations

Slang	Copilot	ChatGPT	Grok	Gemini	DeepSeek	Human
Dem no dey tell person	Some things are better experienced than explained	Some things are better experienced than explained	It's obvious, no one needs to say it	They don't tell a person	They don't give warnings	Experience is the best teacher
You dey cap	You're lying	You're lying	You're lying	You are lying	You're lying	You're lying
Opueh!	Wow or Oh my goodness	Whoa! Damn! Oh my God!	Plenty or a lot	Trouble or problem	Wow or Damn	Sex
Oblee!	Slang for a police officer	The cops	You're on fire/Living it up	Someone stylish or well-dressed	Look at this nonsense	Enjoyment
E choke	It's overwhelming or intense	This is too much!	It's overwhelming	Being overwhelmed	It's too much	It's too much
Soro soke	Speak louder	Speak up!	Talk louder	Speak up	Speak up	Speak up
Idan mi	My mystical self or confidence	My G! The real MVP!	My boss	I'm exceptional	I'm the GOAT	My boss/My Idol
Omo!	Wow or Oh		Dude	Wow	Wow or	Wow

Slang	**Copilot**	**ChatGPT**	**Grok**	**Gemini**	**DeepSeek**	**Human**
	boy	Damn! Wow! Bruh!			Damn	
Send your aza	Send your account details	Drop your account number	Send your account details	Send your bank account details	Send your bank details	Send your account details
Pepper don rest	Situation calmed down	Things are quiet now	The money has stopped flowing	The tension has eased	The hype is over	The money has arrived

Table 2.Relevance Theory Likert Scale.

Slang expressions	**ChatGPT**	**Grok**	**Copilot**	**Gemini**	**DeepSeek**	**Human Translators**
Dem no dey tell person	5	5	5	1	1	5
You dey cap	5	5	5	5	5	5
Opueh!	1	1	1	1	1	4
Oblee!	1	3	1	1	1	5
E choke	5	5	5	3	4	5
Soro soke	5	5	5	5	5	5
Idan mi	5	5	1	1	1	5
Omo!	5	5	5	5	5	5
Send your aza	5	5	5	5	5	5
Pepper don rest	2	5	1	1	1	3

Table 3. Equivalence Theory Likert Scale

Slang	ChatGPT	Grok	Copilot	Gemini	DeepSeek	Human
Dem no dey tell person	5	5	4	3	2	5
You dey cap	5	5	5	5	5	5
Opueh!	1	1	1	1	1	5
Oblee!	1	5	1	1	1	5
E choke	5	5	5	5	5	5
Soro soke	5	5	5	5	5	5
Idan mi	5	5	2	2	2	5
Omo!	5	5	5	5	3	5
Send your aza	5	5	5	5	5	5
Pepper don rest	2	5	1	2	2	4

Table 4. Polysystem Theory Likert Scale

Slang expressions	**ChatGPT**	**Grok**	**Copilot**	**Gemini**	**DeepSeek**	**Human Translators**
Dem no dey tell person	4	4	5	1	1	5
You dey cap	4	5	5	5	5	5
Opueh!	1	1	1	1	1	4
Oscroh/OS	1	5	1	1	1	5
Oblee!	1	5	1	1	1	5
E choke	5	5	5	4	4	5
Soro soke	5	5	5	5	5	5
Idan mi	5	5	2	2	1	5
Omo!	5	5	5	5	5	5
Send your aza	5	5	5	5	5	5
Pepper don rest	1	5	1	1	1	3