Coptic Low Resource Machine Translation (LRMT)

GEORGETOWN UNIVERSITY
McCourt School of Public Policy
MASSIVE DATA INSTITUTE

MDI Scholars: Alexia Guo, B.S, Mathematics and Statistics, 2026
MDI Advisor: Amir Zeldes, Department of Linguistic, Georgetown University

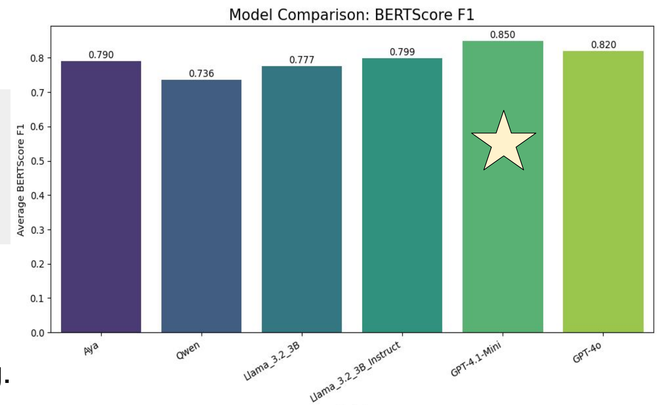
Abstract

We investigate low-resource **Coptic→English translation** using an in-context learning **(ICL)** approach with **GPT-4.1 Mini.** By augmenting prompts with lexical glosses, dictionary entries, syntax and morphological analyses, and grammatical cues from the Coptic Scriptorium corpus, we improve translation accuracy without fine-tuning. Experiments show a **3.78% increase in BERTScore F1**, demonstrating that linguistically informed prompting can meaningfully enhance translation quality for low-resource languages.

Methodology

Figure 1. Different model's Bert Score

- Firstly, we selected **GPT-4.1 Mini** based on translation quality (highest average BERT score)



- **In-Context Learning (ICL):** Provided Coptic→English exemplars to guide translation behavior without fine-tuning.
- **Lexical + Dictionary Information:** Added glosses and word-level meanings for key Coptic terms.
- **Curated Data Source:** Used Coptic–English sentence pairs from the *Coptic Scriptorium* corpus.
- **Prompt Engineering:** Combined linguistic guidance, syntax information, and multiple ICL examples with the target sentence for improved translation.

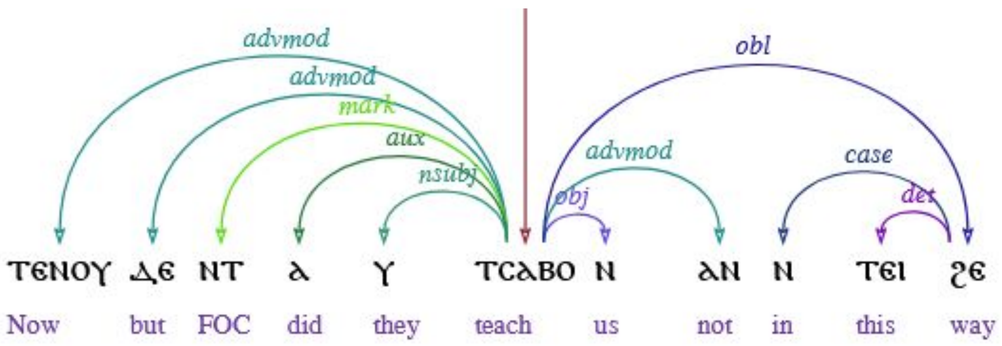


Figure 2. Dependency Syntax Analysis of a Coptic-to-English Sentence

- The syntax-analysis diagram demonstrates how Coptic words relate through dependencies such as nsubj, aux, obj, and obl, giving the model explicit structural guidance. Including these syntactic relations in our prompts helps GPT-4.1 Mini generate more accurate and grammatically aligned translations.

Results

- We developed an evaluation pipeline capable of comprehensively assessing the performance of the language model using various metrics.

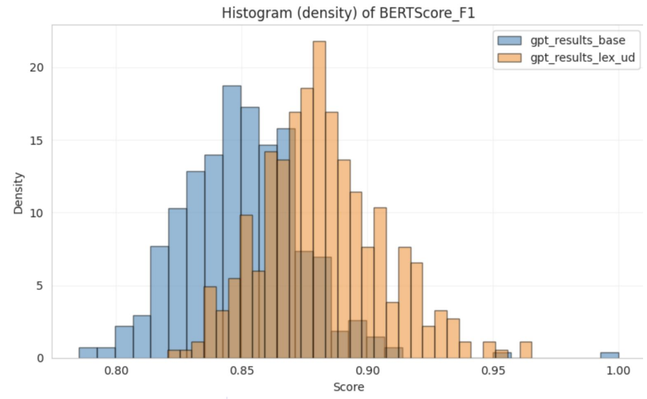


Figure 3. Comparison of BERT score distributions

- The BERT score has improved by 3.78%, with the baseline mean BERTScore F1 increasing from 0.8500 to 0.8821.

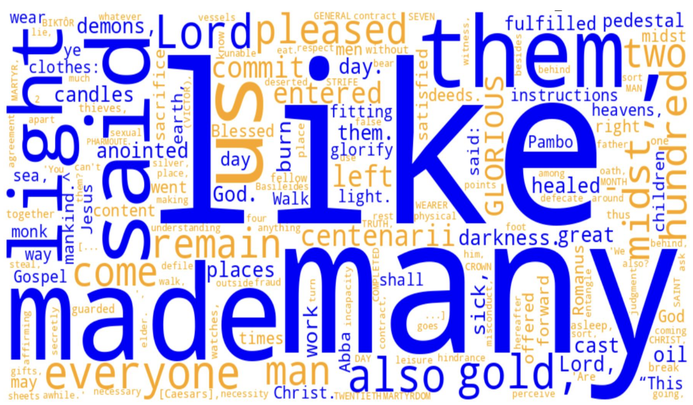


Figure 4. The frequent words in the best 10 and the worse 10 translation sentences from improved gpt model

Conclusion

Our findings show that GPT-4.1 Mini, combined with curated Coptic–English examples and linguistic cues, substantially improves low-resource MT performance. The enhanced prompting strategy leads to higher translation accuracy.

Limitations

- The Coptic–English dataset remains small, which restricts the diversity of grammatical structures and vocabulary available in prompts.
- Performance is tied to the capabilities and quirks of GPT-4.1 Mini; other models may respond differently to the same ICL strategies

Ethical Considerations

- Coptic is a historically and religiously significant language; translations must remain faithful to original meanings to avoid misrepresenting cultural heritage.
- All linguistic resources (Coptic Scriptorium corpora) were used under appropriate research-friendly licenses, ensuring ethical sourcing of data.

Acknowledgements

We would like to express our appreciation to Dr. Ali Arab for his valuable suggestions, as well as to the members of Corpling Lab—Abhishek Purushothama and Emma Thronson—and to the Massive Data Institute and the McCourt School of Public Policy for providing the opportunity to participate in the MDI Scholars program and for their ongoing support of this work.

References

- Oussama Akallouch and Khalid Fardousse. In-context learning for low- resource machine translation: A study on tarifit with large language models. Algorithms, 18(8), 2025.
- M. Saeed et al. Machine translation of coptic texts. In Proceedings of the ACL Workshop, 2024