Journal of Advances in Mathematical & Computational Sciences
An International Pan-African Multidisciplinary Journal of the SMART Research Group
International Centre for IT & Development (ICITD) USA
© Creative Research Publishers
Available online at https://www.isteams.net/ mathematics-computationaljournal.info
CrossREF Member Listing - https://www.crossref.org/06members/50go-live.html

Proficiency of Turing test on Nigerian Major Languages; Yoruba, Igbo and Hausa Using Google Translate.

[1]Abiola O. Arowolo, [2]Adebisi R. Oluwatoyin. & [3]Akinola S. Olalekan
[1,2,3] Department of Computer Science
University of Ibadan
Ibadan, Nigeria.
E-mails: [1]oladimejiarowolo@yahoo.co.uk, [2]abdurrazaq.raheemah@gmail.com,
[3]solom202@yahoo.co.uk

ABSTRACT

Google Translate performance evaluation in translating between English and three Nigerian languages: Yoruba, Igbo, and Hausa, highlighting its strengths and limitations using the Turing Test. The study assessed whether participants could distinguish Artificial Intelligence (AI) - generated translations from human translations which reveals that Google Translate handles basic communication in these languages but struggles with more complex linguistic features. The tool adequately translates straightforward sentences in Yoruba language, but falters with idiomatic expressions, proverbs, and tonal nuances, the same problem refects in Igbo language facing challenges with more complex sentences and dialectal variations, resulting in culturally irrelevant outputs. Hausa performed better due to the availability of a larger training dataset, delivering more accurate translations for common phrases, but still struggled with cultural and contextual nuances.The findings suggest that Google Translate's limitations prevent it from passing a Turing Test with native speakers of Yoruba, Igbo, and Hausa. While understandable translations are produced for basic sentences, native speakers can easily detect inconsistencies, awkward phrasing, and a lack of cultural nuance. These issues underscore the significant gap between AI translation and human-like proficiency for African languages.

Keyword: Turing test, Nigerian Major Languages; Yoruba, Igbo, Hausa Using Google Translate

Abiola, O.A., Adebisi R.O. & Akinola, S.O. (2024): Proficiency of Turing test on Nigerian Major Languages; Yoruba, Igbo and Hausa Using Google Translate. Journal of Advances in Mathematical & Computational Science. Vol. 12, No. 3. Pp 1-12
Available online at www.isteams.net/mathematics-computationaljournal. dx.doi.org/10.22624/AIMS/MATHS/V12N4P1

1. INTRODUCTION

Artificial Intelligence (AI) language models are advanced algorithms designed to understand and produce human language by training on vast amounts of text data, to capture nuances, grammar, meaning, and context.

Effective machine translation, especially whendealing with ambiguous sentences requires critical understanding of the language context. Strength of AI models is their ability to generate natural-sound translations, unlike older rule-based systems. These model adapt to different languages and technical vocabularies by fine-tuning with specific datasets which handle difficultlinguistic elements like idioms and ambiguities. AI models have transformed machine translation by increasing accuracy and ensuring translations while still maintain the original meaning's subtlety. Their speed and scalability make themideal for real-time applications and global communication. The natural-sound outputs enhance user experience, particularly incontexts like chatbots and customer support system. These models also improve accessibility by breaking down language barriers like bias in training data, cultural misunderstandings, and other limitations.

1.2 Turing Test

Turing Test was developed by Alan Turing in 1950 to measure whether a machine can behave intelligently in a way indistinguishable from a human. The original version involved text- based communication between a human assessor and a machine. If the human couldn't reliably differentiate between the machine and another human, the machine passed the test. Turing Test helps assess the quality of machine translation. To pass, the translation must be accurate, sound natural, and reflect context and cultural nuances like a human translator would. Evaluators would compare translations from both a machineand a human to see if they could identify the machine output based on quality alone.

1.3 Translation Problem

Translation of African languages is becoming more critical due to globalization; and effective communication across its many languages which makes it essential for education, business, and governance. Many translation tools, including Google Translate, face significant shortcomings in translating African languages like Yoruba, Igbo, and Hausa. Language translations are vital for cultural preservation and clear communication in area of healthcare, education, and business. Google Translate supports only a limited number of African languages and the translations are often inaccurate due to insufficient training data. African languages have unique grammatical structures and idiomatic expressions that Google Translate struggles to translate correctly. Many dialects within these languages are notaccurately captured leading to further issues.

2. RELATED WORKS

The history of AI-based translators, especially Google Translate, shows how far language processing technologies have come and how important machine translation is becoming in today's world. It all started in the mid-20th century with early machine translation efforts, which relied on rules and dictionaries to translate text. These systems had a hard time dealing with the complexities of natural language, like idioms and slang, because language is not always straightforward. In the late 1980s and early 1990s, statistical machine translation changed things significantly.This approach used large datasets of bilingual text to find patterns and probabilities for language translation. Google Translate was launched in 2006 and originally used these statistical methods. Its popularity grew quickly due to its easy-to-use interface and its supportfor many languages.

Their was a major breakthrough in the mid-2010s with the introduction of neural machine translation. The method used deep learning to analyze whole sentences instead of just translating word by word. With neural networks, Google Translatecould understand the context better, resulting in smoother and more natural translations. In 2016, Google adopted this method, improving its translations significantly. The major advantages of AI in translation is that it learns and improve over time by using data from user corrections to enhance its translations. AL-Jashami, et, al. (2024) reported that current AI systems like Google Translate struggle with African languages due to the complexity of idioms, tone, and dialects. The authors emphasize the importance of including native speakers in the AI development process to improve translation accuracy.

Yoruba, Igbo, and Hausa languages don't perform maximally in AI translations due to limited data and cultural nuances. Community involvement can improve translation systems forunder represented African languages like Yoruba, Igbo, and Hausa. By allowing local speakers to contribute to the development of datasets Nekoto, *et al* (2020). AI-based translators advances from simple rule-based systems to neural network based methods. Though there has been major progress in AI translation tools, yet there is still need for inclusive and improve accurate, especially forunder represented African languages.

Arikpo, & Dickson, (2018) developed a machine translation system for English to Efik using a transfer-based approach. The study bridged communication gaps in Nigeria's multilingual setting. The system was implemented by Java, offering optimal translations for simple sentences but struggled with complex ones. The limitation of the study was that it did not address Turing Test and was also limited to one local language, Efik.

Grounding problem in AI was examined by Alberts (2020) where he emphasized that AI must link language toreal-world experiences to mimic human understanding. Traditional rule-based systems werecritiqued for lacking this ability. The study advocates for AI systems to integrate sensory, perceptual, and motor experiences, drawing on cognitive theories like 4E Cognition to make language processing more human-like. Gams, *et al* (2024) evaluated ChatGPT's performance in the Turing Test, concluding that while it convincingly mimics human behavior, it falls short of true consciousness. The authors applied Integrated Information Theory (IIT) and found that, despite impressive language skills, ChatGPT lacks semantic understanding and conscious experience.

Uchendu, *et al* (2021) introduced TURINGBENCH, a dataset of 200K samples to evaluate AI's ability togenerate human-like text and detect machine-generated content. It revealed that models likeGPT-3 are highly human-like, making detection difficult, emphasizing the need for better AI detection systems to prevent misuse, such as spreading fake news. Bouguesmia, (2020) explored how AI affects translators, particularly in Algeria. The case study found that while educators saw AI as a potential threat to traditional translation roles, many were open to embracing AI in translation. However, the research had limitations due toits small sample size.

2.2 Turing Test in Language Models

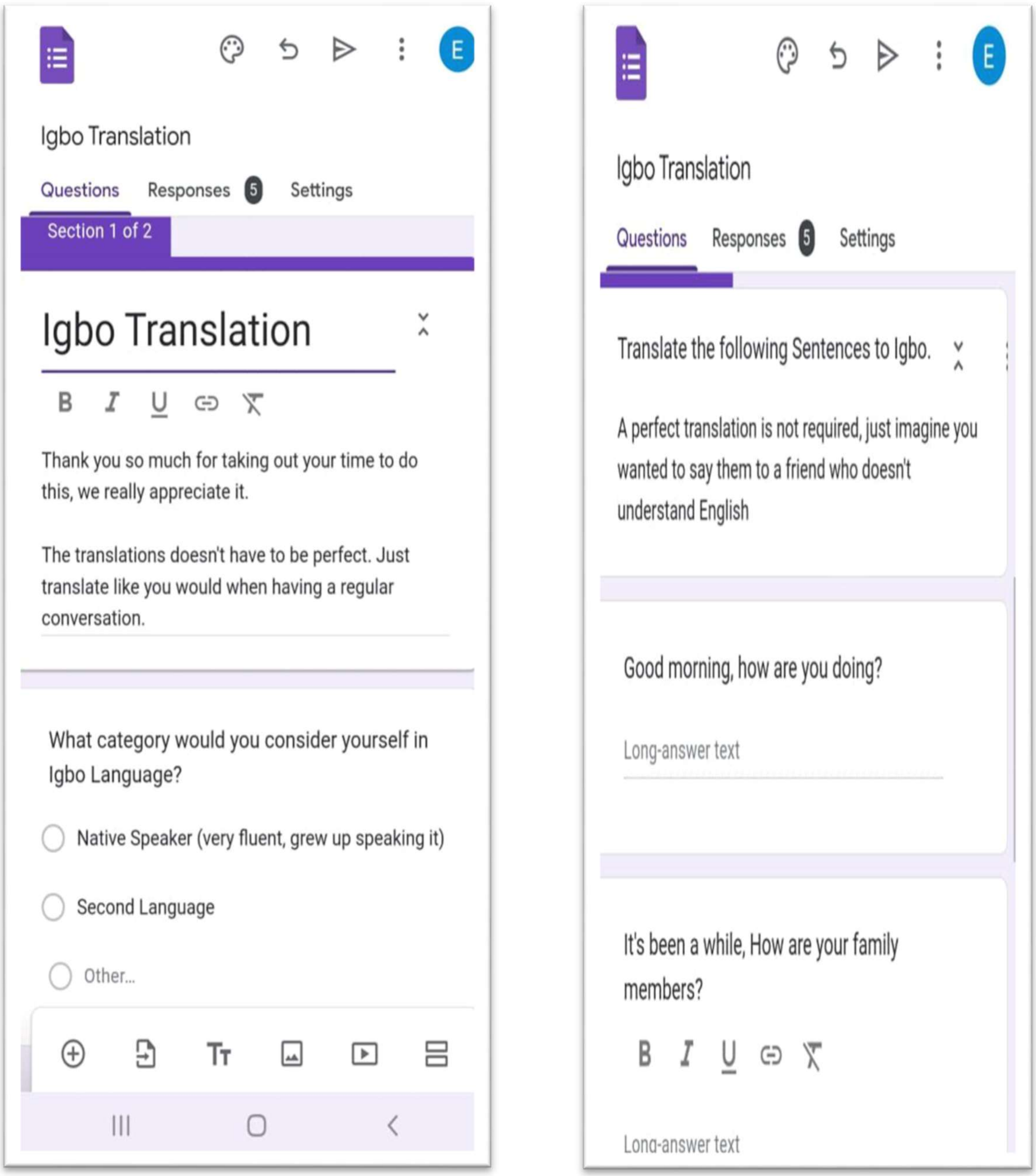
The Turing Test, created by Alan Turing in 1950, is one of the earliest ways to measure if a machine is intelligent. It works like this: If a human can't tell whether they are having a conversation with a machine or another human, then the machine is considered intelligent. Research into the Turing Test has looked at how well AI models can carry on conversations.In these experiments, judges talk to both humans and machines without knowing which is which. Advanced AI models can sometimes give very convincing responses that sound human-like, but they still struggle with subtleties like emotion, humor, and the deeper meanings behind words.

Martinus, et, al., (2019) reviewed existing machine translation systems for African languages, focusing on the need for better datasets. The research uses corpora from South Africa to train models andprovides benchmarks that highlight the current gaps in translation quality for African languages. Another area of research focuses on how AI performs in different languages. AI models tendto work better with languages that have a lot of data available, like English or Spanish. However, they don't perform as well in languages that are less common in digital spaces, such as many African languages. This makes it harder for these models to pass the Turing Test in those languages. Turing Test plays a vital role in designing an improved AI applications. Turing Test is a classic method for evaluating AI and currently researchers are exploring better techniques for testing whether AI truly understands language beyond just mimicking human speech.

3. METHODOLOGY

This study evaluated the performance of Google Translate in translating between English and three Nigerian languages: Yoruba, Igbo, and Hausa. The Turing Test methodology was used to assess whether participants could distinguish AI-generated translations from human translations. This approach allows us to explore the proficiency of Google Translate when applied to African languages, an area that has received limited attention in machine translation research (Eludiora & Odejobi, 2011; Gouws et al., 2012). Byanalyzing translations of everyday conversations, idiomatic expressions, and formal communications, the study tested Google Translate's ability to handle both formal and informal language contexts.

Language fluency can impact how translation quality is perceived, having both native and non-native speakers participate was essential for a more holistic evaluation. Native speakers often have an intuitive understanding of idiomatic andculturally specific language features, which AI translation models frequently fail to grasp (Toral et al., 2018).

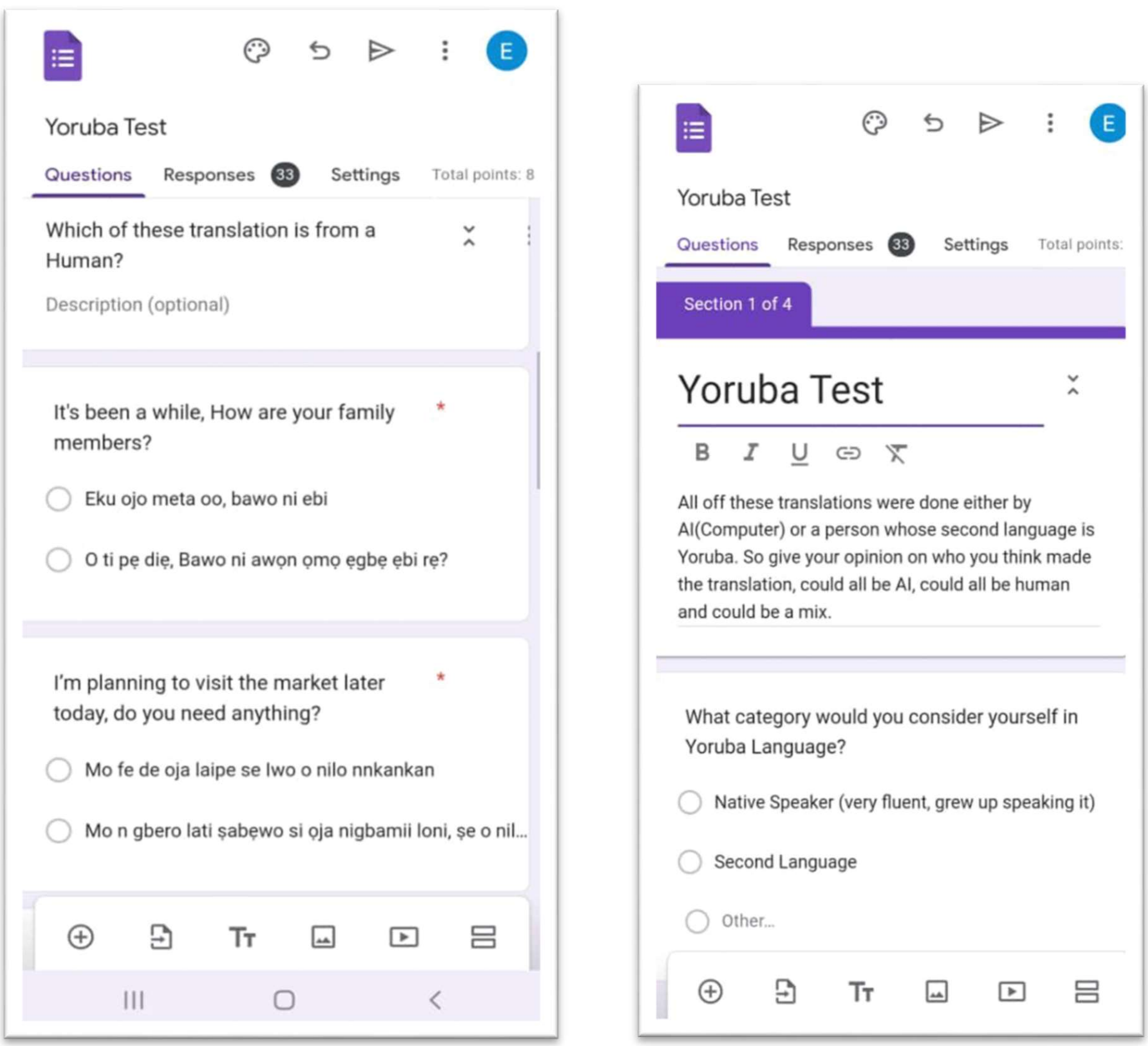


(a) (b)

Figure 1(a and b): (a)Google form for Igbo text, (b) Google form responses

3.2 Participants

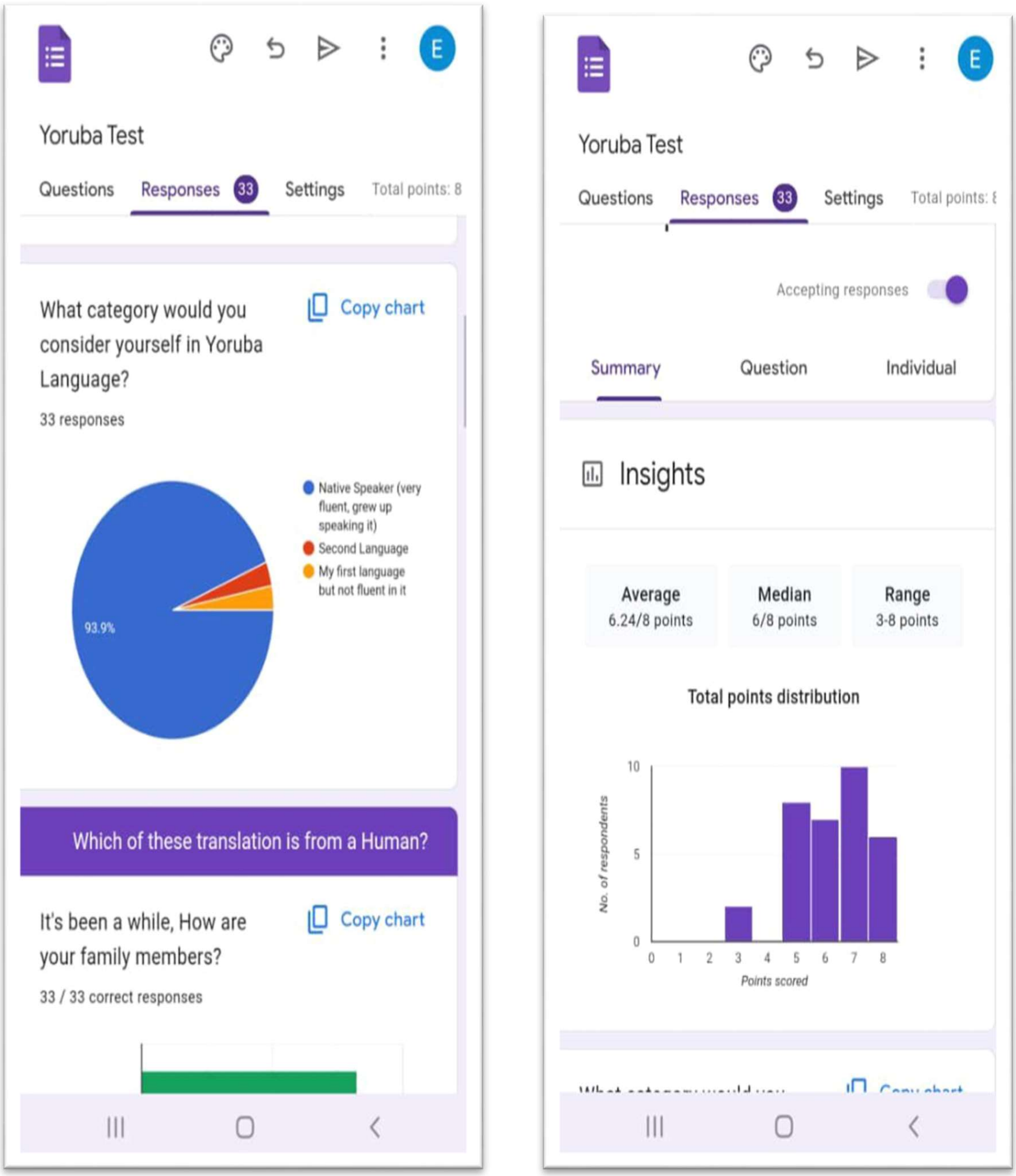
A total of 46 participants took part in this study, divided across Yoruba, Igbo, and Hausa speakers in Figure1(a&b). The participants were a mix of native and non-native speakers, which allowed us to assess how Google Translate performs across different levels of language proficiency. For Yoruba, 94% of the participants were native speakers from Figures 2(a&b), which gave us deeper insight into thesystem's ability to translate nuanced and culturally specific phrases. Hausa participants were evenly split, with 57.1% being native speakers.



(a) (b)

Figure 2(a & b):Google form for Yoruba test and responses

The Igbo group had 83.3% native speakers. This distribution reflects the varied linguistic landscape of Nigeria in Figure 3(a&b), where multilingualism iscommon, and non-native fluency in multiple languages is widespread (Ethnologue, 2021).



(a) (b)

Figure 3(a & b): Distribution Chart for Yoruba language

3.3 Test Design

The test was structured into three sections for each language as seen in Figure 2 , and 3 which was delivered through Google Forms. The test sentences were carefully selected to reflect real-life situations where accuratetranslation is critical. This methodology was informed by prior research that emphasizes the importance of practical, everyday scenarios in assessing machine translation tools (Eludiora& Odejobi, 2011).

3.3.1 Human vs. AI Translation Comparison

Participants were presented with two translations of the same sentence—onedone by a human and the other by Google Translate. They were asked to guess which was human-generated. Sentences ranged from casual greetings like "Good morning, how are you doing?" to more complex expressions such as "I'm planning to visit the market later today, doyou need anything?" This setup allowed us to evaluate Google Translate's ability to handle common conversational sentences. Google Translate often struggles with these types of sentences in non-Western languages due to a lack of extensive training data (Fomicheva et al., 2020).

3.3.2 Guessing Task - Human or AI

The participants were shown isolated translations and asked to guess whether the translation was AI-generated or human. This format eliminated the side-by-side comparison,allowing us to test Google Translate's performance in a more naturalistic setting, without explicit comparison to human translation. The design mirrors approaches used in studies thataimed to evaluate AI's ability to pass the Turing Test (Levesque et al., 2012).

3.3.3 Reverse Translation Task

In this phase of the research, Yoruba, Igbo, and Hausa languages were translated back into English to evaluate theaccuracy of reverse translations. Participants were to assess translations ofsentences such as "I recently bought this phone, and it works great! The battery lasts all day, and the camera quality is fantastic." This task helped highlight whether Google Translate could maintain meaning when translating from African languages into English, an area wheremachine translation often struggles (Hassan et al., 2018). Participants were also asked to suggest idiomatic expressions from their languages.This was crucial for testing Google Translate's ability to handle culturally specific language,which has been identified as a major challenge in prior studies (Gouws et al., 2012; Eludiora& Odejobi, 2011).

3.3.4 Data Collection and Analysis

Quantitative and qualitative data were collected from participants through the survey. Participants' ability to distinguish between human and AI-generated translations was measured using a scoring system. A higher average score indicated that participants found iteasier to identify AI-generated content, which suggests that Google Translate struggled withproducing natural-sounding translations (Fitria, 2020). Qualitative feedback was also gathered to understand why participants believed certain translations were AI-generated, providing insights into specific translation errors such as grammatical mistakes, literal translations, or lack of cultural context (Gouws et al., 2012).

4. RESULTS AND ANALYSIS

Evaluating Google Translate's performance with Yoruba, Igbo, and Hausa shows a mix of strengths and weaknesses. For Yoruba, basic sentences are generally well-translated, but the tool struggles with proverbs, idiomatic expressions, and tonal variations, leading to inaccuracies. In Igbo, simple phrases are often translated correctly, but the tool faced challenges with more complex sentences and dialectal differences. Hausa performs better than Yoruba and Igbo due to a larger amount of training datasets, but even here, the tool struggles with cultural nuances and context. Yoruba Language Results: Google Translate struggled with Yoruba, scoring 6.06/8, showingthat participants easily identified AI translations. Issues arose with cultural and idiomatic expressions, such as the literal translation of "Lálá tó ròkè, ilẹ ló ń bọ" ("Whatever goes up must come down"). Participants also flagged errors in diacritical marks, which play a crucialrole in meaning. The system's failure to capture the fluidity and cultural nuances ofconversational Yoruba was evident.

Hausa Language Results: In contrast, Google Translate performed better in Hausa, scoring 4.43/8. Participants found it harder to distinguish between human and AI translations, particularly in formal contexts like medical instructions. Hausa's simpler grammar and lackof tonal complexity contributed to this success. However, idiomatic expressions like "Wonitsuntsu ya ke gudu rua, awuagwa rua ta ke nema" were still correctly identified as AI generated. Igbo Language Results: Google Translate's performance in Igbo was similar to Yoruba, witha score of 6.67/10. Participants noted overly formal translations and a failure to account for dialect variations. The translation of idioms, such as "Eme ngwa ngwa emeghara odachi" ("Act quickly to avoid disaster"), was also problematic, highlighting Google's difficulty withculturally nuanced language. Comparative Analysis: Overall, Google Translate performed best in Hausa due to its straightforward grammar. Yoruba and Igbo, with their tonal complexities and culturalcontexts, proved more challenging, leading to higher detection rates of AI translations. Thesefindings align with previous research on the limitations of AI in low-resource languages withcomplex structures and underrepresentation in training data.

4.1 Interpretation Of Result

A 50% score in Hausa suggests that while Google Translate can provide basic accuracy, it struggles with consistency and context. Factors such as limited training data and cultural nuances contribute to this. Yoruba and Igbo posed more challenges due to their tonal nature and rich cultural context, making it easy for native speakers to detect AI-generated translations. The lack of data and failure to account for tone and idiomatic expressions furthercomplicate the process, leading to unnatural translations that fail the Turing Test for these languages..

4.2 Challenges In AI Translation

AI translation faces several hurdles with Nigerian languages. A major issue is contextual understanding, as AI often provides literal translations, missing the cultural or broader meaning. Nigerian languages are filled with idiomatic expressions, proverbs, and unique grammar rules, which AI struggles to capture. Furthermore, the scarcity of high-quality training data and multiple dialects, especially in languages like Igbo, complicates accuracy. Tonal differences in languages like Yoruba limits AI models from interpreting tone correctly, thus impacting correct translations.

4.3 Implications

Future AI translation models need to focus on improving contextual understanding and cultural nuance by using diverse datasets, including dialects and idiomatic expressions. Collaborating with linguists and native speakers will be crucial for incorporating cultural insights. Additionally, improving the availability of high-quality training data, recognizing tonal variations, and engaging local communities in AI development are essential to enhancing the accuracy and cultural relevance of AI translations for African languages.

5. CONCLUSION AND FUTURE WORK

Google Translate's performance with Yoruba, Igbo, and Hausa reveals several strengths and weaknesses. It handles simple phrases well but struggles with idiomatic expressions, tonal nuances, and dialectal differences, especially in Yoruba and Igbo. Hausa translations are more accurate due to larger training datasets, but Google Translate still faces challenges withcultural and contextual understanding across all three languages. Overall, while useful for basic communication, Google Translate fails to convincingly pass the Turing Test, as nativespeakers easily identify errors and lack of cultural relevance.

5.1 Future Improvements

Improving AI translation for African languages requires better cultural contextualization, including algorithms that capture idiomatic expressions and proverbs. Tonal awareness is essential for languages like Yoruba. Expanding training data from various sources, such as dialectal variations, and incorporating user feedback would enhance the accuracy. Partnerships with local linguists and adaptive machine learning techniques would also significantly improve performance, fostering AI models that better reflect linguistic diversity.

5.2 Limitations

This study was limited by the small participant pool, which constrained the diversity of insights. A larger, more regionally diverse pool would have enhanced the findings. Additionally, providing remuneration could drive greater participation and engagement, leading to a more comprehensive analysis.

5.3 Challenges

The study faced difficulty securing participation without remuneration. Future studies could explore how dialect variations and cultural context affect translation accuracy, particularly in languages with rich regional differences like Igbo and Hausa. Addressing technological and linguistic challenges and incentivizing user engagement could offer valuable insights into improving machine translation for African languages.

5.4 Googleform Link for the research

Hausa Test: https://forms.gle/Yi1JmMWmERsY6rVj9
Igbo Test: https://forms.gle/6MCr9FqfWupP2zeS9
Yoruba Test: https://forms.gle/N78D6aLiCotL7LbF7

REFERENCES

1. Safiriyu I., & Odejobi, O. A. (2011). Development of an English to Yorùbá Machine Translator.International Journal of Computer Applications.
2. Bojar, O., et al. (2018). Proceedings of the Third Conference on Machine Translation (WMT18).
3. Eludiora, S. I., & Odejobi, O. A. (2011). Development of an English to Yorùbá Machine Translator.International Journal of Computer Applications, 25(1), 1-6.
4. Ethnologue. (2021). Languages of Nigeria.
5. Fitria, T. N. (2020). Penerapan Teknologi Informasi dalam Pembelajaran: Web-Based Learning diIndonesia.
6. Fomicheva, M., et al. (2020). Evaluation of translation in low-resource languages. Journal ofMachine Translation, 45(3), 55-70.
7. Gouws, S., et al. (2012). Contextual translation of idioms. Journal of Computational Linguistics,38(4), 709-727.
8. Hassan, H., et al. (2018). Achieving human parity on automatic Chinese to English news translation.Proceedings of ACL.
9. Levesque, H., et al. (2012). The Winograd Schema Challenge. Proceedings of KR.
10. Toral, A., et al. (2018). Attaining human-level translation: challenges for low-resource languages.Journal of Natural Language Engineering, 24(1), 55-77.
11. Nekoto, W., Marivate, V., Matsila, T., Fasubaa, T., Kolawole, T., Fagbohungbe, T., ... & Bashir, A.(2020). Participatory research for low-resourced machine translation: A case study in african languages.
12. AL-Jashami, R. A. U., Al Abbasi, O. M. M., & Bedu, A. M. (2024). Artificial Intelligence (Ai) Challenges and Opportunities in Translation: An African Experience.Generative Approach to African Linguistics and Literary Studies Afestschrift in Honour of Professor Mohammed Munkaila.
13. Martinus, L., & Abbott, J. Z. (2019). A focus on neural machine translation for african languages.
14. Arikpo, I., & Dickson, I. (2018). Development of an automated English-to-local-language translator using Natural Language Processing. International Journal of Scientific and Engineering Research, 9(7), 378-383.
15. Alberts, L. (2022). Not cheating on the Turing Test: towards grounded language learning in ArtificialIntelligence.
16. Gams, M., & Kramar, S. (2024). Evaluating ChatGPT's consciousness and its capability to pass theTuring test: A comprehensive analysis. Journal of Computer and Communications.
17. Uchendu, A., Ma, Z., Le, T., Zhang, R., & Lee, D. (2021). Turingbench: A benchmark environment forturing test in the age of neural text generation.
18. BOUGUESMIA, M. T. (2020). Using AI in Translation, a Technological Leap, or a Translator'sNightmare. ALTRALANG Journal, 2(02), 78-102.